Comparative effects of lifestyles, childhood conditions and education on health: An investigation with a British cohort

Sandy Tubeuf [a]**, Florence Jusot** [b]**, Damien Bricard** [c]**,**

[a] s.tubeuf@leeds.ac.uk – Academic Unit of Health Economics (University of Leeds)

[b] florence.jusot@dauphine.fr – LEDa-LEGOS (Université Paris-Dauphine) and IRDES (Institut de Recherche et de Documentation en Economie de la Santé).

[c] damien.bricard@ensai.fr – LEDa-LEGOS (Université Paris-Dauphine)

Abstract

This paper focuses on the long-term effects of lifestyles with comparison to social and health-related initial conditions, and education on health status over the life-cycle in a cohort of British people born in 1958. Using the longitudinal follow-up of health and lifestyles at age 23, 33, 42 and 46, it investigates the influence of each determinant on health and the mediating role of education and lifestyles in the relationship between social and health-related initial conditions and health. Self-assessed health is modeled using a dynamic panel Probit specification addressing individual unexplained heterogeneity. Health inequality is explored using the decomposition of the variance. Our results show an impressive contribution of lifestyles to health inequalities about 35% in the baseline specification and 26% in the structural specification. We also found a predominant contribution of childhood conditions when their indirect role on education and lifestyles is taken into account, which underlines the relevance of mediated effects between determinants of health and outperform previous works ignoring childhood characteristics as a relevant determinant of health. Noticeably, average past lifestyles appear to matter more for adult health than transitory lifestyles.

Keywords: circumstances; cohort; health inequality; lifestyles; unobserved heterogeneity

Codes JEL: D63; I12.

Acknowledgements

1. Introduction

Among the various determinants of health inequalities, lifestyles present a controversial approach in comparison with current social determinants and initial conditions. Numerous literature references have agreed the important role played by current individual social characteristics, such as income, education level, wealth, and social status (e.g. van Doorslaer and Koolman 2004; Cutler et al. 2006; Lantz et al. 2010) in the explanation of health inequalities. More recently, several studies have also found childhood conditions and parental health as a relevant determinant of health inequalities (Dias 2009; Galobardes et al. 2006; Jusot et al. 2010; Marmot et al. 2001; Osler et al. 2009; Power et al. 2005; Trannoy et al. 2010; Tubeuf and Jusot 2010; Van de Mheen et al. 1997). However, the importance of lifestyles in the magnitude of health inequalities is less clear. Whereas epidemiological literature concluded until recently that lifestyles make a relatively minor contribution to the social gradient in health (Khang et al. 2009; Lantz et al. 2010; Mokdad et al. 2004; Singh-Manoux et al. 2008; Skalická et al. 2009; van Oort et al. 2005), health economists have shown that differences in lifestyles can explain a relevant part of health and mortality inequalities (Balia and Jones 2008) and few recent epidemiological studies (Laaksonen et al. 2008; Menvielle et al. 2009; Strand and Tverdal 2004; Stringhini et al. 2010; Woodward et al. 2003) have also confirmed that the impact of lifestyles on health disparities would be larger than it was previously estimated, particularly if lifestyles are observed longitudinally.

The study of well-known determinants of health inequality together with lifestyles and the evaluation of their respective contribution to the magnitude of health inequality is particularly relevant for policy makers. According to the determinants that are found important for health inequalities, different public policies will be designed (Bambra et al. 2010b). For instance, the most appropriate interventions to tackle health inequalities related to social determinants would be interventions in the field of housing and the working environment (Bambra et al. 2010a) whereas the most effective ways to tackle risky health behaviours such as smoking, drinking, having a poor diet or not exercising, are those that combine measures aimed at the whole population such as increasing prices or reducing availability, with measures that target the vulnerable and disadvantaged groups who may be at increased risk, such as minimum age or health promotion interventions (Chaloupka and Warner 2000; DiNardo and Lemieux 2001; Powell et al. 2005; WHO Tobacco Free Initiative 2004).

However, even if constraints over the life cycle are important determinants of health-related behaviours, health behaviours might also reflect individual preferences, and individual free choice. The underlying public policy recommendations become then more complicated and so, the legitimacy of the induced differences in health or mortality as well as the implications in terms of health promotion and prevention policies have to be discussed. The issue at stake is therefore to know whether individuals have freely chosen their unhealthy behaviours and when they have, whether policy makers will consider them as responsible for those choices and consider the subsequent health differences as legitimate.

Conversely to health-related beahaviours, social background and family characteristics cannot be chosen by the individual and so, are considered as fully exogenous from individual responsibility. They are thus considered as the most illegitimate sources of health inequalities and would undoubtedly justify policy interventions aiming compensation (Dias and Jones 2007; Trannoy et al. 2010; Jusot et al. 2010).

Several studies provide evidence on the transmission of socioeconomic status over different generations and its relevance in the explanation of health inequalities (e.g. Power et al. 2005; Case et al. 2005; Trannoy et al. 2010; Van de Mheen et al. 1997). Moreover, a range of family factors have been empirically analysed including low parental socioeconomic status (Dias 2009; Göhlmann et al. 2009; Power et al. 2005; Jusot et al. 2010), family issues, such as living in a single parent family or experiencing marital discord (Case and Katz 1991; Dekovic et al 2003), parental health-risk lifestyles (Engels et al. 2004; Francesconi et al 2010; Wickrama et al. 1999; Zhang et al. 1999; Anda et al. 2002). Parents and past conditions would thus be associated with current socioeconomic position and the development of problem behaviours.

It is therefore essential to understand the interrelationships between those various determinants of health and evaluate their respective contribution to the magnitude of health inequalities.

The objective of this study is to explore the long term effects of family and childhood conditions, social status and lifestyles on health over the life-cycle. We consider the interdependence between social and health-related childhood conditions, education and health related behaviours and investigate the magnitude of that relation and its effect on overall health inequality over the life course. This study relies on the British National Child Development Study (NCDS) which provides a description of social and health childhood

characteristics, information on education level and longitudinal follow-up of a broad set of lifestyles and health status variables for a representative cohort of individuals born in 1958.

The findings provide new elements on the determinants of health inequalities which are relevant for policy makers and that remained to be empirically assessed. On the one hand, the role of childhood conditions is explored with a large set of indicators, including parental social and health conditions in addition to the individual's initial health status. On the other hand, education level is considered, and instead of considering current unhealthy lifestyles, this study analyses the evolution of unhealthy lifestyles, their changes over an extended period of time and their consequence on health status. Finally, the longitudinal dimension of those data allows using dynamic panel analysis in order to control unexplained individual heterogeneity and analyses health state dependence.

The structure of the paper is as follows. Section 2 describes the theoretical model that is empirically tested. Section 3 describes the National Child Development Study (NCDS) data and the variables of interest. Section 4 discussed the results and section 5 concludes.

2. Theoretical model

2.1 The model

In contrast to Jusot et al. (2010), who were primarily interested in the role played by lifestyles and childhood conditions on health and their correlation and focused on a reduced-form model, we use a full form model specification including individual socioeconomic status. Our approach also differs from Contoyannis and Jones (2004) and Balia and Jones (2008), as our health production function includes childhood and family conditions as a potential determinant for health in addition to socioeconomic status and lifestyles, as defined by a set of health-related behaviours. We use a dynamic panel Probit specification.

Let us assume that individual health status H can be written using the following health production function:

(Eq. 1) $H = f(C, D, E, B, u)$

The vector of social and health-related childhood conditions C consists of a set of variables beyond individual control which may be related to health status. The literature on health determinants suggests an influence of childhood conditions and family background on

health status in adulthood (see for example Currie and Stabile 2003; Case et al. 2005; Dias 2009; Lindeboom et al. 2009; Trannoy et al. 2009). Moreover, initial health such as birth weight and health problem during childhood and adolescents also significantly influence health in adulthood and the most adverse health risks in adulthood tend to be experienced by people having experienced poor health in childhood and adolescence (Moser and Power 2003; Case et al. 2005). Moreover, the inclusion of social and health childhood conditions in the model can be interpreted as a particular specification of the individual component. A well-explained vector of circumstances strongly contributes to the reduction of the correlation between individual effects and initial conditions as it minimizes unobserved time-invariant characteristics which affect outcomes at each point of time.

The vector E represents individual's education level, as measured by the highest diploma achieved. Literature in many countries has observed a relevant and persistent association between education and health as measured by various health measures (Grossman 2006; Arendt 2005). Moreover, we assume that education level is a reliable proxy of other social outcomes such as employment, housing or income.

The vector of health-related behaviours B captures individual decisions to invest in health capital, such as lifestyles (Balia and Jones 2008; Contoyannis and Jones 2004; Dias 2009; Jusot et al. 2010).

The vector D represents demographic characteristics which are biological determinant of health status, only captured by gender in the cohort.

Finally, a possible interpretation of the residual terms appeals to pure luck and others random factors (accident for instance) which cannot be captured by the other determinants. This unobserved individual heterogeneity is characterised by two components: a time invariant and individual specific random element ω_i and a time variant individual specific error term ε_{it} which is assumed to be normally distributed and uncorrelated across individuals and waves.

Longitudinal data allows avoiding an under-estimation of the impact of lifestyles on health that would be observed using health and the contemporary lifestyles variable as done in Jusot et al. (2010) therefore we consider that efforts in the previous wave are more likely to influence future health status than current health status and introduce $B_{i,t-1}$.

Furthermore, individual effects ω_i may be correlated with the past lifestyles variables. As suggested by Mundlack (1978) and Chamberlain (1984), we can use a random effect

Probit specification allowing ω_i and ε_{it} to be correlated introducing a vector of average individual past lifestyles variables $\overline{B}_i$. Lifestyles are therefore regarded both as a measure of lifestyles shocks on health and a measure of long-term or "permanent" lifestyles on health. Nevertheless, this specification presents a potential endogeneity bias related to the correlation of both health and lifestyles with past health. The endogeneity bias could be ruled out introducing past health $H_{i,t-1}$ into the health production function. Current health is also dependent on previous health (Kenkel 1995). Moreover, the introduction of past health status in our empirical model will allow us to capture the state dependence in health reports and to reduce strongly the impact of individual heterogeneity.

The introduction of dynamic into the random effect model raises a question about the initial condition specification. The difficulty is to deal with two main assumptions concerning a discrete time stochastic process with binary outcomes (Heckman 1981). The first assumption is that the initial observations are exogenous variables. The second is that the process is in equilibrium inasmuch the marginal probabilities have reached their limited values and can therefore be assumed to be time-invariant. These assumptions are tenable if the equation is specified by at least including the first value of the independent variable $H_{i,1}$ as suggested by Wooldridge (2005).

More concretely, let us assume that health of individual i at wave t is measured by a continuous latent variable H_{it}^* which is measured using a binary variable H_{it} as follows:

(Eq. 2) $H_{it}^* \geq 0$ when $H_{it} = 1$

$\qquad H_{it}^* < 0$ when $H_{it} = 0$

The latent variable model that we estimate can be written as follows:

(Eq. 3) $H_{it}^* = \alpha_1 C_i + \alpha_2 D_i + \gamma_1 E_i + \delta_1 B_{i,t-1} + \delta_2 \overline{B}_i + \beta_1 H_{it-1} + \beta_2 H_{i1} + \omega_i + \varepsilon_{it}$

$\qquad$ with $i = 1,...,N$ and $t = 2,...,T_i$

We are particularly interested in the interdependent relationships linking the set of health determinants. Therefore, we consider two subsequent specifications to model self-assessed health: a baseline specification and a structural specification.

The baseline specification corresponds to (Eq. 3) and basically considers all the regressors as they are observed in the data. However the structural specification applies a specific treatment to the correlation that may exist between the regressors using auxiliary equations that are initially estimated.

2.2 The structural specification

The structural aims to identify whether explanatory variables influence health directly or indirectly, that is by affecting or being affected by another explanatory variable. For example, social conditions during childhood may affect health in adulthood through their effect of individual's education attainment. Three sets of auxiliary equations are therefore estimated in order to explore the inter-influence between social and health-related childhood conditions, individual education level, gender, lifestyles and past health. The first set explores the influence of social and health-related childhood conditions on the individual's education level; the second set explores the influence of both social and health-related childhood conditions and education level on lifestyles; and finally the third set explores the influence of both social and health-related childhood conditions and education level on past health. The three sets can be written as follows:

(Eq. 4a) $E_i = a_1 C_i + e_i$

(Eq. 4b) $B_{it} = a'_1 C_i + a'_2 E_i + a'_3 D_i + b_i$

(Eq. 4c) $H_{i,t-1} = a''_1 C_i + a''_2 E_i + a''_3 D_i + h_i$

The estimation of the auxiliary equations of the structural specification relies on the methodology already used by Bourguignon et al. (2007) and Jusot et al. (2010) according to which it is possible to purge each explanatory variable from the effect of other explanatory variables estimating the residual term as a generalised residual (Gourieroux et al. 1987).

(Eq. 5a) $\hat{e}_i = E(e_i / E_i, C_i)$

(Eq. 5b) $\hat{b}_{it} = E(b_{it} / B_{it}, C_i, E_i, D_i)$

(Eq. 5c) $\hat{h}_{i,t-1} = E(h_{i,t-1} / H_{i,t-1}, C_i, E_i, D_i)$

The estimated generalised residuals related to each explanatory variable correspond to the "true" effect for given social and health-related childhood conditions, individual education level, and gender. They are then introduced in the health equation in replacement of the respective observed explanatory variables. The generalised residuals associated to time dependant explanatory variables B_{it} and $H_{i,t-1}$ are estimated at each point of time. The individual mean of health-related variables is substituted by the individual means of the generalised residuals associated to the health-related variables.

The latent variable model that we estimate in the structural specification can be written as follows:

(Eq. 6) $H_{it}^* = \alpha_1 C_i + \alpha_2 D_i + \gamma_1 \hat{e}_i + \delta_1 \hat{b}_{i,t-1} + \delta_2 \bar{\hat{b}}_i + \beta_1 \hat{h}_{it-1} + \beta_2 \hat{h}_{i1} + \omega_i + \varepsilon_{it}$

with $i = 1, ..., N$ and $t = 2, ..., T_i$

2.3 Measurement of inequality

To assess the magnitude of inequality, we use the variance. Shorrocks (1982) showed that if we are interested in an absolute measure of inequality, i.e. a measure invariant to one translation, the variance is a good index and its natural decomposition presents the desired properties. The decomposition of total inequality can be written as follows:

(Eq. 7) $V(H^*) = V(\sum_{k=1}^{K} \eta^k X^k) + V(\omega) + V(\varepsilon)$

where η^k represents the estimated coefficient associated to the explanatory variable X^k in the health equation estimation the baseline specification (Eq. 3) or structural specification (Eq. 6).

In a linear case, the share of variance explained by a set of M explicative variables (with $M < K$) simply consists in the share of the R² of the model which is explained by X^k.

In a non linear context, the share of variance is not straightforward as H_{it}^* can only be measured as a prediction and, ϖ_i and ε_{it} are defined as independent of the set of M explicative variables. A variance estimated from the data is attributed to the time invariant individual error term ω_i whereas the time variant individual error term ε_{it} has a variance normalised to be equal to 1 in the case of a Probit model.

We use the pseudo R² proposed by McKelvey and Zavoina (1975) in order to measure the share of variance explained by the variable X^k, which is based on predictions of the latent endogenous variables:

(Eq. 8) $\hat{H}_{it}^* = \sum \eta^k X_{it}^k$

Assuming that the variance of the error term ϖ_i and ε_{it} follows a normal distribution in this random effect Probit model, we can write:

(Eq. 9) $R^2 = \dfrac{V(\hat{H}^*)}{V(\hat{H}^*) + \sigma_\omega + 1}$

3. The National Child Development Study

The National Child Development Study (NCDS) is a continuing, multi-disciplinary longitudinal study which takes as its subjects all the people born in one week in March 1958 in England, Scotland and Wales. This study was designed to examine the social and obstetric factors associated with stillbirth and death in early infancy among the children born in Great Britain in that one week. Information was gathered from almost 17,500 babies. Following the initial birth survey in 1958, there have been seven attempts to trace all members of the birth cohort in order to monitor their physical, educational, social and economic development. These were carried out in 1965, 1969, 1974, 1981, 1991, 1999/2000 and 2004. For the birth survey, information was obtained from the mother and from medical records by the midwife. For the purposes of the first three NCDS surveys, information was obtained from parents (who were interviewed by health visitors), head teachers and class teachers (who completed questionnaires), the schools health service (which carried out medical examinations) and the subjects themselves (who completed tests of ability and, latterly, questionnaires). The 1981 and later surveys differ, in that information was gathered by professional survey research interviewers. In 1981 information was obtained from cohort members and from the 1971 and 1981 Censuses (from which variables describing area of residence were taken). In the 1991 survey there was a professional interview with the cohort member, but self-completion questionnaires were also used to gather data from NCDS subjects and from husbands, wives, and cohabiters. Data were gathered from the children themselves, and from their mother, or mother-figure (who might be the cohort member, or their spouse or partner), using a series of age-specific assessments of cognitive and behavioural development. These were supplemented by a mother interview, and by interviewer observations of mother-child interaction. For the 1999-2000 sweeps, information was obtained from cohort members by interviewer and self-completion, but using CAPI (Computer-Assisted Personal Interviewing) for the first time. The 2004 survey was administered by telephone.

3.1 Our sample

For the purpose of our study, we focus on the four last sweeps of the cohort in order to have repeated measures of both lifestyles and health status as an adult. Data collected before 23 will be used to inform individual circumstances and childhood conditions. We excluded

cohort members who missed one or more of the fourth first sweeps in order to ensure a description of circumstances with a limited non response.

The balanced sample for which individuals have fully informed both health status and lifestyles in all the sweeps 4 to 7 contains 4,644 individuals whereas the unbalanced sample contains between 5900 and 7900 individuals (see table 1).

Table 1: Comparison of sample sizes from collected sample to selected samples

3.1.1   Health variable

The NCDS includes a repeated measure of the respondent's health in the cohort, namely self-assessed health (SAH). Respondents are asked to rate their own health on a four or five point categorical scale ranging from poor (waves 4, 5 and 6) or respectively very poor (wave 7) to excellent health status. Given the changes in scale in the variable over the different waves, we use SAH as a binary variable which takes the value one if the individual rates his health as good health or higher, and zero if she rates her health less than "good". This dichotomisation results in a loss of information as multiple-category are collapsed into a binary variable but it does not require any assumption on the scale of SAH and so, it enforces homogeneity on health reports over the four sweeps. Moreover, dichotomisation was particularly required as the necessary condition on equality of slope for ordinal Probit models turns out not to be verified when adding the lifestyles variables into our models. Self-assessed health has been shown to be a good predictor of mortality, morbidity and subsequent use of health care (Idler and Benyamini 1997).

Table 2: Distribution of health status in the balanced sample

Table 2 shows the age effect on health status over the life-cycle. From a mean of about 92.68% of good health at 23 years old, the proportion of respondents reporting a good health declines to about 78.18% at 46 years old. Between the first three sweeps the mean is declining by a constant rate of about 4 percentage points. There is a break with a decrease of 6 percentage points between the two last sweeps that are yet separated by only four years. This difference could be explained by an increasing effect of ageing on health when an individual enters her forties and be related to the mid-life crisis. This shift could also come from the change in the categorical scale of self-assessed health between wave 6 and wave 7 but we believe that this latter issue is minimised by the dichotomisation of health.

3.1.2 Education

The NCDS provides several current social characteristics. We assume that education level is a reliable proxy of other social outcomes such as employment, housing or income and consider that unique variable. Education is provided at each wave of the cohort. We use the highest diploma achieved and generated a binary indicator, that equals one if the cohort member achieved a diploma lower than secondary school and zero otherwise.

3.1.3 Lifestyles variables

The NCDS includes a longitudinal follow-up of lifestyles and health records at age 23, age 33, age 42 and age 46. We consider four lifestyles binary variables.

Exercising indicates whether the cohort member is regularly doing exercise or sports; it equals one if the cohort member has exercised at least once in the last four weeks and zero otherwise.

Non smoking informs whether the cohort member is a current smoker at the time of the wave; it equals one if she does not currently smoke and zero otherwise.

Drinking prudently is a gender-specific lifestyles based on the number of units of alcohol drinks taken the week before the interview. Males are considered to drink prudently if they drank between 0 and 21 units of alcohol in a week whereas it is between 0 and 14 units a week for females (Working Party of the Royal College of Physicians UK 2001; Balia and Jones 2008). The binary variable takes the value one if the respondent drinks prudently and zero otherwise. Regarding non-report of alcohol consumption, it might be interpreted therefore we use an additional variable related to the question "how often do you drink?" to identify those individuals and found two alternative types of respondents. The first type is a cohort member who reported no alcohol units in the past week and that he never drinks alcohol, he has therefore a true healthy behaviour and the binary variable drinking prudently takes the value 1. The second type is a respondent who had an "exceptional" non-consumption on that specific week. This type is identified if the individual reported to drink at least once a week and so we impute one unit of alcohol to her consumption in the past week.

A fourth lifestyle that we consider is the absence of obesity. Obesity may appear as an intermediated or genetic outcome of health and not a pure lifestyle. We can control the genetic and the family transmitted effect on obesity using the respondent's obesity status when she was 16. The absence of obesity thus captures aggregated effects of lifestyles. Absence of obesity is constructed using the reported height and weight and calculating individuals' body mass index (BMI[1]). The absence of obesity is a binary variable taken the value one if the cohort member's BMI is strictly lower than 30 and zero otherwise.

The four lifestyle variables are presented in table 3.

Table 3: Descriptive statistics of lifestyles variables in the balanced sample

3.1.4 Social and health-related childhood conditions

The vector of social and health-related childhood conditions that we consider has three main types of variables: social conditions in childhood, parents' health and health-related behaviours, and child and adolescent health.

The social conditions in childhood include the father's social class at the time of birth, the father and the mother's education level, and parental reports of financial hardships when the cohort member was 16. Father's social class is described into five categories: I – Professional or student, II – Managerial/technical, III – Skilled, IV – Partly skilled, V – Unskilled, and a sixth category is added if the mother reported no male figure in the household at the time of birth. Parental education consists in a two categories variable: parents who dropped out from school before or at the minimum age (14-15 years depending on the year of birth) and parents who were still at school after this age.

Parents' health is measured by parental report of chronic illness when the cohort member was 16 years old. Regarding parents' health related lifestyles, they reported their smoking status. We use a parental smoking indicator taking the value 1 if one of the two parents reported to be a smoker and an obesity status variable.

Respondent's health in childhood and adolescence are used both as control and achievement variables as they may represent health-related difficulties from the living environment during childhood. We use the same approach as Case et al. (2005) whose study was based on the same dataset and consider the report of at least one chronic condition at 16

---

[1] BMI in kg/m2= weight/height²

as well as a birth weight below 2500g as health indicators before adulthood. Furthermore, we introduce obesity at 16 years old as a circumstance. Since obesity in adolescence is medically assessed with gender-specific thresholds, we used those specific values.

The social and health-related childhood conditions are presented in table 4.

Table 4: Descriptive statistics of social and health-related childhood conditions variables in the balanced sample

4. Results

4.1 Random effect dynamic panel Probit results: baseline and structural specifications

The results of the random effect dynamic panel Probit for both the baseline and the structural specifications are presented in Table 5.

Table 5: Random effect Probit models coefficients and significance levels

The estimated coefficients in the two specifications are only comparable for social and health-related childhood conditions variables as they are identical variables in two models and the error terms of the two models have the same variance. However, the other explanatory variables are replaced by generalised residuals in the structural specifications[2]. The estimated coefficients related to those explanatory variables are thus only comparable in terms of signs and significance levels.

The results from the baseline model show that several childhood conditions variables have a direct effect on the probability to report good health, namely the father's social class, the experience of financial hardship during childhood, the mother's education level and both parent's illness. The education level of the cohort member appears to be significantly associated with health: a low educational level is negatively associated with the report of a good health over time. Regarding lifestyles variables, none of the lagged variables show a significant effect on health, and their effects on health appear to be concentrated on the means effect: the four lifestyles considered are significantly associated with reports of good health and the significance levels is the highest for the mean effect of exercising and being non-obese. This result suggests therefore a persistent effect of lifestyles on health instead of a

---

[2] The results of the auxiliary equations are not presented and commented but are available upon request.

transitory effect of lifestyles on health. The relevance of the dynamic modelling is shown by the significant effect on health of the past health status and the initial health as well as the important share of individual unexplained heterogeneity addressed by the model (39, 6% of the unexplained heterogeneity).

Result of the structural specifications permit confirming that social and health-related childhood conditions variables as well as current education level directly influence health over the life cycle whereas gender, father's education, mother's smoking status and childhood health variables (chronic conditions at 16, low birth weight and the individual obesity status at 16) are found having an indirect effect on health. Regarding lifestyles, the significant effect of the mean version of those variables remain except for prudent consumption of alcohol. Alcohol consumption therefore appears to influence health only indirectly. The differences between the results of the two specifications confirm that there are mediated effects of lifestyles, education and the state dependence of health status.

4.2 Decomposition of health inequality

Table 6 presents the results of the decomposition of the variance of predicted health in the baseline and in the structural specifications at every point of time. Results are similar between the waves because of the unicity of coefficients over time. They differ when variables are changing over time (e.g. lifestyles or lagged health) and when the changes in the latent health predictions are modifying the covariance between predicted health and the explanatory variables.

Table 6: Decomposition of health inequality at different ages

The decomposition in the baseline specification shows that the most important contribution to health inequalities comes from the state dependence of health and the initial health, which would explain between 40% and 44, 6% of the variance in the predicted health. Lifestyles are directly explaining between 34, 2% and 36, 5% of health inequalities, which confirm that they are important determinants of health inequalities. Childhood conditions explain about 15% of health inequalities and the direct impact of educational level is weaker and would represents about 7% of health inequalities.

The structural specification offers a different distribution among the determining factors of health inequalities. The relative importance of childhood conditions is increasing

and would represent between 26% and 28, 9% of health inequalities. This increase underlines the importance of the mediated effects of social and health-related childhood conditions with other determining factors. On the contrary, the contribution of lifestyles on health inequalities reduces and would represent about a quarter of inequalities. Lifestyles are thus strongly influenced by childhood characteristics and to a lesser extent by educational level whose contribution to health inequalities marginally increases in the structural approach. Health state dependence remains a very strong factor of health inequalities but its contribution is also reduced because of the influence of childhood conditions.

Comparison between the decompositions of the baseline model and the structural model suggests that the correlation between childhood conditions, and respectively lifestyles, past health states and educational level is important; as well as the correlation between education level and lifestyles, are important. When we purge the contribution of lifestyles to health inequalities from their mediated effect with childhood circumstances, we reduce their contribution to health inequalities and emphasise the importance of childhood circumstances for health inequalities over the life cycle.

We thoroughly study the share of social and health-related childhood conditions in health inequalities and explore the relative contribution of social background, parent's health, parents' lifestyles, and childhood health to this vector (Table 7).

Table 7: Decomposition of health inequality due to childhood conditions at different ages

The decomposition in the baseline model suggests that social background variables are the leading contributing factor, representing about 78% of the share of social and health-related childhood conditions in health inequalities. Parent's health and lifestyles represent about 21% whereas the contribution of childhood and adolescent health is very almost negligible. If we add their indirect contributions to health inequalities as done in the structural model, the relative contribution of childhood health increases and would represent 9% of the impact of social and health-related childhood conditions in health inequalities whereas social background and parent's health and lifestyles variables remain having with a substantial influence.

5. Conclusion

We developed a simple model to evaluate the contribution of several essential determinants of health to health inequalities using a representative cohort of individuals born in 1958 and a unique follow-up of health status, lifestyles as well as a good description of childhood conditions and education level. Our results show an impressive contribution of lifestyles to health inequalities, namely about 35% in the baseline specification. Nevertheless, we also found a predominant contribution of childhood conditions when their indirect role on education and lifestyles is taken into account, which underlines the relevance of mediated effects between the determinants of health and outperform previous works ignoring childhood characteristics as a relevant determinant of health inequalities. Among childhood circumstances, social background seems to be the more important determinant of overall health inequalities.

Adult health is also significantly influenced by average past lifestyles. Noticeably, average past lifestyles appear to matter more for adult health than transitory lifestyles. The study shows the indirect effect of the mother's education level, and individual health in younger age on adult health going through lifestyles. This indirect effect dramatically increases the relative contribution of childhood conditions to health inequalities, since their contribution equals 28% in the structural model and thus becomes directly comparable to the contribution of lifestyles.

Finally, the dynamic panel analysis permits controlling a large part of individual unexplained heterogeneity as well as the important effect of health state dependence over time (nearly 40% of the explained variance).

Our study has some limitations. The inequality measure is based on the explained part of the variance that is allowed by the model specifications. According to the pseudo-R^2 that is built using the variance of the latent variable, we would be able to explain 14.4%. Therefore the unexplained health inequality remains very large. Regarding education and lifestyles variables, we used binary indicators for the sake of simplification in order to measure the indirect influence of childhood conditions through those factors, but for some of these variables it would have been possible to carry out regression models with count data using for example the number of years of education, the number of cigarettes smoked per day, the frequency of exercise, etc. The panel data perspective also presents several limits. The first

problem is the presence of attrition due to mortality in the cohort that we have ignored in the analysis. This leads us to an underestimation of the effect of childhood conditions and education on health inequality as we worked on a selected sample of British people still alive at 46 years old. We did investigate on mortality in our data and we found that mortality rate appears to be more important before age 23 than between age 23 and age 46. Finally, the NCDS cohort has a singular structure as the different waves are not equidistant in time. In particular there is a four year interval between the two last sweeps whereas there were about ten years between the past sweeps. We try to catch this effect by introducing a year dummy into the models. Therefore, the estimated coefficients in the models can be interpreted as a mean of the effects of lifestyles, education, and childhood conditions over time.

Despite those limits, this study highlights new avenues for policy decisions. Results suggest recommending policies that would improve childhood conditions and prevent people from starting a poor behaviour or that would ensure permanent changes in lifestyles instead of transitory changes.

6. References

Anda R.F, Whitfield C.L, Felitti V.J, Chapman D, Edwards V.J, Dube S.R, Williamson D.F. Adverse childhood experiences, alcoholic parents, and later risk of alcoholism and depression. Psychiatric Services 2002; 53(8); 1001-9

Arendt, 2005 J.N. Arendt, Does education cause better health? A panel data analysis using school reforms for identification, Economics of Education Review 2005; 24 (2); 149–160

Balia S, Jones A. Mortality, lifestyle and socio-economic status. Journal of Health Economics 2008; 27(1); 1-26

Bambra C, Gibson M., Sowden A, Wright K, Whitehead M., Petticrew M. Tackling the wider social determinants of health and health inequalities: evidence from systematic reviews, J Epidemiol Community Health 2010a; 64; 284-291

Bambra C, Joyce KE, Bellis MA, Greatley A, Greengross S, Hughes S, Lincoln P, Lobstein T, Naylor C, Salay R, Wiseman M, Maryon-Davis A. Reducing health inequalities in priority public health conditions: using rapid review to develop proposals for evidence-based policy, J Public Health. 2010b; doi: 10.1093/pubmed/fdq028

Bourguignon F., Ferreira F., Menendez M. Inequality of Opportunity in Brazil. Review of Income and Wealth 2007; 53(4)

Case A. Katz L. The company you keep: the effects of family and neighborhood on disadvantaged youths, NBER Working Paper 1991

Chaloupka FJ.Warner KE. The economics of smoking. In A. J. Culyer and J. P. Newhouse (eds.), Handbook of Health Economics 2000. Elsevier

Chamberlain G. Panel Data. In Griliches and Intriligator (Ed.), Handbook of Econometrics 1984. Amsterdam: North-Holland.

Contoyannis P., Jones A. Socio-economic status, health and lifestyle. Journal of Health Economics 2004; 23; 965-995.

Contoyannis P., Jones A., Rice N. The dynamics of health in the British Household Panel Survey, Journal of Applied Econometrics 2004; 19(4); 473-503

Currie J, Stabile M. Socioeconomic status and child health: why is the relationship stronger for older children. American Economic Review 2003; 93; 1813-1823

Cutler D., Deaton A., Lleras-Muney A. The Determinants of Mortality, Journal of Economic Perspectives, American Economic Association 2006; 20(3); 97-120

Dekovic, Maja, J. M. Janssens and N. M. Van As (2003) Family predictors of antisocial behaviour in adolescence. Family Process, 42 (2), 223-235

Dias P. R. Inequality of Opportunity in Health: evidence from the UK cohort study. Health Economics 2009; 18(9); 1057-1074

Dias P. R, Jones A. Giving equality of opportunity a fair innings. Health Economics 2007; 16; 109-112

DiNardo J, Lemieux T. Alcohol, marijuana, and American youth: the unintended consequences of government regulation. Journal of Health Economics 2001; 20; 991-1010

Engels Rutger C. M. E., F. Vitaro, E. Den Exter Blokland, R. de Kemp, R. H. J. Scholte (2004), Influence and selection processes in friendships and adolescent smoking behaviour: the role of parental smoking. Journal of Adolescence 2004; 27;531-544

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

Francesconi M. SP. Jenkins ST. The effect of lone motherhood on the smoking behavior of young adults. Health Economics, 2010

Galobardes B, Smith GD, Lynch JW. Systematic review of the influence of childhood socioeconomic circumstances on risk for cardiovascular disease in adulthood. Ann Epidemiol. 2006 Feb; 16(2):91-104.

Grossman, M. Education and nonmarket outcomes. In: E. Hanushek and F. Welch, Editors, Handbook of the Economics of Education 2006; vol. 2, North-Holland, Amsterdam

Gourieroux C, Monfort A, Renault E, Trognon A. Generalised residuals. Journal of Econometrics 1987; 34(1-2); 5-32.

Heckman J. The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In C.F. Manski, & D. McFadden (Eds.), Structural Analysis of Discrete Data with Econometric Applications 1981; 179-195. Cambridge: MIT Press.

Idler E.L., Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. Journal of Health and Social Behaviour 1997; 38(March); 21-37.

Jusot F, Tubeuf S, Trannoy A, Effort or Circumstances: Does the Correlation Matter for Inequality of Opportunity in Health? Mimeo 2010

Kahn JR, Pearlin LI. Financial strain over the life course and health among older adults. Journal of Health and Social Behavior 2006; 47; 17-31

Khang YH, Lynch JW, Yang S, Harper S, Yun SC, Jung-Choi K, Kim HR. The contribution of material, psychosocial, and behavioral factors in explaining educational and occupational mortality inequalities in a nationally representative sample of South Koreans: relative and absolute perspectives, Social Science and Medicine. 2009 Mar; 68(5):858-66

Laaksonen M., Talala K., Martelin T., Rahkonen O., Roos E., Helakorpi S., Laatikainen T., Prättälä R. Health behaviours as explanations for educational level differences in cardiovascular and all-cause mortality: a follow-up of 60 000 men and women over 23 years, European Journal of Public Health 2008; 18(1); 38-43

Lantz PM, Golberstein E, House JS, Morenoff J. Socioeconomic and behavioral risk factors for mortality in a national 19-year prospective study of U.S. adults. Soc Sci Med. 2010; 70(10); 1558-66

Lindeboom M, Llena-Nozal A, van der Klaauw B. Parental education and child health: Evidence from a schooling reform. Journal of Health Economics 2009; 28(1); 109-131

Kenkel D., Should you eat breakfast? Estimates from health production functions. Health Economics 1995; 4(1); 15-29

McKelvey RD., Zavoina W. A statistical model for the analysis of ordinal level dependent variables. Journal of Mathematical Sociology 1975; 4; 103-120.

Marmot M., Shipley M., Brunner E., Hemingway H. Relative contribution of early life and adult socioeconomic factors to adult morbidity in the Whitehall II study, J Epidemiol Community Health 2001; 55(5); 301–307

Menvielle G., Boshuizen H., Kunst A.E., et al., The Role of Smoking and Diet in Explaining Educational Inequalities in Lung Cancer Incidence, JNCI Journal of the National Cancer Institute 2009; 101(5); 321-330

Mokdad AH., Marks JS., Stroup DF., Gerberding JL. Actual causes of death in the United States, 2000. Journal of the American Medical Association 2004; 291(10); 1238–1245

Mundlak,Y. On the pooling of time series and cross section data. Econometrica 1978; 46(1); 69-85

Osler M, Madsen M, Nybo Andersen AM, Avlund K, McGue M, Jeune B, Christensen K. Do childhood and adult socioeconomic circumstances influence health and physical function in middle-age? Soc Sci Med. 2009 Apr; 68(8); 1425-31

Powell, Lisa M., J. A. Tauras, and H. Ross (2005) The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. Journal of Health Economics, 24, 950-68

Power C, Hyppönen E, Smith GD. Socioeconomic position in childhood and early adult life and risk of mortality: a prospective study of the mothers of the 1958 British birth cohort. Am J Public Health. 2005 Aug; 95(8):1396-402

Shorrocks, A. Inequality Decomposition by Factor Components. Econometrica 1982; 50(1); 193-211.

Singh-Manoux A, Nabi H, Shipley M, Guéguen A, Sabia S, Dugravot A, Marmot M, Kivimaki M. The role of conventional risk factors in explaining social inequalities in coronary heart disease: the relative and absolute approaches to risk. Epidemiology 2008; 19(4); 599-605.

Skalická V, van Lenthe F, Bambra C, Krokstad S, Mackenbach J. Material, psychosocial, behavioural and biomedical factors in the explanation of relative socio-economic inequalities in mortality: evidence from the HUNT study. International Journal of Epidemiology 2009; 38(5);1272-84.

Strand BH, Tverdal A. Can cardiovascular risk factors and lifestyle explain the educational inequalities in mortality from ischaemic heart disease and from other heart diseases? 26 year follow up of 50 000 Norwegian men and women. Journal of Epidemiology and Community Health 2004; 58; 705-709

Stringhini S, Sabia S, Shipley M., Brunner E., Nabi H., Kivimaki M., Singh-Manoux A. Association of Socioeconomic Position With Health Behaviors and Mortality, JAMA 2010; 303(12); 1159-1166.

Trannoy A, Tubeuf S, Jusot F, Devaux M. Inequality of Opportunities in Health in France: A First Pass. Health Economics 2010; 19; 921-938.

Tubeuf S, Jusot F, Social health inequalities among older Europeans: the contribution of social and family background. European journal of health economics 2010 doi: 10.1007/s10198-010-0229-3

van Oort F. V A, van Lenthe F. J, Mackenbach J. Material, psychosocial, and behavioural factors in the explanation of educational inequalities in mortality in the Netherlands. Journal of Epidemiology and Community Health 2005; 59(3); 214–220.

Van den Mheen H, Stronks K, Van den Bos J, Mackenbach JP. The contribution of childhood environment to the explanation of socioeconomic inequalities in health in adult life: a retrospective study. Social science and medicine 1997; 44(1); 13-24.

Wickrama, K. A. S., R. D. Conger, L. E. Wallace, G. H. Elder, JR. (1999) The intergenerational transmission of health-risk behaviours: adolescent lifestyles and gender moderating effects. Journal of Health and Social Behaviour, 40, 258-272

WHO Tobacco Free Initiative. Building blocks for tobacco control: a handbook. Geneva, World Health Organization, 2004

Woodward, M. and Oliphant, J. and Lowe, G. and Tunstall-Pedoe, H. Contribution of contemporaneous risk factors to social inequality in coronary heart disease and all causes mortality. Preventive Medicine 2003; 36(5); 561-568.

Wooldridge J. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. Journal of Applied Econometrics 2005; 20(1); 39-54.

Working Party of the Royal College of Physicians UK. Alcohol--can the NHS afford it? Recommendations for a coherent alcohol strategy for hospitals. 2001. London.

Zhang L, Welte J. W, Wieczorek W. F. The influence of parental drinking and closeness on adolescent drinking. Journal of Studies in Alcohol 1999; 60; 245-251

7. Tables

Table 1: Comparison of sample sizes from collected sample to selected samples

	Birth	**NCDS 4**	**NCDS 5**	**NCDS 6**	**NCDS 7**
Year	1958	1981	1991	1999/00	2004
Age	Birth	23 yo	33 yo	42 yo	46 yo
Cross-sectional collected sample	17,416	12,044	10,986	10,979	9,175
Balanced selected sample			4,644		

Table 2: Distribution of health status in balanced sample (N=4644)

	Age 23	**Age 33**	**Age 42**	**Age 46**
Excellent	46.04%	35.75%	31.63%	32.21%
Good	46.62%	52.97%	53.14%	46.04%
Good health	**92.66%**	**88.72%**	**84.77%**	**78.25%**
Fair	6.76%	10.06%	12.75%	15.07%
Poor	0.58%	1.22%	2.48%	5.02%
Very poor				1.66%
Poor health	**7.32%**	**11.28%**	**15.23%**	**21.75%**

Table 3: Descriptive statistics of lifestyles variables in the balanced sample (N=4644)

	Age 23	Age 33	Age 42
Exercising	49.81%	80.06%	75.73%
No smoking	64.21%	71.62%	74.25%
Drinking prudently	87.47%	92.53%	84.99%
No obesity	97.33%	89.92%	85.66%

Table 4: Descriptive statistics of social and health-related initial conditions in the balanced sample

Variables	N=4644	Proportion	
Gender			
Male	2162	46.55	%
Female	2482	53.45	%
Fathers' social class			
I - Professional	220	4.74	%
II - Managerial/technical	664	14.30	%
III - Skilled	2749	59.19	%
IV - Partly skilled	531	11.43	%
V - Unskilled	328	7.06	%
No male head	152	3.27	%
Financial hardship			
Yes	339	7.30	%
No	4197	90.37	%
Non response	108	2.33	%
Father's education			
Minimum schooling age and below	3579	77.07	%
Beyond the min age	1065	22.93	%
Mother's education			
Minimum schooling age and below	3584	77.17	%
Beyond the min age	1060	22.83	%
Parental illness			
Father's illness	325	7.00	%
Mother's illness	233	5.02	%
Parental smoking			
Father's smoking	2526	54.39	%
Non response	305	6.57	%
Mother's smoking	2034	43.80	%
Non response	152	3.27	%
Chronic condition at 16			
Yes	548	11.80	%
No	3595	77.41	%
Non response	501	10.79	%
Low birth weight	221	4.76	%
Obesity at 16			
Yes	60	1.29	%
No	3948	85.01	%
Non response	636	13.70	%
Cohort member's education:			
High educational level	3739	80.51	%
Low educational level	905	19.49	%

Table 5: Random effect Probit models coefficients and significance levels (NT=13932; Rho1=0,396; Rho2=0,397)

Variables	Baseline model[$]		Structural model	
Gender Male	0,037		0,120	***
Fathers' social class (Ref.: V - Unskilled)				
I - Professional	0,287	**	0,423	***
II - Managerial/technical	0,201	**	0,356	***
III - Skilled	0,145	*	0,250	***
IV - Partly skilled	0,033		0,086	
No male head	-0,137		-0,146	
Financial hardship (Ref.: None)				
Yes	-0,221	***	-0,377	***
Non response	0,116		0,027	
Father's education (Ref.: beyond minimum schooling age)				
Before or at the min age	-0,037		-0,108	*
Mother's education (Ref.: beyond minimum schooling age):				
Before or at the min age	-0,152	***	-0,212	***
Parental illness (Ref.: None)				
Father's illness	-0,187	**	-0,235	***
Mother's illness	-0,164	*	-0,209	**
Parental smoking (Ref.: None)				
Father's smoking	0,060		0,015	
Non response	0,013		0,023	
Mother's smoking	-0,068		-0,159	***
Non response	-0,080		-0,121	
Chronic condition at 16 (Ref.: None)				
Yes	-0,009		-0,120	*
Non response	0,109		0,173	
Low birth weight	-0,074		-0,155	*
Obesity at 16 (Ref.: No):				
Yes	0,274		-0,363	**
Non response	-0,137		-0,225	*
Education : Low educational level	-0,225	***	-0,237	***
Lifestyles (mean):				
Exercising	0,573	***	0,346	***
No smoking	0,200	**	0,121	**
Drinking prudently	0,212	*	0,101	
No obesity	0,766	***	0,380	***
Lifestyles (lagged):				
Exercising	-0,043		-0,032	
No smoking	0,084		0,047	
Drinking prudently	-0,030		0,024	
No obesity	-0,070		-0,036	
Health status (lagged)	0,304	***	0,163	***
Health status at 23	1,009	***	0,495	***
Age 42	-0,240	***	-0,255	***
Age 46	-0,587	***	-0,611	***

[$]Significance levels of test of rejecting the hypothesis of the nullity of the coefficient: *** 1%, **5%, *10%.

Table 6: Decomposition of health inequality at different ages: contribution in % (variance)

	Baseline model			Structural model		
	Age 33	**Age 42**	**Age 46**	**Age 33**	**Age 42**	**Age 46**
Sex	0,36%	0,42%	0,40%	1,22%	1,35%	1,34%
	(0,001)	*(0,001)*	*(0,001)*	*(0,004)*	*(0,004)*	*(0,004)*
Childhood conditions	14,27%	15,76%	15,67%	26,04%	28,85%	28,55%
	(0,045)	*(0,046)*	*(0,046)*	*(0,084)*	*(0,084)*	*(0,084)*
Lifestyles	34,16%	36,48%	35,15%	24,92%	27,01%	26,35%
	(0,107)	*(0,106)*	*(0,104)*	*(0,080)*	*(0,079)*	*(0,078)*
Education	6,60%	7,33%	7,20%	7,43%	8,24%	8,14%
	(0,021)	*(0,021)*	*(0,021)*	*(0,024)*	*(0,024)*	*(0,024)*
Health state-dependence	44,62%	40,01%	41,59%	40,38%	34,55%	35,61%
	(0,139)	*(0,116)*	*(0,123)*	*(0,130)*	*(0,101)*	*(0,105)*
$\text{Var}(\hat{H}^{*})$	100%	100%	100%	100%	100%	100%
	(0,312)	*(0,290)*	*(0,296)*	*(0,323)*	*(0,291)*	*(0,294)*

Table 7: Decomposition of health inequality due to childhood conditions at different ages: contribution in % (variance)

	Baseline model			Structural model		
	Age 33	**Age 42**	**Age 46**	**Age 33**	**Age 42**	**Age 46**
Social background	77,56%	77,64%	77,46%	69,52%	69,52%	69,52%
	(0,035)	*(0,036)*	*(0,036)*	*(0,058)*	*(0,058)*	*(0,058)*
Parent's health and lifestyles	21,27%	20,84%	21,10%	21,35%	21,35%	21,35%
	(0,009)	*(0,010)*	*(0,010)*	*(0,018)*	*(0,018)*	*(0,018)*
Initial health	1,16%	1,52%	1,44%	9,13%	9,13%	9,13%
	(0,001)	*(0,001)*	*(0,001)*	*(0,008)*	*(0,008)*	*(0,008)*
Childhood conditions	100%	100%	100%	100%	100%	100%
	(0,045)	*(0,046)*	*(0,046)*	*(0,084)*	*(0,084)*	*(0,084)*