THE UNIVERSITY of York

CENTRE FOR HEALTH ECONOMICS

Measuring income related inequality in health within general practices

Mark Dusheiko

Hugh Gravelle

CHE Technical Paper Series 22

CENTRE FOR HEALTH ECONOMICS TECHNICAL PAPER SERIES

The Centre for Health Economics has a well established Discussion Paper series which was originally conceived as a means of circulating ideas for discussion and debate amongst a wide readership that included health economists as well as those working within the NHS and pharmaceutical industry.

The introduction of a Technical Paper Series offers a further means by which the Centre's research can be disseminated. The Technical Paper Series publishes papers that are likely to be of specific interest to a relatively specialist audience, for example papers dealing with complex issues that assume high levels of prior knowledge, or those that make extensive use of sophisticated mathematical or statistical techniques.

The content and format of Technical Papers are entirely the responsibility of the author, and papers published in the this series are not subject to peer-review or editorial control, unlike those produced in the Discussion Paper series. Offers of further papers, and requests for information should be directed to Frances Sharp in the Publications Office, Centre for Health Economics, University of York.

August 2001

Measuring income related inequality in health within general practices

Mark Dusheiko* **Hugh Gravelle***

Centre for Health Economics, Technical Paper No. 22

02 August 2001

Abstract. The partial concentration index is a measure of income related inequality purged of the effects of variables which are correlated with health and income. We show that indirect standardisation is likely to underestimate the partial concentration index compared with two methods of direct standardisation. Estimates of partial concentration indices are constructed from a sample of individual patients from 60 English general practices, using direct and indirect standardisation. Although estimates are highly correlated, indirect standardisation based estimates are smaller than those based on direct standardisation. There appear to be no significant differences across practice inequality scores, nor do such differences appear to be predominately associated with differences in the distribution of income or the effect of income on health. Practice inequality scores are greater if practices receive deprivation payments or run diabetes clinics and smaller the number of practice staff. Practice characteristics appear to have little associated with the level of individual health or on the effect of income on health.

---

* National Primary Care Research and Development Centre, Centre for Health Economics, Alcuin Block D, University of York, York YO10 5DD; mad105@york.ac.uk; hg8@york.ac.uk. Funding from the Department of Health to NPCRDC is acknowledged. The views expressed are those of the authors and are not necessarily those of the DH. We are grateful to Stephen Campbell for access to the practice data from the QUASAR project and for comments from participants in the York Seminars in Health Econometrics and the Bergen Economics Department seminar, and the International Health Economics Association conference, York July 2001.

1. Introduction

One of the aims of government policy in the British National Health Service is to reduce differences between the health of rich and poor. The NHS is a publicly provided system funded almost entirely from taxation so that, to the extent that the NHS can affect health, reductions in inequality will require changes in the organisation and delivery of services, rather than changes in methods of financing to alter patient use.

Ninety per cent of patient contacts with the NHS are made via primary care. Patients are registered with a general practice and their general practitioners (GPs) act as gatekeepers, controlling non emergency access to the rest of the NHS. Most GPs are independent contractors, rather than employees. Even with recent attempts to introduce greater regulation, GPs have considerable freedom in the services they choose to provide to their patients and in the way they organise their practices to do so. Hence it is of interest to examine whether practice policies and organisation have any effect on inequalities in health between rich and poor. In this paper we report the results of investigations of income related inequality in health within practices and of differences in inequality between practices and the extent to which they are related to practice organisation.

There is *horizontal inequity* in the distribution of health across the population when people with different incomes but otherwise similar characteristics have different health. It is possible to test for the existence of income related inequity by including income as an explanatory variable in regressions of health on patients' characteristics. However, for policy analysis it is useful to have a measure of the amount of inequity, rather than just its existence.

The standard methodology (Wagstaff and van Doorslaer, 2000a) is to measure income related inequality by a concentration index of health on income. The *crude concentration index* C_{hy} may not be relevant for policy purposes since it is determined in part by factors such as age and sex, which affect health, and are correlated with income but are not amenable to policy. It is customary to attempt to remove the effects of such factors, variously referred to as unavoidable, policy irrelevant or standardising, to estimate the standardised health of the population at different income levels. The *partial concentration index* I_{hy} which measures the income related inequality in standardised health can then be calculated.

To date researchers have either used indirect standardisation (based on individual or grouped data) or direct standardisation. It has been suggested (Kakwani et al, 1997) that, since direct standardisation requires the grouping of data and hence loss of information, indirect standardisation is preferable. However, indirect standardisation amounts to the deliberate creation of omitted variable bias and leads to underestimates of income related inequality. It is also possible to apply to individual level data which is equivalent to the epidemiological procedure of direct (Gravelle, 2001). The first aim of the paper is to compare practice inequality scores calculated by direct and indirect standardisation using individual level data.

The second issue we address is whether there are differences in the degree of within practice inequality across our sample of general practices. The samples of patients from each practice are relatively small (mean 42) so that it is important to be able to distinguish genuine differences from those due to sampling variation.

The third set of questions considered are whether differences in practice level inequality are due to differences in the distribution of incomes within practices or to differences in the relationship between income and health which are linked to observable characteristics of practices which may be amenable to policy. We examine the relationship between practice level inequality scores and practice characteristics. The practice level approach loses information by grouping individuals in 60 practices. Hence we also look at the effect of practice characteristics on individual health.

2. Measuring income related inequality

Individual health is related to individual characteristics (age, sex, ethnicity, income, etc.) and possibly to the way in which their general practice delivers health care. Some of these characteristics (age, sex) may be felt to be justifiable or unavoidable sources of variation and others may not. Income and ethnicity are the most obvious candidates as unjustified sources of differences in health. In this paper we are concerned only with income related inequality and treat ethnicity unavoidable. The extent of unjustified differences in health with respect to income may vary across practices and may be related to aspects of practices that are amenable to policy. To examine these issues we need a summary measure of the inequity in the distribution of health: a measure of the amount of variation which is due to variations in income rather than to variations in standardising variables such as age and sex.

2.1 Concentration index

One commonly used inequality measure is the concentration index, a generalisation of the Gini coefficient. To measure income related inequality in health we plot the concentration curve $L(s)$, which graphs the cumulative proportion of health against the cumulative proportion of the population ranked by income (see Figure 1). If there is no income related inequality in health the poor will be, other things equal, as healthy as the rich and the poorest k% of the population will have k% of total population health. The concentration curve will then coincide with the 45^o line. If poorer people are less healthy than the rich, the poorest k% will have less than k% of the total health. Hence the concentration curve will lie below the 45^o line. If health is positively related to income the concentration curve will lie above the 45^o line. In Figure 1 the poor have a disproportionately small share of health and $L(s)$ lies below the 45^o line.

The concentration index C_{hy} summarises the total amount of income related inequality in health and is analogous to the Gini coefficient. It is defined as twice the area between the health concentration curve $L(s)$ and the diagonal:

$$C_{hy} = 1 - 2\int_0^1 L(s)ds$$

When the poor have a disproportionately small share of health the concentration curve $L(s)$ is below the diagonal and C_{hy} is positive.

But suppose that women are both poorer and healthier than men. Suppose also that the true effect of income on health is positive: rich women are healthier than poor and rich men are healthier than poor, but less healthy than rich women. The average health of the rich will be reduced relative to the average health of the poor by the fact that the proportion of healthier women is smaller in the rich. The simple bivariate association between income and health will be contaminated by the systematic variation of the sex ratio with income. Hence cumulating health by income, without allowing for the effects of other factors which may be associated with health and income, will give a misleading impression of the amount of inequality in health which is due to income. van Doorslaer and Koolman (2000) have emphasised the importance of controlling for other factors. They refer to the inequality arising from the association of health with these other factors as "unavoidable" inequality and suggest that it should be removed from the calculation of income related inequality.

2.2 The partial concentration index as an inequality measure

Consider the individual level health production function

$$h = b_0 + b_y y + b_z z + e \tag{1}$$

where h is a measure of health, y is income, and z is another variable affecting health. We assume that there are no other factors affecting health which are correlated with income or z. e is therefore an error uncorrelated with any of the factors affecting health. To simplify notation, z is interpreted as single variable, though the arguments below generalise in an obvious way (Gravelle, 2001).

The concentration index can be written as (Lambert, 1993)

$$C_{hy} = \frac{2}{m_h}\mathrm{Cov}(h, F(y)) = \frac{2}{m_h}\mathrm{Cov}(b_0 + b_y y + b_z z + e, F(y)) \tag{2}$$

where m_h is mean population health, and $F(y)$ is the distribution function for income. Since the covariance is additive and the error e is uncorrelated with y, the concentration index for health against income can be decomposed as

$$C_{hy} = \frac{2}{m_h}\left[b_y \mathrm{Cov}(y, F(y)) + b_z \mathrm{Cov}(z, F(y))\right] = \frac{b_y m_y}{m_h} C_{yy} + \frac{b_z m_z}{m_h} C_{zy} \tag{3}$$

where m_y, m_z are the population means of y, z. C_{yy} is the concentration index of income against income C_{yy}, otherwise known as the Gini coefficient. C_{zy} is the concentration indices of z against income.

The decomposition of the concentration index reveals the potential problem in using C_{hy} as a measure of income related inequality when other variables affect health ($b_z \neq 0$). If z is also correlated with income then the concentration index of z against income (C_{zy}) will be non zero and C_{hy} will reflect non income factors.

Only in cases in which there are no standardising variables, perhaps because we are examining income related inequality within in a highly specific population sub group defined by particular values of the standardising factors, will C_{hy} be a suitable summary measure of inequality. Otherwise, to obtain a policy relevant measure of income related inequality, the effects of the standardising variables must be removed from the overall concentration index. The obvious way to measure income related inequality when there are standardising variables is to deduct the terms involving them from C_{hy}, yielding the *partial concentration index*

$$I_{hy} \equiv \frac{b_y m_y}{m_h} C_{yy} \tag{4}$$

as an inequality measure. I_{hy} is just the first term in (3). Since it is the product of the elasticity of health with respect to income and the Gini, it reflects both the effect of income on health, holding all other factors constant, and the extent of variation in income across the population. Plotting the two components of I_{hy} in (elasticity, Gini) space gives potentially useful devices for cross section comparisons or for showing the time path of inequality (Gravelle and Sutton, 2001).

Note that if the standardising variable z was say age and had a negative effect on health ($b_z < 0$) and was positively correlated with income ($C_{zy} > 0$), then the partial concentration index I_{hy} will be greater than the concentration index of unstandardised health C_{hy}. Conversely if the rich are on average younger than the poor.

2.3 Direct standardisation

There are two methods of direct standardisation to estimate I_{hy}. They differ in their treatment of the residuals from the estimated health equation but are asymptoptically equivalent. The first method, which is most immediately comparable with indirect standardisation, is to

(a) estimate the health production function (1) as $h = b_0 + b_y y + b_z z + e$

(b) use the estimated coefficients b_z on the standardising variable to calculate health after removing any effect of income

$$h^b = b_0 + b_z z \tag{5}$$

(c) calculate the concentration index of h^b directly standardised health against income[1]

$$\hat{C}^b_{hy} = \frac{b_z \bar{z}}{\bar{h}^b} \hat{C}_{zy}, \tag{6}$$

(d) calculate the concentration index of unstandardised health against income $\hat{C}_{hy}$,

(e) multiply the concentration index of $\bar{h}^b$ by $\bar{h}^b / \bar{h}$ and subtract it from the concentration index of unstandardised health to get the directly standardised inequality index[2]

$$I^{D1}_{hy} = \hat{C}_{hy} - \frac{\bar{h}^b}{\bar{h}} \hat{C}^b_{hy} \tag{7}$$

Provided a consistent estimator of the production function is used (and OLS is consistent under the assumptions made so far)

$$\text{plim } I^{D1}_{hy} = C_{hy} - \text{plim } \frac{\bar{h}^b}{\bar{h}} \hat{C}^b_{hy} = C_{hy} - \frac{m_{h^b}}{m_h} \frac{b_z m_z}{m_{h^b}} C_{zy} = \frac{b_y m_y}{m_h} C_{yy} = I_{hy} \tag{8}$$

Thus the directly standardised inequality index I^{D1}_{hy} is a consistent estimator of the partial concentration index.

The second procedure for estimating a directly standardised concentration index is to estimate the health production function (1) to get the coefficient on income b_y, calculate the Gini coefficient $\hat{C}_{yy}$, mean health $\bar{h}$ and income $\bar{y}$ and so get

[1] Since the concentration index of any variable w against income can be written as $2\text{Cov}(w, F(y)) / \bar{w}$, OLS regression of $w[2S_{FF} / \bar{w}]$ on F (where S_{FF} is the sum of squared deviations of F from its sample mean), yields a regression coefficient $\text{Cov}(w[2S_{FF} / \bar{w}], F) / S_{FF} = 2\text{Cov}(w, F) / \bar{w} = \hat{C}_{wy}$. [See Kakwani et al, 1997]

[2] Equivalently we could also have used the result in the previous footnote with w defined as $h - h^D$ and run a single regression of $(h - h^D) S_{FF} / \bar{h}$ on F to get I^D_{hy} immediately, rather than running separate regressions to calculate the two concentration indices.

$$I_{hy}^{D2} = \frac{b_y \bar{y}}{\bar{h}} \hat{C}_{yy} \tag{9}$$

If the health production function is consistently estimated the second version of the direct standardisation procedure also produces a consistent estimate of the partial concentration index.

The two procedures for estimating the partial concentration index by direct standardisaion are asymptotically equivalent. They differ for finite samples only because the first procedure leaves the residuals from the estimated health production function in the estimated inequality index.[3] The first procedure can provide standard errors for the estimate of income related inequality if we use the convenient regression method mentioned in footnote 3. The second has the merit of giving a nice decomposition of the partial concentration index as a product of the Gini coefficient and the income elasticity of health. We present results from both direct standardisation methods.

2.4 Indirect standardisation

Indirect standardisation has been suggested as a simple and convenient method of removing the confounding effects of health affecting policy irrelevant variables which are correlated with income (Kakwani, et al, 1997; Wagstaff and van Doorslaer 2000b, 2000b; van Doorslaer and Koolman, 2000).

Indirect standardisation differs from the first method of direct standardisation only in the first step. Instead of estimating the health equation (1) including income, indirectly standardised health is estimated from a regression of health only on the standardizing variable z: $h = a_0 + a_z z + e^N$. The estimated concentration index for indirectly standardised health

$$h^N = a_0 + a_z z$$

(10)

is

[3] Using the additivity properties of the covariance: $I_{hy}^{D1} = \hat{C}_{hy} - \frac{\dddot{h}^b}{\bar{h}} \hat{C}_{hy}^b = \frac{b_y \bar{y}}{\bar{h}} \hat{C}_{yy} + \hat{C}_{ey} = I_{hy}^{D2} + \hat{C}_{ey}$

where $\hat{C}_{ey} = 2\text{Cov}(e, \hat{F}(Y))/\bar{h}$ and plim $\hat{C}_{ey} = 0$ (Gravelle, 2001).

$$\hat{C}_{hy}^{N} = \frac{2}{\overline{h}^{N}}\mathrm{Cov}\left(h^{N}, F(y)\right) = \frac{2a_z\overline{z}}{\overline{h}^{N}\overline{z}}\mathrm{Cov}\left(z, F(y)\right) = \frac{a_z\overline{z}}{\overline{h}^{N}}\hat{C}_{zy} \tag{11}$$

and the indirectly standardised inequality index is

$$I_{hy}^{N} = \hat{C}_{hy} - \hat{C}_{hy}^{N} = \hat{C}_{hy} - \frac{a_z\overline{z}}{\overline{h}^{N}}\hat{C}_{zy} \tag{12}$$

Since the indirect standardising regression equation omitted income:

$$\text{plim } a_z = b_z + b_{yz}b_y \tag{13}$$

where b_{yz} is the regression coefficient of income on the standardising variable. Hence the indirectly standardised inequality index has

$$\begin{aligned} \text{plim } I_{hy}^{N} &= \text{plim } \left(\hat{C}_{hy} - \hat{C}_{hy}^{N}\right) \\ &= \frac{b_y m_y}{m_h}C_{yy} + \frac{b_z m_z}{m_h}C_{zy} - \frac{(b_z + b_{yz}b_y)m_z}{m_h}C_{zy} = I_{hy} - \frac{b_{yz}b_y m_z}{m_h}C_{zy} \end{aligned} \tag{14}$$

and I_{hy}^{N} is not a consistent estimator of I_{hy}.

If conditional mean of z is linear in income

$$\text{plim } I_{hy}^{N} = \frac{b_y m_y}{m_h}C_{yy} - \frac{b_y b_{yz} m_z}{m_h}\left(\frac{m_y}{m_z}b_{zy}C_{yy}\right) = I_{hy}\left(1 - b_{yz}b_{zy}\right) = I_{hy}\left(1 - r_{yz}^{2}\right) \tag{15}$$

where r_{yz}^{2} is the squared correlation coefficient between income and the standardising variable affecting health. This result also holds for a vector of standardising variables $\mathbf{z}$ (Gravelle, 2001). The only difference is that the squared correlation coefficient between y and z is replaced by the coefficient of determination $R^2(y, \mathbf{z})$ from the multiple regression of income on the vector of standardising variables

The indirectly standardised inequality index tends to underestimate the partial concentration index I_{hy} irrespective of whether the standardising variable z has a positive or negative effect on health or whether the standardising variable is negatively or positively correlated with income.

A stark illustration of the potential problems with indirect standardisation is provided if the "standardising" variable z has no effect on health but is correlated with income.[4] Then there is no problem in using the concentration index C_{hy} as a measure of income related inequality since there are no confounding variables. In these circumstances $C_{hy} = I_{hy}$. But the correlation between income and z means that I_{hy}^N will tend to underestimate C_{hy} which is here an acceptable measure of income related health inequality. The example is extreme but it illustrates that indirect standardisation by regression on non income variables will tend to over correct for confounding. It removes both the direct effect of standardising variables and any indirect effect due to their correlation with income. But by removing the indirect influence of income via the standardising variables it also reduces the direct effect of income on health and hence tends to underestimate income related inequality.

If there are omitted variables in the estimated health equation and the indirect standardising regression then direct standardisation and indirect standardisation may over or underestimate the partial concentration index (Gravelle, 2001). However, the direct standardisation procedures seem preferable on the grounds that they are based on a regression which omits variables only because of lack of data, whereas indirect standardisation also omits a variable, income, by design. In section 4 we compare the estimates of income related inequality using indirect standardisation with those from the two methods of direct standardisation.

3. Data

3.1 GPAS

As part of a project on quality in general practice, General Practice Assessment Survey (GPAS) (*www.gpas.co.uk*) questionnaires were posted to 12104 patients in 60 practices. 4462 completed questionnaires were returned, an overall response rate of 37%. GPAS asks patients about use of their general practice, their views on its accessibility and quality of care. We used an augmented version of GPAS, which had additional socio-economic questions, including income and employment status and various aspects of health There was some item non-response, especially for the income question, and valid responses obtained

[4] One might wish to compare age and sex specific income related health inequalities across different areas and standardise on ethnicity which might not have a direct effect on health but might be

for 2508 individuals. The variables are summarised and defined in Table 1.

3.2 Health variable

GPAS includes the SF-6D instrument which asks about six dimensions of health: physical functioning, role limitation, social functioning, pain, mental health, vitality. Each dimension has between two and six levels. Scores for each state were assigned using results from Brazier, Roberts and Devrill (2001). They asked a representative sample of 836 members of the UK population to value subsets of health states using a standard gamble technique. The value was transformed to be between 1 and 0, where 1 was full health and 0 the worst health state. Their regression of the values on the characteristics of the states yielded a set of coefficients which can be used to construct scores for any health state.

3.3 Equivalised household income

Survey responders reported household income within income bands. We used interval regression to generate a continuous income variable. Interval regression utilises the fact that we know the boundaries of a household's income and other individual specific variables, enabling us to predict their household income within their stated band. Income bands were specified in logarithms. The interval regression used age, age squared, gender, ethnic origin, number of children, type of accommodation, marital status, occupation, car ownership and self reported physical health status from the SF-6D. Expected income conditional on the reported income band and characteristics of the individual was then predicted. The model allowed for the right censoring of the open upper income band (incomes greater than £40,000). Predicted income was equivalised according to the number and composition of members in the household (Cowell 1995). We do not here investigate possible selection bias associated with income arising from unit and item non response.

3.4 Statistical inference allowing for survey design

Because the survey was part of a wider study, its design was complex, with multistage sampling and stratification by subgroup. There were 100 Health Authorities (HAs) within 8 Regional Health Authorities (RHAs). Three RHAs were chosen and within each of them two HAs were randomly selected. Prior to selection HAs were stratified according to the proportion of their practices receiving additional capitation payments for patients from

---

correlated with income.

deprived electoral ward (Jarman payments) and the proportion of their practices receiving rural practice payments. From each of the 6 selected HAs, a random sample of 10 practices was taken, stratified to ensure that the sampled practices were nationally representative according to number of partners, proportion receiving deprivation payments, and proportion having training status. If a sampled practice refused to take part, then the next practice of that 'type' was selected. Within each practice, approximately 200 adult patients were randomly selected from the practice list. Practices varied in population list size, so that the probability of an individual being selected depended on the size of the practice.

Given that the probability of selecting a sampling unit differed across clusters (RHAs, HAs, practices), the sample is not self-weighting and sampling weights were required. The weights used were proportional to the inverse probability of a unit being selected.

The clustering of observations within sampling units implies observations are not independent. Stata 7 survey estimation commands were used to produce variance estimates based only on computations at the primary sampling unit level (PSU). This allows for correlation between sampling units within the PSUs. Variance estimates will be approximately unbiased or biased towards larger standard errors.

Survey estimators were used for the health production model, the regressions explaining variations in practice level inequality and for the imputation regression of income. We did not use survey weights when estimating individual practice inequality scores as within a practice all individuals had the same probability of being sampled. Estimates for population inequality indices using the convenient regression procedure did not use survey weights. Since the regressions to calculate standardised health and the predicted income regression were survey regressions, the population inequality estimates are implicitly weighted.

Models estimated using survey regressions to allow for clustering in the sample selection have the property that only d-1 constraints can be estimated at once, where d is the number of clusters (21 in our case)[5]. The test statistic is distributed $F(k,\ d\text{-}k{+}1)$, where k is the number of constraints. Hence we could not test whether the 60 practice income coefficients estimated by survey regression were equal. The Stata 7 manual also suggests that there may be a problem with estimating the survey model when the number of parameters exceeds the number of clusters, because the standard error calculation is asymptotic in the number of clusters (Stata, 2001). We therefore test whether there are significant differences across practices via standard OLS with probability weights and robust standard errors, which produces slightly smaller standard errors because it assumes independence of errors.

4. Results

4.1 Health equations

The health production function (1) is estimated as

$$h_i = b_0 + \sum_{p=2} b_{0p} D_{ip} + b_y y_i + \sum_{p=2} b_{yp} D_{ip} y_i + \sum_k b_{zk} z_{ik} + e_i \qquad (16)$$

where y_i is the log of equivalised household income, the z_{ik} are standardising variables and D_{ip} (p = 2, ..., 60) are practice dummies. We experimented with a number of functional forms to allow for non-linear income effects, including polynomials in income and in logs of income. The performance of the equation was not greatly sensitive to the specification of non-linearity. We settled on the log of equivalised income since interpretation of results is more intuitive. Column 1 of table 2 reports part of the results for the model with practice constant and income slope dummies and columns 2 and 3 show results without practice income slope dummies and with no practice dummies. The coefficients on the standardising variables are plausible: being female, living in local authority housing, single, being separated or divorced, having smoked for more than a year and not owning a car are all associated with worse health. There are also significant differences in health across ethnic groups. The coefficients on standardising variables are robust to the inclusion of practice constant and income slope dummies.

Columns 2 and 3 show that the overall effect of income on health was significantly positive. In model 1 with practice income slope dummies, the effect of income on health was

[5] The number of clusters is determined by the number of primary sampling units (PSUs) multiplied by the

significantly positive in 2 out of 60 practices (all at 5%). These results were obtained using robust standard errors, with a Survey estimator that allows for clustering we found that 32 practices were significant at 5% (see discussion of validity of estimation methods in section 3.4.).

The model (16) fitted is linear in the (log of) income. If the true health equation is non linear in the income variable it is possible that cross practice differences in the effect of practice income on health could be a reflection of non-linearity coupled with differences in the part of the overall income distribution from which practice populations are drawn. Thus if the underlying relationship with the income variable is concave, practices which have predominantly low income patients will have greater income slopes but smaller constants than practices with predominantly high income patients.

We plotted in a single diagram the predicted health of patients within practices against log income over the interquartile range of income in the practice, holding standardising variables constant. The plots did not seem to be approximations to any underlying stable relationship. We also regressed the practice income coefficient on mean practice log income. The income coefficient was negatively associated with mean practice log income but the t statistic was 1.16 and the R^2 for the regression was 0.02.

The equation estimated for indirect standardisation included the same standardising variables as the health production function and practice constant dummies but had no income variable and no practice income slope dummies. Results are in column 4 of Table 2. The coefficients on the standardising variables have a broadly similar sign and significance pattern compared to the health production function. The main differences are that living in local authority housing, being Bangladeshi, and not owning a car have larger and more significant detrimental effects on health. These are characteristics which are negatively correlated with income.

---

number of strata.

4.2 Overall inequality estimates

Table 3 shows estimated inequality scores over all patients. The first two columns show directly standardised inequality I_{hy}^{D1} estimated by the convenient bivariate regression of $2(h_i - h_i^b)S_{FF}/\bar{h}$ on income rank. S_{FF} is the average squared deviation of relative income ranks, where the individuals are ordered from smallest to highest income and the relative income rank of individual i is $F_i = (2i-1)/2n$. In column 1 h_i^b is estimated using from the health equation reported in column 2 of Table 2 which contains standardising variables, income, and practice constant dummies. Column 2 is derived from the model in column 3 of Table 2 where there are no practice constant dummies. Columns 3 and 4 report $I_{hy}^{D2} = (b_y \bar{y}/\bar{h})\hat{C}_{yy}$ where b_y is the coefficient on income estimated with and without practice constant dummies as reported in columns 2 and 3 of Table 2. Columns 5 and 6 give indirectly standardised inequality estimated by convenient bivariate regression of $2(h_i - h_i^N)S_{FF}/\bar{h}$ on income rank. Column 5 is derived from the model reported in column 4 of Table 2 where practice constant dummies are included. Column 5 is derived from the indirect standardising equation without practice constant dummies (not shown).

All methods indicate pro-rich income related health inequality and the estimates using the convenient regression approach have highly significant coefficients. The direct methods show more inequality than the indirect, supporting the argument in section 2.

4.3 Practice level inequality

Practice level inequality was estimated by using the results from the health production function and indirect standardising regressions. I_{hy}^{D1} was estimate for the p'th practice by the convenient regression of

$$2(h_i - h_i^b)S_{pFF}/\bar{h}_p = c_0 + \sum_{p=2} c_{0p}D_{ip} + c_F F_{i1} + \sum_{p=2} c_{Fp}D_{ip}F_{iP} + e_i \qquad (17)$$

where h_i^b is estimated from the health production function with practice constant and income slope dummies (model 1 of Table 2). $\bar{h}_p$ is mean health in practice p. I_{hy}^{D2} was calculated for the p'th practice as

$$\left[\left(b_y + b_{yp}\right)\bar{y}_p / \bar{h}_p\right]\hat{C}_{pyy} \qquad (18)$$

using the estimated income coefficients from model 1 of Table 2 where $\hat{C}_{pyy}$ is the Gini coefficient for log income in practic $p. I^N_{hy}$ was derived from estimated indirectly standardised health h^N_i from model 4 of Table 2 with practice constant dummies using the same convenient regression (17) as for I^{D1}_{hy} with h^N_i replacing h^b_i

We also calculated the three inequality estimates from the results of separate health production function and indirect standardising regressions for the 54 practices with at least 21 observations, thus allowing the slope coefficients on the standardising variables to differ across practices.

Table 4 gives summary statistics and Table 5 the correlation matrices for the practice inequality scores. Practice inequality scores are highly correlated, both across direct and indirect procedures and across the underlying regression equations. The practice level I^N_{hy} scores tend to show less inequality than the two directly standardised inequality scores. This provides some support for the suggestion in section 2 that indirect standardisation will tend to report less inequality than direct standardisation if there are no omitted variables correlated with income or the standardising variables. The slope coefficients from bivariate regressions of I^N_{hy} on I^{D1}_{hy} and I^{D2}_{hy} are also less than one, though not significantly so.

Tests of whether (c_F + c_{Fp}) from (17) were different from zero suggested that only four practices out of 60 had significant (at the 5% level) indirectly standardised inequality for I^N_{hy} and one of these had negative (pro-poor) inequality. A similar test on I^{D1}_{hy} indicated significant (pro-rich) inequality in 8 practices.

4.4 Testing for practice differences in inequality

To test whether practice inequality scores differ significantly across practices we estimated the convenient regressions (17) for I^{D1}_{hy} and I^N_{hy} imposing the restriction that the slope coefficients on relative income rank were identical: $c_{Fp} = 0$, p = 2,…,60. The F test of the null hypothesis that all slope coefficients for I_{hy}^{D1} were the same has F(59,2388) = 0.96, Prob 0.5724 and for I_{hy}^{N} has F(59, 2388) = 0.99, Prob 0.4883. The tests suggest that there are no significant differences in income related inequality across practices.

4.5 Practice characteristics and practice level inequality

Income related inequality in health depends on the extent to which income affects health and the degree of income inequality. The second method of calculating directly standardised inequality I_{hy}^{D2} has the advantage that it can be used to show the interaction of these two factors. Figures 2 and 3 are scatter plots in (income elasticity, practice income Gini) space. The lines are contours (rectangular hyperbolas) for I_{hy}^{D2} and in the positive quadrant higher contours indicate higher pro rich inequality. In the lower right quadrant practices on lower contours (not shown) have higher levels of pro-poor inequality. Figure 2 plots the I_{hy}^{D2} scores derived from the single health equation regression (16) with practice constant and income slope dummies. Figure 3 has I_{hy}^{D2} scores from the separate practice level regressions. There is no obvious pattern to the observations in either figure so that we cannot assign variations in practice inequality predominantly to variations in within practice income distribution or to differences in income elasticities.

To more fully investigate the cross practice variation in inequality scores we regressed the inequality scores on a number of practice characteristics. The results are shown in Table 6. The dependent variables I_{hy}^{D1} in columns 1 and 2 and I_{hy}^{D2} in columns 3 and 4 are derived from the health production function reported in column 1, Table 2. The dependent variable I_{hy}^{N} in columns 5 and 6 is derived from the model reported in column 4, Table 2. In columns 7 to 9 the dependent variable is the practice income slope coefficient from (17). We have also included the income Gini for each practice as an explanatory variable. There were incomplete observations on some practice characteristics so we report two versions of each regression, with the more comprehensive model fitted on a smaller number of observations.

The most noticeable feature of the results is that few practice characteristics are associated with inequality. The practice deprivation variable is positively associated with inequality in most specifications. When a larger set of explanatory variables is included, the presence of diabetes clinic is positively associated with inequality and the greater the number of professions allied to medicine at a practice the lower the level of inequality.

The practice income Gini is not significant in any of the specifications. This is somewhat surprising since income related inequality in health depends on the effect of income on health and the income distribution.

To see if practice characteristics were associated either with the level of individual health or with the effect of income on health we estimated the individual level health equation

$$h_i = b_0 + b_y y_i + \sum_k b_k z_{ik} + \sum_j d_j g_{ij} + \sum_j d_{jy} g_{ij} y_i + e_i \qquad (19)$$

where g_{ij} is practice characteristic j in the practice to which individual i belongs. The results are in Table 7. Columns 1 and 2 are for models using the smaller set of practice characteristics with and without interactions between income and the practice characteristics. Columns 3 and 4 use a larger set of practice characteristics, with and without income-practice characteristic interactions. Comparing Table 7 with Table 2 we see that replacing practice dummies with practice characteristics has little effect on the pattern of coefficients on standardising variables. The coefficient on income is positive and significant when there are no interactions of the practice characteristics with income but becomes insignificant when interaction effects are allowed for. Though few of the interaction terms are significant their overall effect seems to be to increase the effect of income on health and *ceteris paribus* to increase inequality. None of the practice characteristics have a significant effect on the level of individual health.

5. Conclusions

Our findings can be swiftly summarised

- theoretical arguments suggest that indirect standardisation is likely to underestimate inequality compared with two methods of direct standardisation
- the basic model of health estimated on individual level data had intuitively plausible

associations of health with income, marital status, sex, age, housing tenure, ethnicity, smoking behaviour and car ownership

- estimates of inequality in 60 practices by direct and indirect standardisation are highly correlated but those based on indirect standardisation are generally smaller than those based on direct standardisation
- there appear to be no significant differences across practice inequality scores
- practice inequality scores are not predominantly associated with differences in the distribution of income or the effect of income on health.
- practice inequality scores are greater if practices receive deprivation payments or run diabetes clinics and smaller the greater the number of allied professional staff.
- practice characteristics appear to have little effect on the level of individual health or on the effect of income on health.

This study is, as far as we know, the first investigation of income related inequality in health at practice level. The lack of findings of an effect of practices on inequality may be because practices can in fact do little to change the relationship between income and health for their patients. It may also reflect difficulties in measuring the effect because of the relatively small samples of patients from each practice. This was in part a reflection of the relatively low response rate. The version of the GPAS questionnaire that we used was much longer and more complex than the standard version now in extensive use in primary care. The standard version does not include an income question but it does include education level which is strongly correlated with income. We will be using it to continue our investigation of whether practice characteristics are associated with practice inequality.

We also will be experimenting with the specification of two crucial variables: health and income. Patients were also asked if they had limiting longstanding illness and we will investigate the effects of using their binary responses instead of their SF6 score. Running logit or probit regressions on income and the standardising variables will yield predicted latent health. Holding the standardising variables constant across individuals gives directly standardised latent health and we can then calculate directly standardised inequality by the convenient regression of directly standardised latent health on income rank (Gravelle, 2001).

We will also examine the effects on the directly standardised inequality scores I_{hy}^{D1} of replacing our interval regression estimates of income with dummy variables defined on household income group and family size.

Discussions of the contribution that primary care can make to reducing practice level inequality lack a firm theoretical foundation to guide empirical work. We require models of the impact of practices on health, of their interaction with the effect of income on health, and of the determination of the distribution of income across practice patients. It is straight forward to formulate plausible hypotheses about the relationship between practice characteristics and average patient health. For example, practice training status, which is conferred only on higher quality practices, might be expected to be associated with better health. List size might be expected to be associated with worse patient health, since more patients per GP may mean that patients receive less care. However, the effect of these and other practice characteristics variables on income related inequality is less obvious.

Income related inequality depends both on the relationship between income and the distribution of income within the practice. It is not clear *a priori* whether, for example we should expect wealthier patients to get relatively more care when the total amount of care provided falls, or relatively less care when the quality of the practice is higher. We also have to take account of possible associations of practice characteristics with income distribution within the practice population. Do certain types of characteristic lead to patients of a narrow or wide income range to select practices? Is there a link between GPs selection of patients and practice characteristics? We need to be able to answer these types of questions to determine the policy significance of any association between practice characteristics and inequality.

References

Brazier J., Roberts J., and Devrill M. (2001). "The estimation of a preference-based measure of health from the SF-36", *mimeo*, University of Sheffield.

Cowell F.A. and Jenkins, S. P. (2000). "Estimating welfare indices: household weights and *Institute for Social and Economic Research Working Papers*: No. 2000-23, University of Essex

Gravelle, H. (2001). "Measuring income related inequality in health and health care: the partial concentration index direct and indirect standardisation of?" *Centre for Health Economics, Technical Paper No 21*, University of York.

Gravelle, H. and Sutton, M. (2001). "Using the partial concentration index to examine trends in income related inequality in health: Scotland 19xx-xx", July, *mimeo*.

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

Lambert, P. (1993). *The Distribution and Redistribution of Income*, 2nd Edition, Manchester University Press, Manchester.

Stata Corporation, *Stata 7 Reference Manual*, Stata Press, College Station, Texas, 2001).

Wagstaff, A. and van Doorslaer, E. (2000a). "Equity in health care finance and delivery", in Culyer, A. J. and Newhouse, J. (eds.), *Handbook of Health Economics*, 1804-1862, Elsevier, Amsterdam.

Wagstaff, A. and van Doorslaer, E. (2000b). "Measuring and testing for inequity in the delivery of health care", *Journal of Human Resources*,

van Doorslaer, E. and Koolman, X. (2000). "Income related inequities in health in Europe: evidence from the European Community Household Panel", *Ecuity II Project Working Paper*, No. 1, March.

Table 1: Description and summary statistics of variables used in estimating the health production function.

Variable	Description	Mean	Std. Dev.	Min	Max
Individual characteristics					
HEALTH	SF-6D health state score	0.8117118	0.108176	0.386	0.99999
LnIncome	Log of equivalised household income[1]	23146	17570	973	99146
Age	Age (years)	48.49123	16.06743	17	98
White (ref)	White	0.9381978	0.240796		
Caribbean	Black – Caribbean	0.0103668	0.101309	0	1
African	Black – African	0.0083732	0.09114	0	1
Blackother	Black - Other	0.0051834	0.071823	0	1
Indian	Indian	0.0075758	0.086726	0	1
Pakistan	Pakistani	0.0091707	0.095342	0	1
Bangladesh	Bangladeshi	0.0035885	0.059809	0	1
Chinese	Chinese	0.0023923	0.048863	0	1
Other	Other	0.0151515	0.12218	0	1
Female	Female	0.5769537	0.494141	0	1
Married (ref)	Married/cohabiting	0.7272727	0.445362		
Single	Single	0.1248006	0.330559	0	1
Separated	Separated	0.0857257	0.280014	0	1
Widowed	Widowed	0.062201	0.241568	0	1
Owner (ref)	Owner-occupied	0.803429	0.397405		
Rentpublic	Rented local authority/ housing association	0.1248006	0.330559	0	1
Rentprivate	Rented from private landlord	0.0558214	0.229622	0	1
OtherAccom	Other arrangement	0.015949	0.125303	0	1
Car	Access to $\geq$ 1 car	0.8185805	0.385442	0	1
Neversmoke	Never smoked for $\geq$ 1 year	0.4936204	0.500059	0	1
Practice characteristics					
PMS	PMS contract practice[2]	0.410714	0.496416	0	1
Training	Training status	0.303571	0.463961	0	1
Deprivpay	Deprivation payments	0.553571	0.501621	0	1
GPs	Number of WTE GPs	2.807143	1.767547	1	8
LISTSIZE	List size per WTE GP	2181.618	517.6913	1111	3524
DIABETCLIN	Practice with diabetic clinic	0.535714	0.503236	0	1
STAFFYEARS	Average length of service in practice of staff (years)	7.576366	2.98075	1	16
FAMILYPLAN	Family planning clinic	0.413044	0.497821	0	1
PAMS	Number of attached PAMs[3]	0.413044	0.932764	0	4
Gini	Income Gini	0.044237	0.006552	0.034007	0.0609019

[1] Summary statistics for income are in levels not logs. The m in front of the income and age variable indicates they were entered in mean deviation form in the analysis.
[2] These practices had special contracts with their HA to provide additional services or were salaried.
[3] PAM: professions allied to medicine. EG chiropodists.

Table 2. Health production function and indirect standardisation estimates

	Health production function			Indirect standardisation
	1 Income/practice interactions	2 Practice effects	3 No practice effects	4 Practice effects
Age	-0.01114	-0.01181	-0.0121	-0.00996
	[3.47]**	[2.84]*	[3.06]**	[2.36]*
Age2	0.00021	0.00023	0.00024	0.00019
	[3.33]**	[2.82]*	[3.11]**	[2.31]*
Age3	-0.000002	-0.000002	-0.000002	0.00001
	[3.60]**	[3.15]**	[3.46]**	[2.64]*
Caribbean	-0.00122	0.00343	0.00101	0.00704
	[0.06]	[0.17]	[0.05]	[0.44]
African	0.0853	0.10028	0.07907	0.09084
	[2.81]**	[2.94]*	[2.85]*	[2.89]*
Black-other	0.04522	0.04331	0.04103	0.03814
	[2.11]*	[4.00]**	[2.58]*	[4.66]**
Indian	0.00583	0.00068	-0.00702	0.00016
	[0.27]	[0.03]	[0.26]	[0.01]
Pakistan	-0.13523	-0.12573	-0.10525	-0.12667
	[2.87]**	[2.64]*	[3.10]**	[3.01]**
Bangladesh	-0.08115	-0.08044	-0.07096	-0.09828
	[1.43]	[1.90]	[1.76]	[2.23]*
Chinese	0.03949	0.03688	0.02739	0.03458
	[2.23]*	[3.13]**	[3.22]**	[2.86]*
other	-0.04007	-0.0351	-0.04601	-0.04863
	[1.71]	[2.19]*	[3.83]**	[3.10]**
Female	-0.01336	-0.01442	-0.01552	-0.01625
	[2.66]**	[3.05]**	[4.12]**	[3.35]**
Single	-0.01776	-0.01765	-0.01866	-0.01491
	[1.97]*	[2.09]	[2.38]*	[1.63]
Separated	-0.02782	-0.02948	-0.03315	-0.03003
	[2.45]*	[3.02]**	[3.34]**	[2.97]*
Widowed	-0.00701	-0.01046	-0.01071	-0.00745
	[0.57]	[0.74]	[0.75]	[0.54]
Rentpublic	-0.04188	-0.04149	-0.03882	-0.05138
	[3.87]**	[4.69]**	[3.80]**	[6.10]**
Rentprivate	-0.00596	-0.0044	-0.00684	-0.0087
	[0.54]	[0.44]	[0.70]	[0.85]
OtherAccom	-0.01217	-0.01129	-0.0074	-0.0182
	[0.62]	[0.51]	[0.35]	[0.83]
Car	0.01039	0.01145	0.01812	0.01937
	[1.29]	[1.76]	[3.57]**	[2.72]*
Neversmoke	0.01254	0.01385	0.01442	0.01572
	[2.59]**	[2.33]*	[2.71]*	[2.61]*
LnIncome	-0.01747	0.01774	0.01752	
	[0.55]	[6.02]**	[9.24]**	
Constant	0.78981	0.80093	0.80822	0.79288
	[32.76]**	[222.95]**	[165.69]**	[186.46]**
Observations	2508	2508	2508	2508
R-squared	0.23	0.20	0.18	0.19

Absolute value of t statistics in brackets. * significant at 5%; ** significant at 1%

Table 3. Estimates of population income related health inequality

	Direct standardisation				Indirect standardisation	
	I_{hy}^{D1}		I_{hy}^{D2}		I_{hy}^{N}	
	Practice effects	No practice effects	Practice effects	No practice effects	Practice effects	No practice effects
	1	2	3	4	5	6
Concentration index	0.01013	0.0106	0.00992	0.0098	0.00674	0.00844
	(6.81)**	(7.06)**			(4.45)**	(5.56)**
Constant	0.03049	0.02943			-0.00325	-0.00471
	(32.23)**	(30.76)**			(3.36)**	(4.86)**
Observations	2508	2508	2508	2508	2508	2508
R-squared	0.02	0.02			0.01	0.01

Robust t statistics in parentheses. * significant at 5%; ** significant at 1%

Table 4. Practice inequality scores – descriptive statistics

	Single regression with practice constant and income slope dummies			Separate practice regressions		
	I_{hy}^{D1}	I_{hy}^{D2}	I_{hy}^{N}	I_{hy}^{D1}	I_{hy}^{D2}	I_{hy}^{N}
Mean	0.008691	0.008232	0.006181	0.010869	0.010586	0.005390
Median	0.008547	0.008823	0.006722	0.010234	0.009820	0.005819
Standard Deviation	0.011993	0.012182	0.011965	0.015398	0.015407	0.007691
Kurtosis	3.800597	2.634187	3.611283	1.500904	1.373865	2.390461
Skewness	0.328405	0.107443	0.238472	0.489262	0.540750	0.416061
Minimum	-0.027793	-0.027594	-0.028840	-0.021320	-0.022181	-0.012687
Maximum	0.053024	0.049707	0.049819	0.061121	0.059830	0.033170
Number	54	54	54	54	54	54

Table 5. Practice inequality scores – correlations

		Single regression with practice constant and income slope dummies			Separate practice regressions	
		I_{hy}^{D1}	I_{hy}^{D2}	I_{hy}^{N}	I_{hy}^{D1}	I_{hy}^{D2}
Single regression, practice constant and income slope dummies	I_{hy}^{D1}	1.0000				
	I_{hy}^{D2}	0.9685	1.0000			
	I_{hy}^{N}	0.9963	0.9677	1.0000		
Separate practice regressions	I_{hy}^{D1}	0.6468	0.6073	0.6316	1.0000	
	I_{hy}^{D2}	0.6506	0.6318	0.6397	0.9899	1.0000
	I_{hy}^{N}	0.7152	0.6612	0.7003	0.9386	0.9080

Table 6. Practice level inequality and practice characteristics

	Direct standardisation				Indirect standardisation		Income slope		
	1	2	3	4	5	6	7	8	9
Dependent variable	I_{hy}^{D1}	I_{hy}^{D1}	I_{hy}^{D2}	I_{hy}^{D2}	I_{hy}^{N}	I_{hy}^{N}	$b_y + b_{yp}$	$b_y + b_{yp}$	$b_y + b_{yp}$
PMS	-0.00025	-0.00287	-0.00057	-0.00336	-0.00051	-0.00295	-0.000385	-0.000127	-0.004048
	[0.08]	[0.98]	[0.18]	[1.11]	[0.16]	[0.99]	[0.07]	[0.02]	[0.81]
TRAINING	0.00136	0.00038	0.00067	0.00108	0.00036	0.00092	0.001353	0.001726	0.004355
	[0.36]	[0.12]	[0.15]	[0.31]	[0.08]	[0.29]	[0.25]	[0.32]	[0.68]
DEPRIVPAY	0.00538	0.00871	0.00753	0.01114	0.00699	0.01053	0.010355	0.009777	0.0126
	[1.79]	[1.85]	[2.44]*	[2.54]*	[2.42]*	[2.51]*	[2.24]*	[2.53]*	[2.31]*
GPS	-0.00108	-0.00095	-0.00043	-0.00053	-0.0004	-0.00056	-0.00214	-0.002229	-0.002519
	[0.73]	[0.74]	[0.29]	[0.40]	[0.29]	[0.44]	[0.80]	[0.85]	[0.93]
LISTSIZE	0.00001	0	0.00001	0.00001	0.00001	0.00001	0.000003	0.000004	0.000001
	[0.69]	[0.29]	[0.55]	[0.17]	[0.42]	[0.13]	[0.66]	[0.81]	[0.11]
DIABETCLIN	0.00458	0.00675	0.0032	0.00585	0.0039	0.00645	0.008349	0.008292	0.01453
	[1.43]	[2.23]*	[1.04]	[2.09]	[1.31]	[2.31]*	[1.33]	[1.31]	[2.53]*
HA_2	0.00552	-0.00282	0.00815	0.00804	0.00944	0.00878	0.009797	0.011024	0.010828
	[1.42]	[0.68]	[1.97]	[1.67]	[2.45]*	[1.95]	[1.34]	[1.88]	[1.49]
HA_3	0.00732	0.00647	0.00952	0.01094	0.0102	0.01179	0.012195	0.011257	0.011595
	[1.33]	[1.72]	[1.76]	[3.12]**	[1.92]	[3.41]**	[1.28]	[1.34]	[1.89]
HA_4	0.01731	0.00942	0.02052	0.02414	0.02147	0.02496	0.030597	0.030277	0.034224
	[3.11]**	[2.36]*	[3.82]**	[4.12]**	[4.16]**	[4.38]**	[3.19]**	[3.18]**	[3.57]**
HA_5	0.01088	0.02122	0.01132	0.01545	0.01262	0.0169	0.020772	0.019922	0.027293
	[2.24]*	[3.36]**	[2.49]*	[2.84]*	[2.73]*	[3.09]**	[2.51]*	[3.15]**	[2.72]*
HA_6	0.0137	0.01469	0.01599	0.01467	0.01653	0.01513	0.023451	0.023811	0.022468
	[2.55]*	[2.34]*	[2.89]*	[2.53]*	[2.97]*	[2.69]*	[2.57]*	[2.42]*	[2.92]*
Gini	0.36641	0.01364	0.41609	0.64032	0.39096	0.60714	0.184521		0.471602
	[0.75]	[2.63]*	[0.85]	[1.49]	[0.82]	[1.42]	[0.25]		[0.76]
FAMILYPLAN		-0.00302		-0.00331		-0.00295			-0.006138
		[4.14]**		[0.97]		[0.91]			[0.74]
PAMS		0.00097		-0.00302		-0.0028			-0.006503
		[0.72]		[4.25]**		[3.94]**			[3.92]**
STAFFYEARS		0.55683		0.0009		0.0009			
		[1.27]		[0.72]		[0.73]			
Constant	-0.02362	-0.03938	-0.02814	-0.04381	-0.02957	-0.04545	-0.019706	-0.011923	-0.025325
	[1.14]	[1.64]	[1.34]	[1.79]	[1.41]	[1.87]	[0.60]	[0.92]	[0.79]
Observations	56	46	56	46	56	46	56	56	46
R-squared	0.23	0.31	0.28	0.38	0.29	0.38	0.22	0.21	0.28

Table 7. Individual health production function with practice characteristics

	Subset of practice characteristics		Full set of practice characteristics	
	No inc interaction	Inc interaction	No income	Income interaction
	1	2	3	4
	Main variables	Main variables interacted	All variables	All variables
Age	-0.01247	-0.012	-0.01277	-0.01265
	[2.96]*	[2.84]*	[2.70]*	[2.69]*
Age2	0.000242	0.000234	0.000246	0.000245
	[3.01]**	[2.89]*	[2.70]*	[2.71]*
Age3	-2E-06	-2E-06	-2E-06	-2E-06
	[3.37]**	[3.25]**	[2.99]*	[3.04]*
Female	-0.01628	-0.01657	-0.01993	-0.01964
	[3.47]**	[3.56]**	[3.77]**	[3.82]**
Caribbean	0.007852	0.007924	0.048099	0.047874
	[0.40]	[0.37]	[2.11]	[2.25]*
African	0.093531	0.092571	0.084686	0.073379
	[2.85]*	[2.76]*	[2.31]*	[3.25]**
Black-other	0.043668	0.042126	0.079325	0.077296
	[3.73]**	[3.21]**	[13.38]**	[7.24]**
Indian	-0.0034	-0.00469	-0.03028	-0.02082
	[0.13]	[0.19]	[1.32]	[0.88]
Pakistan	-0.0935	-0.09144	-0.08364	-0.0875
	[2.63]*	[2.48]*	[1.12]	[1.19]
Bangladesh	-0.05937	-0.05565	-0.06747	-0.07384
	[1.49]	[1.37]	[1.38]	[1.29]
Chinese	0.033858	0.035139	0.03311	0.033705
	[2.48]*	[2.70]*	[2.95]*	[3.04]*
Other	-0.03346	-0.03136	-0.02736	-0.0306
	[1.96]	[1.76]	[1.89]	[1.63]
LnIncome	0.018629	0.006968	0.016184	-0.01034
	[6.40]**	[0.60]	[5.30]**	[0.34]
Single	-0.02019	-0.02007	-0.02813	-0.02836
	[2.53]*	[2.53]*	[2.21]*	[2.30]*
Separated	-0.0341	-0.03447	-0.02645	-0.02604
	[3.50]**	[3.46]**	[2.30]*	[2.24]*
Widowed	-0.01204	-0.01075	0.010644	0.012579
	[0.82]	[0.75]	[0.74]	[0.96]
Rentpublic	-0.03897	-0.0388	-0.04765	-0.04853
	[4.28]**	[4.10]**	[4.97]**	[5.27]**
Rentprivate	-0.00338	-0.00191	-0.01089	-0.00999
	[0.34]	[0.19]	[0.97]	[0.87]
OtherAccom	-0.0055	-0.00569	-0.00857	-0.00783
	[0.25]	[0.26]	[0.33]	[0.31]
Car	0.015441	0.015265	0.019893	0.020404
	[2.65]*	[2.59]*	[3.44]**	[3.55]**
Neversmoke	0.014598	0.014573	0.014127	0.014682
	[2.34]*	[2.33]*	[1.73]	[1.71]
PMS	0.000513	0.001063	-0.0005	0.001298
	[0.09]	[0.19]	[0.17]	[0.43]
TRAINING	-0.00559	-0.00679	0.001924	0.000153
	[0.82]	[0.93]	[0.36]	[0.03]
DEPRIVPAY	0.009674	0.00854	0.007039	0.005533
	[1.75]	[1.52]	[1.66]	[1.45]
GPS	0.000526	0.027952	-0.0009	0.052384
	[0.27]	[0.93]	[0.71]	[2.27]*
LISTSIZE	0.000008	-2.1E-05	-0.00001	0.000039

	[0.99]	[0.47]	[1.87]	[0.36]
DIABETCLIN	-0.00454	-0.00598	-0.00446	-0.00323
	[1.18]	[1.44]	[0.90]	[0.61]
ha_2	-0.01343	-0.01377	-0.0114	-0.00897
	[1.36]	[1.25]	[1.47]	[1.67]
ha_3	0.020033	0.021624	0.008186	0.007819
	[3.48]**	[3.85]**	[1.51]	[1.29]
ha_4	0.001676	-0.00145	-0.00344	-0.00552
	[0.30]	[0.27]	[0.73]	[1.01]
ha_5	0.015668	0.015522	0.004594	0.004119
	[2.50]*	[2.18]*	[0.74]	[0.60]
ha_6	-0.0016	-0.00107	-0.012	-0.01213
	[0.23]	[0.15]	[1.70]	[1.71]
(PMS)*LnIncome		0.002165		-0.00295
		[0.46]		[0.68]
(TRAINING)*LnIncome		0.004442		0.006537
		[0.79]		[1.30]
(DEPRIVPAY)*LnIncome		0.006115		0.017492
		[1.54]		[2.93]*
(DIABETCLIN)*LnIncome		0.005694		0.013447
		[0.90]		[2.42]*
Inc_GPS		-0.00279		-0.00542
		[0.92]		[2.32]*
Inc_LISWTEGP		0.000003		-5E-06
		[0.68]		[0.44]
ASTAFFYEARS			-0.001	-0.03518
			[1.54]	[2.15]
PAMS			0.006153	0.003022
			[1.02]	[0.45]
FAMILYPLAN			-0.00195	-0.00522
			[0.66]	[1.51]
ATTACH			-0.00028	-0.00638
			[0.91]	[0.54]
(PAMS)*LnIncome				-0.01602
				[2.16]
(FAMILYPLAN)*LnIncome				-0.00062
				[0.08]
Inc_STAFFYEARS				0.003484
				[2.19]*
Inc_ATTACH				0.000633
				[0.52]
(ha_2)*LnIncome		-0.00039		0.001363
		[0.08]		[0.22]
(ha_3)*LnIncome		-0.0066		-0.00042
		[0.80]		[0.05]
(ha_4)*LnIncome		0.020002		0.043037
		[2.02]		[4.20]**
(ha_5)*LnIncome		0.007886		0.020338
		[1.14]		[2.04]
(ha_6)*LnIncome		0.002169		-0.00059
		[0.40]		[0.08]
Constant	0.781633	0.779272	0.836636	0.835454
	[31.23]**	[32.06]**	[69.10]**	[53.56]**
Observations	2438	2438	1998	1998
R-squared	0.19	0.19	0.18	0.19

Absolute value of t statistics in brackets
* significant at 5%; ** significant at 1%

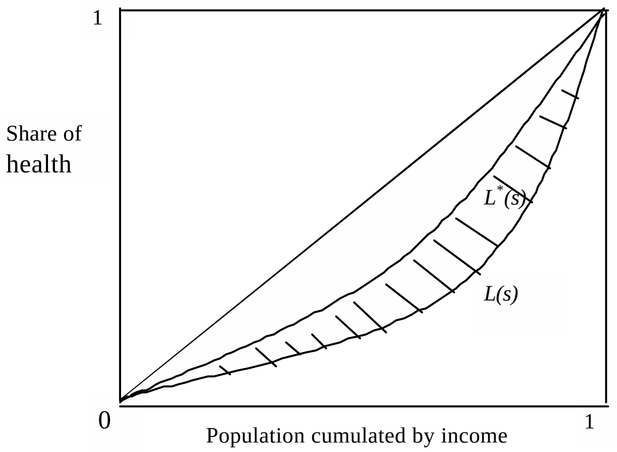

Figure 1. Concentration curves for raw ($L(s)$) and standardised ($L^*(s)$) health. Partial concentration index I_{hy} is twice shaded area and indicates pro-rich inequality if $L^*(s)$ lies above $L(s)$.

Figure 2. Decomposition of directly standardise inequality I_{hy}^{D2} as product of income elasticity and Gini coefficient – single health equation with practice constant and income slope dummies

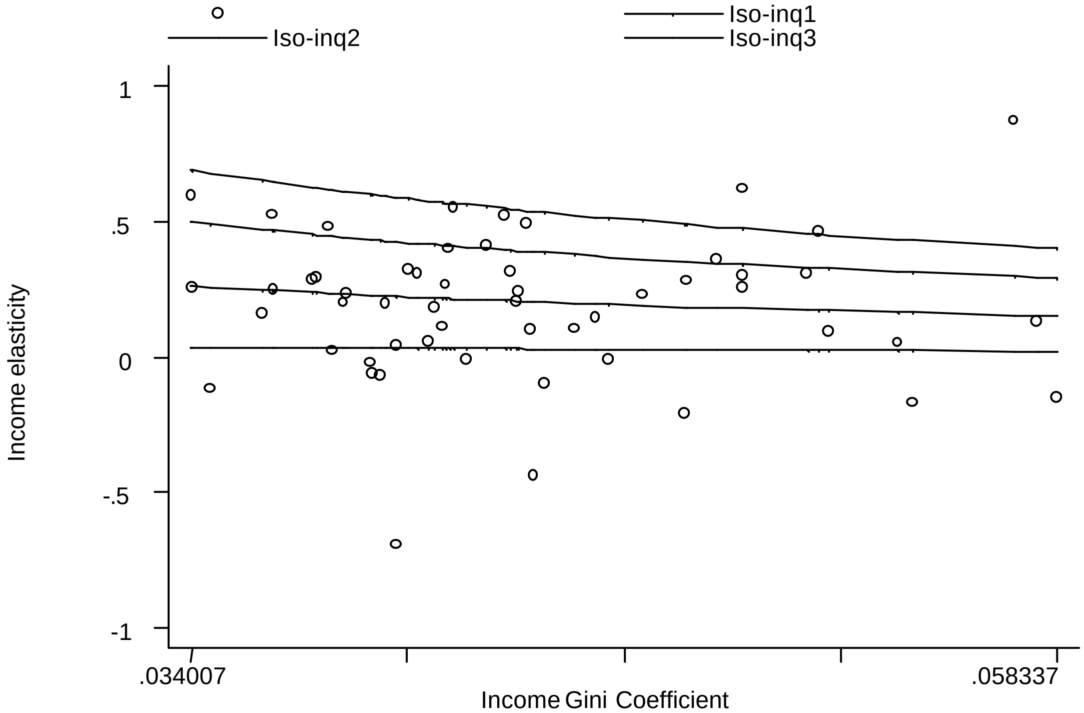

Figure3. Decomposition of directly standardise inequality I_{hy}^{D2} as product of income elasticity and Gini coefficient – separate health equations for each practice.