HEDG Health, Econometrics and Data Group

THE UNIVERSITY of York

HEDG Working Paper 08/06

Health care deprivation profiles in the measurement of inequality and inequity: an application to GP fundholding in the English NHS

Mauro Laudicella
Richard Cookson
Andrew M Jones
Nigel Rice

May 2008

ISSN 1751-1976

http://www.york.ac.uk/res/herc/research/hedg/wp.htm

Health care deprivation profiles in the measurement of inequality and inequity: an application to GP fundholding in the English NHS

Mauro Laudicella[α], Richard Cookson[β], Andrew M. Jones[γ], Nigel Rice[δ]
University of York

Revised 06 May 08

Abstract: This paper proposes a new approach to the measurement of inequality and inequity in the delivery of health care based on contributions from the literature on poverty and deprivation. This approach has some appealing characteristics: 1) inequity is additively decomposable by population subgroups; 2) the approach does not rely on socio-economic ranks; 3) it provides a graphical representation of the distribution of inequity; 4) it offers a range of indices consistent with dominance. An empirical application is provided investigating the effect of the GP fundholding reform on equity in English NHS. The results show that the most equitable GP practices self-selected into the scheme in 1991; evidence of an inequity-reducing treatment effect as well as a self-selection effect are found in 1992 and 1993; the self-selection process reduces and no evidence of a treatment effect is present thereafter.

Keywords: inequality, inequity, health care, poverty, deprivation, dominance, GP fundholding.

Acknowledgements: Dataset assembly for this project was funded by the ESRC Public Services Programme Phase II Award RES-166-25-0055. Hospital episode statistics were provided by the NHS Health and Social Care Information Centre, on license from the Department of Health. Census data were provided by the MIMAS service at the University of Manchester. Census data for 1991 is licensed by the Office of Population Censuses and Surveys, and census data for 2001 is licensed by the Office of National Statistics.

[α] Centre for Health Economics, University of York. Correspondence: Centre for Health Economics, Alcuin "A" Block, University of York, York, YO10 5DD, UK. Tel: +44 1904 321 423. E-mail: ml529@york.ac.uk.
[β] Department of Social Policy and Social Work, University of York.
[γ] Department of Economics and Related Studies, University of York.
[δ] Centre for Health Economics, University of York.

1. Introduction

Horizontal equity in health care is commonly defined as equal treatment for equal need. A large part of the empirical literature in this field focuses only on violations of the principle of horizontal equity that are related to income or socio-economic status and measures inequity using standardized concentration curves and indices (Wagstaff and van Doorslaer, 2000a). These studies fail in providing a measure of inequity associated with variables that do not vary with income or socio-economic status since their tools of analysis do not capture such aspects. Moreover, the magnitude as well as the sign (pro-poor or pro-rich) of inequity is affected by the degree of control for need within the morbidity categories considered. This is particularly evident whenever heterogeneity in unobservable need variables occurs across rich and poor (van Doorslaer and Wagstaff, 1992; Bago d'Uva et al., 2007).

There is evidence that in OECD countries many other variables drive inequity beside income, especially in countries where health care is provided mainly free of charge. Income does not explain the probability of visiting a general practitioner across most of the OECD countries, while it is associated with inequity in visits to a specialist in almost all countries (van Doorslaer et al., 2006). However, income is not always the most important factor in explaining subsequent visits to a specialist (van Doorslaer et al., 2004). Moreover, some authors find evidence that other variables such as ethnicity, education and employment status drive inequity in both primary and secondary care in the English NHS (Morris et al., 2005).

Many of the empirical studies on inequity in health care adopt a regression-based approach (Gravelle et al., 2006). This strategy includes characteristics that are not related to income in the analysis of inequity. Most of these studies model the probability of receiving health care as a function of a set of need variables and a set of other individual characteristics and policy variables. Inequity is found whenever the probability of receiving a treatment depends on variables that do not belong to the set of the need variables. This approach is effective in identifying the presence of inequity as well as the factors driving inequity, but has considerable limitations in quantifying inequity and in making inequity comparisons. For instance, suppose we find a significant difference in the regression coefficient that describes the admission rate of people in region A with respect to people in region B, after controlling for need. Such a result is sufficient to state that the principle of equity has been violated since, on average, people having equal need are not treated equally in the two regions. However, the difference in the admission rate between A and B is a measure of inequity only if we define equity as average equal treatment for average equal need. The regression approach restricts the inequity analysis: the measure of inequity is consistent with social preferences that care only for the group averages captured by the regression analysis. Hence, we lose information on the distribution of health care within groups A and B, which the parameters of the regression are not able to represent. This restriction is also present in empirical work that investigates the determinants of inequity using the Oaxaca-type (1973) decomposition technique. Generally, such work decomposes the difference between the group means of the outcome variable (e.g. health care utilization of poor and rich groups) into differences in the means of the determinants (e.g. need, income, regional location) and differences in the effects of these determinants (O'Donnell et al., 2006). Therefore, inequity is measured as differences in the group means and it is implicitly assumed that giving the same weight to each observation is a desirable way of aggregating the health care distribution into a single index.[*]

The methodological contribution of this paper is to apply some developments in the literature on poverty and deprivation to the analysis of inequity in health care (Jenkins and Lambert, 1997; Shorrocks, 1998). This produces a new approach to the measurement of inequity that enlarges the scope of the analysis and offers advantages with respect to the previous empirical literature. It provides inequity ranks that are consistent with a wide range of social preferences for equity and offers a useful graphical representation of important aspects of the distribution of inequity in the population. Further, a class of indices consistent with dominance is introduced in order to aggregate the distribution of inequity into a single value.

[*] Some empirical studies on discrimination in the labour market adopt a quantile regression technique and then apply the Oaxaca-type decomposition at different points of the wage distribution, in order to incorporate distributive aspects in the inequity analysis. Del Rio et al. (2006) highlight a shortcoming of this approach, i.e. re-rankings of people across quantiles in the counterfactual distribution.

The new approach allows a decomposition of inequity into population subgroups. Inequity measured in the population as a whole can be decomposed into a weighted sum of inequity experienced in any partition of the population. This property permits inequity rankings across specific demographic groups within the population (e.g. education, or region, or ethnic groups) and identification of those which are particularly suffering inequity. Also, it is possible to analyse variation in inequity associated with discrete variations in any characteristics of interest in the groups considered (e.g. level of education, or age, or travel distance from hospital, or socio-economic status). Thus, the additive decomposability property enlarges substantially the scope of the inequity analysis with respect to previous studies based on concentration curves and indices.

Moreover, this property allows us to test for the robustness of the inequity ranks when need is not perfectly observable in the population. We can decompose the measure of inequity for the total population in subgroups that are relatively homogeneous with respect to their unobservable need characteristics. Then, we can test if the inequity ranks for the total population are consistent in these subgroups. For example, if we are interested in ranking inequity across regions, we can test the robustness of our ranks by selecting the poorest deciles of the population and checking whether our ranks for the total population hold in that sub-sample. Since income is likely to be correlated with need, it can be used as a proxy when the latter are unobservable.

Finally, the measure of inequity does not rely on the socio-economic status (SES) of our observations. We are able to obtain a measure of inequity whenever such information is missing or inaccurate (e.g. when SES is a discrete or categorical variable). Also, we are able to analyse sources of inequity which are unrelated or weakly related to income, such as regional disparities in the supply of care, education, ethnicity, age, and gender.

The second contribution of this paper is an empirical application of the new method to the analysis of the GP fundholding reform in the English NHS. We measure the effect of the reform on equity in the utilization of secondary care for elective hip replacements across the electoral wards of England. GPs act as gatekeepers to secondary care in the English NHS and traditionally bore no financial cost in referring patients to secondary care. In allowing GP practices to hold budgets for hospital care, and to re-invest any savings within the practice, the fundholding reform aimed to give a financial incentive to manage access to secondary care to those GPs who joined the scheme. The consequences of this reform for overall access to secondary care have been intensively studied. There is evidence that GP fundholding practices reduced overall elective admission rates (Dusheiko et al., 2006) while at the same time reducing the average waiting time for secondary care admissions (Propper et al. 2002). However, there is little evidence on equity. We provide measures of the impact of the GP fundholding reform on equity that are consistent with a wide range of social preferences for equity. We use elective hip replacements as an indicator of health care inequity and examine inequity in the distribution of elective hip replacements within the most socio-economically deprived quartiles of electoral wards of England. Hip replacement is performed routinely and is a common, effective and low risk surgical procedure. It is often considered one of the most important examples of health care inequality in small area administrative studies, since it shows a clear socio-economic gradient after controlling for need (Chaturvedi and Ben-Shlomo 1995; Dixon et al., 2004; Cookson, et al., 2007).

We find evidence that the most equitable GP practices self-selected into the scheme during the first three "waves" of fundholders (FHs) from 1991 to 1993. We also find evidence of a significant inequity-reducing treatment effect in 1992 and 1993 as well as a self-selection effect: the hypothesis of no treatment effect on equity is rejected in 1992 and 1993, after controlling for self-selection. However, we cannot attribute either self-selection or treatment effects to individual year-by-year waves of FHs. Since effects are observed at ward level, it is not possible to identify which among multiple vintages of fundholding practices serving the same ward contributed which portion of the effect. No evidence of a treatment effect is found in 1994 to 1996, and a reduction in the self-selection process is also found as these later waves of GP fundholding were launched.

2. Health care deprivation profiles as a measure of inequity

This section provides the theoretical framework underlying our approach to the measurement of inequity in the delivery of health care. Its foundations rely on previous work by Jenkins (1997) and Shorrocks (1998) in the literature on poverty and deprivation.

Consider a vector of observations on health care utilization: $h = (h_1, h_2, \ldots, h_n)$. The elements of the vector h might be number of visits to a specialist, or measure of a specific procedure such as number of elective hip replacements, within a given time period.

Define the vector $h^* = (h^*_1, h^*_2, \ldots, h^*_n)$ that for each observation in "h" gives its need-expected level of utilization as a function of a set of need indicators "n" (e.g. age, sex, self assessed health status, presence of limiting long term illness, etc.):

$$h^* = f(n) \tag{1}$$

This approach does not depend on a specific "need adjustment". For instance, h^* can be exogenously set by a policy maker as the minimum amount of health care resources that individuals or communities should receive given their need. In this paper, we follow the tradition of the literature on inequity in health care and define h^* as a function of a set of need indicators and observed health care utilization:

$$h^* = f(n; h) \tag{2}$$

Equation (1) and (2) can be specified at either the individual or small area level. When the units of analysis are individuals, the vector h^* in Equation (2) could be computed using the indirect standardization approach described in Wagstaff and van Doorslaer (2000a). Here, we describe the computation of h^* when the units of analysis are small area observations. This should facilitate the illustration of the empirical application in section 4, in which we use electoral wards as unit of analysis. However, the methods that we describe in Section 2 and 3 also apply at the individual level.

$$h_i^* = \sum_{j=1}^{J} p_{ji} h_j \tag{3}$$

$$h_j = \frac{\sum_{i=1}^{n} h_{ji}}{\sum_{i=1}^{n} p_{ji}} \tag{4}$$

Where p_{ji} is the number of people in area i that belong to need class j (e.g. number of males aged 45-54 suffering from arthritis in electoral ward i); h_{ji} is the observed health care utilization in area i for need class j. The value h_i^* can be interpreted as the amount of health care utilization that we would expect to observe in the area i given its population, if each of its need groups had the same utilization rate as the population as a whole (i.e. the national utilization rate h_j). The computation of a need-expected distribution similar to h^* is a well establish approach in many empirical studies on the measurement of inequity (Wagstaff and van Doorslaer, 2000b). Also, in the epidemiology literature h^* defines the number of expected events in the computation of the indirectly standardized incidence rate (where j usually are based on age and sex variables; Breslow and Day, 1987; Clayton and Hills, 1993).

Now, define the vector of health care gaps (HCG), $x = (x_1, x_2, \ldots, x_n)$, as the vector of distances between need-expected and actual utilization, when the gap is positive, and zero otherwise:

$$x_i = \max\left(h_i^* - h_i; 0\right) \tag{5}$$

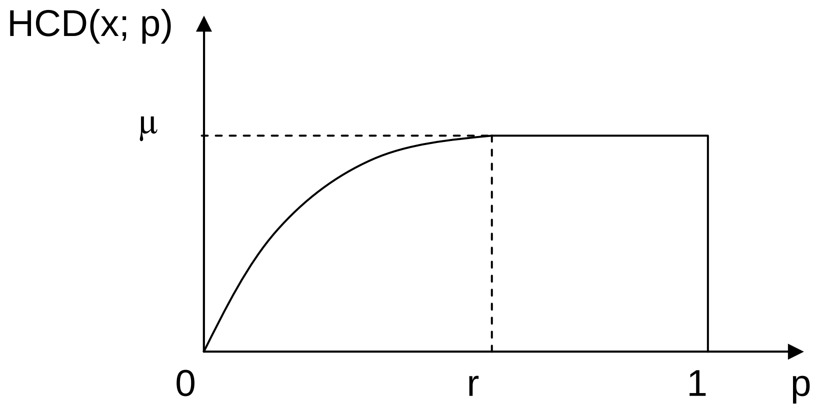
Define the health care deprivation profiles[†] (HCD) as:

$$HCD(x;p) = \int_{F^{-1}(1-p)}^{\xi} x dF(x) \quad \forall \mathrm{p} \in [0,1] \tag{6}$$

[†] The HCD profiles can be considered a special case of the deprivation profiles illustrated in Shorrocks (1998).

Where *F(x)* is the cumulative distribution function of *x*. For each $p \in [0, 1]$ the HCD profile is the cumulative sum of the first 100*p per cent of HCGs ranked from the largest (ξ) to the smallest (0).

Figure 1. The Heath Care Deprivation Profile.



The HCD profile *HCD(x; p)* for *x* is described by a concave curve that rises continuously from the largest to the smallest HCG; the curve becomes flat in the interval [r, 1] where the HCGs equal zero. The HCD profile provides a graphical representation of useful information on incidence, intensity and inequality in the distribution of HCGs across the areas. The length of the interval [0, r] represents the proportion of areas receiving less health care than need expected, i.e. the incidence of areas having positive HCGs:

$$r = F(x = 0) \tag{7}$$

The height of the curve, μ, is equal to the average HCG in the *whole* population of areas, i.e. the intensity of the gap between need expected and observed health care in the population of areas:

$$\mu = \int_0^{\xi} xF(x) \tag{8}$$

The degree of concavity of the HCD is determined by the rate at which HCGs decrease in the areas. This offers a representation of inequality in the distribution of the HCGs: if the HCGs are constant across all those areas receiving less health care then needed, then the curve become a straight line; the more the HCGs are unequally distributed across the areas, the greater is degree of concavity of the curve. When all the areas are not below their need-expected value h_i*, then all HCGs equal zero and the curve lies on the x-axis.

If we define deprivation in health care utilization as receiving less health care than needed, then the HCD profiles offer a graphical representation of the distribution of deprivation in health care across the areas. Moreover, the HCD profiles provide a measure of horizontal inequity in health care for all the areas below the frontier of need expected utilization h^*. We illustrate this argument through three propositions.

> Proposition 1. *No areas are deprived in health care if and only if the observed health care utilization equals the need expected utilization in all areas:*

$$x_i \equiv \max\left(h_i^* - h_i;0\right) = 0 \text{ },\forall i \in N \quad \Leftrightarrow \quad h_i = h_i^*,\forall i \in N \text{ .} \tag{9}$$

The Proof of proposition 1 is given in Appendix A.

Proposition 1 holds for both "health care deprived" and "health care non-deprived areas", i.e. those areas below and above their need expected health care utilization h^*. The HCG vector rules out the possibility of having all the areas above or below the need expected frontier of health care utilization (i.e. $h_i > h_i^* \;\; \forall i$, or $h_i < h_i^* \;\; \forall i$). If some area increases its health care consumption, this will result either in subtracting resources from other areas and/or in raising the need expected health care utilization for all. Intuitively, Proposition 1 is a direct consequence of the definition of the vector h* as a function of the average level of health care utilization observed in the total population (Equations (3) and (4)). Thus, the need expected frontier of health utilization is endogenously determined.

Proposition 2. *No areas are deprived in health care if and only if the share of health care utilization equals the need adjusted share of population in all the areas:*

$$x_i \equiv \max\left(h_i^* - h_i; 0\right) = 0 \ , \forall i \in N \ \Leftrightarrow \ \sum_j \left(\frac{p_{ij}}{p_j} \frac{h_j}{H} \right) = \frac{h_i}{H} \ , \forall i \in N \qquad (10)$$

The proof of Proposition 2 is given in Appendix B.

p_{ij} / p_j is the share of the total population in the need class *j* living in the area *i*; h_j / H is the share of health care utilization of all the people in the need class *j* with respect to the population as a whole; h_{ij} / H is the share of utilization of people in the need class *j* living in the area *i*. The need adjusted share of the population is obtained from indirect standardization of the population of every area with respect to the *J* need classes (Equation (3) and (4)). The right hand side of Equation (10) contains a condition of horizontal equity:

$$\text{If } \sum_j \frac{p_{ij}}{p_j} \frac{h_j}{H} = \sum_j \frac{p_{i'j}}{p_j} \frac{h_j}{H} \ , \forall i, \ i' \in N \ \Rightarrow \ \frac{h_i}{H} = \frac{h_{i'}}{H} \ \ \forall i, \ i' \in N \qquad (11)$$

If two areas have an equal need adjusted population share, then they should have an equal share of health care utilization. This condition is similar to the one proposed in Wagstaff and van Doorslaer (1991) with respect to population income groups, and widely accepted in the literature. Proposition 2 states an unequivocal relationship between the condition of *the absence of deprivation* and the condition of *horizontal equity* in the whole population of areas. Conversely, the relationship between *the presence of deprivation* and *horizontal inequity* is weaker:

Proposition 3: *If an area is deprived in health care, then it is also unequally treated with respect to all other areas*:

$$x_i \equiv \max\left(h_i^* - h_i; 0\right) > 0 \;\Rightarrow\; \sum_j \left(\frac{p_{ij}}{p_j} \frac{h_j}{H} \right) \neq \frac{h_i}{H} \tag{12}$$

The proof of Proposition 3 is given in Appendix C.

By Proposition 3, an area can be unequally treated and not necessarily be deprived in health care. Specifically, all the areas above the frontier of need expected utilization are not deprived in health care and may well have unequal share of health care for equal need. However, if any area is above the frontier of need expected utilization, then at least one area will be deprived in health care by Proposition 1, and the whole population will suffer inequity by Proposition 2 and 3. Thus, the presence of areas above the frontier of health care utilization implies inequity for the population as a whole.

Now, we are able to interpret the HCD profile *HCD(x; p)* as a measure of both deprivation and inequity in health care:

i. *HCD(x; p)* is a continuous, concave and monotonically increasing function defined over *x*.

ii. *HCD(x; p)* lies on the horizontal-axis if and only if all the HCGs are zero; thus if and only if there is no deprivation and no inequity *in the whole population of areas*.

iii. *HCD(x; p)* is monotonically increasing in the average HCG μ.

iv. *HCD(x; p)* is monotonically increasing in the variance of the distribution of the HCGs.

Properties (i) and (ii) derive directly from the definition of *HCD(x; p)*; (iii) and (iv) are implied by condition (i).

The HCD profile offers a measure of deprivation in the utilization of health care in the whole population of areas, defined as receiving less health care than needed. The reliability of this measure depends on the extent to which the need expected health care utilization h^* is a correct estimator of the population need for health care. h^* is defined as a function of the average utilization in the population as a whole (Equation (3)). Thus, it is assumed that observed and needed health care on average are equal in the total population. Specifically, the HCD profile assumes that on average there is not an over or under utilization of health care in the population with respect to each of the J need classes considered. If a large part of the demand for some health care services is not met by the supply, this assumption may not hold. In this case, the measure of health care deprivation captured by the HCD profiles will be under-estimated.

The HCD profile also provides a measure of horizontal inequity for those areas deprived in health care relatively to the population as whole. Proposition 3 allows an interpretation of the distance between h^* and h as a measure of both deprivation and horizontal inequity for the areas below the frontier of need expected health care utilization. Although, the HCD profile measures inequity experienced only in this part of the areas, their measure of inequity takes into account the whole distribution of health care in all the areas. For all those areas above the frontier of health care utilization, the HCG vector returns a zero entry. Hence, the HCD profile is not able to measure inequity experienced by the non-deprived in health care areas. However, if some areas are above the frontier of need expected utilization, then at least one area will be deprived in health care (Proposition 1). Thus, the HCD profile will be indirectly able to measure this source of inequity in the population.

Finally, Proposition 2 states an unequivocal relationship between *the absence of deprivation* and *horizontal equity* in the whole population of areas. Therefore, the absence of deprivation in health care is a desirable condition since it guarantees horizontal equity for *all* the areas.

3. HCD-dominance and HCD-indices

In the inequality literature, the Lorenz dominance criterion requires a minimal set of value judgements[‡] in order to compare inequality in the distribution of income of different populations. This results in robust, but incomplete rankings. Inequality indices (e.g. the Gini, Theil or Atkinson index) play an important role whenever the Lorenz dominance criterion is not capable of ordering two distributions. These indices incorporate a larger number of value judgements, but are able to produce complete rankings. In this section, we follow the same axiomatic approach and illustrate the normative assumptions which characterize the dominance relation between HCD profiles and the class of indices consistent with HCD-dominance rankings.

Consider the class of indices Δ defined over the HCG vector, x, that satisfies the following normative properties:

1. Symmetry or anonymity axiom: Δ(x) is not influenced by permutations in the values of x.

2. Replication invariance axiom: if we can obtain x' from x by replication of the population, then Δ(x') = Δ(x).

3. Strict monotonicity axiom: if x' is obtained from x by reducing some HCGs, then Δ(x') < Δ(x').

4. Equality preferring axiom: if x' is obtained from x by mean preserving equalizations, then Δ(x') < Δ(x). This means that if we transfer health care from some health care deprived area to some area that is more deprived, the value of Δ(.) increases. The same result holds if we transfer health care from areas non-deprived to areas deprived in health care.

5. Modified focus axiom: if we obtain x' from x transferring health care across the non-deprived in health care, then Δ(x') = Δ(x).

[‡] These are the symmetry or anonymity axiom and the Pigou-Dalton principle of transfers.

Then:

> For any vector x' and x, x' HCD-dominates x if and only if $\Delta(x') < \Delta(x)$ for all $\Delta(.) \in \Delta$. (Shorrocks, 1998; Jenkins and Lambert, 1997 and 1998)

Axioms 1 to 4 are widely accepted in the inequality literature and so we will not comment on these further (Lambert, 2001). Conversely, the focus axiom is chiefly used in the literature on poverty and has important implications for the inequity measure that we obtain. First, it implies that inequity among health care deprived areas should not be offset by inequity among health care non-deprived areas. Specifically, greater variance in the distribution of the gap of the former is not compensated by less variance in the distribution of the gap of the latter. Second, in ranking inequity between populations of areas *having equals average HCG* what should matter is inequity among the deprived in health care, not among the non-deprived in health care. This implies HCD ranks give zero weight to inequity to all those areas using more health care than need expected. However, this is not equivalent to excluding the top of the distribution of health care from the inequity measure that we obtain. For instance, if the utilization of health care non-deprived areas increases, while utilization of health care deprived areas remains unchanged, then the HCD profile will rise since the value of h^* for each observation is now slightly higher. Only transfers among health care non-deprived areas (so long as this does not result in a non-deprived area becoming deprived in health care) will not be captured by the HCD profile. This characteristic distinguishes the modified focus axiom defined above from the usual focus axiom delineated in the literature on poverty: an exogenous change in the utilization of health care non-deprived areas is captured by our class of indices.

Analysis of the distribution of health care across areas non-deprived in health care is achievable under this approach. We just need to invert the sign in the definition of the HCGs (i.e. $x_i = \max\left(h_i - h_i^*; 0\right)$) in order to capture the upper part of the health care distribution. The need-expected distribution of health care, h^*, is the same as that for the deprived in health care. Thus, we can easily define health care affluence profiles

(HCAs) and obtain a measure of inequity as unequal access for equal need for the non-deprived in health care.

The HCD-dominance relation produces ranks that satisfy all the normative axioms listed above and are robust to any specification of the indices in the class Δ. This means that the HCD ranks are consistent with any social preferences for inequity and deprivation that satisfy the axioms. In particular, HCD-dominance ranks are consistent with preferences concerned exclusively with those areas having the largest HCG (Rawlsian preferences), and with preferences that make trade-offs between the largest and the average HCG in the population of areas. When the HCD profiles cross there is no dominance relation between distributions, and thus it is necessary to specify an index or a set of indices in the class Δ(.) in order to obtain a complete ranking. This is equivalent to assuming some particular specification of social preferences. Complete rankings are obtained at the price of a loss of robustness, since they are no longer consistent with the entire preference spectrum defined above.

The Foster, Greer and Thorbecke (1984) class of poverty indices belong to the set Δ(.) and are valuable candidates for the inequity analysis[§]. They can be defined as a function of HCGs and used as indices of horizontal inequity and deprivation in health care. Following the framework delineated in Section 2, we obtain:

$$\Delta_{\alpha}^{FGT}(x) = \int_{0}^{\xi} x^{\alpha} dF(x) \qquad (13)$$

Where α is a parameter that defines the social preferences for the distribution of inequity among the areas deprived in health care. When $\alpha = 1$, the FGT-index equals μ and we obtain inequity ranks consistent with preferences that care only about the average HCG in the population of areas. When $\alpha > 1$, the transfer axiom holds and the inequity ranks are consistent with preferences that allow trade-offs between the distribution of the HCG and the value of μ. When $\alpha \to \infty$, no trade-off is allowed and inequity is ranked according to the area having the largest HCG in the population

[§] The FGT indices are not new to health economics. Madden (2006) has applied this class of indices in the analysis of the distribution of Body Mass Index scores in Ireland, but to our knowledge they have not been applied to equity in the delivery of health care.

(Rawlsian preferences). A useful characteristic of the FGT family of indices is their additive decomposability property: if the population of areas is partitioned into K subgroups, then the overall value of the index can be expressed as a weighted sum of the value of the index in each subgroup. Therfore, the measure of inequity is subgroup consistent: a reduction in the inequity of one group will result in a reduction of inequity in the total population of areas. This property can be expressed as:

$$\Delta_{\alpha}^{FGT}(x) = \sum_{k=1}^{K} \pi_k \Delta_{\alpha}^{FGT}(x_k) \qquad (14)$$

Where π_k is the proportion of areas in the subgroup k and x_k is the HCG vector for the group k.

4. Analysis of the GP fundholding reform

This section uses HCD profiles to measure the effect of GP fundholding on equity in the English NHS. The fundholding scheme was part of the “Internal Market” reforms implemented in the NHS during the years 1991-1999. One objective of the reform was to improve cost-efficiency and quality in the delivery of health care by creating local public payers responsible for purchasing care from local hospital providers and thereby introducing “quasi-market” hospital competition into the NHS. Two types of payer were introduced: geographically defined District Health Authorities and GP fundholders (GPFHs). GP fundholding was a voluntary scheme: the GPs who decided to become FHs received a budget to purchase a wide range of secondary care for the patients on their practice list. They could reinvest any savings in their practice, including improvements to premises and the range of services offered, both of which would potentially increase the capital value of the practice owned by the GP principals. Those GPs who did not join the scheme bore no cost in referring patients to hospital.

Many empirical studies have investigated the consequences of the scheme for secondary care admissions (Coulter and Bradlow, 1993; Propper, et al., 2002; Dusheiko et al., 2006). Virtually all these studies use a regression approach to evaluate the effect of fundholding and do not focus on the consequences for equity. The main contribution of our study is to analyse the effect of fundholding on equity, and to do so in a way consistent with a wide range of social preferences.

4.1. Data

We use a small area panel dataset composed of over 8,000 electoral wards covering the whole of England from 1989 to 2001. Aggregation and record linkage were performed using the frozen 1991 ward codes provided by the Hospital Episode Statistics (HES). This improves comparability and minimises the risk that inequality effects are an artefact of electoral ward boundary changes. The main variables are described below.

Hospital utilization

Data on hospital utilisation was obtained from the Hospital Episode Statistics (HES) for financial years 1989/90 through 2001/2. This includes data on all NHS hospital inpatient episodes in England, but excludes private hospital admissions. In our analysis we use data on all finished first consultant episodes for elective primary total prosthetic replacement of hip joint, excluding revisions, for all adults aged over 44 by age, sex and frozen 1991 ward code.

Population

Data on the characteristics of ward resident populations in 1991 and 2001 was obtained from the respective National Censuses (Office of Population Censuses and Surveys 1991, Office of National Statistics 2001). Census 2001 data was allocated to frozen 1991 wards at the level of 2001 Census Output Areas, which each contain around 125 households. This was realized using geographical software (MapInfo), by matching each Output Area to the ward which contained the "population centroid" of that Output Area (i.e. the point of maximal population density). Census data was then linearly interpolated between 1991 and 2001 endpoints to estimate population characteristics in the years 1992-2000. Demographic and socio-economic characteristics in 1989 and 1990 were assumed to be unchanged with respect to the 1991 census point.

Socio-economic status

The Census-based Townsend deprivation Z-score is used as an indicator of socio-economic status (Carstairs, 2000). This is constituted by four components: unemployment, overcrowding, non-car ownership, non-home ownership. Each component is linearly interpolated between 1991 and 2001 endpoints, and normalised to zero over the period 1991-2001 by subtracting the mean and dividing by the standard deviation (following a log-transformation for unemployment and overcrowding). These four component Z-scores are then summed. The resulting Townsend score ranges from about –8 to about +13, with higher values indicating greater deprivation.

Need variables

The ward population size aged over 44, split by five-year bands (from 44 up to over 90) and by sex is used as indicator of need. Few elective hip procedures are performed on patients under 45 (Dixon *et al.*, 2004). In addition, we consider a standardized illness ratio: the proportion of residents of households aged 60 and over having limiting long standing illness, standardized for age and sex.

GP fundholding index

We compute an index of GP fundholding penetration in order to capture the effect of the reform on secondary care admissions. The index is time varying and is defined as the proportion of the ward population registered with a fundholding practice. The sources of data for this index includes Prescription Pricing Authority data from 1998/9 on all practices in England including practice code, fundholding type and start/end date; Attribution Data Set data on practice registrations for 1999 2000; and record linkage data from the Codes Service Database 2001.

Table 1. Descriptive statistics; number of observations in each year: 8326 electoral wards.

year	Variables	Mean	Std. Dev.	Min	Max
1989	Elective hip replacements	2.504	2.622	0	33
	Population over 44	2099	1450	235	13778
	Townsend index score	1.021	3.371	-7.296	13.877
1990	Elective hip replacements	2.631	2.649	0	28
	Population over 44	2099	1450	235	13778
	Townsend index score	1.021	3.371	-7.296	13.877
1991	Elective hip replacements	3.156	2.823	0	30
	Population over 44	2099	1450	235	13778
	Townsend index score	1.021	3.371	-7.296	13.877
	Proportion of people in FHs	0.057	0.149	0	0.9997
1992	Elective hip replacements	3.383	3.083	0	34
	Population over 44	2117	1451	234.2	13821
	Townsend index score	0.873	3.354	-7.582	13.644
	Proportion of people in FHs	0.102	0.198	0	0.9997
1993	Elective hip replacements	3.549	3.094	0	31
	Population over 44	2135	1453	233.4	13863
	Townsend index score	0.722	3.336	-6.881	13.459
	Proportion of people in FHs	0.199	0.262	0	0.9997
1994	Elective hip replacements	3.766	3.35	0	39
	Population over 44	2153	1456	232.6	13906
	Townsend index score	0.562	3.327	-6.722	13.255
	Proportion of people in FHs	0.278	0.296	0	0.9998
1995	Elective hip replacements	3.896	3.404	0	34
	Population over 44	2171	1459	231.8	13948
	Townsend index score	0.394	3.324	-7.001	13.029
	Proportion of people in FHs	0.317	0.308	0	0.9998
1996	Elective hip replacements	3.6	3.218	0	37
	Population over 44	2189	1463	231	13991
	Townsend index score	0.215	3.328	-7.399	12.777
	Proportion of people in FHs	0.411	0.326	0	1

4.2 Methods of analysis

Elective hip replacements for all adults aged over 44 are used as an indicator of health care utilization. The vector of need-expected utilization, h^*, is computed following Equations (3) and (4):

$$h_i^* = \sum_{j=1}^{J} p_{ji} h_j \tag{3}$$

Where J = 10 are need classes (five age classes for men and five for women aged from 44 up to over 90); p_{ji} is the number of people in the ward i in the age and sex stratum j; h_j is the national utilization rate for age and sex stratum j, i.e. the national count for England in the current year divided by the national population in that year. This specification of h^* facilitates inequity comparison between different years since it allows for the increase in national utilization over the period considered.

The HCG vector defined in section 1 is obtained as: $x_i = \max(h_i^* - h_i; 0)$, where h_i is the observed count for elective hip replacements in ward i. We account for potential bias due to heterogeneity in the ward population size using indirect standardization (O'Donnell et al., 2007):

$$x_i^{is} = x_i - \hat{x}_i + \overline{x} \tag{15}$$

$$\hat{x}_i = \hat{\beta}_0 + \hat{\beta}_1 p_i + \hat{\beta}_2 \overline{d} + \hat{\beta}_3 \overline{g} \tag{16}$$

The HCGs, x_i, are regressed against population size p_i controlling for level of deprivation d_i and fundholding penetration g_i. Equation (15) gives the indirectly standardized HCGs at the average population size effect. This variable can be interpreted as the HCGs that we would observe if the population in each ward was equal to the average population[**].

[**] Also, we attempt to control for unobservable heterogeneity of the HCGs in the need for hip replacements that is not captured by demographic variables. The standardized illness ratio defined in section 4.1 is used as a proxy of the need for hip replacements. However, we find that the latter is not significant in explaining the variation in the HCGs after controlling for the ward population aged over 44.

Trends in inequity over time were tracked for all the health care deprived wards. The FGT indices are computed for the whole population of English wards from 1989 to 1991 and decomposed by level of *socio-economic deprivation*, as measured by the Townsend index, in four quartiles (most affluent 1st, most affluent 2nd, most affluent 3rd, most deprived quartile). The parameter α is set: $\alpha = 1$, producing ranks consistent with social preferences that care only for the average HCG; $\alpha = 2$, producing ranks consistent with social preferences that trade-off inequality in the distribution of HCGs with the magnitude of the average HCG. Robust standard errors are computed accounting for heterogeneity in the ward population size using statistics developed for the index in the literature on poverty[††] (Kakwani, 1993; Joliffe et al., 2004).

The impact of the GP fundholding reform on equity is analysed by focusing on the most socio-economically deprived quartile of electoral wards deprived in health care. We expect that, if the reform has an impact on equity this is more likely to be evident in those wards where inequity has consistently been present over the period of time considered (Figure 2 and Figure 3). Also, selecting the most socio-economically deprived quartile of wards has two advantages. First, we achieve some additional control for unobserved heterogeneity in the need for hip replacement across the wards: if unobservable need for hip replacement is positively correlated with the level of deprivation of the wards, selecting the most socio-economically deprived wards would reduce such heterogeneity in the analysis. Second, we exclude that part of the population that is more likely to select out from the sample by opting for private health care treatment, which is not observable in our dataset. This is a particularly relevant issue for hip replacement, since one quarter of all the procedures recorded in England during the period examined were performed by the private sector (Williams and Rossiter, 2004). Thus, potential bias coming from sample selection should be considerably reduced.

The most socio-economically deprived quartile of wards is initially partitioned into two groups: the control group composed of wards having a proportion of people registered in a FH practice lower than 20%, and the treatment group having a proportion of people

[††] These statistics are validated using bootstrapped standard errors.

registered in a FH practice larger than 20%. In order to check the robustness of our findings further subgroups are considered, dividing the treated group according to increasing fundholding penetration[‡‡] (20-40%, 40-60%, more than 60%, 40-100%). We compute HCD profiles for each group from 1991 to 1996 and test for HCD-dominance implementing the statistical test developed by Davidson and Duclos (2000). The HCD profiles are displayed in Figures 4 and the results of the tests are in Table 2.

FGT indices (parameters α=1 and α=2) are computed for the groups of wards defined above from 1991 to 1996. An index of relative inequity is computed dividing within-group FGT indices by the FGT index for the sum of the groups:

$$r_{\alpha}^{FGT}(x_k) = \frac{\Delta_{\alpha}^{FGT}(x_k)}{\Delta_{\alpha}^{FGT}(x_K)} = \frac{\Delta_{\alpha}^{FGT}(x_k)}{\sum_{k=1}^{K} \pi_k \Delta_{\alpha}^{FGT}(x_k)} \quad ,\forall k \in K \subseteq N \qquad (17)$$

K is the number of subgroups of wards defined by increasing degree of fundholding penetration in the most socio-economically deprived quartile of wards (i.e. a subset of the total number of wards *N*). "The relative FGT index" shows the inequity share of each group according to the level of fundholding penetration: a ratio above 1 indicates more inequity than the average, i.e. the inequity measured in the sum of the groups. Results are shown in Table 3 and in Figure 5 and Figure 6. We test for the hypothesis of no significant difference in the share of inequity between all groups from 1991 to 1996 using bootstrapped standard errors.

The test for the hypothesis of no treatment effect is implemented using a difference-in-differences approach. First, we compute the difference between treated groups before and after the fundholding reform was implemented, then we subtract from this value the difference between the control group before and after the reform. 1989/90 was set as benchmark financial year, since it falls two years before the start of the FH scheme, allowing us to account for potential strategic behaviour of GPs preparing to become

[‡‡] Further stages of decomposition are not achievable for 1991 and partially achieved for 1992 wards, because of insufficient observations having fundholding penetration higher than 40% in those early years (only 58 wards in 1991 and 130 wards in 1992).

FHs in 1991 (Croxson, et al., 2001). This difference-in-differences approach can be formulated as:

$$DID = (\Delta_{f,t}^{FGT} - \Delta_{f,'89}^{FGT}) - (\Delta_{nf,t}^{FGT} - \Delta_{nf,'89}^{FGT}) \qquad (18)$$

where *f* is the treated group and *nf* is the controll group. The DID allows us to control for time confounders and for self-selection. On the one hand, comparing the FGT indices for the treated groups before and after the reform alone will not identify the treatment effect because of unobservable heterogeneity in policy incentives and other temporal factors that might have occurred along the two periods. On the other hand, comparing the FGT indices between the non-treated and treated groups alone will not identify the treatment effect because of unobservable heterogeneity in relevant characteristics of the practices that self select into the scheme. The DID identifies the treatment effect under the assumption that the effect of time and policy confounders follow a common time trend for fundholding and non fundholding practices (Jones, 2007). The results are reported in Table 4.

4.3 Results

HCD profiles and FGT indices can be interpreted both as measures of deprivation in the access to elective hip replacements and as measures of inequity with respect to those wards which result to be deprived in health care. We will follow the latter interpretation in explaining the results of our analysis.

Figures 2 and 3 show that among those wards deprived in health care, the most socio-economically deprived quartile experiences the greatest inequity (or equivalently the greatest deprivation in health care) over all periods considered. The remaining quartiles have no significant differences in their values of the FGT indices. This result is robust to two alternative specifications of the social preference for the distribution of inequity captured by the parameters α=1 and α=2 of the index. Our findings are consistent with a large part of the literature on socio-economic inequalities and inequity in elective hip replacement in the NHS (Chaturvedi, 1993; Cookson et al., 2007). The difference between the average HCG of the most socio-economically deprived quartile and the average HCG of other quartiles ranges from 0.66 elective hip replacements in 1989 to 1.54 in 1993 (Figure 2). This is a quite substantial disparity if we consider that the (population-over-44 weighted) average number of elective hip replacements for each ward was 3.57 in 1989 and 4.86 in 1993.

Table 2. HCD-dominance ranks in the most socio-economically deprived quartile of wards; all observations are grouped by level of fundholding penetration (0-100%); test for dominance from Davidson and Duclos (2000).

Population subgroups	Years					
	1991	1992	1993	1994	1995	1996
A vs B	SSD**	SSD**	no dom	no dom	no dom	no dom
A vs C	-	no dom	no dom	no dom	no dom	no dom
A vs D	-	-	SSD**	no dom	no dom	no dom
A vs E	-	-	SSD**	no dom	SSD**	SSD*
A vs F	-	SSD**	SSD**	no dom	no dom	no dom
C vs D	-	-	no dom	no dom	no dom	no dom
C vs E	-	-	SSD**	no dom	SSD*	SSD**
C vs F	-	no dom	SSD**	no dom	no dom	no dom
D vs E	-	-	no dom	no dom	no dom	no dom

Note: A = 0 to 20% of people registered in a GPFH practice; B = 20% to 100%; C = 20% to 40%; D = 40% to 60%; E = 60% to 100%; F = 40% to 100%. SSD** = second order stochastic dominance significant at 5%; SSD*= significant at 10%; no dom = not significant.

In the most socio-economically deprived quartile of wards, higher FH penetration is associated with less inequity for all those wards deprived in health care (Table 2, Figure 4). There is a clear dominance rank by increasing degree of FH penetration in 1991 to 1993. However, this dominance relationship holds only with respect to the top FH penetration group (i.e. 60-100% FH penetration) in 1995 and 1996, and is not statistically significant in 1994. Turning to the FGH indices, inequity falls with increasing degree of FH penetration over all the years of the reform (Table 3, Figures 5 and 6). In 1993, people in socio-economically deprived wards with over 60% FH penetration have an average HCG 29% smaller than the average for wards in the same most socio-economically deprived quartile of electoral wards (Table 3, Figure 5). The positive relationship between FH penetration and equity reduces in intensity after 1993 and holds only for the top FH penetration groups, confirming the dominance ranks in Table 2. We can interpret this change in the trend as a break in the characteristics of the FHs: early waves of FHs were more responsive to equity in the distribution of access to elective hip replacements, while late waves of FHs were less responsive. Previous studies support the hypothesis of a break in the characteristics of the GPFHs that joined the scheme after 1993 (Baines and Whynes, 1996). The authors bring evidence that FHs in the first three waves were more likely to meet a set of quality and efficiency targets than the FHs in subsequent waves. An alternative hypothesis is that the change in the trend is a consequence of a change in the behaviour of the early wave FHs. However, neither the former nor the latter hypothesis can be tested in our study. Since our unit of analysis is the electoral ward, we are not able to adequately disentangle the effects of individual waves of FHs.

Table 3. Differences in inequity share in most socio-economically deprived quartile of electoral wards of England; all observations are grouped by level of fundholding penetration (0-100%); relative FGT indices; bootstrapped standard errors.

Population subgroups	Preference parameter	Years					
		1991	1992	1993	1994	1995	1996
A – B	α = 1	16.8%**	19.2%**	14.5%**	8.8%**	4.3%	1.6%
	α = 2	21.9%**	22.9%**	14.8%**	9.0%	4.4%	-3.0%
A – C	α = 1	-	18.5%**	8.6%	4.9%	-0.2%	-4.0%
	α = 2	-	21.2%**	6.8%	2.6%	-0.9%	-16.1%*
A – D	α = 1	-	-	21.6%**	13.3%**	8.8%	2.2%
	α = 2	-	-	23.3%**	16.9%*	6.5%	1.5%
A – E	α = 1	-	-	29.1%**	13.6%**	9.8%*	10.3%*
	α = 2	-	-	36.4%**	15.8%*	15.7%**	13.2%*
A – F	α = 1	-	20.5%**	23.9%**	13.4%**	9.2%*	5.7%
	α = 2	-	26.5%**	27.3%**	16.5%**	10.3%	6.6%
C – D	α = 1	-	-	13.1%*	8.4%	9.0%	6.3%
	α = 2	-	-	16.4%	14.3%	7.4%	17.6%*
C – E	α = 1	-	-	20.5%**	8.8%	9.9%*	14.3%**
	α = 2	-	-	29.5%**	13.3%	16.6%**	29.3%**
C – F	α = 1	-	2.0%	15.3%**	8.6%	9.4%*	9.8%*
	α = 2	-	5.3%*	20.4%**	13.9%	11.2%	22.7%**
D – E	α = 1	-	-	7.4%	0.3%	0.9%	8.0%
	α = 2	-	-	13.1%	-1.0%	9.2%	11.7%

Note: A = 0 to 20% of people registered in a GPFH practice; B = 20% to 100%;
C = 20% to 40%; D = 40% to 60%; E = 60% to 100%; F = 40% to 100%.
** = significant at 5%; * = significant at 10%.

The difference-in-differences analysis shows that a large part of the observed equity associated with FH penetration is due to self-selection. In 1991, among the most socio-economically deprived quartile of wards, more equitable practices tended to self-select into the scheme (Table 4). In 1992 and 1993, there is evidence of a positive treatment effect of the FH reform on equity in the distribution of hip replacements. The magnitude of the effect is quite substantial: in 1993 the average HCG in the top FH penetration group reduces by 1.12 elective hip replacements with respect to the population weighted average number of elective hip replacements of 4.86 for each ward. However, we are not able to identify to what extent the effect is explained by the individual waves of FHs that join the scheme in 1991, 1992 or in 1993, since we are not able to control for FH waves. The treatment effect is not statistically significant after 1993. This finding is consistent with evidence that early wave FHs were more "enthusiastic" and "active" than late wave FHs (Le Grand, Mulligan and Mays 1997). It suggests that the former may have been willing to increase hospital referral rates for severe osteoarthritis among "high-need" individuals living in socio-economically deprived wards, relative to the rest of the population. However, other possible explanations cannot be ruled out in the absence of practice level data that could disentangle selection and treatment effects by FH waves and in the absence of data on GP referral rates as opposed to hospital admission rates.

Table 4. Difference in differences estimates from FGT indices; population subgroups defined by level of fundholding penetration (0-100%); non-treated group: 0-20% FH penetration; reference year: 1989.

Treated groups	Preference parameter	Years					
		1991	1992	1993	1994	1995	1996
B	α = 1	-0.285	-0.461**	-0.620**	-0.222	0.062	0.226
	α = 2	-2.02	-2.985*	-3.9**	-1.193	0.52	2.710*
C	α = 1	-	-0.304	-0.493**	-0.267	0.100	0.116
	α = 2	-	-2.069	-2.958	-1.37	1.072	2.928
D	α = 1	-	-	-0.689**	-0.101	-0.063	0.335
	α = 2	-	-	-4.304*	-0.704	0.059	2.637
E	α = 1	-	-	-1.117**	-0.279	0.143	0.270
	α = 2	-	-	-7.827**	-1.423	-0.278	2.470*
F	α = 1	-	-0.788**	-0.819**	-0.168	0.022	0.306
	α = 2	-	-4.891*	-5.379**	-0.973	-0.085	2.559*

Note: A = 0 to 20% of people registered in a GPFH practice; B = 20% to 100%;
C = 20% to 40%; D = 40% to 60%; E = 60% to 100%; F = 40% to 100%.
** = significant at 5%; * = significant at 10%.

4.4 Robustness checks

The analysis described in the previous section has been repeated in a subset of the wards considered. We select the most socio-economically deprived deciles of wards and obtain very similar predictions both for the treatment effect and the cross sectional analysis. In this case the results should gain some additional control for unobservable heterogeneity in the need for hip replacement if this is correlated with the level of socio-economic deprivation.

The treated groups were merged and split in the following sets: (20-100%), (20-40%, 40-100%), (20-40%, 40-60%, and 60-100%). Our predictions are robust across these sets of groups. Moreover, the larger the proportion of people registered in FH practices, the lower is the value of the FGT indices and thus the lower is inequity for that subgroup.

Finally, we check for potential bias coming from heterogeneity in the presence of GP practices preparing to become FHs in the follower year across the groups considered. Previous studies show evidence of strategic behaviour: pending FHs raised their admission rates the year before joining the scheme in order to assure themselves a larger budget in the following year (Croxson, Propper et al., 2001). This may result in lowering or raising the HCGs, the larger or the smaller is the presence of the next wave of FHs across the groups considered. Since we do not observe single GP practices but only the proportion of people in such practices for each ward, we are not able to control for this variable directly. However, we provide some indirect control for it: we define a binary indicator identifying those wards having more than 20% of people registered in GP practices preparing to become FHs in the next year. This indicator was cross tabulated with the treated and non-treated groups. The proportion of wards having more than 20% of forthcoming FHs penetration is larger in the control group than in the treated groups from 1991 to 1993; it becomes almost equal in the following years. Thus, if forthcoming waves of FH were artificially increasing admission to secondary care for elective hip replacements, this should result in a disproportionately lower value of the FGT index for the control group and in an upward biased estimate of the treatment effect in 1992 and 1993 - i.e. the treatment effect should be larger in absolute terms.

5. Conclusions

The methodological contribution of this paper is to introduce some developments in the literature of poverty and deprivation to the analysis of inequality and inequity in health care. The HCD profiles define deprivation in health care as receiving less care than needed. This definition is strictly related to inequity, since a common interpretation of equity in health care is that resources have to be allocated with respect to individuals' needs rather than individuals' preferences. Thus, if some minorities in the population have inadequate access with respect to their need, this is generally regarded as an issue of social justice. In this context, the definition and the measurement of need for health care assumes a crucial role. The HCD profiles do not depend on a specific definition of need. In this paper, we have proposed a function of the average utilization observed in the total population as need estimator, following the tradition of the literature on health care inequality. In turn, this has led us to determine a link between horizontal equity and deprivation. However, alternative definitions of need can be specified, and thus alternative view on equity can be included in the HCD profiles.

A wide class of indices consistent with the HCD-dominance ranks can be introduced from the literature on poverty to provide measures of inequity and deprivation in health care. The FGT indices are members of this class and have some appealing properties. First, these indices are additive decomposable: inequity measured in the total population is a weighted average of inequity in the population subgroups. Hence, the measure of inequity is subgroup consistent: a reduction in the inequity of one group will result in a reduction of inequity in the total population. This property enables us to calculate different subgroups' contributions to overall inequity and to identify those groups suffering the largest inequity in the population. Often, the latter represent valuable information for the policy maker. Second, the FGT indices permit an explicit specification of social preferences for inequity and deprivation into a single parameter. Therefore, it is possible to test the robustness of the inequity ranks to alternative views about equity which endorse different trade-offs between the average HCG and inequality in the distribution of the HCGs.

The HCD profiles disregard information on inequity of one part of the distribution, i.e. those units having more utilization than their need-expected level. On one hand, focusing sharply in this way results in a graphical representation of useful information on the inequity experienced by those who are below their need expected utilization (i.e. the incidence the intensity and the inequality in the distribution of the HCGs). On the other hand, the HCD profiles represent inequity experienced only by one part of the population. However, we have shown in Section 3 that this measure of inequity takes into account information on the whole distribution of health care in the population of areas. Furthermore, it is possible to reverse the focus and look only at that part of the distribution receiving more than their need expected level of health care - i.e. an analysis of health care affluence profiles.

Section 4 illustrates an empirical application of the HCD profiles to measure the effect of the GP fundholding reform in the English NHS. Inequity in the distribution of elective hip replacements is measured as unequal treatment for equal need among those electoral wards deprived in health care. Inequity trends over time are analysed across socio-economic quartiles of wards. Measures of inequity are decomposed according to increasing level of FH penetration in the most socio-economically deprived quartile of wards. We produce HCD-dominance ranks associated with FH penetration, which are consistent with a large range of social preferences. The effect of the FHs reform on equity is identified using a difference-in-differences approach under two alternative specification of social preference for inequity embedded in the parameter α of the FGT index. The additive decomposability property of the FGT indices is crucial for the implementation of this estimation method, since it permits a decomposition of inequity measures by treated and non treated groups.

Our results show that in 1991 the lower level of inequity associated with fundholding is due to self-selection of the most equitable practices into the scheme. There is evidence of self-selection and also a positive treatment effect on equity in 1992 and 1993. The treatment effect is considerable. In 1993/4, the average health care gap between the actual and need-expected number of elective hip replacements in the top fundholding penetration wards fell by 1.12 hip replacements, compared to a population-weighted average number of elective hip replacement of 4.86 for each ward. However, we are not able to identify such effect by waves of FHs. Finally, in 1994 to 1996 the positive effect on equity is explained only by self-selection and is significant only in the group with the highest penetration of fundholding.

References

Bago d'Uva, T., A. M. Jones and E. van Doorslaer, 2007, Measurement of horizontal inequity in health care utilisation using European Panel Data. *HEDG working paper 07/17, University of York*.

Baines, D. L. and D. K. Whynes, 1996, Selection bias in GP fundholding. *Health Economics* 5, 139-140.

Breslow, N. E. and N. E. Day, 1987, *Statistical Methods in Cancer Research. Vol. II, The Design and Analysis of Cohort Studies* (International Agency for Research on Cancer, Lyon, France).

Carstairs, V., 2000, Socio-economic factors at area level and their relationship with health, in: P. Elliot, W. J.C., B. N.G. and B. D.J., eds., Spatial Epidemiology: Methods and Applications (Oxford University Press, Oxford) 51-67.

Chaturvedi, N. and Y. Ben-Shlomo, 1995, From the surgery to the surgeon: does deprivation influence consultation and operation rates? *British Journal of Medical Practice* 45, 127-131.

Clayton, D. and M. Hills, 1993, *Statistical Models in Epidemiology* (Oxford University Press, Oxford).

Cookson, R., M. Dusheiko and G. Hardman, 2007, Socio-economic inequality in small area use of elective total hip replacement in the English National Health Service in 1991 and 2001. *Journal of Health Services Research and Policy* 12 Suppl 1, S1-10-7.

Coulter, A. and J. Bradlow, 1993, Effect of NHS reforms on general practitioners' referral patterns. *British Medical Journal* 306, 433-437.

Croxson, B., C. Propper and A. Perkins, 2001, Do doctors respond to financial incentives? UK family doctors and the GP fundholder scheme. *Journal of Public Economics* 79, 375-398.

Davidson, R. and J.-Y. Duclos, 2000, Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica* 68, 1435-1464.

del Río, C., C. Gradín and O. Cantó, 2006, The measurement of gender wage discrimination: The distributional approach revisited. *ECINEQ Working Paper Series* WP 2006 – 25.

Dixton, T., M. Show, S. Ebrahim and P. Dieppe, 2003, Trends in hip and knee joint replacement: socioeconomic inequalities and projections of need. *Ann Rheum Dis* 63, 825-830.

Dusheiko, M., H. Gravelle, R. Jacobs and P. Smith, 2006, The effect of financial incentives on gatekeeping doctors: evidence from a natural experiment *Journal of Health Economics* 25, 449-478.

Foster, J., J. Greer and E. Thorbecke, 1984, A class of decomposable poverty measures. *Econometrica* 52, 761-766.

Gravelle, H., S. Morris and M. Sutton, 2006, Economic studies of equity in the consumption of health care, in: A. M. Jones, ed., The Elgar Companion to Health Economics (Edward Elgar Publishing Limited, Cheltenham) 193-204.

Jenkins, S. P. and P. J. Lambert, 1997, Three I's of poverty curves, with an analysis of UK poverty trends. *Oxford Economic Papers* 49, 317-327.

Jenkins, S. P. and P. J. Lambert, 1998, Three 'I's for poverty curves and poverty dominance: TIPs for poverty analysis, in: D. J. Slottje, ed., Research on Economic Inequality (Jai Press Inc., London) 39-56.

Jolliffe, D., G. Datt and M. Sharma, 2004, Robust Poverty and Inequality Measurement in Egypt: Correcting for Spatial-price Variation and Sample Design Effects. *Review of Development Economics* 8, 557-572.

Jones, A. M., 2007, Panel data methods and applications to health economics. *HEDG working paper 07/18, University of York*.

Kakwani, N., 1993, Statistical inference in the measurement of poverty. *Review of Economics and Statistics* 75, 632-639.

Lambert, P., 2001, *The Distribution and Redistribution of Income* (Manchester University Press, Manchester).

Le Grand, J., N. Mays and K. Mulligan, 1998, *Learning from the NHS internal market: a review of the evidence* (King's Fund Publishing, London).

Madden, D., 2006, Body mass index and the measurement of obesity. *HEDG working paper 06/11*, *University of York*.

Morris, S., M. Sutton and H. Gravelle, 2005, Inequity and inequality in the use of health care in England: an empirical investigation. *Social Science & Medicine* 60, 1215-1266.

Oaxaca, R., 1973, Male-female differentials in urban labor markets. *International Economic Review* 14, 693-709.

O'Donnell, O., E. v. Doorslaer and A. Wagstaff, 2006, Decomposition of inequalities in health and health care, in: A. M. Jones, ed., *The Elgar Companion to Health Economics* (Edward Elgar Publishing Limited, Cheltenham) 179-192.

O'Donnell, O., E. van Doorslaer, A. Wagstaff and M. Lindelow, 2008, *Analysing Health Equity Using Household Survey Data* (World Bank Publications, Washington, DC).

Propper, C., B. Croxson and A. Shearer, 2002, Waiting times for hospital admissions: the impact of GP fundholding. *Journal of Health Economics* 21, 227-252.

Shorrocks, A., 1998, Deprivation profiles and deprivation indices, in: S. P. Jenkins, ArieKapteyn and V. P. B. M. S., eds., *The Distribution of Welfare and Household Production: International Perspectives* (Cambridge University Press, London) 250-267.

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

van Doorslaer, E., C. Messaria and X. Koolman, 2006, Inequalities in access to medical care by income in developed countries. *Canadian Medical Association Journal* 174, 177-183.

van Doorslaer, E. and A. Wagstaff, 1992, Equity in the delivery of health care: some international comparisons. *Journal of Health Economics* 11, 389-411.

Wagstaff, A., E. van Doorslaer and P. Paci, 1991, On the Measurement of Horizontal Inequity in the Delivery of Health Care. *Journal of Health Economics* 10, 169-205.

Wagstaff, A. and E. v. Doorslaer, 2000b, Equity in health care finance and delivery, in: A. J. Culyer and J. P. Newhouse, eds., *Handbook of Health Economics* (Elsevier Science B. V., Amsterdam) 1804-1862.

Wagstaff, A. and E. van Doorslaer, 2000a, Measuring and testing for inequity in the delivery of health care. *Journal of Human Resources* 35, 716-733.

Williams, J. and A. Rossiter, 2004, *Choice: the evidence* (Social Market Fundation, London).

Figure 2.

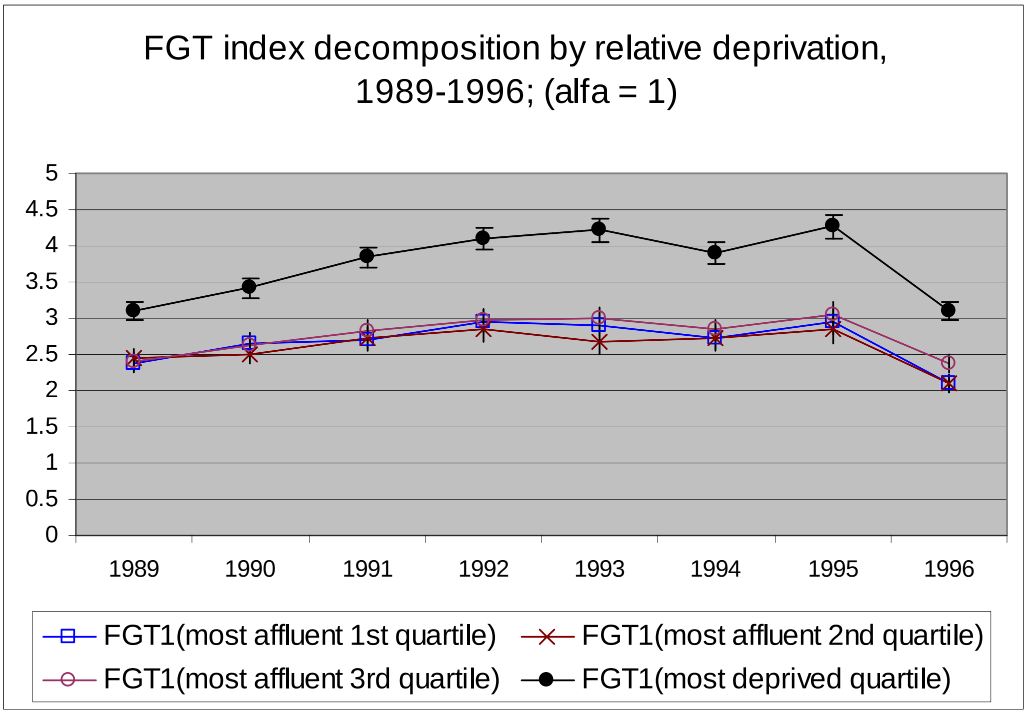


Figure 3.

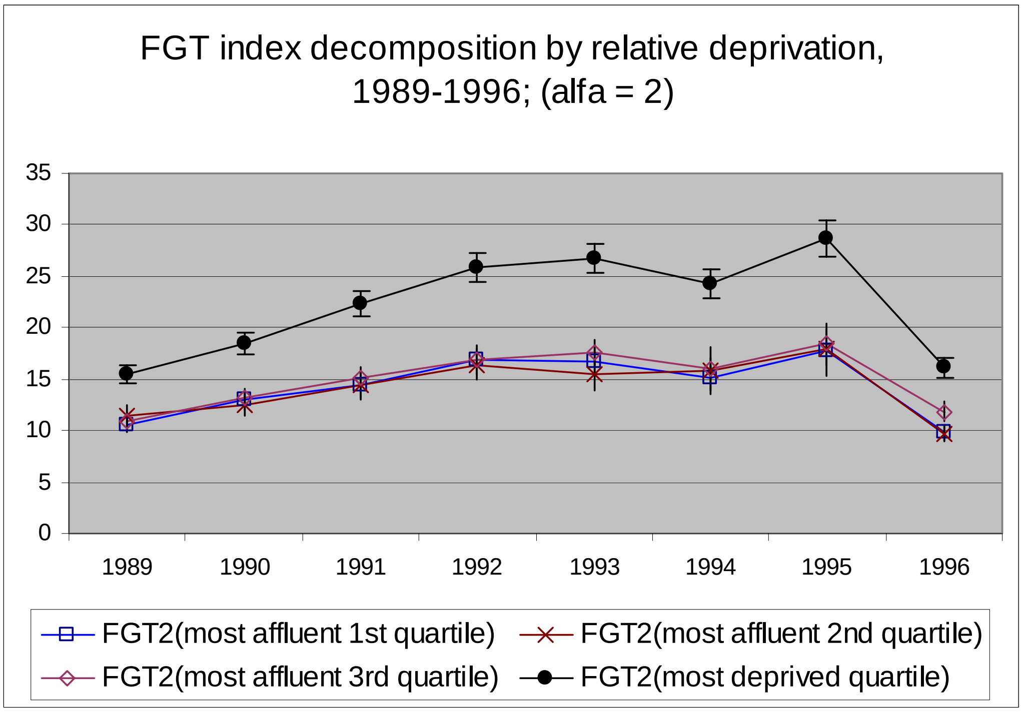


Figure 4. HCD profiles by presence of FHs in most socio-economically deprived quartile of electoral wards.

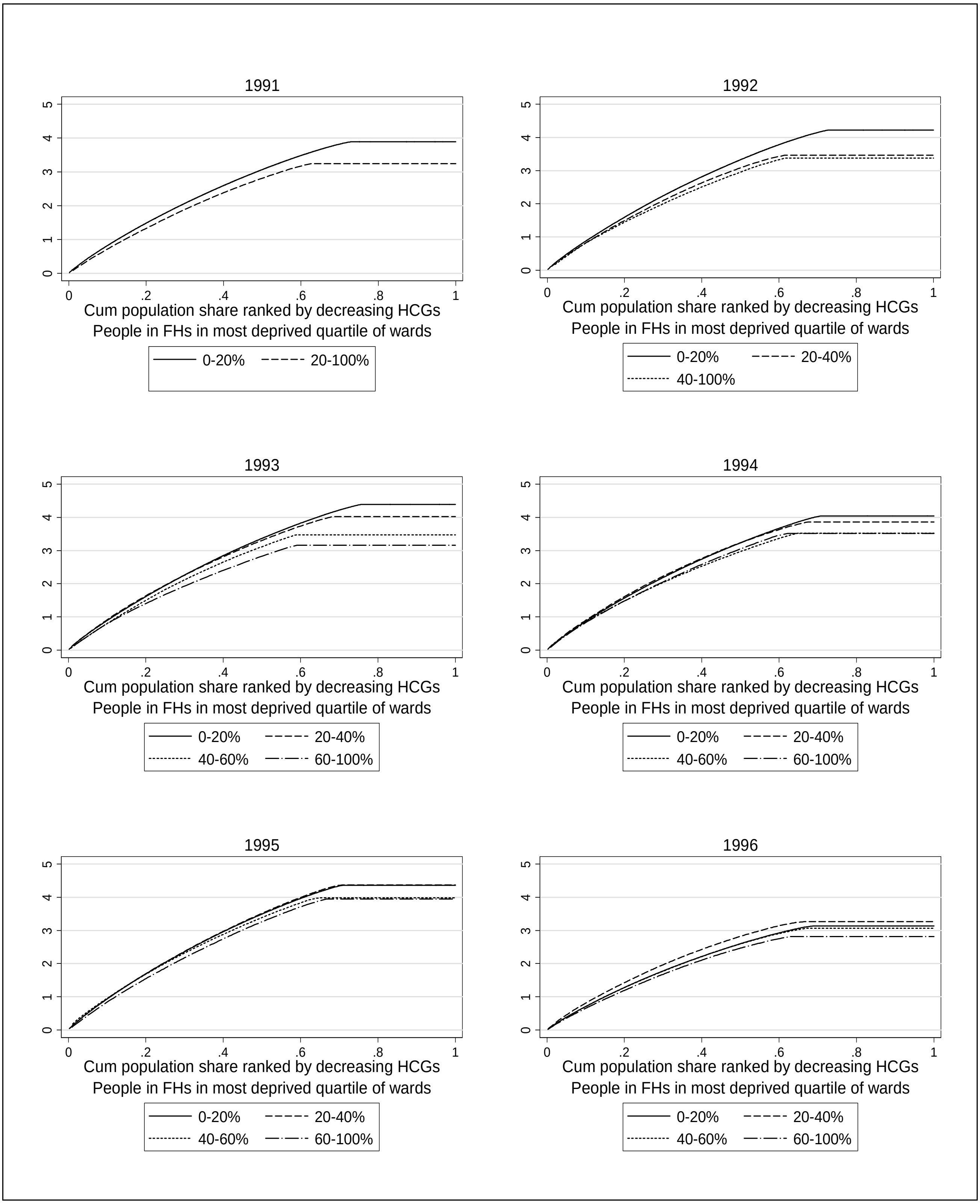


Figure 5. Relative FGT index decomposed by presence of FHs (average inequity in the total population = 1; social preference parameter α = 1).

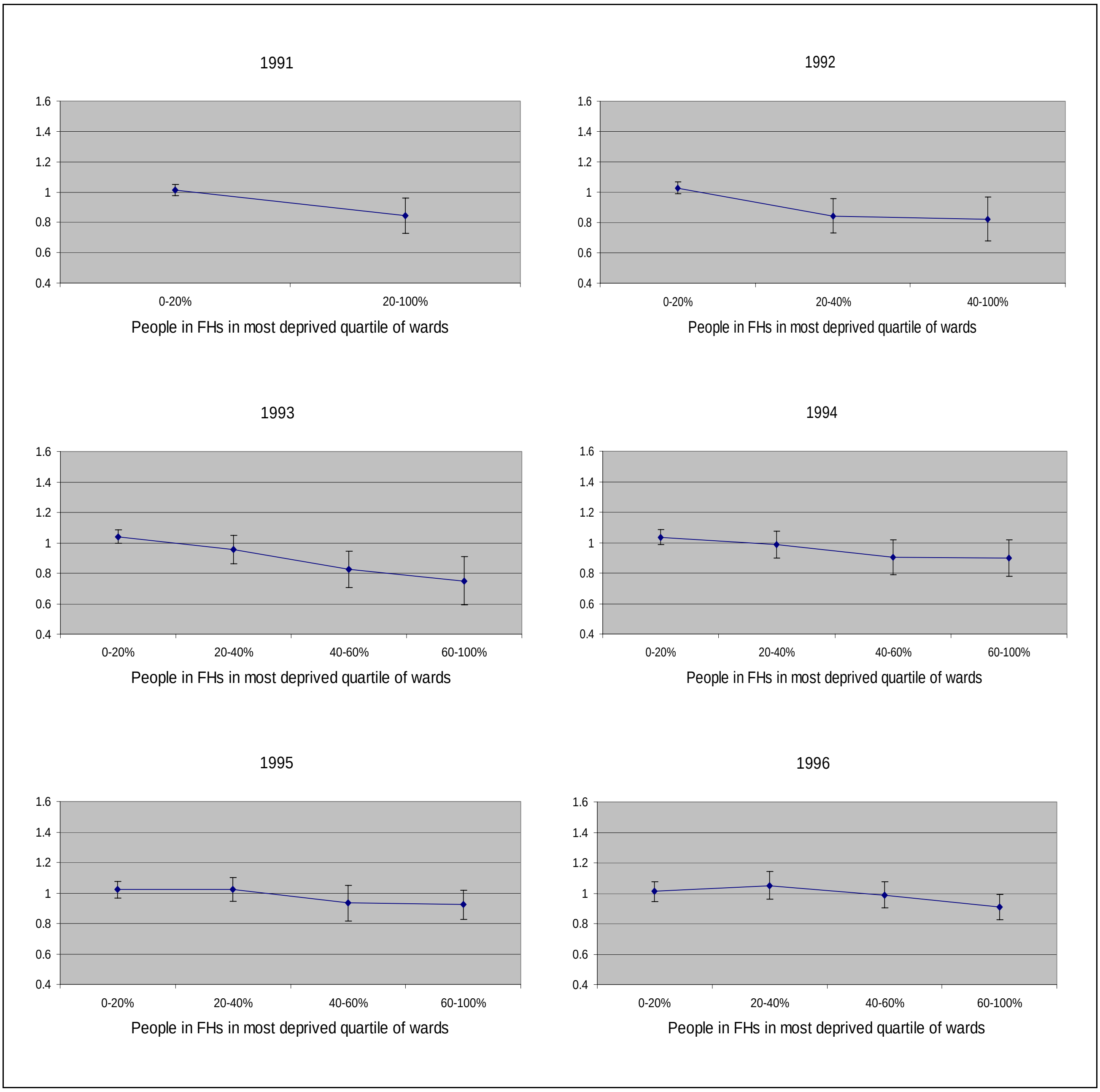
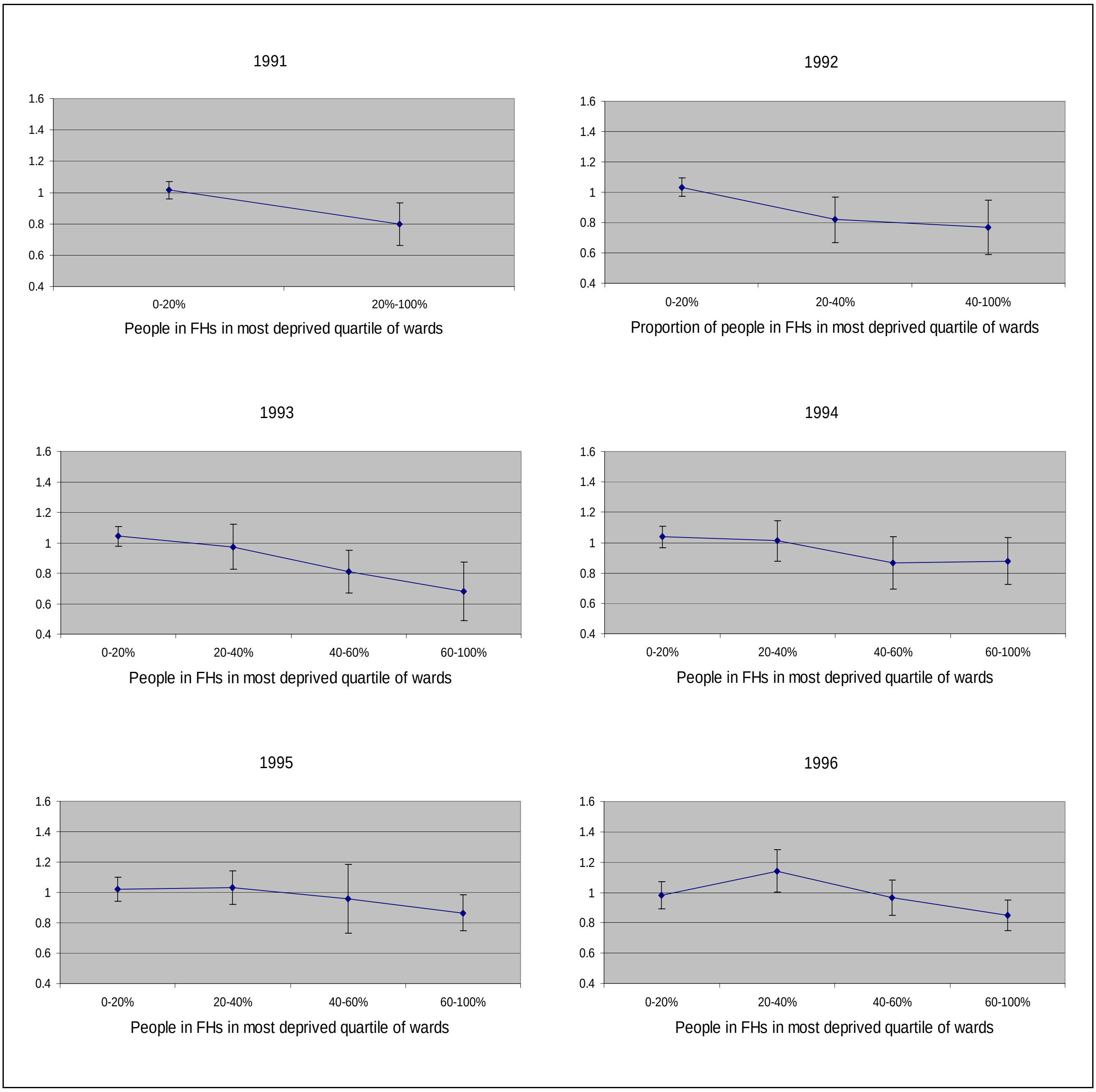


Figure 6. Relative FGT index decomposed by presence of FHs (average inequity in the total population = 1; social preference parameter α = 2).



Appendix A.

Proof of Proposition 1:

$(a)\ x_i \equiv \max\left(h_i^* - h_i; 0\right) = 0 \quad \forall i \in N \quad \Leftrightarrow \quad \text{(b) } h_i^* = h_i \quad \forall i \in N$

The proof is by contradiction. Assume (a) holds and condition (b) does not: $h_i^* \neq h_i$, then: $\exists i : h_i^* > h_i\ , i \in N$ cannot hold since violate assumption (a), then we just have to disproof: (c) $\exists i : h_i^* < h_i\ , i \in N$

From condition (c) and the definition of $h*$ in Equation (2) and (3):

$\Rightarrow\ \exists i:\ h_i^* \equiv \sum_j p_{ij} \frac{h_j}{p_j} < h_i$, where: $h_j = \sum_i h_{ij}$ and $p_j = \sum_i p_{ij}$, $j \in J$ need classes

Consider $J = 1$

$\Rightarrow\ \exists i:\ p_i \frac{h}{p} < h_i \quad \Rightarrow \quad \frac{h_i}{p_i} > \frac{h}{p} \quad \Rightarrow \exists i':\ p_{i'} \frac{h}{p} > h_{i'}$ since $h = \sum_i h_i$ and $p = \sum_i p_i$

$i, i' \in N$ (in fact, x_i is just a deviation from the mean!); this violates assumption (a) $\Rightarrow$ (b) $h_i^* = h_i\ \forall i \in N$

Consider $J > 1$

For simplicity assume that $\frac{p_{ij}}{p_j}$ is constant across all the area i (but not between the J need classes!) and normalize the total population $p = \sum_i p_i = 1$. Then, from condition (c): $\Rightarrow\ \exists i:\ h_i^* \equiv \sum_j \frac{h_j}{p_j} < h_i$ since $\frac{p_{ij}}{p_j} = \frac{1}{p_j}$ for every area *i*.

$$\Rightarrow\ \exists \varepsilon_i > 0:\ \sum_j \frac{h_j}{p_j} = \sum_j h_{ij} + \sum_j \varepsilon_{ij} \quad , \varepsilon_i = \sum_j \varepsilon_{ji} \qquad \text{(d)}$$

ε_i is the sum of all the residuals between observed and need expected health care utilization for all the J need class in the area i: $\varepsilon_{ji} = \frac{h_j}{p_j} - h_{ij}\ \ , \forall j \in J$. Notice that ε_{ji} can be positive or negative under condition (b) and (c). However, the sum $\varepsilon_i = \sum_j \varepsilon_{ji}$ must be positive under condition (c) and would be zero under condition (b).

Expression (d) can be rewritten as:

$$\sum_{j}^{J-1}\frac{h_j}{p_j}=\sum_{j}^{J-1}h_{ij} \quad ,\forall j\neq j' \text{ and } \frac{h_{j'}}{p_{j'}}=h_{ij'}+\varepsilon_i \quad , j'\in J$$

$$\Rightarrow \quad \exists i': \sum_{j}^{J-1}\frac{h_j}{p_j}=\sum_{j}^{J-1}h_{i'j} \quad ,\forall j\neq j' \text{ and } \frac{h_{j'}}{p_{j'}}=h_{ij'}+\varepsilon_{i'} \quad , j'\in J \text{ , but now } \varepsilon_{i'}=\sum_j \varepsilon_{ji'}<0$$

since: $h_{j'}=\sum_i h_{ij'}$ and $p_{j'}=\sum_i p_{ij'} \quad i,i'\in N$

this violates assumption (a) $\Rightarrow$ (b) $h_i^*=h_i \ \forall i\in N$.

The reverse proof of proposition one is trivial:

(b) $h_i^*=h_i \ \forall i\in N \Rightarrow (a)\ x_i\equiv \max\left(h_i^*-h_i;0\right)=0 \ \ \forall i\in N$ by definition of x_i.

Appendix B.

Proof of Proposition 2:

(a) $x_i\equiv \max\left(h_i^*-h_i;0\right)=0 \ ,\forall i\in N \ \Leftrightarrow$ (b) $\sum_j\left(\frac{p_{ij}}{p_j}\frac{h_j}{H}\right)=\frac{h_i}{H} \ ,\forall i\in N$

Assume (a):

$x_i\equiv \max\left(h_i^*-h_i;0\right)=0 \ \ \forall i\in N \ \Leftrightarrow \ h_i^*=h_i \ \forall i\in N$ from Proposition 1.

$\Rightarrow \ h_i^*\equiv \sum_j p_{ij}\frac{h_j}{p_j}=h_i$ from Equation (2) and (3).

$\Rightarrow \ \sum_j\left(\frac{p_{ij}}{p_j}\frac{h_j}{H}\right)=\frac{h_i}{H} \ ,\forall i\in N$ dividing both sides by H.

The reverse proof of proposition one is trivial:

$\sum_j\left(\frac{p_{ij}}{p_j}\frac{h_j}{H}\right)=\frac{h_i}{H} \ ,\forall i\in N \ \Rightarrow \ h_i^*=h_i \ ,\forall i\in N$

$\Rightarrow x_i\equiv \max\left(h_i^*-h_i;0\right)=0 \ ,\forall i\in N$ by definition of x_i.

Appendix C.

Proof of Proposition 3: If $x_i \equiv \max\left(h_i^* - h_i; 0\right) > 0 \;\Rightarrow\; \sum_j \left(\frac{p_{ij}}{p_j}\frac{h_j}{H}\right) \neq \frac{h_i}{H}$

$x_i \equiv \max\left(h_i^* - h_i; 0\right) > 0 \;\; i \in N \quad \Rightarrow h_i^* > h_i$ by definition of x_i

$\Rightarrow \; h_i^* \equiv \sum_j p_{ij} \frac{h_j}{p_j} > h_i$ from Equation (2) and (3).

$\Rightarrow \; \sum_j \left(\frac{p_{ij}}{p_j}\frac{h_j}{H}\right) > \frac{h_i}{H} \;, \forall i \in N$ dividing both sides by H.