Soc Choice Welf (2017) 49:499–544
DOI 10.1007/s00355-017-1044-x

ORIGINAL PAPER

Inequality of income acquisition: the role of childhood circumstances

Paul Hufe[1] · **Andreas Peichl**[2] · **John Roemer**[3] · **Martin Ungerer**[4]

Received: 23 December 2015 / Accepted: 21 March 2017 / Published online: 29 March 2017

Abstract Many studies have estimated the effect of circumstances on income acquisition. Perhaps surprisingly, the fraction of inequality attributable to circumstances is usually quite small—in the advanced democracies, approximately 20%. One reason for this is the lack of data on circumstance variables in empirical research. Here, we argue that all behaviors and accomplishments of children should be considered the consequence of circumstances: that is, an individual should not be considered to be responsible for her choices before an age of consent is reached. Using two data sets that contain data on childhood accomplishments, other environmental circumstances and the income as an adult, we calculate that the fraction of income inequality due

We are grateful to the Russell Sage Foundation for support. We would like to thank Francisco Ferreira, Dirk Van de gaer as well as the participants in seminars in Mannheim and workshops in London and at Duke University for their helpful comments and suggestions.

✉ Andreas Peichl
peichl@zew.de

Paul Hufe
hufe@zew.de

John Roemer
john.roemer@yale.edu

Martin Ungerer
ungerer@zew.de

1 ZEW and the University of Mannheim, Mannheim, Germany

2 ZEW, the University of Mannheim, IZA and CESifo, Mannheim, Germany

3 Yale University, New Haven, USA

4 ZEW and the University of Cologne, Mannheim, Germany

to circumstances in the US rises from 27 to 43% when accounting for childhood circumstances. In the UK it rises from 18 to 27%.

1 Introduction

There is now a large theoretical and empirical literature in economics on inequality of opportunity (IOp),[1] which evolved out of writings in political philosophy, beginning with John Rawls and extending to the present day. In one prominent formulation (Roemer 1993, 1998), outcomes that individuals enjoy (such as income) are the consequence of two sorts of factor: *circumstances*, those characteristics of a person and her environment that are beyond her control or for which she cannot be held responsible, and *effort,* which comprises those choices within her realm of control. Equality of opportunity is said to hold if individual positions in an outcome distribution are independent of individual circumstances, and only influenced by personal effort. The empirical literature measures the extent of IOp for various outcomes (income, wages, health) in many countries.[2]

Almost all empirical studies estimate that the extent to which income inequality is dependent on circumstances is quite small.[3] Since it is this part of inequality that is ethically troubling, the conclusion might be drawn that any existing income inequality is ethically acceptable, being largely dependent on differential effort. Indeed, Kanbur and Wagstaff (2016) have recently argued that equality-of-opportunity studies may be doing more harm than good, because they could be used by policymakers in developing countries to argue that most income inequality is ethically acceptable, and that social policy need not be concerned with reducing it.

We believe that the equal-opportunity approach based on the dichotomy between circumstances and effort is valuable, as it is based upon sound philosophical principles. Moreover, surveys routinely find that most people agree with the principle that inequalities due to circumstances should be rectified by social policy. Indeed preferences for redistribution are systematically correlated with beliefs about the relative importance of effort and luck in the determination of outcomes (Alesina and Giuliano 2011). Individuals are more willing to accept income differences that are dependent on effort (or laziness/industriousness) rather than on circumstances beyond individual control (e.g., Fong 2001). Furthermore, the experimental literature has shown convincingly that people do not merely endorse this fairness ideal in theory, but are willing to act on it even if their own material interests are at stake (Cappelen et al. 2007, 2010).

However, we also believe that previous measurements of IOp are inadequate. Many important circumstances that play a role in income determination have been ignored

[1] See also the recent 'The Equality of Opportunity Project' for the US: http://www.equality-of-opportunity.org/.

[2] For recent survey articles on both the theoretical and empirical literature, see Fleurbaey and Peragine (2013), Roemer and Trannoy (2016), Ferreira and Peragine (2016), or Ramos and Van de gaer (2016).

[3] Notably, the share of inequality explained by circumstances appears to be higher in developing as opposed to industrialized countries (see Brunori et al. 2013). Furthermore, relative measures of IOp are higher for consumption expenditures than for income measures. Yet as discussed by Ferreira and Gignoux (2011) this difference is attributable to the larger transitory component of the latter measure of economic advantage.

in the empirical literature. The effects of these circumstances appear statistically as effort, because effort is often measured as the residual cause of income variation after explicit circumstances have been accounted for. Hence, the measurement of IOp is biased downward, perhaps considerably so (see the simulations in Bourguignon et al. 2007 and the discussions in Ferreira and Gignoux 2011; Niehues and Peichl 2014).

In this paper we make use of high-quality micro-panel data to correct this shortcoming. In particular, we use the National Longitudinal Survey of Youth 1979 (NLSY79) and the 1970 British Cohort Study (BCS70) to construct fine-grained circumstance sets that take account of both the social environment of children and their cognitive and non-cognitive achievements during childhood.

The central issue we must confront is what aspects of the child's environment and performance should be deemed as comprising, or due to, circumstances. We take what some might find to be a radical position: that all measurable achievements and behaviors of children, before an age of consent is attained, are the result of their circumstances. We believe that children should not be held responsible for any of their accomplishments before that age.[4] Indeed, we could take a cue from the law and use the sexual age of consent, or the age at which a child is judged to be an adult in a court of law to be the age of consent for responsible choice. Ideally, if we had a complete biography of the child at the age of, say, sixteen, we would consider that to comprise the child's circumstances.

In particular, we need not distinguish between the effects of nature and nurture: a child's genetic and somatic make-up is certainly a circumstance. Some may object to this, believing that the child deserves to benefit from her innate traits. We demur—at least we do not believe a person *deserves a higher income* because she has valuable inborn traits. This does not mean we begrudge naturally talented people the satisfaction they enjoy from being beautiful, intelligent or charming. But our study here concerns equality of opportunity for income, not life satisfaction, and we do not countenance the view that such desirable traits should result in more generous material conditions. Naturally, this would imply that equalizing income opportunity must—at least to some degree—conflict with the reward structure of market economies.

Our analysis shows a significant increase in IOp measures when we expand the set of circumstances to include the attributes of the individual and her environment as a child. We find that the fraction of income inequality explained by circumstances rises from 26.8 to 43.5% using the NLSY79 and from 17.9 to 26.9% in the BCS70.[5]

In Sect. 2, we present our conceptual framework and methodology. Section 3 provides some intuition on the potential role of circumstances in explaining income determination, Sect. 4 describes the data sets, Sect. 5 displays our results, and Sect. 6 concludes.

[4] Richard Arneson writes that one can simplify the true process of the development of responsibility in a person by thinking of a canonical moment at which children become responsible for their choices. The "canonical moment" simplifying abstraction of the equal opportunity principle is motivated by the thought that there is a non-arbitrary and morally significant line between childhood and adulthood and that children are not morally responsible for their preferences in the way that adults are deemed to be (Arneson 1990).

[5] Note that these baseline estimates would increase further when allowing for heterogeneous effects of circumstances on income. See Hufe and Peichl (2015) for a discussion.

2 Conceptual framework and methodology

The main outcome of interest in this study is individual gross income y. One measurement of the extent of income inequality due to circumstances is defined as follows. Consider the mean logarithmic deviation (MLD) of an income distribution $F(y)$. $MLD(F)$ captures the total inequality of outcomes (IO). Let us suppose we have partitioned the population into types, each type corresponding to the set of individuals with a given set of circumstances. Each type is characterized by its own income distribution. Let the type distributions be $\left\{F^t(y), t \in \Omega\right\}$, where Ω is the set of types, and let type t have mean income μ^t and frequency f^t in the population. Define the *smoothed distribution* Φ with respect to this decomposition of F as a (counterfactual) distribution which assigns every member of type t the mean income μ^t. The graph of the cdf of Φ is a step function with as many steps as there are types. The *MLD* is a convenient measure of inequality because the following decomposition holds:

$$MLD(F) = MLD(\Phi) + \sum_{t=1}^{T} f^t MLD\left(F^t\right). \tag{1}$$

Therefore, we can interpret the first term on the right-hand side of (1) as the inequality due to circumstances, and the second as the inequality due to differential effort. After all, were the true distribution Φ, then differential effort within types would make no contribution to income since by construction all members of a given type have the same income in Φ. The ratio

$$r = \frac{MLD(\Phi)}{MLD(F)}$$

is therefore a measure of the extent to which income inequality is due to circumstances.

The disturbing result we mentioned earlier is that, using popular data sets, which record information on a limited set of circumstances, one finds that the measured value of r is quite small—far less than one half—especially in highly developed countries (Brunori et al. 2013). Is it really the case that much more than half of income inequality is due to differential effort, as these results would suggest, or are we seriously underestimating the effect of circumstances due to poor data sets?

The approach that we have just summarized, using the MLD decomposition, is non-parametric: it partitions the population into types defined by their circumstances, and takes as data the type distributions and the aggregate distribution of the outcome of interest. This non-parametric estimation of the role of circumstances in causing inequality is, unfortunately, of limited use, because it is only feasible if we have meaningful distributions of income by type. That requires either a very large data set, or a small set of types. Suppose, for example, we had 20 circumstances, each of which could take on three values (low, middle, high). Then the set of types would have 3^{20} elements. Even if one-third of these were empty, we would still have 3^{19} types. To get statistically meaningful distributions for all types, we would need, say, $50 * 3^{19}$, or about 84 billion observations. To circumvent this problem, practitioners use parametric estimations of Φ, in which regression analysis replaces the partition of the population into a typology. Using a parametric approach, we can estimate the impact of numerous circumstance variables even with the presence of small sample and cell sizes.

We follow Ferreira and Gignoux (2011) and Niehues and Peichl (2014) who use a parametric specification to estimate lower bounds of IOp. The empirical specification reads

$$ln\ y_i = \boldsymbol{\alpha C_i} + u_i, \tag{2}$$

and can be estimated by OLS to derive the fraction of variance that is explained by circumstances. In this reduced form, the estimates measure the overall effect of circumstances on earnings, including the indirect effect of type-specific effort heterogeneity. Based on this estimation, we can construct a parametric estimate of the smoothed distribution Φ defined earlier by replacing earnings outcomes by their predictions:

$$\hat{y}_i = \exp\left(\boldsymbol{\hat{\alpha} C_i}\right). \tag{3}$$

We then let Φ be the distribution of these estimated incomes. In this counterfactual, all individuals with the same circumstances necessarily have the same incomes. Thus, in the case where all income differences are due to circumstances (and so the error terms in (2) are all zero), the ratio r would be unity. Thus r can be rewritten as:

$$r = \frac{MLD\left(\left\{\hat{y}_i\right\}\right)}{MLD\left(\left\{y_i\right\}\right)}$$

Practitioners recognize that this procedure leads to lower bound estimates of the true share of inequalities due to circumstances. The intuition for this is just like that of an R^2-measure (Ferreira and Gignoux 2011): adding another circumstance variable to the analysis increases the explained variation (or at least does not decrease it in the case of orthogonality), and hence the share of inequality due to circumstances cannot decrease, although coefficients might be upward or downward biased. However, not all potential circumstances are usually observable in the data. Therefore, the extent of this underestimation bias is unclear.[6]

Moreover, circumstances, taken from typical data sets, often appear to explain very little of the inequality in the aggregate distribution of income. Roemer (2017) attempts to explain this fact by raising the following question: given an aggregate distribution of income F, and T types with frequencies $f = f^1, \ldots, f^T$ and mean incomes $\mu = \mu^1, \ldots, \mu^T$, what is the *maximum* value of r that could be attained, if we were able to choose the T-component (type) distributions F^t subject only to the conditions that F^t possess mean μ^t, types t have frequency f^t in the population, and, of course, that the convex combination of the component distributions equals F?

Roemer (2017) shows that *the supports of the type distributions* resulting in a maximal r are mutually disjoint.[7] However, this is typically not the case in reality.

[6] Note that this methodology is reminiscent of the inequality decomposition suggested by Cowell and Jenkins (1995). However, their decomposition does not strictly rely on the division between circumstances and effort.

[7] The same question has been studied by Elbers et al. (2008). Their solution, however, does not generalize to continuous distributions.

Instead, the supports of the type distributions are overlapping—and very far from being disjoint.[8] This observation suggests that to get relatively high values of r, we need circumstances that define types with the property that there are many subsets of types that share very little income mass. Usually this is not the case when we use the common circumstances of parental education, occupation, race, or region of the country. Put another way, market economies do a pretty good job of equalizing opportunities for income acquisition, if we define the typology to be sufficiently coarse.

3 Data

For our empirical analysis, we use two data sets, NLSY79 and BCS70, which are described in turn after a short overview of the different circumstance sets that we are using.

3.1 Sets of circumstances

The empirical analysis comprises several scenarios including different sets of explanatory variables beyond the individual's control. We grouped the explanatory variables into meaningful subsets by topics. There is a base scenario and four further specifications. The base scenario is chosen to include the circumstances most commonly used in the literature (such as parental background and ethnic origin), whereas the other scenarios include more detailed childhood outcomes unique to the data at hand and novel to this literature. While scenarios one to four feature a certain degree of comparability between NLSY79 and BCS70, the fifth circumstance set comprises variables unique to the respective data sets.

Table 1 provides an overview of the circumstance sets we consider (see also Appendix Tables 4 and 5 for more details on the respective variables). In particular, sets two to four contain information on attributes and achievements of the individual as a child, before the age of consent.

3.2 The National Longitudinal Survey of Youth 1979 (NLSY79)

The NLSY79 is a longitudinal micro-study sponsored by the US Department of Labor, the first wave of which was collected in 1979 from a nationally representative sample of individuals aged 14–21 on December 31, 1978. It makes available a wealth of information on respondents' educational, income and employment biographies, family processes, health-related behaviors as well as psychological dispositions and cognitive abilities. At the time of the first round the sample consisted of 12,686 respondents covering the cohorts 1957–1964. This implies that respondents were aged between 47 and 56 in 2012, the year of the latest available survey round. As of 1986 the NLSY79 has been accompanied by the Child and Young Adults supplement (NLSY79 Child/YA),

[8] See Assaad et al. (2015) for an example calculation using Egyptian data.

Table 1 Overview of circumstance scenarios

Scenario	Circumstance set	Circumstance var.
First	Base	Sex, country of birth, ethnic affiliation, cohort, academic achievement mother, occupation code mother, rural/urban, height, family income
Second	Child–parent relationship	Childcare, play w/ parents, perceived quantity of time w/ mother, parents split, schoolwork support from parents
Third	Health-related behavior	Smoking during pregnancy, smoking habits mother, drinking habits mother, school absence due to health, restrictions in school work due to health, inability to play due to health, age of mother at birth, vaccination
Fourth	Ability	AFQT mother, standardized math and reading assessment
Fifth	Survey specifics	Specific to NLSY79 and BCS70. See text and appendix Tables 4 and 5 for more information

which tracks the lives of all biological children of female NLSY79 respondents. It thus greatly expands the scope of child information collected. Interviews are conducted on a biennial basis, where separate questionnaires are administered to children below the age of 15 and young adults above this age. The former collects detailed information from both mothers and children on psychological and physiological child development, socio-economic background characteristics, family interactions and educational assessments. The latter is based on the NLSY79 questionnaire and provides a host of information on outcome variables, such as income and educational achievements. As of the 2012 wave, 11,512 descendants of the NLSY79 cohort have been interviewed covering the age range 14–41. The breadth of available information on mothers originating from NLSY79 as well as the detailed records on living conditions and socialization processes of children before the age of consent originating from the child questionnaires, make this study particularly suitable to construct rich circumstance sets for the estimation of IOp. The NLSY79 oversampled Hispanic and African American respondents, which is also reflected in the raw sample of the Child and Young Adults supplement. Furthermore, given the long time frame of interest, non-random sample attrition may call the representativeness of our results into question. In order to address both issues, we use sample weights at the level of the respective outcome variables in order to turn our raw sample into a population that is nationally representative for children born to mothers who were aged 14–21 on December 31, 1978.

Our base scenario comprises a set of basic demographic characteristics of the respondents. In particular, we include dummies for sex, country of birth, ethnic group and the respondent's cohort in order to take generational effects into account. Furthermore, we control for maternal educational achievement by including indicator variables for different academic degrees grouped in three categories of increasing rank. To be sure, we restrict ourselves to degrees attained before the year of child birth. Similarly, we introduce a battery of occupation dummies for the mother, which are measured in the year of child birth. To further refine our account of the child's socio-economic background we employ the following circumstances which are observed at different age thresholds of the child. Neighborhood characteristics are introduced by dummy variables for whether the child lived in a Metropolitan Statistical Area (MSA), i.e. a core urban area with a population of 50,000 or more, and if yes, whether its residence is located in the center of such an area. Lastly, we include the net family income and the child's height.

Scenario 2 extends the scope of circumstances to the child-parent relationship. In particular we take account of the family status of parents by controlling for whether parents lived in the same household subject to the condition that both were alive. Moreover, we construct a binary variable from the child's responses on whether their parents spent time with them engaging in games and activities and whether they supported them in their schoolwork. Lastly, we measure the child's desire to spend more or less time with each parent in three categorical variables of increasing wish intensity. All variables in this group are measured at different age thresholds of the child.

Subsequently, we focus on health-related information for both children and parents in scenario 3. As regards the former, we make use of the mothers' assessment on whether her child's school attendance, school work or leisure activities were restricted due to a medical condition. Again we measure these reports at different points of time in the child's biography. Furthermore, we record the mother's age at birth as another circumstance variable related to the child's health status. With respect to parental behavior we are confined to maternal information. Therefore, we include a dummy variable for whether the respective child has ever been exposed to a mother smoking on a daily basis. Additionally, we take account of the consumption of alcoholic beverages by including indicator variables for monthly drinking frequencies measured at the child's age of eight. It is noteworthy that these latter variables on smoking and drinking behaviors yield important sample size reductions. Therefore, we will also consider an alternative reduced set of health-related information by exclusively focusing on the restrictions placed on the child as a result of medical conditions and the mother's age at birth.

The fourth scenario makes use of the availability of academic achievement tests in the NLSY79 Child/YA in order to serve as a proxy for the child's ability. Ability at age 16 or younger is assumed to be beyond personal control and hence can be interpreted as a circumstance. A common approximation of ability is the use of standardized test scores. Specifically, at this stage we include the standardized score of the Peabody Individual Achievement Test (PIAT) in the areas of mathematics, reading/recognition and reading/comprehension measured at different age thresholds. PIATs are widelyused measures of academic achievement credited with high test-retest reliability and concurrent validity. We furthermore exploit the fact that all NLSY79 respondents were subjected to an Armed Forces Qualification Test (AFQT) at the beginning of the study. Thus, we are able to include the mother's AFQT-score as a proxy variable for maternal intelligence.

Lastly, we augment the circumstance set considered thus far with a host of variables that do not immediately correspond to either of the outlined categories and have no analogue in the BCS70. First, for educational background we include a binary indicator for whether a child attended a public as opposed to a private school. Moreover, four variables are introduced that measure the number of people with a certain educational level in the household at different age cut-offs. The considered levels are "Less than 12 years", "12–13 years", "13–15 years" and "$>$ 15 years". Second, we introduce a series of psychological assessment scores. For the child we include the total percentile score of the Behavioral Problems Index (BPI). The BPI is an aggregate measure of child behavior and attitudes constructed from a series of 28 questions posed to mothers of children between four and 14 years of age. Again, we make use of the availability of test scores for each child at different ages. Similarly, NLSY79 conducted psychometric assessments with every respondent at the beginning of the study. As a result, we are able to include the Pearlin Mastery Scale, Rotter's Locus of Control Scale and the Rosenberg Self-Esteem Scale for the mothers. The first two scores measure the extent to which respondents perceive themselves to be in control of forces that impact their lives. As its name suggests, the Rosenberg score can be interpreted as a measure of self-esteem. Lastly, socio-economic background variables of the circumstance set are enriched by a binary indicator on whether the mother was ever convicted of a crime.

3.3 The 1970 British Cohort Study (BCS70)

The BCS70 is a longitudinal survey funded by the Economic and Social Research Council and managed by the Centre for Longitudinal Studies. It follows the lives of more than 17,000 individuals born in England, Scotland, and Wales in a single week in 1970. Since the first survey wave in 1970, there have been eight follow-up interviews of all cohort members at ages 5, 10, 16, 26, 30, 34, 38, and 42. The latest survey was carried out in 2012. In addition to the main interviews, there have been five sub-studies where additional data has been collected from samples of cohort members selected for their particular characteristics or circumstances. The survey is supplemented by interviews with the parents and headteachers, standardized test scores, health records, nutrition and activity diaries as well as labor market histories. Thereby, the BCS70 has collected information on health, physical, educational and social development, and economic circumstances. The data set contains detailed information on early childhood and parental background. In contrast to NLSY79, questionnaires are filled by both parents, revealing broader information on parental background. Moreover, similar and identical questions on family and social situation are addressed to parents and children. While the BCS70 starts with 16,569 individuals giving full or at least partial response in 1970, the sample size decreases to 9354 until the most recent sweep in 2012 (Mostafa and Wiggins 2014). Deaths account for 853 cases while 7077 initial cohort members are classified as unproductive due to relocation and non-response. Both missing information and non-response are a potential cause of bias in standard errors and point estimates if those patterns are non-random. Considering the estimation of a smoothed distribution, if individuals living in poor circumstances are more likely to drop out of the survey, the dispersion in the smoothed distribution is lower and we underestimate IOp. Mostafa and Wiggins (2014) state that individuals from the bottom of the distribution are in fact more likely to drop out and show that non-response within one sweep can be accounted for by using sampling weights. However, non-response weights can only adjust for non-response in one wave of the survey at a time. As we use variables from multiple points in time, these weights would not work properly. Considering metadata on interviewers and further information regarding the process of data collection would improve weights. Unfortunately, this data is not available in the BCS70.

The baseline scenario covers basic demographic and parental background variables. We include dummies for gender and foreign origin, defined by the birthplace of the mother. Furthermore, we define four categories of academic achievement of the mother at birth, "No degree", "Secondary", "Intermediate", and "College". In the same way, we include six occupational categories for the mother at birth. In order to account for the socio-economic background of the child, we use the urbanization degree of the child's neighborhood and parental income as explanatory variables. Urbanization is measured by three dummies, grouping rural, suburbs/towns, and inner urban areas at the age of 10. Parental income is measured at ages 10 and 16 and classified into seven (10) and eleven (16) groups. Finally, we use the height of the individual measured at age 10.

Scenario 2 covers detailed information on the family background and the child-parent relationship. The time spent with the parents might affect the character of the individual. Therefore, we use the average time spent with the parents in a week as stated by the child at ages 10 and 16. It is classified into 5 groups at age 16, from "Most days a week" to "Little or never" and into three groups at age 10. In addition, we utilize the questions on common indoor or outdoor hobbies shared with the parents at both age 10 and age 16. In order to account for the potential effects of childcare, we include a dummy for whether the child spent one year or more in pre-school childcare until age 5. Additionally, we use variables on missing fathers and mothers as well as the death of parents before age 5. As a final part of this scenario, we include the marital status of the parents at birth, categorized into "Single", "Partnership", "Separated", and "Widowed".

Scenario 3 deals with health and medical conditions of individuals and parents during childhood. The BCS70 contains detailed information on the smoking and drinking habits of the mother, such that we can use smoking behavior at ages 10 and 16, as well as alcohol consumption measured at age 16. Related to smoking behavior during childhood, we also observe smoking behavior during pregnancy. In addition, we include the age of the mother at birth of the child as well as the birthweight of the child. As severe sickness might affect school attendance and hence education, we construct a dummy for missing school days due to sickness at age 10. Finally, we construct a variable indicating any vaccination of the child until age 5.

Scenario 4 covers additional ability measures during childhood. The BCS70 provides information on standardized vocabulary test scores as well as standardized math test scores at the age of 10. The test used for the assessment of reading ability is the

Edinburgh Reading Test, while the Friendly Math Test is used to account for ability in mathematics.

Finally, scenario 5 consists of further variables available in the BCS70 that could not be classified into the previous scenarios. For the BCS70, this scenario consists of the education of the father, a dummy for singleton children, and information on whether the child was firstborn. The BCS70 offers a huge variety of variables that are of potential interest. However, in view of the small sample sizes and the ensuing limitations in the available degrees of freedom, we refrain from using this information in the following analysis.

4 Results

4.1 NLSY79

As outlined previously, the observational units covered by the NLSY79 span a wide time range. While the first children of NLSY79 mothers were born as early as 1970, the latest birth we observe dates 1997. Therefore, a choice has to be made of how to treat the variance in age. First, we use appropriate sample weights to maintain the representativeness of the sample for each year of observation respectively.[9] Second, we restrict the sample to all subjects aged 25-30 in the year of observation. Ideally, one would conduct the analysis separately for each cohort within the sample. However, this approach leads to very small sample sizes, especially towards the age threshold of 30, which is commonly assumed to be a strong predictor of long-term earning potential (Chetty et al. 2014). Therefore, our approach of considering a restricted range of six cohorts, strikes a balance between the ambition to maintain a reasonably sized sample and to cushion the effect of outlier incomes of younger cohorts who are at the beginning of their careers. Lastly, within the range of cohorts we take account of the remaining variance in age by including the year of birth in the set of circumstance variables (see Section 3.1). To further address the influence of transitory income components, we consider incomes averaged over the reporting period 2010-2012 as an outcome variable. While in the following we also report all results for the years 2008, 2010 and 2012 individually, we will focus our discussion on this average measure.

Figure 1 gives an overview of how our IOp estimates vary as we sequentially introduce the circumstance sets laid out in Table 1. First of all, it is noteworthy that inequality in the US is higher according to this sample than in comparison to other works. Pistolesi (2009) uses the PSID to calculate a Theil index of permanent labor earnings of 0.25 in 2001. Similarly, Niehues and Peichl (2014) rely on the same data source to calculate an MLD of 0.24 with respect to permanent gross earnings. The NLSY79 sample used in this study yields an MLD of approximately 0.60 for the average measure of income. It is important, however, not to overemphasize these differences. Niehues and Peichl (2014), for instance, use a sample of individuals aged between 25 and 55. Their latest period of observation for the US dates back to 2007

[9] Recall that the sample is not representative for the entire US population in the respective year, but nationally representative for the subpopulation of children born to mothers aged 14–21 on December 31, 1978. See Appendix Table 4 for a change in summary statistics when restricting the sample to the age range 25-30 and applying the respective sample weights.

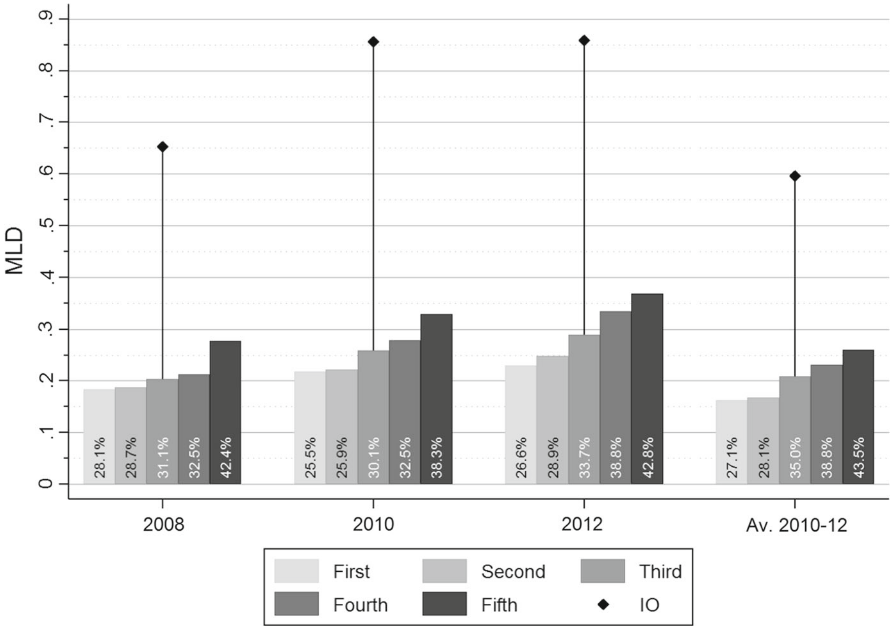

Fig. 1 IOp over Time (NLSY79). The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower bound absolute measure of IOp. The number in each *bar* is the measure of r, the fraction of inequality due to circumstances. For example, the height of the *second bar* in the 2008 cluster is the measure of $MLD\,(\Phi)$ where the set of circumstances is the second set in Table 1, and the outcome variable is income in 2008. In this example, $r = \frac{MLD(\Phi)}{MLD(F)} = 28.7\%$. All circumstances are measured either at birth or age 16 with no circumstance repeatedly measured at different age thresholds

and they average incomes over at least five consecutive years. Given these differences in samples and the timeframe of analysis, significant differences in inequality are to be expected. To be sure, when we discuss the impact of expanding the set of circumstances to include the attributes of the individual and her environment as a child, we conduct an *internal* comparison. The reference point is the IOp measure calculated in the base scenario on the NLSY79 sample, not those of previous works on IOp in the US. Although the circumstances used in the base scenario are comparable to these contributions, the results are hardly comparable in view of the discussed sampling differences.

In terms of IOp, we find a MLD of the smoothed distribution of 0.16 ($r = 27.1\%$) for the base scenario. In spite of the sample differences the share of inequality explained by circumstances is comparable to Niehues and Peichl's (2014) estimate of 28% and Pistolesi's (2009) estimate of about 20%. Adding more circumstances substantially increases the estimate of IOp to an MLD of 0.23 ($r = 38.8\%$) in the fourth scenario. The estimate further increases to 0.26 ($r = 43.5\%$) in the fifth scenario, which includes a host of circumstances that do not correspond to either of the above categories, such as various test scores on psychological dispositions and further information on educational background.

To test the statistical significance of our results, we rely on a bootstrapping procedure with 100 draws. For each scenario, the p-values in the last column of Table 2 refer to the comparison of absolute IOp to its respective predecessor scenario. Based on this comparison, only the inclusion of health-related circumstances yields a significant increase of IOp in average income (Table 2) at conventional 5% levels. The difference between the base scenario and the full circumstance set, however, is statistically significant at the 1% level.

Note that the results presented in scenarios 1–4 are not strictly comparable to the results of BCS70. In particular, we exclude all variables on drinking and smoking behaviors of the mother, which are sparsely populated in the NLSY79. Excluding these variables causes the sample size to double. Furthermore, we use circumstances that lie within the realm of the defined categories, but are not available in BCS70. In Table 2 we therefore present results from circumstance sets that closely match the circumstance sets used in BCS70.[10] In this restricted sample, outcome inequality reaches a level of 0.73, while IOp increases from 0.30 ($r = 41.0\%$) in the base scenario to 0.39 ($r = 53.5\%$) in the most extensive circumstance set.[11]

Some people may disagree with our approach of treating the entire child biography up to age 16 as a circumstance. In order to test the robustness of our results to the exact specification of the age threshold, we partition our circumstance set by the age-cut offs “At birth”, “Age 12”, and “Age 16”. The results are presented in Fig. 2.

First, we separate all circumstances determined prior to birth. These include the child’s sex, its country of birth, ethnic identity, cohort and its mother’s age at birth. Furthermore, information regarding the mother’s educational achievement, her occupation as well as her scores on the AFQT, the Pearlin Mastery Scale, Rotter’s Locus of Control Scale and the Rosenberg Self-Esteem Scale are only recored prior to child birth and thus are also represented in the leftmost bar of Fig. 2. Thus, if we abstracted from childhood circumstances we would obtain an IOp estimate of 0.109 ($r = 19.5\%$) for incomes averaged over the period 2010–2012.

All remaining circumstances discussed in Sect. 3.2 are recorded at age 12 and an older age. The income of the child’s household, for example, is recorded at ages 12 and 16. For the construction of the results presented in Fig. 1, however, we have only made use of household income at age 16. To investigate the sensitivity of our results to the exact age cut-off we now replicate our previous analysis by replacing all variables measured at age 16 with their age 12 analogues. The results are shown in the central bar of Fig. 2. Including childhood circumstances measured at age 12 in addition to circumstances determined prior to birth increases the MLD of the smoothed distribution from 0.109 ($r = 19.5\%$) to 0.250 ($r = 44.6\%$). This difference is significant at the 1% level. Thus, our conclusion that accounting for childhood circumstances leads to a substantial upwards correction of the lower bound IOp measure is robust to

[10] Specifically, we use the following set of NLSY79 (BCS70) circumstances: sex, country of birth, ethnic identity, cohort, academic achievement mother, occupation code mother, rural/urban, height, family income, play w/ parents, perceived quantity of time w/ mother, parents split, smoking habits mother, drinking habits mother, school absence due to health, age mother at birth, standardized math and reading assessment.

[11] See Appendix Table 6 in which we decompose the differences between the NLSY-specific and the comparable results into changes due to variable exclusions/inclusions and the respective sample size adjustments.

Table 2 Results overview (NLSY79)

	Period	Scenario	N	IO	IOp	r	Difference	p-value
NLSY-specific	2008	First	811	0.653	0.183	0.281		
NLSY-specific	2008	Second	811	0.653	0.188	0.287	0.004	0.541
NLSY-specific	2008	Third	811	0.653	0.203	0.311	0.016	0.165
NLSY-specific	2008	Fourth	811	0.653	0.212	0.325	0.009	0.303
NLSY-specific	2008	Fifth	811	0.653	0.277	0.424	0.065	0.015
NLSY-specific	2010	First	1091	0.856	0.218	0.255		
NLSY-specific	2010	Second	1091	0.856	0.222	0.259	0.004	0.643
NLSY-specific	2010	Third	1091	0.856	0.258	0.301	0.036	0.089
NLSY-specific	2010	Fourth	1091	0.856	0.278	0.325	0.020	0.089
NLSY-specific	2010	Fifth	1091	0.856	0.328	0.383	0.050	0.032
NLSY-specific	2012	First	1077	0.860	0.229	0.266		
NLSY-specific	2012	Second	1077	0.860	0.248	0.289	0.019	0.080
NLSY-specific	2012	Third	1077	0.860	0.289	0.337	0.041	0.026
NLSY-specific	2012	Fourth	1077	0.860	0.333	0.388	0.044	0.022
NLSY-specific	2012	Fifth	1077	0.860	0.368	0.428	0.035	0.094
NLSY-specific	2010–2012	First	707	0.597	0.162	0.271		
NLSY-specific	2010–2012	Second	707	0.597	0.168	0.281	0.006	0.407
NLSY-specific	2010–2012	Third	707	0.597	0.209	0.350	0.041	0.041
NLSY-specific	2010–2012	Fourth	707	0.597	0.232	0.388	0.023	0.059
NLSY-specific	2010–2012	Fifth	707	0.597	0.259	0.435	0.028	0.221
Comparison BCS70	2008	First	358	0.704	0.166	0.236		
Comparison BCS70	2008	Second	358	0.704	0.213	0.303	0.047	0.132
Comparison BCS70	2008	Third	358	0.704	0.280	0.398	0.067	0.032
Comparison BCS70	2008	Fourth	358	0.704	0.300	0.427	0.020	0.226
Comparison BCS70	2010	First	531	0.988	0.276	0.279		
Comparison BCS70	2010	Second	531	0.988	0.278	0.281	0.002	0.891
Comparison BCS70	2010	Third	531	0.988	0.394	0.399	0.116	0.002
Comparison BCS70	2010	Fourth	531	0.988	0.414	0.419	0.020	0.247
Comparison BCS70	2012	First	498	0.919	0.319	0.347		
Comparison BCS70	2012	Second	498	0.919	0.355	0.386	0.036	0.166
Comparison BCS70	2012	Third	498	0.919	0.373	0.406	0.018	0.439
Comparison BCS70	2012	Fourth	498	0.919	0.427	0.465	0.054	0.038
Comparison BCS70	2010–2012	First	367	0.725	0.297	0.410		
Comparison BCS70	2010–2012	Second	367	0.725	0.307	0.424	0.010	0.400
Comparison BCS70	2010–2012	Third	367	0.725	0.358	0.494	0.051	0.059
Comparison BCS70	2010–2012	Fourth	367	0.725	0.388	0.535	0.030	0.186
Responsibility cut-off	2008	Birth	439	0.648	0.137	0.212		
Responsibility cut-off	2008	Birth + age 12	439	0.648	0.287	0.442	0.149	0.017
Responsibility cut-off	2008	All	439	0.648	0.401	0.619	0.115	0.032
Responsibility cut-off	2010	Birth	692	0.837	0.178	0.213		

Table 2 continued

	Period	Scenario	N	IO	IOp	r	Difference	p-value
Responsibility cut-off	2010	Birth + age 12	692	0.837	0.297	0.355	0.119	0.006
Responsibility cut-off	2010	All	692	0.837	0.396	0.473	0.099	0.013
Responsibility cut-off	2012	Birth	657	0.852	0.192	0.226		
Responsibility cut-off	2012	Birth + age 12	657	0.852	0.333	0.391	0.141	0.001
Responsibility cut-off	2012	All	657	0.852	0.444	0.521	0.111	0.001
Responsibility cut-off	2010–2012	Birth	428	0.559	0.109	0.195		
Responsibility cut-off	2010–2012	Birth + age 12	428	0.559	0.250	0.446	0.141	0.009
Responsibility cut-off	2010–2012	All	428	0.559	0.329	0.588	0.079	0.008
Ability	2008	w/o Ability	811	0.653	0.270	0.413		
Ability	2008	w/ Ability	811	0.653	0.277	0.424	0.008	0.368
Ability	2010	w/o Ability	1091	0.856	0.317	0.371		
Ability	2010	w/ Ability	1091	0.856	0.328	0.383	0.011	0.252
Ability	2012	w/o Ability	1077	0.860	0.334	0.389		
Ability	2012	w/ Ability	1077	0.860	0.368	0.428	0.034	0.066
Ability	2010–2012	w/o Ability	707	0.597	0.243	0.407		
Ability	2010–2012	w/ Ability	707	0.597	0.259	0.435	0.016	0.144

This table lists inequality of outcomes (IO), absolute inequality of opportunity (IOp) and relative inequality of opportunity (r) for each of the considered scenarios. p-values are calculated for the difference in absolute inequality of opportunity (IOp) between each scenario and its predecessor scenario. The respective standard errors are based on a bootstrapping procedure with 100 draws. Note that the t-statistics and the associated p-values are derived from a paired t-test. The results from Figs. 1, 2 and 3 are shown in panels 1, 3, and 4, respectively

specifying the age of consent threshold at an earlier age. Furthermore, it is noteworthy that the magnitude of the result is strinkingly similar to our preferred estimate of 0.26 ($r = 43.5\%$), which is constructed by using the exact same set of circumstances but measured at the later age cut-off.

In a last step, we *additionally* include the same set of circumstances but now measured at age 16. The results as represented in the rightmost bar of Fig. 2 indicate another substantial upward correction of IOp to 0.33 ($r = 58.8\%$). Note that the underlying normative premise is equivalent to our baseline estimate of 0.26 ($r = 43.5\%$), as we implicitly treat the entire child biography until the age of 16 as a circumstance. The fact that we observe a strong upward correction as compared to our baseline estimate is consistent with the lower-bound nature of our approach to opportunity measurement: The consideration of childhood circumstances measured at both age thresholds increases the mere number of circumstances under consideration.[12] Therefore our baseline measure is a conservative measure of IOp when treating 16 as the age of

[12] Comparing the central bar of Fig. 2 with our baseline estimate, we hold the number of circumstances constant while varying the permissible age of consent. To the contrary, when comparing the rightmost bar of Fig. 2 with our baseline estimate, we hold the age of consent constant, while varying the number of circumstances.

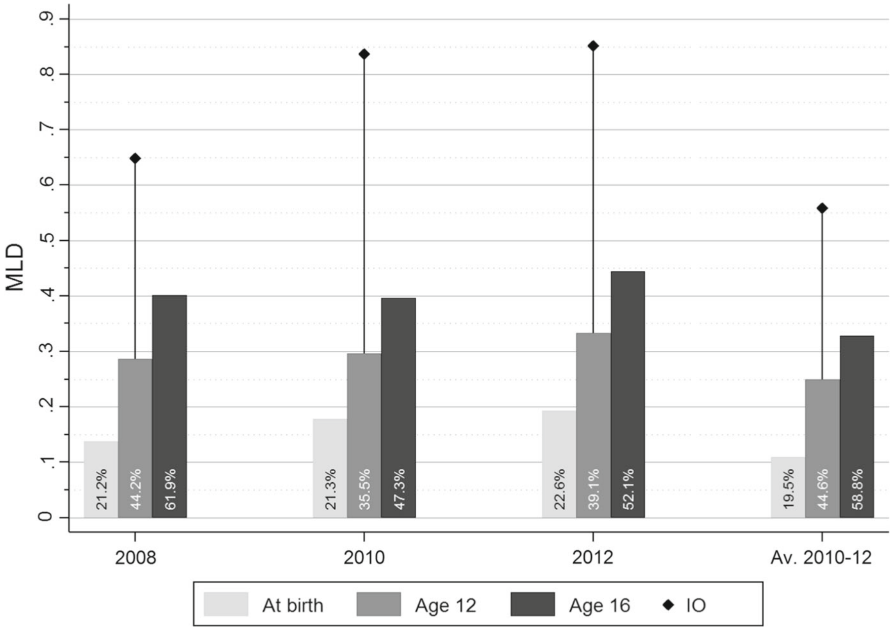

Fig. 2 IOp with Different Ages of Consent (NLSY79). The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower bound absolute measure of IOp. The number in each *bar* is the measure of r, the fraction of inequality due to circumstances. In the *first bar* only circumstances measured at birth are included. The *second bar* includes circumstances measured at birth and additionally those measured at age 12. The *third bar* additionally includes all circumstances measured at age 16. Note that the sample size differs between Figs. 1 and 2 (see Table 2 and footnote 12 which explains why the numbers differ)

consent since we do not account for any information from circumstance variables measured at age 12.[13]

Some may disagree with our approach because we characterize ability measures as a circumstance. In so far as ability is due to differences in genetic endowment, some may be reluctant to accept redistribution on such grounds if they believe a person has a right to receive a higher income because of her genetic endowment (see for instance the Rawlsian account of justice (1971) for a contrasting view). The position we take stands in contrast to the idea of *meritocracy,* the view that a person deserves to benefit from her skills, regardless of their genesis. In contrast, we believe persons rightly benefit only from the portion of their skills attributable to effort. Of course, markets are meritocratic, and so cannot be expected to implement perfectly in market economies the kind of opportunity egalitarianism we envision here.

13 The reason is that the data provides only few respondents whose circumstance information is available both at age 12 and at age 16. When accounting for circumstances measured at both ages the sample is almost cut in half (Table 2). Therefore, we only consider circumstances measured at age 16 in our baseline estimations.

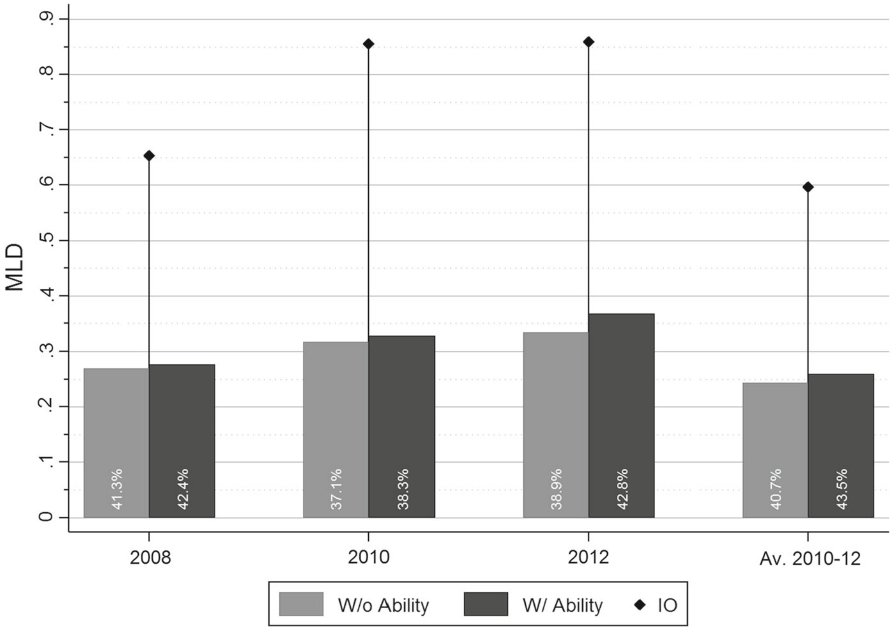

Fig. 3 IOp and Ability (NLSY79). The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower bound absolute measure of IOp. The number in each *bar* is the measure of r, the fraction of inequality due to circumstances. All circumstances are measured either at birth or age 16 with no circumstance repeatedly measured at different age thresholds

As ability is at least partially affected by a child's upbringing, a conservative approach to address these normative objections is to compare the IOp measures once with and once without information on ability in the circumstance sets.

Figure 3 shows that our previous results on the magnitude of IOp within the sample of analysis remain largely unaltered. The MLD of the smoothed distribution of incomes averaged over the period between 2010 and 2012 is reduced by roughly 0.02 whereas the relative measure r decreases from 43.5 to 40.7%.

4.2 The British Cohort Study

In contrast to the NLSY79, the BCS70 only observes one cohort of individuals. Therefore, there is no variance in age for this sample. We observe individual gross and net earnings from 2004 to 2012 (age 34–42) in 4-year periods. Using the annual information as well as average earnings over 2008 and 2012, we are able to cover a period of 8 years.

Checchi et al. (2010) find a MLD in net income of 0.204 as well as IOp in levels of 0.041 ($r = 20.5\%$) using EU-SILC data from 2005. Similarly, OECD data from 2010 indicate a MLD of 0.201 for net (disposable) income in the mid-2000s (OECD 2010). Generally, our measures for IO are somewhat higher than these estimates, which may be attributed to the fact that we observe one cohort instead of a representative sample

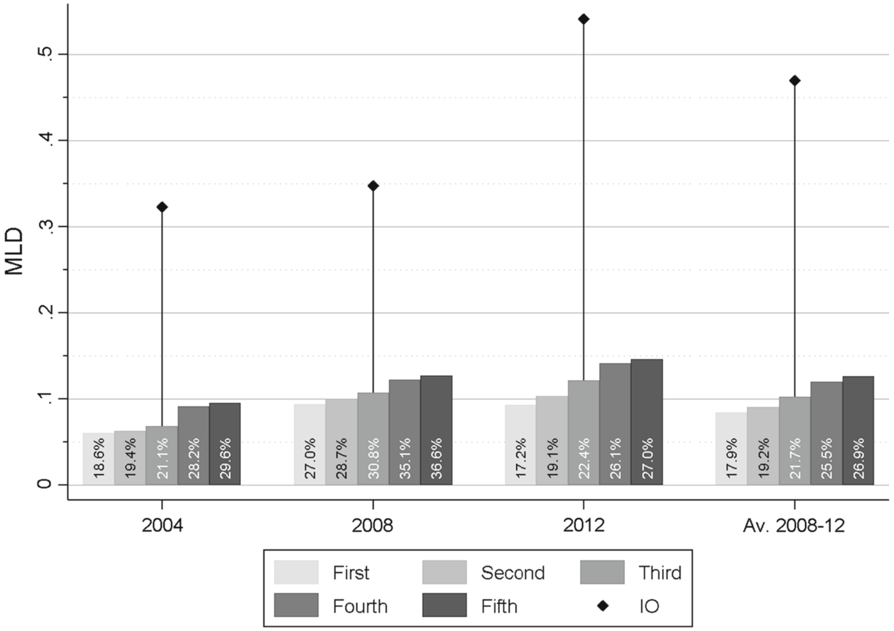

Fig. 4 IOp across Time (BCS70), Gross Income. The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower bound absolute measure of IOp. The number in each *bar* is the measure of r, the fraction of inequality due to circumstances. All circumstances are measured either at birth or age 16 with no circumstance repeatedly measured at different age thresholds

for the entire population. IOp for average income over 2008 to 2012 takes values from 0.126 ($r = 26.9\%$) in gross and 0.086 ($r = 29.2\%$) in net income. Hence, our measure is substantially higher compared to previous studies.

Figure 4 displays IOp estimates in gross income for 2004–2012 and average income over 2008 and 2012. As the BCS70 carries information on gross and net income, we use these numbers to compare IOp before and after taxes and transfers. The figures for net income are available in Appendix Figs. 7, 8 and 9.[14] As we only observe one cohort, income inequality tends to increase over individual's lifetime. To limit the influence of transitory income components, we will again focus our discussion on average income. By expanding the set of explanatory variables using the circumstance sets defined in Table 1, we find IOp levels up to 0.126 ($r = 26.9\%$) in gross income. Interestingly, the highest IOp in levels is not always in the same year as the highest IOp as a share. The last two columns of Table 3 again show tests on the significance of the difference between the absolute IOp measure and the respective predecessor scenario. Only the inclusion of ability measures (Scenario 4) yields a statistically significant upwards correction of absolute IOp. However, as is the case in the US, the total difference

[14] Generally, we find lower income inequality (in terms of the MLD) in net income when compared to gross income. At the same time, the level of IOp is higher in gross income. However, IOp as a percentage of total inequality is higher in net income. One interpretation of these findings is that the tax and transfer system in the UK is equalizing in terms of income inequalities rather than opportunities.

between the base scenario and the most extensive circumstance scenario is significant at the 1% level.

As these results are not comparable due to the specific circumstance sets, we perform the analysis with restricted circumstance sets that are similar for both datasets. While such a comparison might not yield results representative for the populations of the US and the UK, it is still interesting to examine to what extent the findings are consistent in both samples. In Table 3 we present results for gross income using circumstance sets comparable to the NLSY79.[15] We find IOp in gross average income to be 0.116 ($r = 24.6\%$) in the UK and 0.370 (51.1%) in the US. Generally IOp is lower in the UK sample compared to the US, both in levels and in terms of the r-ratio.

As previously mentioned, the correct responsibility cut-off is subject to debate regarding the question of whether children should be considered responsible for their achievements. Therefore, as in the analysis of the NLSY79, Fig. 5 shows three scenarios for gross income: “At birth”, “Age 10” and “Age 16”. Focusing on average income, we find that circumstances available at birth already account for an IOp measure of 0.105 ($r = 21.8\%$) in gross income. The maximum value of IOp using all available information at age 16 yields a lower bound MLD of the smoothed distribution of 0.133 ($r = 27.9\%$) in average gross earnings.

Interestingly, point estimates of IOp with circumstances measured at 16 are only marginally higher compared to those at age 10, with the difference being statistically insignificant at the 10% level. Thus, as in the US, accounting for childhood circumstances leads to a substantial upwards correction of the lower bound IOp measure irrespective of specifying the age of responsibility at 10 or 16.

As already discussed, the standpoint that all accomplishments and attributes of the child up to an age of consent be considered circumstances is debatable. Therefore, we exclude all information related to test scores and schooling as part of a sensitivity analysis. We find that excluding these variables has a statistically significant effect on the absolute IOp measure. Yet with a downward correction from $r = 26.9\%$ to $r = 23.6\%$, the magnitude of this effect is rather small. Therefore, our general conclusions remain intact. As in the case of the US, we can conclude that ability is not the major determining factor for our results.

5 Conclusion

We have argued that important circumstances that play a role in income determination have been ignored in the empirical literature on IOp. From our perspective, all behaviors and accomplishments of children should be considered the consequence of circumstances: that is, an individual should not be held responsible for her choices before an age of consent is reached, in so far as these choices affect her future income. In wealthy societies, it is appropriate to determine the age of consent as occurring no earlier than adolescence. Ideally, if we had a complete biography of the child at the age of, say, sixteen, we would consider that to comprise the child’s circumstances.

[15] See footnote 10 for the precise set of circumstances.

Table 3 Results overview—gross income (BCS70)

	Period	Scenario	N	IO	IOp	r	Difference	p-value
BCS70-specific	2004	First	478	0.323	0.060	0.186		
BCS70-specific	2004	Second	478	0.323	0.063	0.194	0.003	0.446
BCS70-specific	2004	Third	478	0.323	0.068	0.211	0.006	0.168
BCS70-specific	2004	Fourth	478	0.323	0.091	0.282	0.023	0.000
BCS70-specific	2004	Fifth	478	0.323	0.096	0.296	0.005	0.137
BCS70-specific	2008	First	389	0.348	0.094	0.270		
BCS70-specific	2008	Second	389	0.348	0.100	0.287	0.006	0.138
BCS70-specific	2008	Third	389	0.348	0.107	0.308	0.007	0.499
BCS70-specific	2008	Fourth	389	0.348	0.122	0.351	0.015	0.021
BCS70-specific	2008	Fifth	389	0.348	0.127	0.366	0.005	0.169
BCS70-specific	2012	First	524	0.542	0.093	0.172		
BCS70-specific	2012	Second	524	0.542	0.103	0.191	0.010	0.220
BCS70-specific	2012	Third	524	0.542	0.121	0.224	0.018	0.058
BCS70-specific	2012	Fourth	524	0.542	0.141	0.261	0.020	0.019
BCS70-specific	2012	Fifth	524	0.542	0.146	0.270	0.005	0.349
BCS70-specific	2008–2012	First	524	0.470	0.084	0.179		
BCS70-specific	2008–2012	Second	524	0.470	0.090	0.192	0.006	0.171
BCS70-specific	2008–2012	Third	524	0.470	0.102	0.217	0.012	0.101
BCS70-specific	2008–2012	Fourth	524	0.470	0.120	0.255	0.018	0.012
BCS70-specific	2008–2012	Fifth	524	0.470	0.126	0.269	0.006	0.207
Comparison to NLSY	2004	First	478	0.323	0.060	0.186		
Comparison to NLSY	2004	Second	478	0.323	0.062	0.191	0.002	0.586
Comparison to NLSY	2004	Third	478	0.323	0.065	0.200	0.003	0.429
Comparison to NLSY	2004	Fourth	478	0.323	0.087	0.270	0.022	0.001
Comparison to NLSY	2008	First	389	0.348	0.094	0.270		
Comparison to NLSY	2008	Second	389	0.348	0.099	0.286	0.006	0.134
Comparison to NLSY	2008	Third	389	0.348	0.104	0.299	0.005	0.560
Comparison to NLSY	2008	Fourth	389	0.348	0.120	0.345	0.016	0.018
Comparison to NLSY	2012	First	524	0.542	0.093	0.172		
Comparison to NLSY	2012	Second	524	0.542	0.103	0.191	0.010	0.236
Comparison to NLSY	2012	Third	524	0.542	0.112	0.206	0.008	0.204
Comparison to NLSY	2012	Fourth	524	0.542	0.134	0.247	0.022	0.012
Comparison to NLSY	2008–2012	First	524	0.470	0.084	0.179		
Comparison to NLSY	2008–2012	Second	524	0.470	0.090	0.191	0.006	0.275
Comparison to NLSY	2008–2012	Third	524	0.470	0.096	0.205	0.007	0.251
Comparison to NLSY	2008–2012	Fourth	524	0.470	0.116	0.246	0.019	0.009
Responsibility cut-off	2004	At birth	398	0.341	0.068	0.201		
Responsibility cut-off	2004	Age 10	398	0.341	0.103	0.302	0.034	0.002
Responsibility cut-off	2004	Age 16	398	0.341	0.107	0.315	0.005	0.252
Responsibility cut-off	2008	At birth	322	0.319	0.111	0.346		

Table 3 continued

	Period	Scenario	N	IO	IOp	r	Difference	p-value
Responsibility cut-off	2008	Age 10	322	0.319	0.136	0.426	0.025	0.003
Responsibility cut-off	2008	Ages 16	322	0.319	0.142	0.446	0.007	0.130
Responsibility cut-off	2012	At birth	437	0.564	0.104	0.184		
Responsibility cut-off	2012	Age 10	437	0.564	0.142	0.252	0.038	0.011
Responsibility cut-off	2012	Age 16	437	0.564	0.145	0.258	0.003	0.606
Responsibility cut-off	2008–2012	At birth	437	0.480	0.105	0.218		
Responsibility cut-off	2008–2012	Age 10	437	0.480	0.132	0.275	0.027	0.007
Responsibility cut-off	2008–2012	Age 16	437	0.480	0.134	0.279	0.002	0.693
Ability	2004	w/o Ability	478	0.323	0.076	0.235		
Ability	2004	w/ Ability	478	0.323	0.096	0.296	0.020	0.001
Ability	2008	w/o Ability	389	0.348	0.115	0.330		
Ability	2008	w/ Ability	389	0.348	0.127	0.366	0.013	0.024
Ability	2012	w/o Ability	524	0.542	0.127	0.235		
Ability	2012	w/ Ability	524	0.542	0.146	0.270	0.019	0.036
Ability	2008-12	w/o Ability	524	0.470	0.111	0.236		
Ability	2008-12	w/ Ability	524	0.470	0.126	0.269	0.015	0.038

This table lists inequality of outcomes (IO), absolute inequality of opportunity (IOp) and relative inequality of opportunity (r) for each of the considered scenarios. p-values are calculated for the difference in absolute inequality of opportunity (IOp) between each scenario and its predecessor scenario. The respective standard errors are based on a bootstrapping procedure with 100 draws. Note that the t-statistics and the associated p-values are derived from a paired t-test. The results from Figs. 4, 5 and 6 are shown in panels 1, 3, and 4, respectively

The credulous reader might well ask, “Well, if you take the complete biography of the child at the age of consent as the same as her circumstances, where does effort come into play? After all, the choices she makes as an adult will be strongly influenced by her ‘biography’ at age sixteen.” We agree, and that is why we believe circumstances account for a far larger fraction of outcome inequality than studies to date have calculated. Nevertheless, we would resist any suggestion to decrease the age of consent to something like 4 or 6 years of age. Perhaps thirteen, the age of majority according to the Jewish faith, would be acceptable—although we must also bear in mind that thirteen was designated the beginning of adulthood at a time when life expectancies were barely one third of what they are now, and the resources society had to allocate to children were far less abundant.

Using the NLSY79 and the BCS70, we construct fine-grained sets of circumstances that include both the social environment of children and their childhood accomplishments in order to calculate the fraction of income inequality due to circumstances in the US as well as in the UK. Our analysis shows a significant increase of IOp measures when we expand the set of circumstances to include the attributes of the individual and her environment as a child. We find IOp to rise from 0.162 (r = 27.1%) to 0.259 (r = 43.5%) in the US and from 0.084 (r = 17.9%) to 0.126 (26.9%) in the UK. The magnitude of our results remains intact when using the same set of circumstances

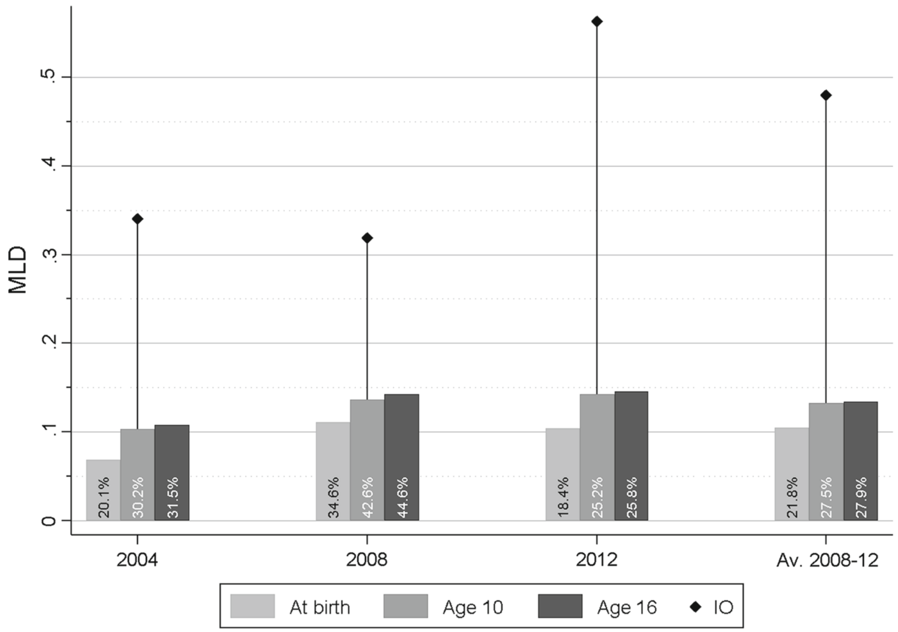

Fig. 5 IOp with Different Ages of Consent (BCS70), Gross income. The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower bound absolute measure of IOp. The number in each bar is the measure of r, the fraction of inequality due to circumstances. In the *first bar* only circumstances measured at birth are included. The *second bar* includes circumstances measured at birth and additionally those measured at age 10. The *third bar* additionally includes all circumstances measured at age 16. Note that the sample size differs between Figs. 4 and 5 (see Table 3 and footnote 12 which explains why the numbers differ)

measured at the earlier age cut-offs of of 12 and 10 in the US and the UK, respectively. Thus, we demonstrate that accounting for childhood circumstances provides a significant upward correction of IOp measures even when setting the responsibility cut-off to an earlier age. Furthermore, we show that our results are robust to excluding ability measures from the set of circumstances.

Our findings invite further study through revisiting the sets of circumstances used in previous studies, given that the results we obtained indicate substantially higher IOp when taking additional childhood information into account. In fact, extending circumstance sets to include childhood achievements up until a particular age of consent addresses some of the concerns regarding the policy relevance of the concept by providing substantial upward corrections of lower-bound measures of IOp. Obviously, in many national contexts data limitations impose considerable restrictions on researchers' ability to conduct analyses as detailed as ours for the US and the UK. To address this problem, one avenue for future research could be to combine different data sets for calculations of IOp. For instance, one might use a first data set with detailed information on circumstances to predict childhood accomplishments of different types. In a second step, one could then use these intermediate types as circumstances in a second data set to calculate a measure of IOp. Such a procedure, which was already

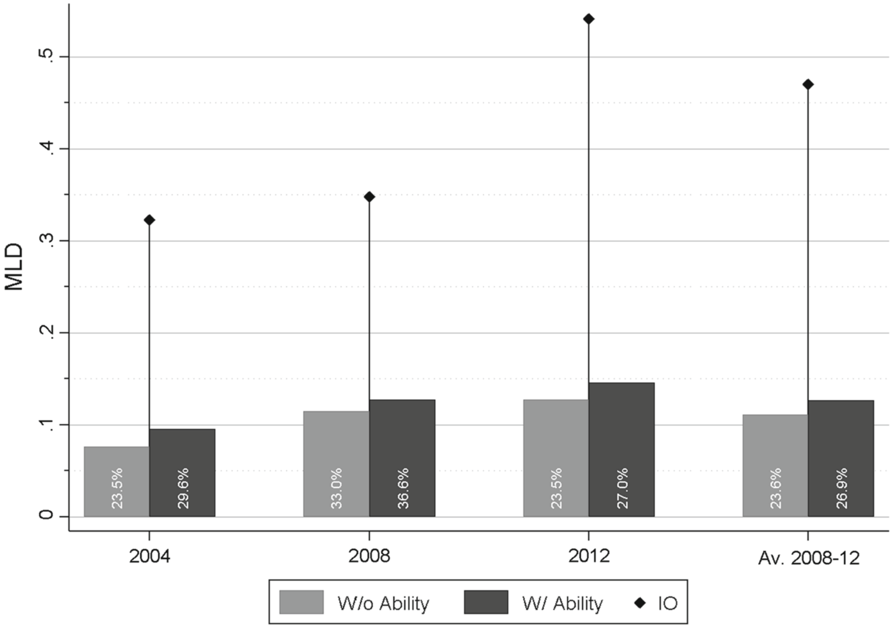

Fig. 6 IOp and Ability (BCS70), Gross Income. The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower-bound absolute measure of IOp. The number in each *bar* is the measure of r, the fraction of inequality due to circumstances. All circumstances are measured either at birth or age 16 with no circumstance repeatedly measured at different age thresholds

implemented in the context of intergenerational mobility (Björklund and Jäntti 1997), would be one promising route to overcome data limitations and enhance the data basis for analyses on IOp.

But clever statistical techniques using existing data sets will have a limited value. To have an accurate estimate of inequality of opportunity in terms of income, we must advocate the creation of panel studies that incorporate both detailed information on childhood achievements and attributes and later income. As we are currently lacking any such data set for developing countries, we can only put poorly estimated lower bounds on the extent of inequality of opportunity for most countries in the world. The results from the US and UK that we have presented suggest that, were we able to calculate the extent to which circumstances account for income inequality in developing countries, the number would be well over 50% in most countries. Thus, contrary to Kanbur and Wagstaff (2016), who advise down-playing IOp analysis, we recommend strengthening it. The political value of showing the true extent of inequality of opportunity in terms of income in developing countries could have immense implications for government policy.

6 Appendices

See Tables 4, 5, 6, 7.

Table 4 Variables NLSY79

Scenario	Var. name	N	Mean	SD	Min	Max	N (2010–2012)	Mean (2010–2012)	Mean (2010–2012) weighted	Study	Ref. age	Question name
Outcome	Gross Inc. (2008)	6201	13,843.95	17,608.77	9	238,232	983	16,474	17,292	NLSY79 Child/YA	–	
Outcome	Gross Inc. (2010)	6014	14,061.45	18,968.46	7	215,000	1030	19,800	21,857	NLSY79 Child/YA	–	
Outcome	Gross Inc. (2012)	5713	16,934.47	22,728.01	9	321,483	1030	25,166	28,610	NLSY79 Child/YA	–	
Outcome	Avg. Gross Inc. (2010–2012)	4539	14,047.2	17624.83	54	232036	1030	22483	25233	NLSY79 Child/YA	–	
1	Cohort	7999	1985.687	5.611	1970	1997	1030	1984	1985	NLSY79 Child/YA	0	CYRB
1	Male	7999	0.513	0.5	0	1	1030	0	0	NLSY79 Child/YA	0	CSEX
1	Hisp.	7999	0.333	0.471	0	1	1030	0	0	NLSY79 Child/YA	0	CRACE
1	Non-black/Non-Hisp.	7999	0.447	0.497	0	1	1030	0	1	NLSY79 Child/YA	0	CRACE
1	Black	7999	0.221	0.415	0	1	1030	0	0	NLSY79 Child/YA	0	CRACE
1	Born in US	7999	0.923	0.266	0	1	1030	1	1	NLSY79	0	FAM-2A
1	Highschool Drop-out (0)	7874	0.254	0.435	0	1	1030	0	0	NLSY79	0	Q3-10B
1	Secondary (0)	7874	0.544	0.498	0	1	1030	1	1	NLSY79	0	Q3-10B
1	Intermediate (0)	7874	0.066	0.247	0	1	1030	0	0	NLSY79	0	Q3-10B

Table 4 continued

Scenario	Var. name	N	Mean	SD	Min	Max	N (2010–2012)	Mean (2010–2012)	Mean (2010–2012) weighted	Study	Ref. age	Question name
1	College (0)	7874	0.137	0.344	0	1	1030	0	0	NLSY79	0	Q3-10B
1	Mom: Blue-collar (0)	6626	0.417	0.493	0	1	1030	1	1	NLSY79	0	Q6-56; CPSOCC70
1	Mom: Farmer (0)	6626	0.001	0.025	0	1	1030	0	0	NLSY79	0	Q6-56; CPSOCC70
1	Mom: White-collar (0)	6626	0.038	0.192	0	1	1030	0	0	NLSY79	0	Q6-56; CPSOCC70
1	Mom: Professional (0)	6626	0.064	0.245	0	1	1030	0	0	NLSY79	0	Q6-56; CPSOCC70
1	Mom: Self-Employed (0)	6626	0.035	0.185	0	1	1030	0	0	NLSY79	0	Q6-56; CPSOCC70
1	Mom: Govt. Sctr. (0)	6626	0.085	0.28	0	1	1030	0	0	NLSY79	0	Q6-56; CPSOCC70
1	Height in inches (14)	6272	63.523	3.664	38	83	1030	64	64	NLSY79 Child/YA	14	–
1	Height in inches (12)	6144	59.012	3.849	23	79	993	59	59	NLSY79 Child/YA	12	–
1	SMSA: Not Center (16)	7460	0.49	0.5	0	1	1030	1	1	NLSY79	16	SMSARES
1	SMSA: Center (16)	7460	0.09	0.287	0	1	1030	0	0	NLSY79	16	SMSARES
1	SMSA: Ambiguous (16)	7460	0.265	0.441	0	1	1030	0	0	NLSY79	16	SMSARES
1	SMSA: Not Center (12)	7399	0.436	0.496	0	1	1017	0	0	NLSY79	12	SMSARES

Table 4 continued

Scenario	Var. name	N	Mean	SD	Min	Max	N (2010–2012)	Mean (2010–2012)	Mean (2010–2012) weighted	Study	Ref. age	Question name
1	SMSA: Center (12)	7399	0.157	0.364	0	1	1017	0	0	NLSY79	12	SMSARES
1	SMSA: Ambiguous (12)	7399	0.234	0.423	0	1	1017	0	0	NLSY79	12	SMSARES
1	SMSA: Ambiguous (12)	7399	0.234	0.423	0	1	1017	0	0	NLSY79	10	SMSARES
1	Net Fam. Inc. (16)	6288	0	1	−1	15	1030	0	0	NLSY79	16	TNFI_TRUNC
1	Net Fam. Inc. (12)	6358	0	1	−1	14	875	0	0	NLSY79	12	TNFI_TRUNC
2	Prnts. tgthr (14)	6825	0.502	0.5	0	1	1030	0	1	NLSY79 Child/YA	14	DADHM (year)
2	Prnts. tgthr (12)	7005	0.518	0.5	0	1	1000	1	1	NLSY79 Child/YA	12	DADHM (year)
2	Activ. (14): Games/play	6071	0.419	0.493	0	1	1030	0	0	NLSY79 Child/YA	14	–
2	Activ. (12): Games/play	6055	0.511	0.5	0	1	911	1	1	NLSY79 Child/YA	12	–
2	Mom (14): Enough time	5554	0.166	0.372	0	1	1030	0	0	NLSY79 Child/YA	14	–
2	Mom (14): More time	5554	0.77	0.421	0	1	1030	1	1	NLSY79 Child/YA	14	–
2	Mom (14): Less time	5554	0.065	0.246	0	1	1030	0	0	NLSY79 Child/YA	14	–

Table 4 continued

Scenario	Var. name	N	Mean	SD	Min	Max	N (2010–2012)	Mean (2010–2012)	Mean (2010–2012) weighted	Study	Ref. age	Question name
2	Mom (12): Enough time	5127	0.196	0.397	0	1	852	0	0	NLSY79 Child/YA	12	–
2	Mom (12): More time	5127	0.753	0.431	0	1	852	1	1	NLSY79 Child/YA	12	–
2	Mom (12): Less time	5127	0.051	0.221	0	1	852	0	0	NLSY79 Child/YA	12	–
2	Activ. (14): Schlwrk	6071	0.29	0.454	0	1	1030	0	0	NLSY79 Child/YA	14	–
2	Activ. (12): Schlwrk	6058	0.4	0.49	0	1	914	0	0	NLSY79 Child/YA	12	–
3	Health (14): School	6781	0.02	0.143	0	2	1030	0	0	NLSY79 Child/YA	14	–
3	Health (12): School	6925	0.02	0.144	0	2	1000	0	0	NLSY79 Child/YA	12	–
3	Non-smoker Mthr	4566	0.217	0.412	0	1	554	0	0	NLSY79	1-16	DS5; DS6
3	Drinks almost every day	4426	0.034	0.182	0	1	742	0	0	NLSY79	8	Q12-4
3	Drinks 3–4 Times/week	4426	0.036	0.186	0	1	742	0	0	NLSY79	8	Q12-5
3	Drinks 1–2 Times/week	4426	0.365	0.482	0	1	742	0	0	NLSY79	8	Q12-5
3	Drinks 2–3 Times/month	4426	0.262	0.44	0	1	742	0	0	NLSY79	8	Q12-5
3	Drinks once/Mth	4426	0.155	0.362	0	1	742	0	0	NLSY79	8	Q12-5

Table 4 continued

Scenario	Var. name	N	Mean	SD	Min	Max	N (2010–2012)	Mean (2010–2012)	Mean (2010–2012) weighted	Study	Ref. age	Question name
3	Drinks never	4426	0.148	0.355	0	1	742	0	0	NLSY79	8	Q12-5
3	Age of mother at birth	7999	25.257	5.308	12	40	1030	24	24	NLSY79	0	
3	Health (14): School Work	6730	0.035	0.188	0	2	1030	0	0	NLSY79 Child/YA	14	–
3	Health (12): School Work	6891	0.038	0.195	0	2	1000	0	0	NLSY79 Child/YA	12	–
3	Health (14): Play	6683	0.028	0.165	0	1	1030	0	0	NLSY79 Child/YA	14	–
3	Health (12): Play	6868	0.029	0.168	0	1	1000	0	0	NLSY79 Child/YA	12	–
4	PIAT (14): Math	6224	99.957	14.698	0	135	1030	100	103	NLSY79 Child/YA	14	MATHZ (year)
4	PIAT (12): Math	6478	101.008	14.853	65	135	970	102	105	NLSY79 Child/YA	12	MATHZ (year)
4	PIAT (14): Reading	6230	103.03	16.504	0	135	1030	103	105	NLSY79 Child/YA	14	RECOGZ (year)
4	PIAT (12): Reading	6471	102.837	15.592	65	135	971	103	105	NLSY79 Child/YA	12	RECOGZ (year)
4	PIAT (14): Compreh.	6176	96.324	13.6	65	135	1030	97	99	NLSY79 Child/YA	14	COMPZ (year)
4	PIAT (12): Compreh.	6392	98.466	14.063	65	135	959	99	101	NLSY79 Child/YA	12	COMPZ (year)
4	AFQT Score	7656	34.552	27.266	1	99	1030	37	47	NLSY79	0	AFQT-1
5	No Pub. School (14)	6708	0.129	0.335	0	1	1030	0	0	NLSY79 Child/YA	14	–

Table 4 continued

Scenario	Var. name	N	Mean	SD	Min	Max	N (2010–2012)	Mean (2010–2012)	Mean (2010–2012) weighted	Study	Ref. age	Question name
5	No Pub. School (12)	6877	0.162	0.368	0	1	978	0	0	NLSY79 Child/YA	12	-
5	# in HH w/ Educ < 12 (16)	7405	0.471	0.741	0	5	1030	0	0	NLSY79 Child/YA	16	NAHGC0 (year)
5	# in HH w/ Educ < 12 (12)	7488	0.451	0.734	0	7	1022	0	0	NLSY79 Child/YA	12	NAHGC0 (year)
5	# in HH w/ Educ 12-13 (16)	7406	1.031	0.916	0	6	1030	1	1	NLSY79 Child/YA	16	NAHGC1 (year)
5	# in HH w/ Educ 12-13 (12)	7491	0.942	0.848	0	6	1022	1	1	NLSY79 Child/YA	12	NAHGC1 (year)
5	# in HH w/ Educ 13-15 (16)	7406	0.278	0.52	0	4	1030	0	0	NLSY79 Child/YA	16	NAHGC2 (year)
5	# in HH w/ Educ 13-15 (12)	7503	0.247	0.49	0	3	1022	0	0	NLSY79 Child/YA	12	NAHGC2 (year)
5	# in HH w/ Educ > 15 (16)	7404	0.305	0.618	0	4	1030	0	0	NLSY79 Child/YA	16	NAHGC3 (year)
5	# in HH w/ Educ > 15 (12)	7506	0.285	0.601	0	3	1022	0	0	NLSY79 Child/YA	12	NAHGC3 (year)
5	No Convict.	7735	0.98	0.14	0	1	1030	1	1	NLSY79	0	POLICE_3
5	BPI-Score (14)	6540	608.309	279.888	82	1000	1030	626	620	NLSY79 Child/YA	14	BPIP (year)
5	BPI-Score (12)	6611	607.75	280.026	62	1000	940	631	624	NLSY79 Child/YA	12	BPIP (year)

Table 4 continued

Scenario	Var. name	N	Mean	SD	Min	Max	N (2010–2012)	Mean (2010–2012)	Mean (2010–2012) weighted	Study	Ref. age	Question name
5	Pearlin Scale (Mother)	7609	493.839	92.37	51	891	1030	488	485	NLSY79	0	PEARLIN_ ZSCORECW
5	Rotter Scale (Mother)	7913	8.95	2.405	4	16	1030	9	9	NLSY79	0	ROTTER_ SCORE
5	Rosenberg Scale (Mother)	7736	488.829	104.284	14	941	1030	492	491	NLSY79	0	ROSENBERG_ ZSCORECW

Table 5 Variables BCS70

Scenario	Var. name	N	Mean	SD	Min	Max	Ref. age	Question name
Outcome	Net earnings	19,559	20,250.5	21,439.3	1.0	877,200	–	B9USLA/B9NETP/ b7cnetpy/b7cnetpd/ b8cnetpy/b8cnetpd
Outcome	Gross earnings	18,075	29,746.9	50,542.1	1.0	2,732,400	–	B9GROA/B9GROP/ b7cgropy/b7cgropd/ b8cgropy/b8cgropd
Outcome	Net earnings av	18,409	21,684.8	19,519.3	13.0	720,000	–	
Outcome	Gross earnings av	18,023	31,667.2	47,837.3	20.0	2,374,188	–	
	Year	19,559	2008.1	3.4	2004.0	2012	–	
1	Female	18,101	0.509	0.500	0	1	0	a0255
1	Ethnicity Mom: European	15,982	0.003	0.056	0	1	0	e246a
1	Ethnicity Mom: Indian/Pakistani/ Asian	15,982	0.974	0.159	0	1	0	e246a
1	Ethnicity Mom: African/Other	15,982	0.023	0.149	0	1	0	e246a
1	Foreign Origin	17,801	0.075	0.264	0	1	0	a0007a
1	Educ Mom: No education	16,110	0.483	0.500	0	1	0	c1_13–c1_19
1	Educ Mom: Secondary	16,110	0.143	0.350	0	1	0	c1_13–c1_19
1	Educ Mom: Intermediate	16,110	0.338	0.473	0	1	0	c1_13–c1_19
1	Educ Mom: College	16,110	0.035	0.185	0	1	0	c1_13–c1_19

Table 5 continued

Scenario	Var. name	N	Mean	SD	Min	Max	Ref. age	Question name
1	Cccup Mom: Housework	16,620	0.503	0.500	0	1	0	a0017
1	Occup Mom: Farmer	16,620	0	0.013	0	1	0	a0017
1	Occup Mom: White-collar	16,620	0.479	0.500	0	1	0	a0017
1	Occup Mom: Professional	16,620	0.014	0.119	0	1	0	a0017
1	Occup Mom: Self-employed	16,620	0	0.021	0	1	0	a0017
1	Occup Mom: Civil Servant	16,620	0.003	0.056	0	1	0	a0017
1	Rural	8,311	0.347	0.476	0	1	16	c16_3
1	Town	8,311	0.528	0.499	0	1	16	c16_3
1	Urban	8,311	0.124	0.330	0	1	16	c16_3
1	Rural	12,566	0.294	0.456	0	1	10	m304–m307
1	Suburbs	12,566	0.635	0.482	0	1	10	m304–m307
1	Inner Urban	12,566	0.071	0.257	0	1	10	m304-20m307
1	Body Height (16)	16,260	0.050	0.987	−3	4	16	BD9HGHTM
1	Parental Income (10)	15,856	4.117	1.247	1	7	10	c9_1–c9_7
1	Parental Income (16)	10,195	4.951	2.471	1	11	16	oe2
2	Time With Parents: Most Days A Week	8,733	0.152	0.359	0	1	16	gb8_3

Table 5 continued

Scenario	Var. name	N	Mean	SD	Min	Max	Ref. age	Question name
2	Time With Parents: Some Days aAWeek	8,733	0.226	0.418	0	1	16	gb8_3
2	Time With Parents: Once A Week	8,733	0.093	0.291	0	1	16	gb8_3
2	Time With Parents: Occasionally	8,733	0.296	0.456	0	1	16	gb8_3
2	Time With Parents: Little Or Never	8,733	0.233	0.423	0	1	16	gb8_3
2	Active With Parents: Rarely	8,386	0.201	0.401	0	1	16	c5r6/c5r2
2	Active With Parents: Sometimes	8,386	0.132	0.338	0	1	16	c5r6/c5r2
2	Active With Parents: Often	8,386	0.051	0.220	0	1	16	c5r6/c5r2
2	Time With Parents: Rarely	15,639	0.009	0.097	0	1	10	k055
2	Time With Parents: Sometimes	15,639	0.370	0.483	0	1	10	k055
2	Time With Parents: Often	15,639	0.621	0.485	0	1	10	k055

Table 5 continued

Scenario	Var. name	N	Mean	SD	Min	Max	Ref. age	Question name
2	Active With Parents: Rarely	17,158	0.018	0.132	0	1	10	m107/m108
2	Active With Parents: Sometimes	17,158	0.380	0.485	0	1	10	m107/m108
2	Active With Parents: Often	17,158	0.602	0.489	0	1	10	m107/m108
2	At Birth: Married	18,082	0.950	0.219	0	1	0	a0012
2	At Birth: Partnership	18,082	0.037	0.188	0	1	0	a0012
2	At Birth: Divorced/Seperated	18,082	0.013	0.112	0	1	0	a0012
2	At Birth: Widowed	18,082	0.001	0.030	0	1	0	a0012
2	Familiy Situation: No Change	19,559	0.036	0.187	0	1	5	e010/e011/e010a/e011a
2	Familiy Situation: Death Of Parents	19,559	0.954	0.210	0	1	5	e010/e011/e010a/e011a
2	Familiy Situation: Divorce Of Parents	19,559	0.001	0.031	0	1	5	e010/e011/e010a/e011a

Table 5 continued

Scenario	Var. name	N	Mean	SD	Min	Max	Ref. age	Question name
2	Familiy Situation: Missing Mom/Father	19,559	0.009	0.094	0	1	5	e010/e011/e010a/e011a
2	Childcare	16,435	0.877	0.328	0	1	10	e158
5	Singleton	19559	0.087	0.282	0	1	5	e006
3	Smoke Mom: No Smoke	19,128	0.981	0.138	0	1	10	e10_1/e9_3
3	Smoke Mom: Less Than 5 Years	19,128	0	0.022	0	1	10	e10_1/e9_3
3	Smoke Mom: More Than 5 Years	19,128	0.019	0.136	0	1	10	e10_1/e9_3
3	Smoke Mom: No Smoke	19,559	0.797	0.402	0	1	16	og2_11
3	Smoke Mom: Smoker	19,559	0.203	0.402	0	1	16	og2_11
3	Alcohol Mom: Never	15,745	0.440	0.496	0	1	16	pg8_3
3	Alcohol Mom: Once A Month	15,745	0.082	0.275	0	1	16	pg8_3
3	Alcohol Mom: 2 Or 3 Times A Month	15,745	0.113	0.317	0	1	16	pg8_3
3	Alcohol Mom: Once Or Twice A Week	15,745	0.252	0.434	0	1	16	pg8_3

Table 5 continued

Scenario	Var. name	N	Mean	SD	Min	Max	Ref. age	Question name
3	Alcohol Mom: 3 Or 4 Times A Week	15,745	0.075	0.263	0	1	16	pg8_3
3	Alcohol Mom: Every Or Most Days	15,745	0.037	0.189	0	1	16	pg8_3
3	Age Mother At Birth	18,086	25.991	5.314	14	53	0	BD1MAGE
3	No Immunisation Until Age 5	19,237	0.013	0.112	0	1	5	e021
3	Smoke Pregnancy: Never	19,462	0.410	0.492	0	1	0	a0043b
3	Smoke Pregnancy: Stop Pre Preg	19,462	0.116	0.320	0	1	0	a0043b
3	Smoke Pregnancy: Stop During Preg	19,462	0.046	0.210	0	1	0	a0043b
3	Smoke Pregnancy: Smoked	19,462	0.428	0.495	0	1	0	a0043b
3	Birthweight	18,091	0.089	0.903	−4	5	0	a0278
3	Sick At School: 7 Or More	12,285	0.369	0	0	1		pb5_1

Table 5 continued

Scenario	Var. name	N	Mean	SD	Min	Max	Ref. age	Question name
4	Reading Score (16)	15,745	−0.489	0.853	−3	3	16	BD4READ/BD4RDAGE
4	Reading Score (10)	18,194	−0.176	0.947	−3	2	10	BD3READ/BD3RDAGE
4	Math Score (10)	14,766	0.186	0.938	−3	2	10	mathscore
5	Firstborn	19,559	0.340	0.474	0	1	5	e006
5	Educ Father: No Education	15,334	0.344	0.475	0	1	0	c1_1/c1_2/c1_3/c1_6/c1_9
5	Educ Father: Secondary	15,334	0.395	0.489	0	1	0	c1_1/c1_2/c1_3/c1_6/c1_9
5	Educ Father: Intermediate	15,334	0.108	0.311	0	1	0	c1_1/c1_2/c1_3/c1_6/c1_9
5	Educ Father: College	15,334	0.152	0.359	0	1	0	c1_1/c1_2/c1_3/c1_6/c1_9
5	Singleton	19,559	0.087	0.282	0	1	5	e006

Table 6 Adjustment NLSY79-BCS70

	Period	Scenario	N	IO	IOp	r	Difference	p-value
Adjustment NLSY79-BCS70	2008	NLSY/BCS	358	0.704	0.300	0.427		
Adjustment NLSY79-BCS70	2008	Adj. Health, Sample Const.	358	0.704	0.273	0.388	0.027	0.135
Adjustment NLSY79-BCS70	2008	Adj. Health, Adj. Sample	811	0.653	0.208	0.319	0.065	0.382
Adjustment NLSY79-BCS70	2008	NLSY (16)	811	0.653	0.277	0.424	0.069	0.011
Adjustment NLSY79-BCS70	2010	NLSY/BCS	531	0.988	0.414	0.419		
Adjustment NLSY79-BCS70	2010	Adj. Health, Sample Const.	531	0.988	0.371	0.376	0.043	0.071
Adjustment NLSY79-BCS70	2010	Adj. Health, Adj. Sample	1091	0.856	0.268	0.313	0.103	0.227
Adjustment NLSY79-BCS70	2010	NLSY (16)	1091	0.856	0.328	0.383	0.060	0.021
Adjustment NLSY79-BCS70	2012	NLSY/BCS	498	0.919	0.427	0.465		
Adjustment NLSY79-BCS70	2012	Adj. Health, Sample Const.	498	0.919	0.416	0.453	0.011	0.489
Adjustment NLSY79-BCS70	2012	Adj. Health, Adj. Sample	1077	0.860	0.312	0.362	0.105	0.227
Adjustment NLSY79-BCS70	2012	NLSY (16)	1077	0.860	0.368	0.428	0.057	0.030
Adjustment NLSY79-BCS70	2010–2012	NLSY/BCS	367	0.725	0.388	0.535		
Adjustment NLSY79-BCS70	2010–2012	Adj. Health, Sample Const.	367	0.725	0.357	0.492	0.031	0.186
Adjustment NLSY79-BCS70	2010–2012	Adj. Health, Adj. Sample	707	0.597	0.208	0.349	0.148	0.095
Adjustment NLSY79-BCS70	2010–2012	NLSY (16)	707	0.597	0.259	0.435	0.051	0.051

This table lists inequality of outcomes (IO), absolute inequality of opportunity (IOp) and relative inequality of opportunity (r) for each of the considered scenarios. *p*-values are calculated for the difference in absolute inequality of opportunity (IOp) between each scenario and its predecessor scenario. The respective standard errors are based on a bootstrapping procedure with 100 draws. Note that the t-statistics and the associated *p*-values are derived from a paired *t*-test

Table 7 Results overview—net income (BCS70)

	Period	Scenario	N	IO	IOp	r	Difference	*p*-value
BCS70-Specific	2004	First	519	0.202	0.041	0.204		
BCS70-Specific	2004	Second	519	0.202	0.044	0.217	0.003	0.251
BCS70-Specific	2004	Third	519	0.202	0.047	0.231	0.003	0.315
BCS70-Specific	2004	Fourth	519	0.202	0.060	0.298	0.014	0.002
BCS70-Speci fic	2004	Fifth	519	0.202	0.062	0.305	0.001	0.402
BCS70-Specific	2008	First	459	0.306	0.075	0.243		
BCS70-Specific	2008	Second	459	0.306	0.081	0.263	0.006	0.152
BCS70-Specific	2008	Third	459	0.306	0.085	0.278	0.005	0.179
BCS70-Specific	2008	Fourth	459	0.306	0.100	0.326	0.015	0.003
BCS70-Specific	2008	Fifth	459	0.306	0.103	0.335	0.003	0.231
BCS70-Specific	2012	First	525	0.376	0.056	0.148		
BCS70-Specific	2012	Second	525	0.376	0.063	0.167	0.007	0.217
BCS70-Specific	2012	Third	525	0.376	0.072	0.192	0.009	0.133
BCS70-Specific	2012	Fourth	525	0.376	0.094	0.250	0.022	0.028
BCS70-Specific	2012	Fifth	525	0.376	0.095	0.254	0.002	0.647
BCS70-Specific	2008–2012	First	525	0.294	0.053	0.179		
BCS70-Specific	2008–2012	Second	525	0.294	0.058	0.199	0.006	0.193
BCS70-Specific	2008–2012	Third	525	0.294	0.065	0.222	0.007	0.205
BCS70-Specific	2008–2012	Fourth	525	0.294	0.083	0.283	0.018	0.003
BCS70-Specific	2008–2012	Fifth	525	0.294	0.086	0.294	0.003	0.345
Comparison to NLSY	2004	First	519	0.202	0.041	0.204		
Comparison to NLSY	2004	Second	519	0.202	0.043	0.211	0.001	0.431

Table 7 continued

	Period	Scenario	N	IO	IOp	r	Difference	p-value
Comparison to NLSY	2004	Third	519	0.202	0.045	0.220	0.002	0.482
Comparison to NLSY	2004	Fourth	519	0.202	0.058	0.288	0.014	0.000
Comparison to NLSY	2008	First	459	0.306	0.075	0.243		
Comparison to NLSY	2008	Second	459	0.306	0.080	0.262	0.006	0.080
Comparison to NLSY	2008	Third	459	0.306	0.083	0.272	0.003	0.320
Comparison to NLSY	2008	Fourth	459	0.306	0.099	0.322	0.015	0.002
Comparison to NLSY	2012	First	525	0.376	0.056	0.148		
Comparison to NLSY	2012	Second	525	0.376	0.062	0.165	0.006	0.294
Comparison to NLSY	2012	Third	525	0.376	0.066	0.175	0.004	0.388
Comparison to NLSY	2012	Fourth	525	0.376	0.088	0.233	0.022	0.031
Comparison to NLSY	2008–2012	First	525	0.294	0.053	0.179		
Comparison to NLSY	2008–2012	Second	525	0.294	0.057	0.195	0.005	0.209
Comparison to NLSY	2008–2012	Third	525	0.294	0.061	0.208	0.004	0.328

Table 7 continued

	Period	Scenario	N	IO	IOp	r	Difference	*p*-value
Comparison to NLSY	2008–2012	Fourth	525	0.294	0.080	0.272	0.019	0.006
Responsibility cut-off	2004	At Birth	430	0.203	0.047	0.234		
Responsibility cut-off	2004	Age 10	430	0.203	0.069	0.340	0.021	0.000
Responsibility cut-off	2004	Age 16	430	0.203	0.072	0.355	0.003	0.226
Responsibility cut-off	2008	At birth	380	0.334	0.088	0.264		
Responsibility cut-off	2008	Age 10	380	0.334	0.111	0.333	0.023	0.003
Responsibility cut-off	2008	Age 16	380	0.334	0.119	0.357	0.008	0.371
Responsibility cut-off	2012	At birth	438	0.382	0.062	0.163		
Responsibility cut-off	2012	Age 10	438	0.382	0.103	0.269	0.040	0.007
Responsibility cut-off	2012	Age 16	438	0.382	0.120	0.315	0.018	0.028
Responsibility cut-off	2008-12	At birth	438	0.305	0.062	0.202		
Responsibility cut-off	2008-12	Age 10	438	0.305	0.091	0.300	0.030	0.006
Responsibility cut-off	2008-12	Age 16	438	0.305	0.100	0.329	0.009	0.125

Table 7 continued

	Period	Scenario	N	IO	IOp	r	Difference	*p*-value
Ability	2004	w/o Ability	519	0.202	0.049	0.242		
Ability	2004	w/ Ability	519	0.202	0.062	0.305	0.013	0.002
Ability	2008	w/o Ability	459	0.306	0.089	0.291		
Ability	2008	w/ Ability	459	0.306	0.103	0.335	0.014	0.009
Ability	2012	w/o Ability	525	0.376	0.076	0.202		
Ability	2012	w/ Ability	525	0.376	0.095	0.254	0.019	0.028
Ability	2008–2012	w/o Ability	525	0.294	0.070	0.240		
Ability	2008–2012	w/ Ability	525	0.294	0.086	0.294	0.016	0.001

This table lists inequality of outcomes (IO), absolute inequality of opportunity (IOp) and relative inequality of opportunity (r) for each of the considered scenarios. p-values are calculated for the difference in absolute inequality of opportunity (IOp) between each scenario and its predecessor scenario. The respective standard errors are based on a bootstrapping procedure with 100 draws. Note that the t-statistics and the associated p-values are derived from a paired t-test

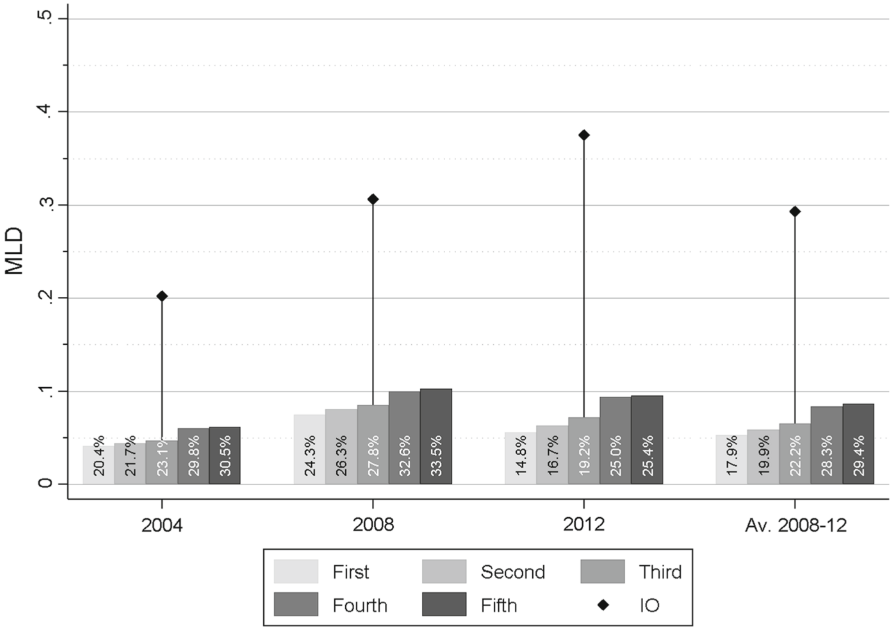

Fig. 7 IOp across Time (BCS70), Net Income. The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower-bound absolute measure of IOp. The number in each *bar* is the measure of *r*, the fraction of inequality due to circumstances. All circumstances are measured either at birth or age 16 with no circumstance repeatedly measured at different age thresholds

6.1 Figures for net income (BCS70)

See Figs. 7, 8 and 9.

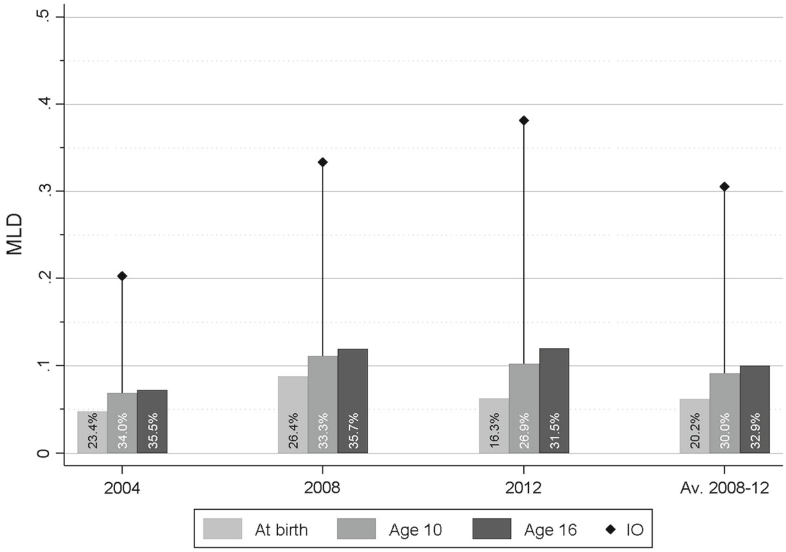

Fig. 8 IOp with Different Ages of Consent (BCS70), Net Income. The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower bound absolute measure of IOp. The number in each *bar* is the measure of r, the fraction of inequality due to circumstances. In the *first bar* only circumstances measured at birth are included. The *second bar* includes circumstances measured at birth and additionally those measured at age 10. The *third bar* additionally includes all circumstances measured at age 16

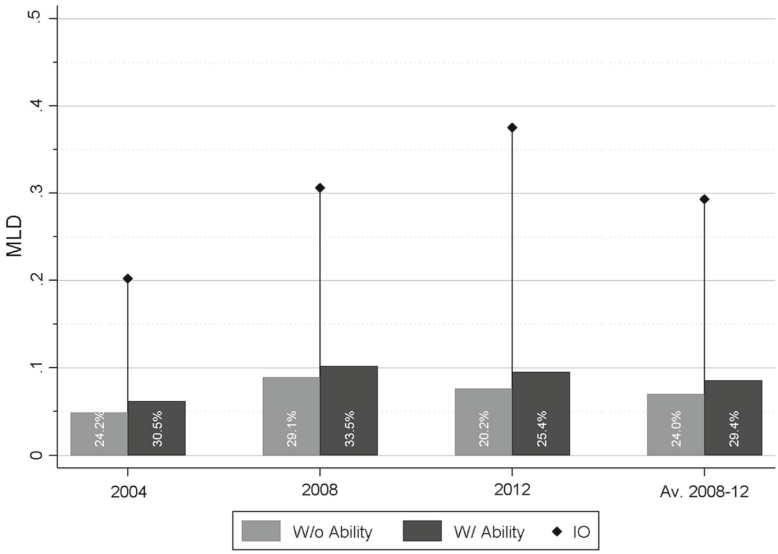

Fig. 9 IOp and Ability (BCS70), Net Income. The spike yields the extent of outcome inequality IO. The *gray bar* yields inequality attributed to circumstances, i.e. the lower-bound absolute measure of IOp. The number in each *bar* is the measure of r, the fraction of inequality due to circumstances. All circumstances are measured either at birth or age 16 with no circumstance repeatedly measured at different age thresholds

References

Alesina A, Giuliano P (2011) Preferences for redistribution. In: Benhabib J, Jackson MO, Bisin A (eds) Handbook of social economics. North-Holland, Amsterdam, pp 93–132

Arneson R (1990) Liberalism, distributive subjectivism, and equal opportunity for welfare. Philos Public Affair 19(2):158–194

Assaad R, Krafft C, Roemer JE (2015) Inequality of opportunity for income and consumption in Egypt. Mimeo, New York

Björklund A, Jäntti M (1997) Intergenerational income mobility in Sweden compared to the United States. Am Econ Rev 87(5):1009–1018

Bourguignon F, Ferreira FHG, Menéndez M (2007) Inequality of opportunity in Brazil. Rev Income Wealth 53(4):585–618

Brunori P, Ferreira FHG, Peragine V (2013) Inequality of opportunity, income inequality, and economic mobility: some international comparisons. In: Paus E (ed) Getting development right: structural transformation, inclusion, and sustainability in the post-crisis Era. Palgrave Macmillan US, New York, pp 8–115

Cappelen AW, Hole AD, Sørensen EØ, Tungodden B (2007) The pluralism of fairness ideals: an experimental approach. Am Econ Rev 97(3):818–827

Cappelen AW, Sørensen EØ, Tungodden B (2010) Responsibility for what? Fairness and individual responsibility. Eur Econ Rev 54(3):429–441

Checchi D, Peragine V, Serlenga L (2010) Fair and unfair income inequalities in Europe. IZA Discussion Paper, No. 5025

Checchi D, Peragine V (2010) Inequality of opportunity in Italy. J Econ Inequal 8(4):429–450

Chetty R, Hendren N, Kline P, Saez E (2014) Where is the land of opportunity? The geography of intergenerational mobility in the United States. Quart J Econ 129:1553–1623

Cowell FA, Jenkins SP (1995) How much inequality can we explain? A methodology and an application to the United States. Econ J 105(429):421–430

Elbers C, Lanjouw P, Mistiaen J, Özler B (2008) Reinterpreting between-group inequality. J Econ Inequal 6(3):231–245

Ferreira FHG, Gignoux J (2011) The measurement of inequality of opportunity: theory and an application to Latin America. Rev Income Wealth 57(4):622–657

Ferreira FHG, Peragine V (2016) Individual responsibility and equality of opportunity. In: Adler MD, Fleurbaey M (eds) Oxford handbook of well-being and public policy. Oxford University Press, Oxford, pp 746–784

Fleurbaey M, Peragine V (2013) Ex ante versus ex post equality of opportunity. Economica 80(317):118–130

Fong C (2001) Social preferences, self-interest, and the demand for redistribution. J Public Econ 82(2):225–246

Hufe P, Peichl A (2015) Lower Bounds and the Linearity Assumption in Parametric Estimations of Inequality of Opportunity. IZA Discussion Paper, No. 9605

Kanbur R, Wagstaff A (2016) How useful is inequality of opportunity as a policy construct? In: Basu K, Stiglitz JE (eds) Inequality and growth: patterns and policy: volume I: concepts and analysis. Palgrave Macmillan, London, pp 131–150

Mostafa T, Wiggins D (2014) Handling attrition and non-response in the 1970 British Cohort Study. CLS Working Paper 2014/2

Niehues J, Peichl A (2014) Upper bounds of inequality of opportunity: theory and evidence for Germany and the US. Soc Choice Welf 43(1):73–99

OECD (2010) OECD Factbook 2010: Economic, environmental and social statistics. OECD Publishing, Paris

Pistolesi N (2009) Inequality of opportunity in the land of opportunities, 1968–2001. J Econ Inequal 7(4):411–433

Ramos X, Van de gaer D (2016) Empirical approaches to inequality of opportunity: principles, measures, and evidence. J Econ Surv 30(5):855–883

Rawls J (1971) A theory of justice. The Belknap Press of Harvard University Press, Cambridge

Roemer JE (1993) A pragmatic theory of responsibility for the egalitarian planner. Philos Public Aff 22(2):146–166

Roemer JE (1998) Equality of opportunity. Harvard University Press, Cambridge

Roemer JE (2017) On the importance of circumstances in explaining income inequality. Rev Econ 68(1):35–56

Roemer JE, Trannoy A (2016) Equality of opportunity: theory and measurement. J Econ Lit 54(4):1288–1332