I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14925

Model-Based Recursive Partitioning to Estimate Unfair Health Inequalities in the United Kingdom Household Longitudinal Study

Paolo Brunori
Apostolos Davillas
Andrew M. Jones
Giovanna Scarchilli

DECEMBER 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14925

Model-Based Recursive Partitioning to Estimate Unfair Health Inequalities in the United Kingdom Household Longitudinal Study

Paolo Brunori
LSE and University of Bari

Andrew M. Jones
University of York

Apostolos Davillas
University of East Anglia and IZA

Giovanna Scarchilli
University of Trento and University of Modena and Reggio Emilia

DECEMBER 2021

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Model-Based Recursive Partitioning to Estimate Unfair Health Inequalities in the United Kingdom Household Longitudinal Study*

We measure unfair health inequality in the UK using a novel data-driven empirical approach. We explain health variability as the result of circumstances beyond individual control and health-related behaviours. We do this using model-based recursive partitioning, a supervised machine learning algorithm. Unlike usual tree-based algorithms, model-based recursive partitioning does identify social groups with different expected levels of health but also unveils the heterogeneity of the relationship linking behaviors and health outcomes across groups. The empirical application is conducted using the UK Household Longitudinal Study. We show that unfair inequality is a substantial fraction of the total explained health variability. This finding holds no matter which exact definition of fairness is adopted: using both the fairness gap and direct unfairness measures, each evaluated at different reference values for circumstances or effort.

JEL Classification: I14, D63

Keywords: inequality of opportunity, health equity, machine learning, unhealthy lifestyle behaviours

Corresponding author:
Apostolos Davillas
Health Economics Group
Norwich Medical School
University of East Anglia
Norwich, NR4 7TJ
United Kingdom
E-mail: a.davillas@uea.ac.uk

* We are grateful to participants of the 2nd APHEC Workshop held at University of Genova for valuable comments and suggestions. An early version of this manuscript benefited from comments of participants at the Conference on Opportunities, Mobility and Well-being, jointly organised by the Institute of Economics Polish Academy of Sciences and ifo Institute in Munich, held online on April 2021. We also thank Achim Zeileis for useful clarifications about the MOB algorithm he provided us. Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service. The funders, data creators and UK Data Service have no responsibility for the contents of this paper.

1 Introduction

According to Fleurbaey and Schokkaert (2009), differences in health status can originate from either fair or unfair sources. They argue that unfair health inequalities are differences in health status determined by circumstances beyond individual control such as sex, ethnicity or socioeconomic background in childhood. Under this distinction a society that wishes to eliminate unfair health inequality should compensate individuals suffering a poorer health status due to unfavourable biological, social and economic circumstances in childhood. On the contrary, a society may not want to compensate individuals for differences in their health that arise from choices and behaviours they can control and are held responsible for. This conception is not new in egalitarian theory. The idea that fairness can be achieved by removing inequality due to circumstances while letting individuals facing the rewards and costs of their responsible choice is rooted in the moral philosophical literature and in the economic social justice theory: see among others Cohen (1989); Dworkin (1981); Fleurbaey (1995, 2008); Rawls (1958, 1971); Roemer (1998); Sen (1980). The distinction between legitimate and illegitimate sources of inequality is well established in the health economics literature, in particular through the distinction between need-related and non-need-related variation in defining equity in the use of health care (Wagstaff and Van Doorslaer, 2000).

Merging the goals of equality and individual responsibility, Fleurbaey and Schokkaert (2009) drew on two distributive principles to be met in order to realize a fair distribution of health: *reward* and *compensation*. When both principles are satisfied, on the one hand, individuals characterized by identical circumstances face the benefits and the costs of their choices, on the other, individuals behaving in the same way all achieve the same health status independently from their circumstances.[1] In this perspective these two principles define a fair distribution of health, measuring unfair inequality in health means to measure violations of both principles: an ideal measure of unfair inequality should be sensitive to inequality within individuals who make the same choices (compensation) and should also be insensitive to any inequality observed between individuals characterized by the same circumstances who make different choices (reward). The first property captures horizontal equity, with respect to effort, and the second reflects judgements about vertical equity in the reward for effort.

A possible empirical approach to measuring unfair inequality consists of deriving a counterfactual distribution that fully reflects only these unfair inequalities and then applying a suitable inequality index to that distribution. However, Fleurbaey (2008) has discussed the impossibility of constructing a distribution which is consistent with both principles, unless the effects of choices and circumstances are independent from each other; that is, the process generating health is additively separable in circumstances and choices. In the general case, to solve

[1] In what follows we consider the terms 'unfair health inequality' and 'inequality of opportunity in health' as if they were interchangeable. Roemer and Trannoy (2015) discuss the near perfect overlap of the two definitions.

this incompatibility problem, Fleurbaey and Schokkaert (2009) proposed two families of measures of health equity. Each of these is fully consistent with only one principle, reward or compensation, and partially satisfies the other principle at some reference value. The two measures are the *direct unfairness*, fully consistent with the reward principle and only partly consistent with the compensation principle, and the *fairness gap* which fully satisfies the compensation principle but is partly inconsistent with the reward principle. In practice, these measures parallel the concepts of direct and indirect standardisation used in the measurement of equity in the use of health care (Wagstaff and Van Doorslaer, 2000).[2]

In this paper we implement the Fleurbaey and Schokkaert (2009) measurement approach using an innovative statistical tool, Model-based recursive partitioning (MOB). MOB is a tree-based supervised learning algorithm developed by Zeileis et al. (2010) and its use to measure unfair inequalities contributes to the growing methodological literature that uses data-driven techniques in the study of inequality of opportunity (Brunori et al., 2019; Carrieri et al., 2020; Li Donni et al., 2015). These data-driven techniques offer a compromise between the data-hungry nonparametric approach, which partitions the sample into all unique combinations of circumstances and, hence, often suffers from a curse of dimensionality, and the parametric approach which assumes that the relationship between observed circumstances and the outcome can be captured by a linear regression model. Tree-based approaches allow the selection of relevant circumstances, and the way that they interact with each other, to be data-driven.

The model we adopt allows the relationship between health outcomes and health-related behaviours (effort) to be estimated, allowing it to vary according to circumstances that are beyond individual control. The MOB algorithm first estimates a parametric link between health status and lifestyle on the entire sample. Then recursively tests whether partitioning the population based on circumstances and re-estimating the model on population sub-samples can reject the null hypothesis of parameters' stability and obtain a better interpolation of the data. The output of the MOB algorithm is a partition of the sample into socioeconomic groups that are homogeneous in terms of their circumstances, what Roemer (1998) calls "types". Such groups are heterogeneous both in terms of expected health and in terms of the relationship between health-related behaviours and the health outcome. This machine learning approach to estimate health inequalities represents an innovative contribution to the literature and, provided that proxies for relevant responsibility variables are observed, could be straightforwardly extended to other welfare domains such as education or income.

We apply the MOB algorithm to estimate the level of unfair health inequality. We base our estimate on the nationally representative UK Household Longitudi-

[2]This literature recognises the importance of reference values, embodied in the notion that "on average the system gets it right", and the implied tension between measuring horizontal and vertical inequity with respect to need (Gravelle, 2003; Sutton, 2002; Wagstaff and Van Doorslaer, 2000).

nal Study (UKHLS) to present estimates of the two unfair inequality measures introduced by Fleurbaey and Schokkaert (2009): direct unfairness and the fairness gap. We show that unfair inequality is a substantial fraction of the total explained health variability. This finding holds no matter which exact definition of fairness is adopted: using both the fairness gap and direct unfairness measures. These are evaluated at different reference values across the full distributions of types and of degrees of effort.

The paper is structured as follows, in Section 2 the metrics proposed by Fleurbaey and Schokkaert (2009) are introduced. Section 3 explains how the MOB algorithm can be used to estimate unfair inequalities. Section 4 presents the data and the empirical results. Section 5 concludes.

2 Fleurbaey-Schokkaert model and measures

Consider a population of N individuals over which a distribution of the health outcome H is defined. We assume that individual health is determined by three types of traits: a finite set of lifestyle related factors over which individuals have control ($\mathbf{E}$), which are called "effort" variables, a set of social factors for which individuals cannot be held responsible ($\mathbf{C}$), which are called "circumstances", and age ($\mathbf{A}$). We use an age-adjusted measure of health so we can abstract from $\mathbf{A}$. The individual health outcome is generated by a function of circumstances and effort variables:

$$H = g(\mathbf{C}, \mathbf{E}) \tag{1}$$

All the possible combinations of circumstance values, taken one at a time from $\mathbf{C}$, define a partition of the population into *types*. Individuals belonging to the same type are characterized by identical circumstances. Similarly, all the possible combinations of values taken one at time from $\mathbf{E}$ define a partition of the population into *tranches*. Individuals belonging to the same tranche exert exactly the same effort.

An important normative and empirical issue concerns the definition of the responsibility variables. While Fleurbaey and Schokkaert (2009) do not explain how responsible choices can be measured, considering it a normative choice that belongs to the political decision-maker, John Roemer goes a little further suggesting that the degree of effort exerted must always be orthogonal to circumstances. In Roemer's view, if individuals belonging to different types face different incentives and constraints in exerting effort, this is to be considered a characteristic of the type and should be included among circumstances beyond individual control.

For example, consider the frequency of eating fruit as a measure of effort. An individual with more educated parents may find it much easier to eat regularly fruit, while an individual who grew up in a less favourable environment may find it harder to eat fruit and avoid junk food. Roemer believes that the distribution of effort is, indeed, a characteristic of the type:

> "Thus, in comparing efforts of individuals in different types, we should somehow adjust for the fact that those efforts are drawn from distributions which are different, a difference for which individuals should not be held responsible."
>
> Roemer (2002) p. 458

Roemer therefore distinguishes between the 'level of effort' and the 'degree of effort' exerted by an individual. The latter is the morally relevant variable of effort and is identified with the quantile of the effort distribution for the type to which the individual belongs. In the example of effort exerted by an individual, the relevant measure is not the number of fruit portions eaten but rather the quantile of the type-specific distribution of fruit portions eaten.[3] Other authors have suggested that when measuring unfair health inequality individuals should be held fully responsible for their choices (see Roemer and Trannoy (2015) for a discussion). However, following the prevalent approach in this literature we will define the degree of effort exerted consistently with Roemer's proposal (the empirical difference between the two approaches is discussed by Jusot et al. (2013)).

In our model, health is determined solely by observable circumstances and effort. We are therefore ignoring health variability within cells, groups of individuals sharing the same observed efforts and circumstances. Empirically we easily observe individuals sharing the same circumstances and exerting the same effort, but obtaining a different health outcome. How then should we consider such unexplained variation? Is it more likely that this inequality arises from unobservable effort or unobservable circumstances? Is it simply the randomness inherent in many health outcomes? Or is it a reflection of measurement error which is convenient to ignore, that is replacing all outcomes in the cell with their mean? The answer depends on our beliefs about the observability of circumstances and effort; Lefranc et al. (2009) consider within-cell inequality to be due to randomness or "luck", a source of unfair inequality. On the contrary, the majority of the empirical studies of income inequality consider variation within cell as due to effort. Checchi and Peragine (2010), for example, claim that this inequality is due to limited observability of effort and therefore should be attributed to effort.

In what follows we explicitly recognize that, to a large extent, health variability cannot be predicted by observable variables. We focus solely on the part of the limited health variability that can be predicted by observable circumstances and efforts and are agnostic about the unexplained variation. We will assign to each individual in type k exerting effort j the average outcome of cell k, j. To evaluate whether within-cell inequality is or is not to be considered unfair health inequality is beyond the scope of this approach.

[3] An alternative way of addressing this issue, purging the influence of circumstances on effort, is to replace the observed level of effort with the residuals from a regression of effort on circumstances (e.g., Carrieri et al. (2020); Jusot et al. (2013)).

Using this framework Fleurbaey and Schokkaert (2009) have proposed two types of measure to quantify Unfair Inequality (UI)[4]. To quantify UI the authors suggest a two-step method: first, starting from a distribution of health outcome (H), a counterfactual distribution ($\tilde{H}$) is derived, which reproduces only unfair inequality and does not reflect any inequality arising from choice and effort of individuals; second, inequality is measured for this counterfactual distribution.

In order to construct a measure of inequality in health that is sensitive to the problem of responsibility, Fleurbaey and Schokkaert (2009) present two conditions:

> Condition 1 (Reward, no influence of legitimate differences). A measure of unfair inequality should not reflect legitimate variation in outcomes, i.e. inequalities which are caused by differences in the responsibility variable.
>
> Condition 2 (Compensation). If a measure of unfair inequality is zero, there should be no illegitimate differences left, i.e. two individuals with the same value for the responsibility variable should have the same outcome.
>
> Fleurbaey and Schokkaert (2009) p. 75.

Putting together both of these requirements, we can state that a counterfactual distribution consistent with the compensation and the reward principles is a distribution that:

1) fully reflects the inequality in outcomes between individuals with the same effort (within-tranche inequality);

2) does not reflect any inequality in outcomes between individuals characterized by same circumstances (within-type inequality).

Any inequality measure applied to such distribution would be a measure of unfair inequality consistent with both the reward and the compensation principle. Fleurbaey and Schokkaert (2009) address the potential conflict between the principles of compensation and reward. They propose two UI measures, each one fully consistent with one of the two principles and maintaining consistency with the other at a reference degree of effort or a reference type, respectively:

> *Direct unfairness (UI_{DU})*: choose a reference value for the the vector of responsibility variables $\tilde{E}$, with $\tilde{h}_i^{k,j} = g(C, \tilde{E})$. In the counterfactual distribution the health of an individual i belonging to type k is the health

[4]Their proposal originates from a number of contributions on fair allocation and distributive justice (Fleurbaey, 2008; Fleurbaey and Maniquet, 2012). In these contributions the authors developed a theory of "responsibility-sensitive egalitarianism" whose ambition is to generalize the egalitarian ideal allowing individuals to be held responsible, to some degree, for their achievements.

attained by an individual in type k that exerts the reference degree of effort. Inequality in the counterfactual distribution, $\tilde{H}_{DU}$, is unfair inequality.

Fairness gap (UI_{FG}): choose a reference type $\tilde{C}$, with $\tilde{h}_i^{k,j} = g(\tilde{C}, E)$. Then $\tilde{H}_{FG}$ is obtained by taking the difference between the individual's health in the initial distribution and the health of individuals who exert the same effort but who have the reference circumstances. Unfair inequality is inequality in $\tilde{H}_{FG}$.[5]

UI_{DU} measures inequality in a counterfactual distribution obtained by removing any inequality due to effort. All individuals belonging to the same type have the same value in $\tilde{H}_{DU}$. Hence UI_{DU} is a measure of unfair inequality fully consistent with the principle of reward (no influence of legitimate differences). On the other hand, UI_{DU} is consistent with the principle of compensation for the reference degree of effort: if all individuals with the reference level of effort obtain the same outcome inequality in $\tilde{H}_{DU}$ is zero. However, UI_{DU} fails to satisfy the principle of compensation for all other effort tranches.

Symmetrically, UI_{FG} measures inequality in a counterfactual distribution obtained by isolating inequality within tranches. It is a measure fully consistent with the principle of compensation: inequality in $\tilde{H}_{FG}$ is zero only if all individuals in the same tranche obtain the same outcome. Moreover, UI_{FG} is consistent with the principle of reward for the reference circumstance; UI_{FG} is insensitive to changes in inequality within individuals characterized by reference circumstances. However, UI_{FG} fails to satisfy the principle of reward for individuals not belonging to the reference type.[6]

Summing up, we can estimate two sets of measures: compensation consistent measures (UI_{FG}), and reward consistent measures (UI_{DU}). These measures depend on either a reference effort or a reference combination of circumstances therefore we estimate a range of measures and we discuss their sensitivity to different reference values.

[5]This index is equivalent to the measure of horizontal equity, based on indirect standardisation, that is typically used in the literature on equity in the delivery of health care (Wagstaff and Van Doorslaer (2000))

[6]Note that these measures differ from the ex-ante and ex-post inequality of opportunity measures inspired by Roemer (1998) and often adopted in empirical studies (Checchi and Peragine, 2010; Roemer and Trannoy, 2015). Ex-ante UI is a reward-consistent measure of UI obtained imposing: $\tilde{h}^{k,j} = \lambda_k = \mu_k$, where μ_k is the average outcome of individuals in type k (see Property 1). Ex-post UI is a compensation-consistent measure of UI obtained imposing: $\gamma_j = \mu_j$, where μ_j is the average outcome of individuals in tranche j (see Property 2). Ex-ante and ex-post UI fail to satisfy both the principle of compensation and the principle of reward respectively, unless g is additively separable in E and C. However, because they are relatively easier to estimate and to decompose, they are very popular in the empirical literature about inequality of opportunity in income and consumption as well as applications to health inequality (Davillas and Jones, 2021; Jusot et al., 2013; Rosa Dias, 2009).

3 Empirical definition of UI_{DU} and UI_{FG} using Model-based Recursive Partitioning

Estimation of UI_{DU} and UI_{FG} requires relevant circumstances beyond individual control to be observed and types to be defined. Ideally, a measure of unfair inequality should consider all the potential sources outside individual control. However, this would require considering a wide and complex set of circumstances, which brings with it the risk of noisy and upwardly biased estimates (Brunori et al., 2019). Traditionally, in empirical studies on unfair inequalities the relevant circumstances have been included in the model through normative decisions. In the nonparametric approach the population is partitioned into a parsimonious number of types and in the parametric approach the relationship between circumstances and the outcomes have been implicitly modelled as additive and fixed using linear regression. For these reasons, coupled with the fact that some circumstances may be unobserved, estimates have been interpreted as a *lower-bound* estimate of the real level of unfair inequality (Carrieri and Jones, 2018; Jusot et al., 2013; Li Donni et al., 2015; Rosa Dias, 2009).

A number of more recent empirical applications instead rely on data-driven semiparametric techniques to explore the information on social groups which is relevant to the formation of unfair inequalities. These are semiparametric in the sense that relationship between health outcome and effort is assumed to take a (linear) parametric form, while the definition of types is nonparametric. On one side, finite mixture models (FMM)[7] have been adopted to study the latent type membership of each individual given their observed circumstances (Brunori et al., 2021; Carrieri et al., 2020; Li Donni et al., 2015). The FMM approach relies on an *a priori* selection of the circumstance variables that influence the probability of belonging to each type. On the other side, tree-based methods have been adopted to perform a data-driven selection of the relevant circumstances and the interactions between them on the basis of model fit (Brunori and Neidhöfer, 2020; Brunori et al., 2018). The estimation approach proposed in this paper, model-based recursive partitioning (MOB), is an extension of the tree-based techniques applied with a specification of types that echoes the semiparametric mixture approach (Carrieri and Jones, 2018; Carrieri et al., 2020).

Consider again equation (1): individual health outcomes, h_i, are attributed to two sets of observable variables: a number of behaviours and a set of circumstances for which individuals are not held responsible, respectively **E** and the **C**. The isolation of the unfair health inequality requires the estimation of a model for health. For the sake of simplicity, and following Carrieri and Jones (2018), assume that behaviours can be summarized by a scalar index of lifestyle (e) and that its effect on health can be modelled using a linear regression:

$$h_i = \beta_0 + \beta_1 e_i + \epsilon_i \tag{2}$$

We can assume that this simple relationship is not independent from **C**.

[7] Mixture models in statistics are a broad family of probabilistic models for observing latent subgroups in a population, including latent class analysis (LCA) as a specific case.

The relationship linking efforts and health can be affected by the circumstances though two channels: the intercept, β_0, and the slope, β_1.[8] A different intercept can be interpreted as the direct contribution of circumstances to health: independently from the choices made having favourable circumstances may improve individuals' health. Heterogeneity in the slope instead means that the contribution of lifestyle to health outcomes may be also affected by circumstances.

The final model can be represented as a weighted sum of sample splits performed to derive $k = 1, ..., K$ different models associated with each subgroup parameters $\beta_{(k)}$:

$$g(h_i|\mathbf{c_i}, e_i, \boldsymbol{\beta_{(1)}}, ..., \boldsymbol{\beta_{(K)}}) = \sum_{k=1}^{K} \pi_k(\mathbf{c_i}) \cdot g(h_i|e_i, \boldsymbol{\beta_{(k)}}) \tag{3}$$

Note that this representation of the individual health model as a function of efforts and circumstances can be either associated with both the FMM and the MOB approaches to estimation. Depending on which of the two methodologies is chosen, the weight $\pi_k(c_i)$ and the K subgroups will be identified with a different estimator.

We opt for the use of the MOB to estimate the indirect relation between circumstances and behaviours, and to allow the health response to effort be estimated varying across meaningful social groups. Tree-based techniques are data-driven and rely on *decision trees* which, in statistics, can be used to visually represent the "decisions", or if-then rules, that are used to generate predictions of a single outcome variable or a model. Moreover, tree-based methods tend to be more parsimonious then FMM in terms of parameters resulting in less conservative (more fine grained) partitions in types. There are essentially two key components to build a decision tree: the *features* to split on the prediction sample, and the *rule* to stop splitting the sample. The MOB is a particular tree-based method which takes as input a set of partitioning variables and whose splitting rule relies on the estimated parameters of a model.

This model is initially estimated on the entire sample, afterwards, a statistical test is performed to verify whether there are any possible sample splits on the partitioning variables which achieve a better fit of the model. The outcome of this process is a set of models estimated on K sub-samples of the original population (terminal nodes).

We briefly summarize here how a MOB is obtained from data (see Zeileis and Hornik (2007); Zeileis et al. (2008) and Zeileis et al. (2010) for details). The MOB uses the vector $\mathbf{C}$ to search for ways of splitting the sample into non-overlapping subgroups. If estimating the response of health to lifestyle into two sub-samples yields statistically different parameters and improve out-of-sample

[8] In the empirical application we consider higher order polynomials for effort, with the chosen specification selected by cross validation. So, although this is the parametric part of the specification, the estimation does allow for a considerable degree of flexibility. Note also that the MOB specification allows for interactions with circumstances through the heterogeneity of parameters across types.

prediction, then the split is performed. The procedure is then repeated in the resulting sub-samples.

The parameter instability is detected by means of Generalised M-fluctuation tests. The test is based on a partial sum process of the estimation scores which captures instabilities (Hothorn and Zeileis, 2015; Zeileis and Hornik, 2007). It can be understood as a generalization of the type of test used to detect structural breaks in time series analysis. In the case of the MOB algorithm, the test is performed on the partial sum of residuals across the space defined by partitioning variables. The fluctuation test statistic is distributed as a χ^2 and we can compute the Bonferroni-adjusted p-value for testing its significance. If the fluctuation test statistic is higher than a certain threshold, the hypothesis of stability of the model parameters is rejected and algorithm splits the sample and re-estimates the model on the distinct subgroups.

Schematically, Zeileis et al. (2010) illustrate the steps of the MOB algorithm as follows:

1. Set a confidence level $(1 - \alpha)$ to be used as tuning parameter;
2. Fit the model - for example: $h_i = \beta_0 + \beta_1 e_i$ - on the entire sample;
3. Test whether there is any partitioning variable causing parameter estimates for the model to be unstable;
4. If the nulla hypothesis of parameters stability across possible sub-sample cannot be rejected stop;
5. If the $p - value$ of the fluctuation test statistics is instead lower than the critical Bonferroni-adjusted α, select the variable associated with the most statistically significant source of instability, otherwise stop;
6. Compute the exact splitting point which optimises the objective function of the estimation according to the selected partitioning variable;
7. Split the node into child nodes and restart the procedure from (2) on the two subsamples.

The depth of the estimated tree depends on the tuning parameter α which determines the $p - value$ threshold for rejecting the null hypothesis in the instability test. The value of α can be set to a specific value or can be selected by a machine-learning technique ensuring that MOB stops splitting the sample when no further split would result in a better out-of-sample fit of the data.

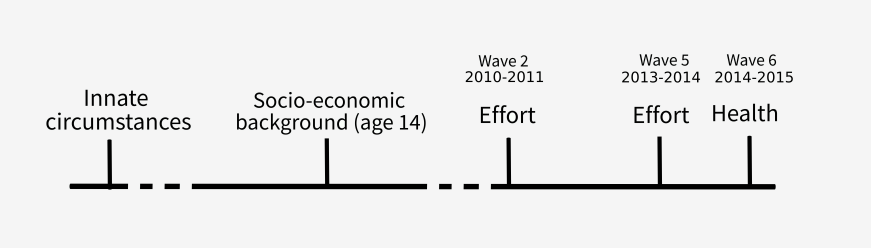
The outcome of the algorithm is a partition of the population into types according to the composition of the terminal nodes. Individuals belonging to each type share the same circumstances and the same parameters for equation (2). The partition into types and the associated set of parameters allows the counterfactual distributions $\tilde{H}_{DU}$ and $\tilde{H}_{FG}$ to be computed. The counterfactual distribution $\tilde{H}_{DU}$ is obtained by choosing a reference degree of effort $\tilde{e}$ and then predicting $\tilde{h}_i^{k,j} = \hat{\beta}_0^k + \beta_1^k \tilde{e}$. The counterfactual distribution $\tilde{H}_{FG}$ is obtained by

choosing a reference type (R) and then predicting $\tilde{h}_i^{k,j} = (\hat{\beta}_0^k + \beta_1^k e_j) - (\hat{\beta}_0^R + \beta_1^R e_j)$. UI_{DU} and UI_{FG} are then obtained by computing a suitable inequality measure of the counterfactual distributions.

4 Data and estimates

The data comes from three waves of the UK Household Longitudinal Study (UKHLS). UKHLS contains information about demographic characteristics, a rich set of information about individuals socioeconomic background in childhood, ethnicity, and place of birth among other things. These provide our measures of circumstances that are used to construct types by the MOB algorithm. Moreover, the survey contains questions about health-related behaviours, that are used to construct the scalar index of lifestyle, and a number of measures of health outcomes. Figure 1 shows the study design and indicates at what moments in time and to which waves the observations of the different variables used in the analysis correspond. Circumstances relate to fixed individual characteristics and to measures of parental background, health-related behaviours are measured at Waves 2 and 5, and the health outcomes are measured in the subsequent follow-up at Wave 6.

Figure 1: Timeline for the study design



Note: Circumstances may be observed in multiple waves.

Our chosen health outcome (H) is measured at UKHLS Wave 6 (2014-2015). We use the Short Form 12 (SF-12), a well validated, self-administered health measure based on a set of 12 questions on respondent's health (Ware et al, 1995). For this study, we use the Physical Component Score (PCS-12), to capture respondents' physical health. The PCS-12 score has values between 0 and 100, and it has been standardized in order to have a mean of 50 and a standard deviation of 10; higher values indicate better physical health functioning. The PCS-12 is a reliable instrument developed to measure physical health in large surveys with higher values of sensitivity and specificity compared to other brief health scales (Ware et al., 2001; Ziebarth, 2010). It has been used in the literature as a robust self-reported measure of physical health (e.g., Eibich (2015); Guber (2019); Schmitz (2011); Ziebarth (2010)). The health measure has been adjusted for individual age (at the time of the interview) in order to control

for the age-specific variability in health. The age-adjustment is performed by regressing individual health status on 5-year age classes between 14 and 100. To remove all the age-class fixed effects from total health variability we use the residuals as our measure of health status.

The full set of observed circumstances (C) beyond individual control that are considered as candidate variables in the MOB algorithm are: ethnic groups (the relevant categories have been summarised into the following levels: UK white; Irish white; other white; mixed: white with Asian/African/Arab; Asian: East and Middle East; Black: African, Caribbean, other; other ethnic groups), place of birth (a dichotomous variable indicating whether born in the UK or not), father and mother's skill levels in the main occupation (unemployed or four skill levels in occupation), mother and father's education (did not go to school, left school without qualifications, some qualification, post-school qualifications, university degree or higher), mother and father's activity status (working, unemployed, deceased, not living in the household). Note that all information about parents relate to when the respondent was 14 years old. We include sex as an additional source of unfair health inequality. The tree structure implicit in the MOB algorithm allows for a full set of interactions between the categories of these circumstance variables. However, as it is a data-driven technique, it guards against the curse of dimensionality and the risk of over-fitting that would be likely with a fully saturated nonparametric specification.

Table 1 shows the frequencies of each circumstance category in the sample. Figure A.2 in the Appendix shows the most frequent patterns of missing values for circumstances and the health outcome. The most frequent missing information is parental education but note that for 4,567 observations of the potential maximum sample to be used in our analysis, the only missing information is the SF-12 Physical Component Score.

Table 1: Descriptive statistics: circumstances

Circumstance Category	Frequency (%)
Ethnic group	
UK white	82.06
Irish white	2.11
other white	2.52
Mixed (white with Asian/African/Arab)	1.87
Asian (East and Middle East)	7.05
Black (African, Caribbean, other)	3.04
other ethnic group	0.26
missing	1.10
Female	
yes	55.95
no	44.04
missing	0.01
Born in the UK	
yes	86.27
no	11.32
missing	2.42
Mother education	
did not go to school	1.15
left school without qualifications	25.73
some qualification	13.59
post-school qualification	15.79
university degree or higher	6.86
unknown	3.94
missing	32.95
Father education	
did not go to school	1.88
left school without qualifications	29.29
some qualification	19.16
post-school qualification	10.94
university degree or higher	4.54
unknown	0.95
missing	33.24
Mother's occupational skill level	
unemployed	38.24
high skill	6.65
up-mid skill	5.84
mid skill	17.97
low skill	9.35
unknown	2.00
missing	19.94
Father's occupational skill level	
unemployed	5.24
high skill	10.94
up-mid skill	26.56
mid skill	15.70
low skill	5.93
unknown	6.93
missing	28.70
Mother activity status	
working	53.36
unemployed	38.24
deceased	1.33
not living in hh	0.67
missing	6.40
Father activity status	
working	80.47
unemployed	5.24
deceased	3.75
not living in hh	3.18
missing	7.37

To implement the specification in equation (2), a composite scalar index of lifestyle is created. Specifically, all our lifestyle indicators are summarised by a scalar index obtained by principal component analysis (PCA). For those lifestyle indicators that respondents are observed in both Waves 2 and 5 (and different responses are obtained) the more risky level of health behaviour is used in the PCA. The choice of using a summary measure of lifestyle is based on two main considerations. The first is to keep the MOB as parsimonious as possible and to avoid over-fitting the data. Second, we consider lifestyle as an intrinsically unobservable latent pattern of behaviour. On the one hand, each specific behaviour we observe is correlated with this lifestyle, on the other, specific behaviours may be a rather imperfect measures of the overall pattern.

The following indicators of health-related behaviours are included in our analysis to proxy efforts: current smoking status (non-smoker, up to 10 cigarettes per day, 10-19 cigarettes per day, 20+ cigarettes per day), a dummy variable for ex-smoker, number of days each week eating fruits (never, 1 - 3 days, 4 - 6 days, every day), number of days each week eating vegetables (never, 1 - 3 days, 4 - 6 days, every day), days per month walked at least 10 minutes (28 categories based on the frequency of walking habits during the days of a month), a dichotomous variable for drinking alcohol five or more days per week. We also account for a self-assessed measure of sports activity, which is an eleven categories scale from 0 to 10, with 0 being "doing no sport at all"' to 10 being "very active through sport".

As shown in Table 2, a non-negligible share of missing information concerns alcohol intake (about 23% in Wave 2, and 17% in Wave 5). Figure A.1 in the Appendix shows the most frequent combinations of missing data for effort variables. Interestingly about half of the missing information concerns only that aspect of lifestyle. Therefore, for respondents reporting complete information about all other effort dimensions we impute drinking behaviour by multiple imputation using observed behaviours as imputers (Van Buuren and Groothuis-Oudshoorn, 2011). The final sample includes all respondents with complete information, obtained by merging the three UKHLS waves and, after imputation, this is made up of 18,016 adults. Although the final sample size is large relatively to similar empirical analysis, the item non response represents an issue and caution should be exercised in generalising the results to the entire UK population.

Figure 2 summarizes the results of the PCA. The first and second component are shown in the horizontal and vertical axis respectively. Because all measures of behaviours are categorical the PCA has been conducted after computing the polychoric transformation of the mixed data to obtain a meaningful covariance matrix (see Drasgow (1986) for detail and Fox (2019) for the implementation in R). The resulting first component of the PCA (Figure 2) accounts for almost 44% of the total variability of all effort dimensions. Moreover, the sign of the correlation of behaviours with the first component appears to be coherent.[9]

[9]Given the positive correlation of the first PCA component with the risky behaviours, the lifestyle variable has been multiplied by (-1) in order to obtain a measure associated with having a healthier lifestyle.

Table 2: Descriptive statistics: life-style behaviours

Life-style categories	Freq (%)
Sport activity	
No sport at all	30.76
1	8.25
2	8.98
3	8.85
4	7.27
5	9.74
6	5.54
7	5.68
8	3.79
9	1.50
Very active through sport	2.04
missing	7.58
Current smoking status	
Not smoking	79.21
Up to 10 cigarettes per day	6.56
Up to 20 cigarettes per day	8.61
More than 20 cigarettes per day	5.62
missing	7.55
Ex-smoker	
No	62.02
yes	30.44
missing	7.55
Fruit per week	
Never	8.61
1-3 days per week	31.83
4-6 days per week	17.69
everyday	34.36
missing	7.51
Vegetables per week	
never	2.59
1-3 days per week	24.23
4-6 days per week	27.22
everyday	38.47
missing	7.51
Days per month walked at least 10 minutes	
0	19.57
[1-10)	26.14
[10-20)	12.94
[20-30)	33.81
missing	7.55
Alcoholic drink ≥ 5 days per week	
yes	7.26
no	76.36
missing	16.37

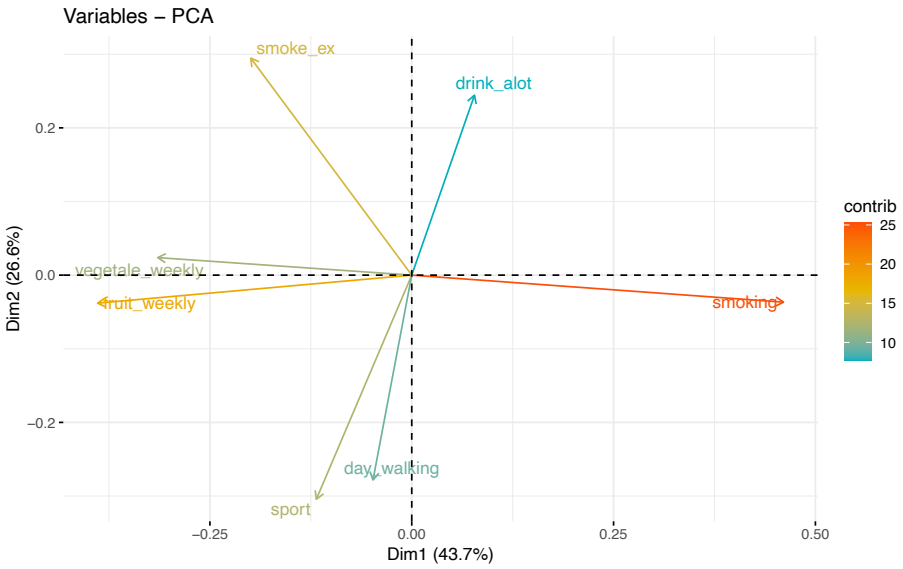
Source: UKHLS Waves 2 and 5

Note: missing values before the imputation of missing values on drinking behaviour.

Table 3 shows the correlation of the lifestyle variable with the observed behavioural variables involved in the analysis. The sign of the correlation is positive for healthy habits such as non-sedentary lifestyle and healthy diet, whilst it is negative for heavy drinking and intensity of smoking.

All of the circumstances and the scalar index of lifestyle are then used to estimate the model-based tree. The algorithm is tuned by 5-fold cross validation. We tested different critical values for the Bonferroni-adjusted p-value

Figure 2: PCA for lifestyle and observed behaviours



Source: UKHLS Waves 2 and 5

Table 3: Spearman correlation with effort

Behaviours	ρ
Sport activity	0.5275***
Smoke intensity	-0.5593***
Ex smoker	0.1545***
Fruit per week	0.6456***
Vegetables per week	0.5410***
Days walked at least 10 minutes	0.6033***
Drink $\geq$ 5 days per week	-0.0139***
% of total variance explained	43.7

Source: UKHLS Waves 2 and 5.
Note: Signif. values: *** ($p < 0.001$).

($\alpha = 0, \alpha = 0.001, \alpha = 0.01, \alpha = 0.05, \alpha = 0.1$) and different health-effort polynomial link specifications (degree 1 to 4). Moreover, in order to guarantee sufficient degrees of freedom for each type, we impose a minimum number of 200 observations per terminal node. The output of the MOB specification with the smallest out-of-sample prediction error is shown in Figure 3, it is obtained

with $\alpha = 0.1$ and assuming a linear relationship between our measure of lifestyle and physical health (PCS-12) rather than higher order polynomials.

The selected tree is made of 11 splits and 12 types. Circumstances used to partition the population are: ethnic group, sex , father's activity, mother's activity, mother's education, father's education, place of birth. Each terminal node contains a scatter plot in which lifestyle is on the horizontal axis and health outcome is on the vertical axis. All type-specific regression models have highly significant regression coefficients and a positive slope (the healthier the lifestyle the higher the expected health). The fitted model explains about 10% of the total health variance in the sample. In what follows we estimate how much of this explained variability is to be considered unfair.

Table 4 reports for each type: the average health status, the average effort exerted, the two parameters (β_0 and β_1) and the population share of each type.

Table 4: Types description

Type	Av. h	Av. eff	% Pop.	β_0	SE	β_1	SE
1	-4.728	3.153	3.96	-9.991***	(0.991)	1.668***	(0.290)
2	-2.606	3.093	2.02	-6.310***	(1.169)	1.197***	(0.346)
3	-2.400	3.042	6.97	-8.306***	(0.702)	1.940***	(0.204)
4	-0.755	3.695	1.76	-6.082***	(1.634)	1.441***	(0.418)
5	-0.608	3.542	1.12	-8.405***	(1.651)	2.201***	(0.434)
6	-0.063	3.587	3.84	-3.702***	(0.966)	1.014***	(0.249)
7	0.082	3.172	17.19	-7.077***	(0.428)	2.257***	(0.120)
8	0.380	3.494	15.20	-8.067***	(0.534)	2.417***	(0.140)
9	0.487	3.480	25.48	-5.737***	(0.371)	1.788***	(0.097)
10	1.172	3.351	1.59	-3.302***	(1.218)	1.335***	(0.334)
11	1.494	3.424	13.57	-5.095***	(0.459)	1.924***	(0.122)
12	2.871	3.584	7.26	-1.725***	(0.485)	1.282***	(0.123)

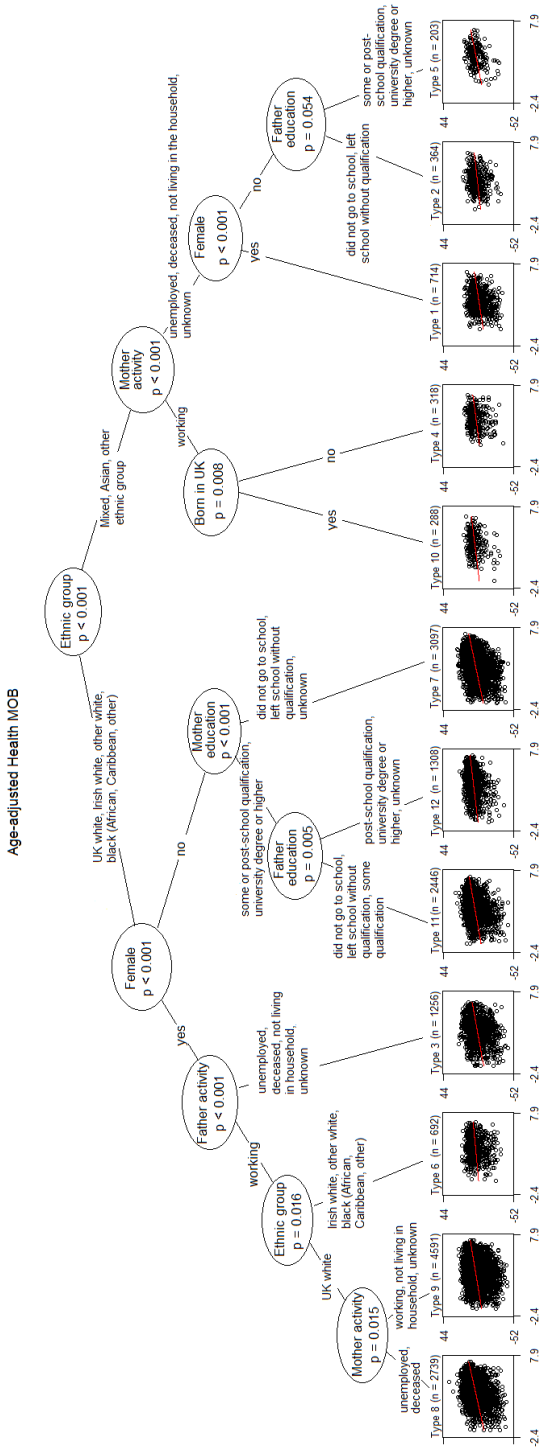
Source: UKHLS Waves 2, 5 and 6.

Note: In the first column types rank is determined by their average health (second column), the third column reports the average effort and the fourth the share of observations in each type. The other columns contain models' parameters. Signif. values: *** ($p < 0.001$)

In terms of average health, the worst-off type is type 1 made up of mixed race, other ethnic and Asian women whose mother did not work. This group represents about 4% of the sample and has an expected health outcome of -4.728 (not far from the 25th percentile of the entire PCS-12 distribution). The best-off type is type 12 made up of white or black men whose mother left school with at least some qualification and whose father has at least a post-school qualification (or for a few respondents is unknown). This type represents slightly more than 7% of the sample and their average health is 2.871 (clearly above the population mean 0.1964).

In general, the splitting rules selected by the MOB algorithm are consistent with what might be expected: ethnicity, place of birth, sex and parental background all play some role. A more advantaged socioeconomic background,

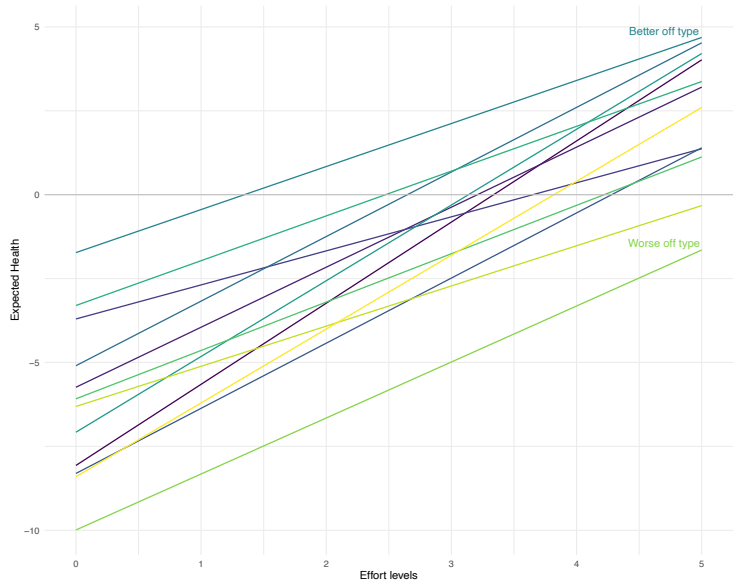
Figure 3: MOB tree diagram



mother's labour force participation, being born in the UK, and being white are predictive of a better health outcome. Less obviously, being either a white or black male is predictive of a better outcome. In terms of the parameters estimated type 1 and 12 are also the types with the lowest and highest intercepts. Type 6 has the lowest return to effort ($\beta_1 = 1.014$). This type is made of women that define themselves as non-UK white or black and whose father was working during their adolescence. Women that define themselves as UK white whose father was working, but whose mother was not (type 8), have the highest return ($\beta_1 = 2.417$), a gradient that is two-and-a-half times that of type 6. Note that slopes heterogeneity is a source of clash between compensation and reward discussed in Section 2 that justifies the need of considering two families of unfair inequality measures.

Figure 4 shows the fitted regression lines for each type. These can be interpreted as the opportunity set (or health constraint) faced by individuals belonging to different types.

Figure 4: Opportunity sets by types: health - level of effort profiles



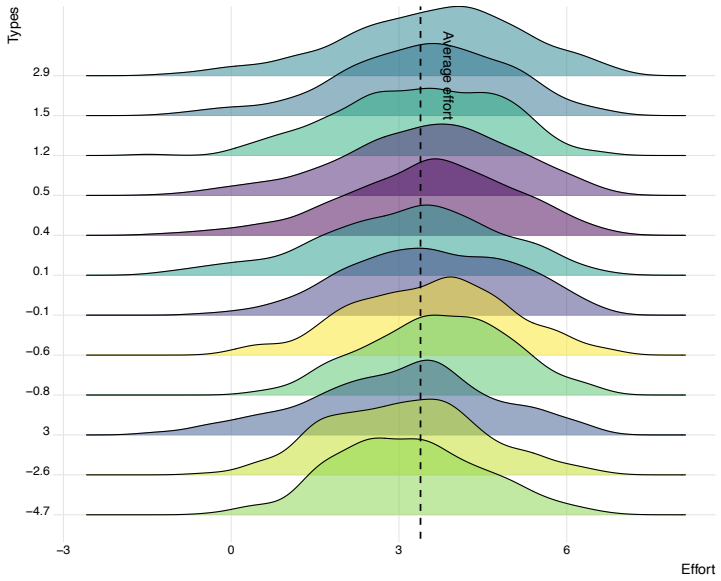
Source: UKHLS Waves 2, 5 and 6.

Note: health-effort relationship is shown on the entire effort range of variation.

What emerges is that having favourable circumstances will produce a fixed advantage (higher intercept) but it will not necessarily imply a higher return to a healthy lifestyle (higher slope). That is, there is a correlation between the intercept and the types' rank in terms of expected health. But there is not a monotonic relationship between slopes and intercepts nor between slopes and expected outcome.

Having estimated the opportunity sets individuals face is not sufficient to obtain the two counterfactual distributions necessary to estimate UI. The counterfactual distributions will depend on these parameters and also on the the type-specific distributions of effort that define the degree of effort that corresponds to the observed levels of effort for each type. An initial intuition regarding the role of effort in determining the different type-specific health outcomes is provided by Figures 5 and 6a. Figure 5 shows the distribution of effort in the 12 types, ranked according to their average health. The effort distribution in better-off types is more dispersed and higher than the overall average (dashed vertical line). The between-type variability of effort is limited ranging between 3.040 and 3.695 (the 39th and 55th percentile of the distribution in the population). There is also a moderate negative correlation between the average effort exerted and return to effort (-0.1478). So both individuals with more favourable circumstances and with lower return to effort tend to have healthier lifestyles.

Figure 5: Distribution of effort across types



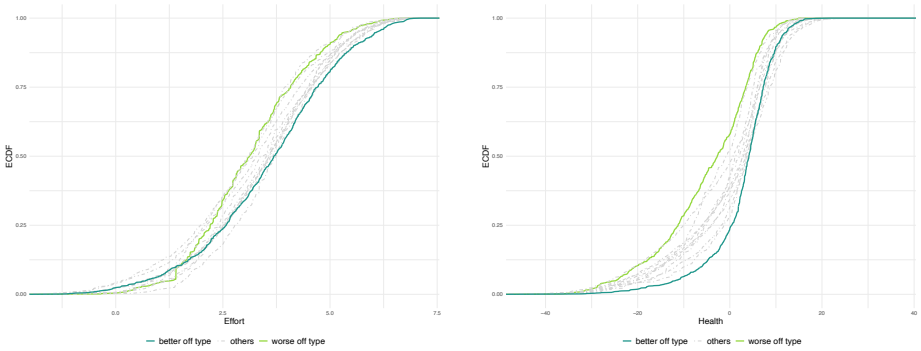
Source: UKHLS Waves 2, 5 and 6

However, focusing on the type-specific empirical cumulative distribution function (ECDF) of effort and health what is striking is the clear dominance in terms of expected health condition for better off types accompanied by absence of dominance in terms of effort.

Consider for example Figures 6a and 6b where both ECDFs are shown for the the two extreme types. Type 1 made of women with Asian or mixed origin, and an absent or non-working mother, and type 12 made of white men with both parents with at least post-school qualification.

While the effort ECDFs cross, with individuals in the least favourable type behaving better at the bottom of the distribution (6a), health ECDFs show a clear dominance of type 12 over type 1, with a particularly marked difference in expected health especially in the left tail of the distribution (6b).

Figure 6: Empirical Cumulative Distribution Functions



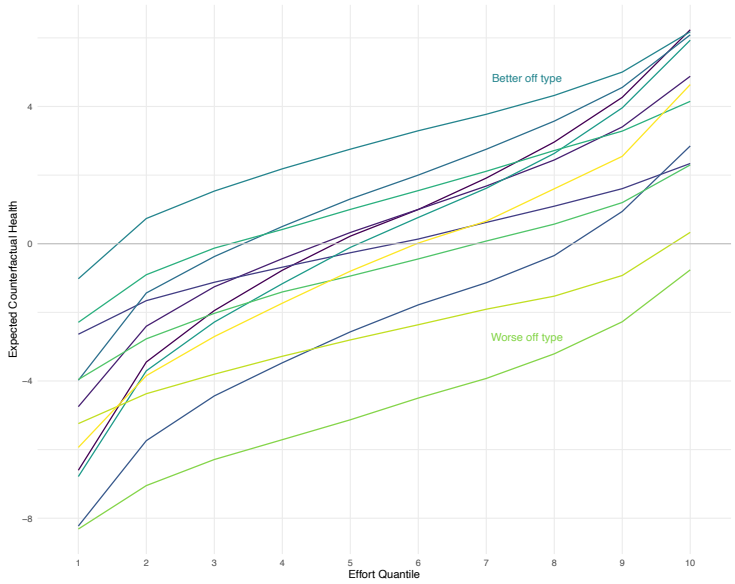
(a) Effort (b) Health

Source: UKHLS Waves 2, 5 and 6

Finally, adopting John Roemer's view about what is the morally-relevant measure of effort, we remove the variation of effort systematically correlated with types by comparing individuals considering the degree of effort they exerted. $\tilde{H}_{DU}$ and $\tilde{H}_{FG}$ are therefore constructed ignoring the absolute level of effort (first component of the PCA) and comparing instead individuals belonging to the same quantile of the type-specific distribution of the same variable. This transforms Figure 4 into Figure 7. The distribution described by these segmented lines together with the types' population shares provides all the information needed to estimate UI_{DU} and UI_{FG}.

The two measures of health unfair inequality are calculated for the 12 possible reference types and for 10 possible reference responsibility values (effort tranches) defined by the deciles of the scalar lifestyle index within each type. For both measures we calculate confidence intervals by bootstraping observations by types. This implies fixing the structure of the tree and then resampling each type 200 times. This procedure is likely to underestimate the level of uncertainty about point estimates. A more robust approach would consist in estimating a different MOB for each sample. However, the need to set a reference type to calculate UI_{FG} requires to fix the structure in types. Figure 8a reports our estimates for UI_{FG} based on the 12 reference types. Types are ordered according to their average health status (labelled below) but the expected outcome does not affect the value of UI_{FG}. Its value is entirely determined by the slope of the regression line estimated for the reference type. The flatter the regression line the more health variability is reproduced in the counterfactual distribution. In the extreme case in which the line is flat, health is independent

Figure 7: Distribution of health inequality: health - degree of effort profiles



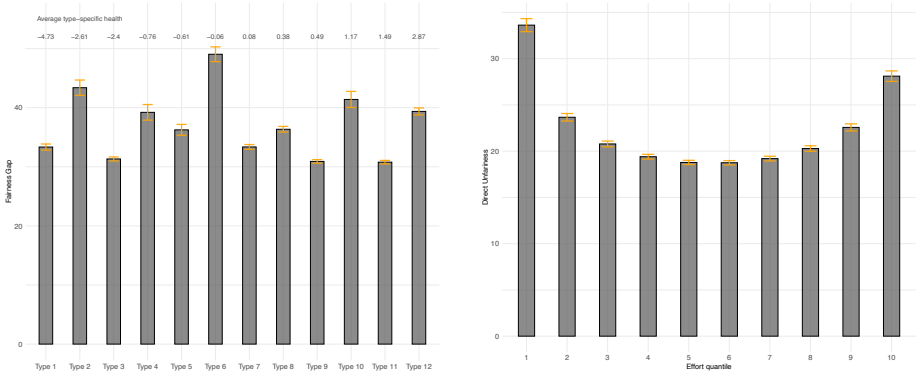
Source: UKHLS Waves 2, 5 and 6

from the degree of effort in the reference type and all health inequality is to be considered unfair. After all, if choices do not play a role, what sort of inequality can be justified? In our case, when type 6 is the reference ($\beta_1 = 1.014$) close to 50% of the explained variability is to be considered unfair inequality. Moreover, no matter what reference type is selected UI_{FG} is never lower than 30%.

Figure 8b reports estimates for direct unfairness for ten reference effort tranches (deciles in ascending order). The ten unfairness measures are significantly smaller than the compensation-consistent measures and their value follows a U-shaped pattern. Unfair inequality is higher when the reference effort is at the two extremes of the lifestyle spectrum (close to 30% and 25% of the explained variance respectively). Figure 7 shows that this pattern is driven by the outcomes for the worse-off types converging on those of the better off types as effort increases from the lower deciles to a more healthy pattern of behaviour in the middle deciles. This is due to the less dispersed distribution of effort in the worse-off types, who appear to catch-up with more advantaged types simply because the average effort exerted in the left tail of the distribution increases more quickly. This pattern is then reversed for individuals in the highest effort tranches. For individuals that adopt the healthiest lifestyle a clear social gradient is visible with two types lagging behind (1 and 2) in terms of health status. The comparison between the two extreme types is striking; no matter how healthily they behave, individuals in type 1 have a predicted health outcome below that of the worst-behaving individuals who have the most favourable

Figure 8: Unfair health inequality

(a) UI_{FG} (b) UI_{DU}



Source: UKHLS Waves 2, 5 and 6.

Note: In 8a reference types are sorted by increasing type-specific expected health. Confidence intervals are obtained from 200 stratified bootstrap samples.

circumstances (type 12). For type 1 there is no level of effort that could compensate for their adverse circumstances (no matter how badly an individual in type 12 behaves she has a higher predicted health).

5 Conclusions

This study aims to provide both a methodological innovation for the measurement of unfair health inequality, as well as new evidence on health inequalities measured in the UKHLS. The methodological innovation is the adoption of the MOB algorithm to estimate the health-to-lifestyle relationship while considering the different socioeconomic backgrounds in childhood. Moreover, a normatively defined responsibility-sensitive framework is adopted to measure Direct Unfairness and the Fairness Gap *à là* Fleurbaey and Schokkaert (2009). Among the main features of the use of MOB in the measurement of unfair health inequality is its ability to capture those socioeconomic characteristics which are fundamental to determine a change in the conditional distribution of the outcome in the health-to-lifestyle model.

The empirical application uses data from the UK Household Longitudinal Study (Waves 2, 5 and 6) considering all observations for which data on physical health status, relevant circumstances beyond individual control, and health-related behaviours are observed. We show that circumstances beyond individual control are a clear source of unfair health inequality. However, this is mostly driven by a fixed advantage for better-off types. Moreover, while on average individuals characterised by more favourable circumstances tend to have a healthier

lifestyle, this seems not to be due to systematic heterogeneity in the return to effort across types.

The estimated UI_{DU} and UI_{FG} show that, when a compensation-consistent approach is adopted, unfair inequality varies in a non-monotonic way depending on the reference type considered. Poorer socioeconomic conditions tend to be associated with lower expected health outcomes more because of a direct contribution (intercept) than due to an indirect contribution through a lower return to efforts (slope). This echoes the findings of Carrieri and Jones (2018) and Carrieri et al. (2020). When adopting a reward-consistent approach, and measuring UI_{DU}, a clear pattern emerges; when the reference degree of effort is at the two extremes the level of unfairness detected is higher. This result is driven by the interactions of types' direct contribution to health (the intercept), the return to a healthier lifestyle (the slope) and the type-specific distribution of effort being more compressed for less advantaged types. The combined effect makes between-type inequality lower for individuals exerting an intermediate degree of effort.

Overall, our results show that the variation in physical health can only be partially explained by observed lifestyle and childhood socioeconomic background in the UKHLS. Indeed, there are many aspects which are not included in the model even though they have an impact on health status. Some of these are likely to remain unobservable, such as genetic endowments, others, however, could fit in the Fleurbaey and Schokkaert (2009) framework and, given suitable data, could be taken into account, such as healthcare consumption and the role of public healthcare services.

References

P. Brunori and G. Neidhöfer. The evolution of inequality of opportunity in germany: A machine learning approach. *ZEW-Centre for European Economic Research Discussion Paper*, (20-013), 2020.

P. Brunori, P. Hufe, and D. G. Mahler. *The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees*. Policy Research Working Papers. The World Bank, Feb. 2018.

P. Brunori, V. Peragine, and L. Serlenga. Upward and downward bias when measuring inequality of opportunity. *Social Choice and Welfare*, 52(4):635–661, 2019.

P. Brunori, A. Trannoy, and C. F. Guidi. Ranking populations in terms of inequality of health opportunity: A flexible latent type approach. *Health Economics*, 30(2):358–383, 2021.

V. Carrieri and A. M. Jones. Inequality of opportunity in health: A decomposition-based approach. *Health Economics*, 27(12):1981–1995, 2018.

V. Carrieri, A. Davillas, and A. M. Jones. A latent class approach to inequity in health using biomarker data. *Health Economics*, 29(7):808–826, 2020.

D. Checchi and V. Peragine. Inequality of opportunity in italy. *The Journal of Economic Inequality*, 8(4):429–450, 2010.

G. A. Cohen. On the currency of egalitarian justice. *Ethics*, 99(4):906–944, 1989.

A. Davillas and A. M. Jones. The first wave of the covid-19 pandemic and its impact on socioeconomic inequality in psychological distress in the uk. *Health Economics*, 30(7):1668–1683, 2021.

F. Drasgow. Polychoric and polyserial correlations. In *The Encyclopedia of Statistics*, volume 7, pages Pp. 68–74. Wiley, 1986.

R. Dworkin. What is equality? part 1: Equality of welfare. *Philosophy & Public Affairs*, pages 185–246, 1981.

P. Eibich. Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics*, 43:1–12, 2015.

M. Fleurbaey. Three solutions for the compensation problem. *Journal of Economic Theory*, 65(2):505–521, 1995.

M. Fleurbaey. *Fairness, responsibility, and welfare*. Oxford University Press, 2008.

M. Fleurbaey and F. Maniquet. *Equality of opportunity: the economics of responsability*. World Scientific Pub Co Inc, 2012.

M. Fleurbaey and E. Schokkaert. Unfair inequalities in health and health care. *Journal of Health Economics*, 28(1):73–90, 2009.

J. Fox. Package 'polycor'. *R package version 3.3. 0, URL http://CRAN. R-project. org/package= polycor*, 2019.

H. Gravelle. Measuring income related inequality in health: standardisation and the partial concentration index. *Health Economics*, 12(10):803–819, 2003.

R. Guber. Making it right? social norms, handwriting and human capital. *Labour Economics*, 56:44–57, 2019.

T. Hothorn and A. Zeileis. partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research*, 16:3905–3909, 2015. URL `https://jmlr.org/papers/v16/hothorn15a.html`.

F. Jusot, S. Tubeuf, and A. Trannoy. Circumstances and efforts: how important is their correlation for the measurement of inequality of opportunity in health? *Health Economics*, 22(12):1470–1495, 2013.

A. Lefranc, N. Pistolesi, and A. Trannoy. Equality of opportunity and luck: Definitions and testable conditions, with an application to income in france. *Journal of Public Economics*, 93(11-12):1189–1207, 2009.

P. Li Donni, J. G. Rodríguez, and P. R. Dias. Empirical definition of social types in the analysis of inequality of opportunity: a latent classes approach. *Social Choice and Welfare*, 44(3):673–701, 2015.

J. Rawls. Justice as fairness. *The Philosophical Review*, 67(2):164–194, 1958.

J. Rawls. *1971: A theory of justice.* Cambridge, MA: Harvard University Press, 1971.

J. E. Roemer. *Theories of distributive justice*. Harvard University Press, 1998.

J. E. Roemer. Equality of opportunity: A progress report. *Social Choice and Welfare*, 19(2):455–471, 2002.

J. E. Roemer and A. Trannoy. Equality of Opportunity. In *Handbook of Income Distribution*, volume 2, pages 217–300. Elsevier, 2015.

P. Rosa Dias. Inequality of opportunity in health: evidence from a uk cohort study. *Health Economics*, 18(9):1057–1074, 2009.

H. Schmitz. Why are the unemployed in worse health? the causal effect of unemployment on health. *Labour Economics*, 18(1):71–78, 2011.

A. Sen. Equality of what? *The Tanner lecture on Human Values*, 1:197–220, 1980.

M. Sutton. Vertical and horizontal aspects of socio-economic inequity in general practitioner contacts in scotland. *Health Economics*, 11(6):537–549, 2002.

S. Van Buuren and K. Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, 45:1–67, 2011.

A. Wagstaff and E. Van Doorslaer. Equity in health care finance and delivery. *Handbook of Health Economics*, 1:1803–1862, 2000.

J. Ware, M. Kosinski, and S. Keller. *SF-36 physical and mental health summary scales, 2nd edition*. Boston: The Health Institute, New England Medical Center, 2001.

A. Zeileis and K. Hornik. Generalized M-fluctuation tests for parameter instability. *Statistica Neerlandica*, 61(4):488–508, 2007.

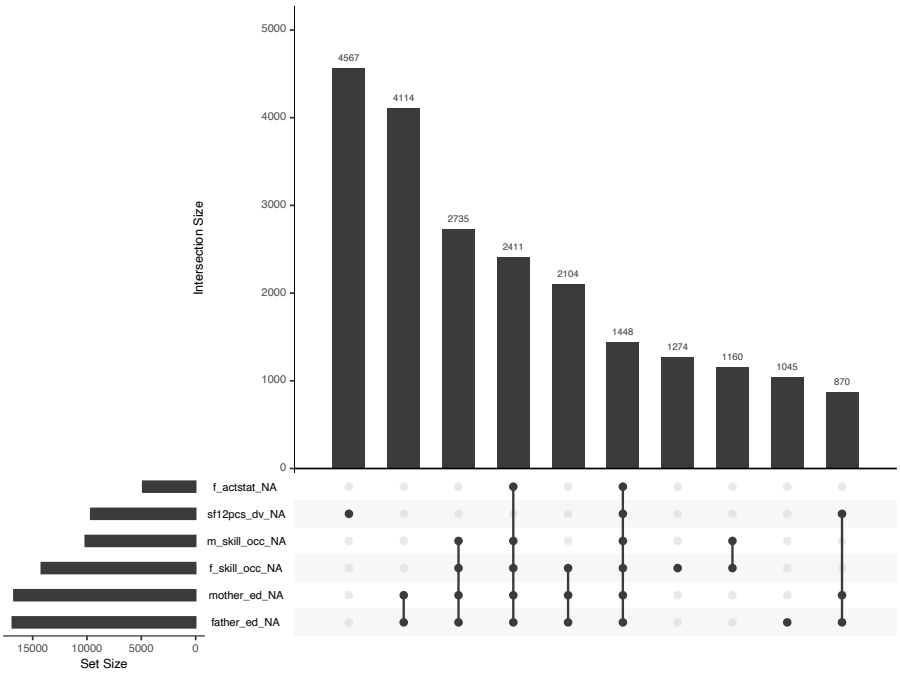
A. Zeileis, T. Hothorn, and K. Hornik. Model-based recursive partitioning. *Journal of Computational and Graphical Statistics*, 17(2):492–514, 2008.

A. Zeileis, T. Hothorn, and K. Hornik. party with the mob: Model-based recursive partitioning in r. *R package version 3.0.0*, 2010.

N. Ziebarth. Measurement of health, health inequality, and reporting heterogeneity. *Social Science & Medicine*, 71(1):116–124, 2010.

A Appendix

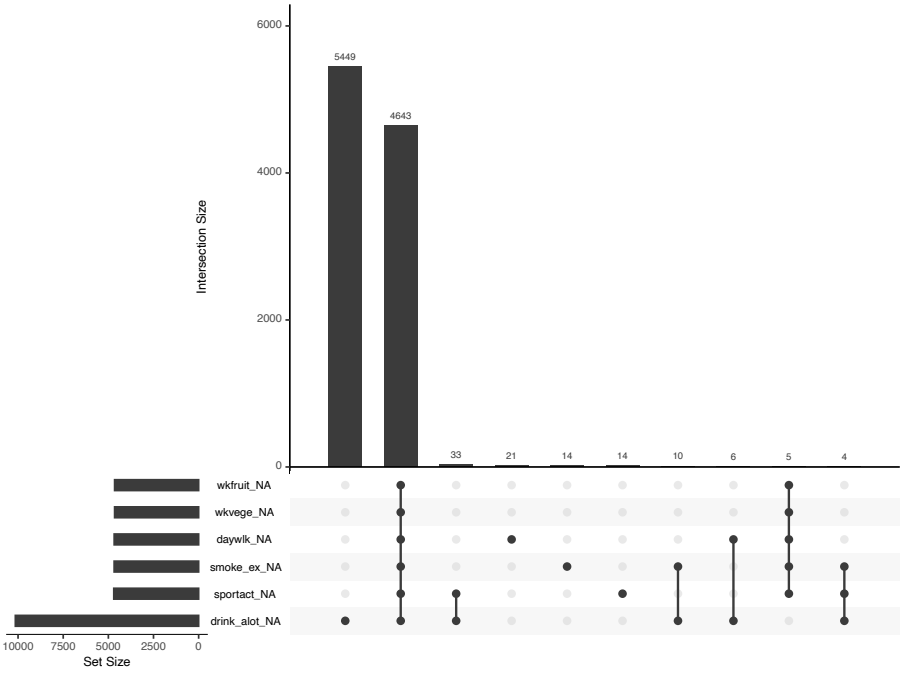
Figure A.1: Missing circumstances and outcome



Source: UKHLS Wave 6

Note: The missing values (NA) are shown for the following variables: health (sf12pcs_dv), father activity status at respondent's age of 14 (f_actstat), mother and father skill in occupation (m_skill_occ, f_skill_occ), mother and father education (mother_ed, father_ed).

Figure A.2: Missing efforts



Source: UKHLS Waves 2 and 5

Note: The missing values (NA) are shown for the following variables: fruit units eaten per week (wkfruit), vegetable units eaten per week (wkvege), days walked at least 10 minutes (daywlk), ex-smoker (smoke_ex), sport activity (sportact), drinking alcohol at least 5 days per week (drink_alot).