ifo BEITRÄGE
zur Wirtschaftsforschung

93
2021

Essays on Equality of Opportunity

Paul Hufe

ifo INSTITUTE
Leibniz Institute for Economic Research
at the University of Munich

ifo
BEITRÄGE
zur Wirtschaftsforschung

93
2021

Essays on Equality of Opportunity

Paul Hufe

Herausgeber der Reihe: Clemens Fuest
Schriftleitung: Chang Woon Nam

ifo INSTITUTE
Leibniz Institute for Economic Research
at the University of Munich

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über
http://dnb.d-nb.de
abrufbar.

ISBN: 978-3-95942-092-1

Druck: ifo Institut, München

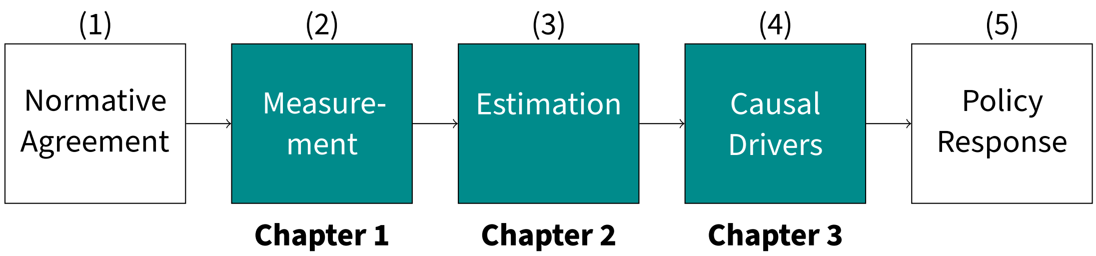
ifo Institut im Internet:
http://www.ifo.de

Preface

Equality of opportunity is a principle of justice that is built on two fundamental ideas. On the one hand, outcome differences across individuals are unacceptable if they are rooted in factors that are beyond individual control. Examples of such *circumstance* characteristics are the biological sex, race, and the socio-economic status of one's parents. On the other hand, if individual outcomes were the result of *effort*, proponents of an equal-opportunity ethic would accept outcome differences across individuals as fair.

While these principles have their origin in the philosophical discourse on distributive justice, they are nowadays widely referenced by public and political actors when discussing inequality in various important domains of life including health, education and income.

Advancing from philosophical reasoning to a society that complies with the deliberated principles can be conceived as involving the following five steps:



After agreeing on the normative principles that underpin a just distribution of resources (Step 1), we need to devise measures that capture these normative intuitions (Step 2). Using appropriate estimators for these measures, we may detect divergences of the status quo from the normative bliss point (Step 3). Having assessed the need for policy intervention based on these estimates, we need to identify the causal factors that drive the detected divergences (Step 4) and design policy responses to move the distribution of resources in the desired direction (Step 5). In this dissertation I contribute to steps 2–4 of the outlined process.

Measurement. The first chapter of this thesis is dedicated to the development of inequality measures that reflect opportunity-egalitarian principles.

Empirical evidence on distributional preferences shows that people do not judge inequality as problematic per se but that they take the underlying sources of income differences into account. In contrast to this evidence, standard measures of inequality do not adequately reflect these normative preferences.

In this chapter, which is based on joint work with Ravi Kanbur and Andreas Peichl, we propose an alternative way of measuring inequality that corresponds more strongly to general principles of justice and the normative preferences upheld by the larger public. In particular, the proposed measures acknowledge that equality of opportunity is important but individually insufficient to define a fair distribution of resources. For example, many people would subscribe to the moral imperative of addressing hunger, homelessness, and material deprivation regardless of how these outcomes came about. However, such a preference stands in contrast to the opportunity-egalitarian doctrine according to which we should accept outcomes if they were the result of individual responsibility and effort exertion. In response, we propose the first family of measures for unfair inequality that incorporate the principles of equality of opportunity and freedom from poverty in a co-equal fashion. We therefore take seriously the idea that equity is not represented by the absence of any inequality in outcomes, but that it requires life success to be orthogonal to exogenous factors outside of individual's control and that everybody should have enough to make ends meet.

Furthermore, we provide two empirical applications of our measure that yield important insights for the inequality debate and the design of appropriate policy responses. First, we analyze the development of inequality in the US over the time period 1969–2014 from a normative perspective. Our results show that the US trend in unfair inequality has mirrored the marked increase of total inequality since the beginning of the 1980s. However, beginning with the 1990s unfair inequality follows a steeper growth curve than total inequality. We illustrate that this trend is mainly driven by a less equal distribution of opportunities across people that face different circumstances beyond their individual control. Second, we provide a corresponding international comparison between the US and 31 European countries in 2010. We find that unfairness in the US shows a remarkably different structure than in societies with comparable levels of unfairness in Europe. Our evidence suggests that inequality in the most unfair European societies is largely driven by poverty increases that followed the financial crisis of 2008. To the contrary, unfairness in the US is driven by marked decreases in social mobility.

Estimation. The second chapter of this thesis is dedicated to the estimation of inequality of opportunity measures.

Measures of inequality of opportunity quantify the extent to which individual outcomes are determined by circumstance characteristics. This idea is commonly operationalized by using a set of circumstances to predict an outcome of interest and calculating inequality in the distribution of predicted outcomes: the more predicted outcomes diverge, the more circumstances beyond individual control influence outcomes, and the more inequality of opportunity there is. However, in standard practice researchers are left to their own devices in specifying the prediction function. This leads to downward biases in inequality of opportunity estimates if the prediction function is too restrictive to capture the dependence of life outcomes on circumstance characteristics. To the contrary, it leads to upward biases if an overly flexible prediction function overfits the data and the relevant parameters are noisily estimated.

In this chapter, which is based on joint work with Paolo Brunori and Daniel Gerszon Mahler, we propose the use of machine learning methods – and regression trees and forests in particular – to overcome the issue of ad-hoc model selection. Machine learning methods allow for flexible models of how unequal opportunities come about while imposing statistical discipline through criteria of out-of-sample replicability. These features serve to establish inequality of opportunity estimates that are less prone to upward or downward bias.

To showcase the advantages machine learning methods we compare them to existing estimation approaches in a cross-sectional dataset of 31 European countries. We demonstrate that current estimation approaches overfit (underfit) the data which in turn leads to upward (downward) biased estimates of inequality of opportunity. These biases are sizable. For example, some standard methods overestimate inequality of opportunity in Scandinavian countries by close to 300%, whereas they underestimate the extent of inequality of opportunity in Germany by more than 40%. Hence, cross-country comparisons based on standard estimation approaches yield misleading recommendations with respect to the need for policy intervention in different societies.

Causal Drivers. The third chapter of this thesis is dedicated to the identification of causal factors that drive the existence of unequal opportunities.

Throughout the post-World War II period, the convergence of wages and labor market participation rates of men and women has been a shared element of labor markets in many industrialized societies. In response to changing economic incentives, heterosexual couples with children have adjusted their time-use and spending patterns, henceforth leading to marked changes in the way they invest into the skill formation of their children. While these long-run trends are well-documented, there is currently no study that causally links the convergence of labor market opportunities across gender groups in the parental generation to the skill formation of children in the following generation.

In this chapter, I study how changes in the parental wage gap influence children's formation of socio-emotional skills as measured by the Big Five personality inventory. I investigate this question causally by constructing a sample of 6,070 German siblings aged 2–17 for whom I observe measures of the Big Five inventory at the same age but in different calendar years. In addition, I match this sibling sample to wage measures that reflect variation in the sex- and education-specific labor demand across commuting zones in Germany. As a result, I can analyze within-family changes in time-use and monetary resources that follow from plausibly exogenous changes in the relative labor market incentives for mothers and fathers, and how these changes affect the socio-emotional development of their children.

I find that decreases in the parental wage gap lead to i) an increase of household's total financial resources, ii) an increase of financial resources controlled by mothers, and iii) an increase in the use of informal care providers. In spite of these changes, I find no effect on the socio-emotional development of children as measured by the Big Five inventory. These null effects are precise enough to exclude various effect sizes from other quasi-experimental interventions studied in the existing literature. In sum, my findings suggest that strides towards gender equality in the labor market do not necessarily come at the cost of detrimental effects on child development.

Keywords: Inequality; Equality of Opportunity; Poverty; Fairness; Measurement; Machine Learning; Random Forests; Family Decision-Making; Gender Wage Gap; Skill Formation

JEL-No: C38; D31; D63; I32; J13; J16; J22; J24

Acknowledgement

I am indebted to many people without whose continued advice and support this thesis would not have been possible.

First and foremost, I want to express the deepest gratitude to my supervisor Andreas Peichl for giving me the opportunity to roam free in the pursuit of my intellectual interests. Throughout my search for a suitable graduate program I had received many skeptic responses to my intention to work at the intersection of economics and normative theory. Today it is hard for me to imagine that I could have found a better academic home.

Second, I want to thank my second and third supervisor, Uwe Sunde and Dominik Sachs, as well as my collaborators Ingvild Almås, Benjamin Arold, Katrin Auspurg, Philipp Dörrenberg, Eric Hanushek, Fabian Kosse, Marc Piopiunik, Laila Schmitt, Lars Schwettmann, John E. Roemer, Paul Schüle, Marc Stöckli, Martin Ungerer and Daniel Weishaar: I have learned a lot from all of you! A special thanks is dedicated to Paolo Brunori, Ravi Kanbur and Daniel Gerszon Mahler without whom Chapter 1 and Chapter 2 of this thesis would not exist.

Third, I am indebted to Marc Fleurbaey, Ravi Kanbur and Bertil Tungodden who have kindly hosted me at Princeton, Cornell and NHH in Bergen. Furthermore, I have benefited from financial support granted by the DFG (PE 1675/5-1 and PE 1675/6-1) and the Fritz Thyssen Foundation.

This thesis is dedicated to Bernhild, Emma, Frieda, Jürgen, Paula, Peter and Susanne from whom I received love and unconditional support from day one!

Lastly, I am eternally grateful to Leo, Felix B., Raggn and Donny Don for keeping me afloat all these years: #20 Karat.

Essays on Equality of Opportunity

Inaugural-Dissertation

Zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)

eingereicht an der
Ludwig-Maximilians-Universität München
2020

vorgelegt von

Paul Hufe

Referent: Prof. Dr. Andreas Peichl
Korreferent: Prof. Dr. Uwe Sunde
Promotionsabschlussberatung: 03.02.2021

Contents

List of Figures

List of Tables

1 Measuring Unfair Inequality: Reconciling Equality of Opportunity and Freedom from Poverty

This chapter is based on the paper "Measuring Unfair Inequality: Reconciling Equality of Opportunity and Freedom from Poverty" and has been co-authored with Ravi Kanbur and Andreas Peichl.

1.1 Introduction

Rising income inequality in many countries around the world has led to intense debates – both in academia and in the public. Calls for more redistribution are often countered by pointing out that outcome inequalities are i) necessary to incentivize individuals and ii) may reflect the just deserts of people in a market economy. However, standard measures of inequality are inappropriate to inform the fairness debate because they neither correspond to standard principles of distributive justice nor to the distributional preferences upheld by the larger public. In this paper, we propose a new measure of (unfair) inequality that reconciles two widely-held normative principles, namely equality of opportunity and freedom from poverty, into a joint indicator. Bringing this new measure to the data, we provide important insights about the fairness of inequality, both over time (in the US) and across countries (in 2010).

Following the seminal work by Piketty and Saez (2003), the literature has documented a continued increase of income inequality since the beginning of the 1980s in many Western societies.[1] This evidence has strongly influenced public discourse. For example, the Occupy Wall Street movement's slogan – "*We are the 99%*" – directly follows from research on the income share of the top 1%. Among other interest groups, this movement has fiercely advocated for more redistribution. To the contrary, free-market pundits emphasize that through trickle-down effects everybody benefits from growth among the job creators at the top. As a consequence, more redistribution would dis-incentivize those individuals and lead to lower welfare for everybody in the long-run. While the equity-efficiency trade-off dominates public

[1] See, among others, Atkinson and Piketty (2007), Guvenen and Kaplan (2017), Leigh (2007), Piketty et al. (2018), and Roine and Waldenström (2015).

discourse on inequality, an explicit discussion of what we understand by an equitable distribution of income is mostly absent. To the contrary, the implicit assumption in much of public discourse as well as in the recent economics literature seems to be that less inequality by necessity implies a more equitable distribution. However, it is highly questionable whether our conception of equity is adequately represented by an inequality measure that invokes perfect equality as the normative benchmark. For instance, is it really the case that everybody receiving the same income (i.e. a Gini coefficient of 0) represents the most equitable distribution when people exert different levels of effort?

In contrast, most theories of distributive justice argue that we should not be concerned by outcome inequality per se, but that we should rather focus on the sources and structure of inequality. To do so, these theories differentiate between fair (justifiable) and unfair (unjustifiable) sources of inequality. Unfair inequality shall be eliminated completely while fair inequalities ought to persist.[2] For example, according to responsibility-sensitive egalitarian theories of justice, outcome inequalities are unfair if they are rooted in factors beyond individual control. These factors could not have been influenced by individual choice and therefore people should not be held responsible for the (dis-)advantages that follow from them.[3] In line with this reasoning, individuals are more willing to accept income differences which are due to effort and preferences rather than exogenous circumstances (Alesina and Giuliano, 2011; Alesina et al., 2018; Cappelen et al., 2007; Fong, 2001).[4] Yet, in spite of its wide acceptance, invoking the notion of individual responsibility alone is insufficient to define fairness (e.g. Konow, 2003; Konow and Schwettmann, 2016). For example, when an outcome is such that it brings deep deprivation to an individual, questions of how it came about seem secondary to the moral imperative of addressing the extremity of the outcome, be it hunger, homelessness, violence or insecurity (Bourguignon et al., 2006).[5] Hence, while outcome differences based on

[2] In the social choice literature these two intuitions are formally represented by *compensation* and *reward* principles (Fleurbaey, 2008; Fleurbaey and Maniquet, 2011).

[3] A non-comprehensive list of works emphasizing this distinction includes Arneson (1989), Cohen (1989), Dworkin (1981a,b), Lippert-Rasmussen (2001, 2011), Rawls (1971), Roemer (1993, 1998), and Sen (1980).

[4] Moreover, the literature branches on intergenerational mobility (see, e.g., Björklund and Jäntti, 1997; Black and Devereux, 2011; Chetty et al., 2014b,c; Corak, 2013; Solon, 1992), the gender pay gap (see, e.g., Blau and Kahn, 2017; Kleven et al., 2018) and also on racial disparities (see, e.g., Kreisman and Rangel, 2015; Lang and Lehmann, 2012) are concerned with inequalities that are in each case rooted in one specific factor beyond individual control. The volume of academic research on these topics is a further indication that circumstance-based inequalities are of foremost public interest.

[5] To illustrate this point, Kanbur and Wagstaff (2016) suggest the following thought experiment: Imagine yourself serving on a soup line. The indigents move forwards and you hand out hot soup. But in one case a new piece of information is given to you. You are told that the outcome of the person in front of you was not due to

exogenous circumstances imply violations of fairness, the reverse statement does not hold true. To the contrary, in addition to the responsibility criterion there are many reasons why a given outcome distribution could be considered unfair – one of them being that not everybody has enough to make ends meet.

In this paper, we propose the first family of measures for unfair inequality that incorporate the principles of equality of opportunity (EOp) and freedom from poverty (FfP) in a co-equal fashion. In line with the previous discussion, we therefore take seriously the idea that equity is not represented by the absence of any inequality in outcomes, but that it requires life success to be orthogonal to exogenous circumstances (EOp) *and* that everybody should have enough to make ends meet (FfP).

To do so, we build on the norm-based approach towards inequality measurement (Cowell, 1985; Magdalou and Nock, 2011). In a first step, we construct a fair income distribution that complies with both the principles of EOp and FfP as the benchmark.[6] In a second step, we measure unfair inequality as the divergence between this norm distribution and the observed income distribution. We show that our proposed measure is easily interpretable and exhibits desirable properties identified in the measurement literature. It furthermore nests standard measures of both equality of opportunity and poverty.

Our paper makes two main contributions. First, we develop the first measure of unfair inequality that reconciles EOp and FfP in a co-equal fashion. Both EOp and FfP have a vast theoretical and empirical literature. Yet, characterizations of unfairness that have relied on separate application of either principle have been criticized concerning their theoretical scope as well as their policy implications (Kanbur and Wagstaff, 2016). Moreover, previous attempts to reconcile the two principles are scant and subject to important drawbacks. For example, existing works give priority to either EOp or FfP, while treating the second principle as a mere weighting factor (Brunori et al., 2013). We address these shortcomings by treating EOp and FfP as co-equal principles conveying different grounds for compensation. That is, we develop an inequality measure that detects unfairness emanating from unequal opportunities or poverty even if one of the two guiding principles is satisfied.

circumstances but a lack of effort. Would you withdraw your soup holding hand because her outcome is morally justifiable according to the responsibility criterion? If not, clearly some other principle is cutting across the power of the responsibility-sensitive egalitarian argument.

[6] Note that standard measures of inequality, such as the Gini index, can also be understood as norm-based measures, in which the norm vector requires perfect equality. The explicit construction of a norm distribution lays bare the normative assumptions that underpin the respective inequality measure.

Second, our measure yields important insights for the inequality debate and the design of appropriate policy responses. We provide two empirical applications of our measure. First, we analyze the development of inequality in the US over the time period 1969-2014 from a normative perspective. Our results show that the US trend in unfair inequality has mirrored the marked increase of total inequality since the beginning of the 1980s. However, beginning with the 1990s unfair inequality follows a steeper growth curve than total inequality. We illustrate that this trend is mainly driven by a less equal distribution of opportunities across people that face different circumstances beyond their individual control. Second, we provide a corresponding international comparison between the US and 31 European countries in 2010. We find that unfairness in the US shows a remarkably different structure than in societies with comparable levels of unfairness in Europe. Our evidence suggests that inequality in the most unfair European societies is largely driven by poverty increases that followed the financial crisis of 2008. To the contrary, unfairness in the US is driven by marked decreases in social mobility. Finally, we acknowledge that the exact definition of the categories "fair" and "unfair" is a normative choice and hence open to debate. We therefore provide extensive sensitivity analyses in which we probe our baseline results against alternative normative assumptions.

The remainder of this paper is organized as follows. In section 1.2 we clarify the underlying normative principles of EOp and FfP. In section 3.2 we develop our measure of unfair inequality. Section 3.5 provides the two empirical applications describing unfair inequality in the US over time and in an international comparison. Sensitivity analyses with respect to alternative normative assumptions are provided in section 1.5. Lastly, section 1.6 concludes.

1.2 Normative Principles

Equality of Opportunity. Equality of opportunity (EOp) is a popular concept of fairness that is used to evaluate distributions of various outcomes, including health, education or income. Following the seminal contributions by Fleurbaey (1995), Roemer (1993, 1998), and Van de gaer (1993), a vivid theoretical and empirical literature evolved that weaves the idea of personal responsibility into inequality research.[7] Opportunity egalitarians deem inequalities ethically acceptable to the extent that they are rooted in factors of individual responsibility. To the contrary, they condemn inequalities that follow from factors beyond

[7] See Ferreira and Peragine (2016), Ramos and Van de gaer (2016), and Roemer and Trannoy (2016) for recent overviews.

individual control. Prominent examples of the latter are, for example, the biological sex, race, or the socioeconomic status of parents. If individual responsibility factors were the sole determinants of the observed outcome distribution, the EOp ideal would be realized to its full extent.

To operationalize the opportunity-egalitarian idea, the literature draws on the concepts of *circumstances* and *efforts*, where circumstances are those outcome determinants for which individuals shall not be held responsible. On the contrary, efforts belong to the realm of personal responsibility. To the extent that the former rather than the latter are stronger (weaker) determinants of the empirical outcome distribution, a society is considered less (more) fair than otherwise. Measures of EOp are underpinned by two fundamental ideas. First, people should be compensated for unequal circumstances. A prominent formulation of this idea is the principle of *ex-ante compensation* which postulates that opportunity sets ought to be equalized across people with differential circumstances. The principle is *ex-ante* because opportunity sets are evaluated prospectively without regard to the individual level of effort exertion. Second, people should be appropriately rewarded for their efforts. While there are again different formulations of this idea, one prominent version is the principle of *utilitarian reward*. Utilitarian reward states that effort should be rewarded in a way that maximizes the aggregate outcome of people with the same circumstances. It entails that outcome differences between individuals with the same circumstances are a matter of indifference. *Ex-ante utilitarian* measures of EOp therefore boil down to measures of between-group inequality where groups are defined by their respective circumstance characteristics.[8] The precise cut between circumstances and efforts is normatively contentious. For example, some argue in favor of including genetic endowments into the set of circumstances (Lefranc et al., 2009) while others deny that outcomes flowing from advantageous natural endowments are less praiseworthy than outcomes flowing from effort (Miller, 1996). Similarly, it is widely debated whether the correlation between effort levels and circumstances constitutes a ground for compensation or not. While some argue in favor of holding people responsible for their preferences regardless of how they are formed (Barry, 2005), others allocate such correlation to the circumstances that demand compensation (Roemer, 1998). In our baseline empirical application in section 3.5, we draw on commonly accepted circumstance characteristics and allocate the correlation between circumstances and efforts to the unfair determinants of

[8] For a comprehensive discussion of different compensation and reward principles see the works of Fleurbaey and Peragine (2013) and Ramos and Van de gaer (2016).

inequality. However, we provide sensitivity analyses for different responsibility cuts in section 1.5.

Beyond theoretical reasoning, there is compelling empirical evidence that people indeed disapprove of inequalities that are rooted in factors beyond individual control. Alesina et al. (2018) use information treatments to show that policy preferences with respect to taxation and spending on opportunity-equalizing policies are robustly correlated with perceptions of social mobility. The lower social mobility within a society, the more people are willing to remedy existing inequalities by appropriate policy interventions. Faravelli (2007) demonstrates that perceptions of justice tend to more equal distributions when income differences originate from contextual factors that could not have been influenced by individuals. The works of Cappelen et al. (2007) and Krawczyk (2010) confirm that people uphold the equal-opportunity ideal even if it adversely affects their own material interests.

Freedom from Poverty. Poverty is an important focal point in public debates about the appropriate distribution of material resources. In the philosophical literature the focus on the least advantaged has been defended by reference to sufficientarian conceptions of justice (Frankfurt, 1987) and arguments that consider material deprivation as a violation of the undeniable rights we have in virtue of being humans (Fleurbaey, 2007).[9] Akin to the literature on EOp, the normative concern for deprivation operates on a principle of compensation: Deprived people are entitled to be compensated so as to attain the material conditions to live a life of reasonable comfort.

While there is wide-spread appreciation for the multidimensionality of poverty (Aaberge and Brandolini, 2015), much of the empirical poverty literature focuses on the income dimension only. In general, poverty measurement follows a two-step process. First, set a threshold that partitions the population into its deprived and non-deprived fractions. All else equal, the more lenient the definition of the deprivation threshold, the larger the group to which compensation is owed. Second, choose a function to aggregate the gaps between observed incomes and the deprivation threshold for those whose income falls below the threshold. In analogy to the cut between circumstances and effort, the appropriate setting of the poverty

[9] Some object that freedom from poverty does not belong to the theoretical realm of fairness or even justice *although* it is morally objectionable. Such moral objections could be raised from a humanitarian or human rights perspective. In this paper we use the term “unfair” in a colloquial sense to indicate that a distribution of some good is unfair if it raises moral objection.

line is a widely debated issue in the literature (among others Decerf, 2017; Foster, 1998). In our baseline empirical application in section 3.5, we draw on an internationally comparable absolute poverty threshold but provide sensitivity analyses for this choice in section 1.5.

The concern for poverty alleviation is strongly reflected in the distributional preferences of the general public. The evidence summarized in Konow (2003) and Konow and Schwettmann (2016) indicates that fairness preferences are sensitive to individual needs and reflect a concern for everybody having enough to make ends meet. Cappelen et al. (2013b) use an international dictator game to show that transfers increase if the recipient comes from a poorer country, while Fisman et al. (2018) show that inequality aversion goes hand in hand with a preference for increasing the incomes of the worst-off in society.

Reconciling EOp and FfP. In this work we treat EOp and FfP as co-equal principles conveying different grounds for compensation. Our approach is philosophically inspired by the recognition that both EOp and FfP are individually insufficient to characterize what a fair distribution of resources requires (Anderson, 1999; Vita, 2007). These theoretical insights are bolstered by empirical evidence that distributional preferences are sensitive to i) ex-ante inequalities that are determined by exogenous circumstances and ii) ex-post inequalities that are insensitive to responsibility considerations. For example, the experiments of Cappelen et al. (2013a) show that people largely endorse an ex-ante equal-opportunity ethic, however, they also correct for ex-post inequalities that are the result of luck. Andreoni et al. (2019) suggest that social preferences are a mix of ex-ante and ex-post considerations where the latter gain in importance once the outcome is observed. Consistent with these findings Gaertner and Schwettmann (2007) show that people tend to compensate extreme outcomes irrespective of whether they are the result of individual responsibility factors or not. In Figure A.5 we furthermore show survey evidence on public support for four principles of justice in 18 European countries that are part of our empirical application. A consistent pattern emerges: People are not perfect outcome egalitarians. Instead, they most strongly endorse a distribution of income that is sensitive to individual need (FfP) and rewards individual effort but not family background characteristics (EOp).

In spite of this evidence, previous attempts to reconcile the (ex-ante) EOp principle with the (ex-post) FfP principle are scant. First, Brunori et al. (2013) propose an "opportunity-sensitive poverty measure" according to which identical incomes below the poverty line receive less weight the more advantageous the circumstances of the poor individuals that are compared. However, since EOp serves as a mere weighting factor in the evaluation of incomes below the deprivation threshold, their measure does not detect any unfairness in societies that are free from poverty but that are characterized by severe violations of EOp. The measure is therefore informative if one aims to prioritize poor individuals based on the responsibility criterion. However, it does not allow to quantify the overall level of fairness in an observed income distribution. Second, Ferreira and Peragine (2016) suggest the construction of "opportunity-deprivation profiles" where members of circumstance types are considered opportunity-deprived if their average outcome falls below a pre-specified deprivation threshold. Effectively, this amounts to applying standard poverty measures to circumstance types instead of individuals. As a consequence the measure is informative for the identification of particularly opportunity deprived types. However, just as the "opportunity-sensitive poverty measure" it does not allow to quantify the overall level of fairness in an observed income distribution.

1.3 Measuring Unfair Inequality

In this section we describe how we construct measures of unfair inequality that – in contrast to previous work – treat EOp and FfP as co-equal principles conveying different grounds for compensation.

1.3.1 Notation

Consider the society $\mathcal{N} = \{1, 2, .., N\}$ and an associated vector of non-negative incomes $Y^e = (y^e_1, y^e_2, ..., y^e_N)$. Y^e corresponds to the empirical income distribution. Let us furthermore define a minimum income threshold y_{min} that is required to make ends meet. Based on Y^e and y_{min} we can partition the population into a poor and a non-poor fraction:

$$\mathcal{P} = \{i \in \mathcal{N} \mid y^e_i \leq y_{\text{min}}\}, \tag{1}$$

$$\mathcal{R} = \{i \in \mathcal{N} \mid y^e_i > y_{\text{min}}\}. \tag{2}$$

Individual incomes at all levels are the result of two sets of factors: First, a set of *circumstances* beyond individual control $\Omega \subseteq \mathbb{R}^C$. Second, a set of individual *efforts* $\Theta \subseteq \mathbb{R}^E$. We define the vector $\omega_i \in \Omega$ as a comprehensive description of the circumstances with which $i \in \mathcal{N}$ is endowed. Analogously we define the vector $\theta_i \in \Theta$ as a comprehensive description of the efforts that are exerted by $i \in \mathcal{N}$. Based on the realizations of circumstances we can partition the population into T circumstance *types* that are defined as follows:

$$\mathcal{T}(\omega) = \{i \in \mathcal{N} : \omega_i = \omega\}. \tag{3}$$

Similarly, we can partition the population into S effort *tranches* that are defined as follows:

$$\mathcal{S}(\theta) = \{i \in \mathcal{N} : \theta_i = \theta\}. \tag{4}$$

For any subgroup $\mathcal{X} \subseteq \mathcal{N}$ of individuals, we denote by $N_{\mathcal{X}} = \mathrm{card}(\mathcal{X})$ the number of individuals in this subgroup and by $\mu^e_{\mathcal{X}} = \frac{1}{N_{\mathcal{X}}} \sum_{i \in \mathcal{X}} y^e_i$ their average income. For ease of notation, we let hereafter $N = N_{\mathcal{N}}$ and $\mu^e = \mu^e_{\mathcal{N}}$.

Next to the empirical income distribution Y^e, consider a norm (or reference) income distribution $Y^r = (y^r_1, y^r_2, ..., y^r_N)$ that describes the fair distribution of incomes. It is the normatively desirable income distribution for which the society should strive in absence of incentive constraints and behavioral responses to redistribution. While Y^e is given in the data, Y^r must be constructed based on explicit normative principles.[10] Before outlining the construction of a Y^r that is consistent with the normative intuitions of EOp and FfP in section 1.3.3, we will now describe how to aggregate the differences between Y^e and Y^r into a scalar measure of inequality.

1.3.2 Measuring Divergence

Endowed with both Y^e and Y^r one must decide how to aggregate the discrepancies between both vectors into a scalar measure of unfair inequality. Prominent divergence measures include the works by Almås et al. (2011), Cowell (1985), and Magdalou and Nock (2011), each of which generalizes standard measures of inequality. While Cowell (1985) and Magdalou and Nock (2011) provide generalizations of the entropy class of inequality measures, Almås et al. (2011) generalize the Gini index. The key difference to standard measures of inequality is that these generalized measures do not decrease (increase) with progressive (regressive) transfers from rich (poor) to poor (rich) but rather with transfers that reduce (increase) the distance between the empirical and the reference distribution. Note that this requirement

[10] Note that standard measures of inequality such as the Gini coefficient adhere to the norm of outcome egalitarianism, i.e. this norm distribution is the perfect equality distribution where $y^r_i = \mu^e, \forall\, i \in \mathcal{N}$.

is equivalent to the standard Pigou-Dalton principle of transfers if and only if the reference distribution is equivalent to the sample mean μ^e. Otherwise, transfers from poor to rich can be desirable if the income of the poor exceeds its reference value, while the income of the rich falls short of it.

In our baseline, we use the measures proposed by Magdalou and Nock (2011) yielding the following aggregator for the divergence between Y^e and Y^r:[11]

$$D(Y^e||Y^r) = \sum_{i \in \mathcal{N}} \left[\phi(y_i^e) - \phi(y_i^r) - (y_i^e - y_i^r)\phi'(y_i^r)\right], \tag{5}$$

where

$$\phi(z) = \begin{cases} -\ln z, & \text{if } \alpha = 0, \\ z \ln z, & \text{if } \alpha = 1, \\ \frac{1}{\alpha(\alpha-1)} z^{\alpha}, & \text{otherwise}. \end{cases} \tag{6}$$

As in the family of generalized entropy measures, α is indicative of different value judgments: The higher α, the more weight is attached to positive divergences of empirical income y_i^e from its respective norm income y_i^r. The lower α, the more weight is attached to shortfalls from y_i^r. In the baseline we choose $\alpha = 0$. This choice is guided by the fact that the MN-measure with $\alpha = 0$ nests the mean log deviation (MLD) if we set $y_i^r = \mu^e, \forall\, i \in \mathcal{N}$. As such we ensure close proximity to the empirical literature on EOp, in which the use of the MLD is prevalent (among others Ferreira and Gignoux, 2011; Hufe et al., 2017). Furthermore, attaching a higher weight to shortfalls from y_i^r is consistent with recent experimental evidence showing a preference for overcompensating the undeserving instead of failing to compensate the deserving (Cappelen et al., 2018).[12] Thus, our baseline measure of unfair inequality aggregates divergences between Y^e and Y^r as follows:

$$D(Y^e||Y^r) = \frac{1}{N} \sum_{i \in \mathcal{N}} \left[\ln \frac{y_i^r}{y_i^e} - \frac{y_i^r - y_i^e}{y_i^r}\right].^{13} \tag{7}$$

[11] We abbreviate this class with *MN* in the following. The MN-family of divergence measures is characterized by the properties of *scale invariance*, the *principle of population*, and *subgroup decomposability*. These properties carry directly over to our measures of unfair inequality. Robustness checks using the measures of Almås et al. (2011) and Cowell (1985) are provided in section 1.5.4.
[12] Robustness checks using alternative specifications of α are provided in section 1.5.4.
[13] Note that we can scale the measure by $1/N$ to satisfy the *principle of population* without further adjustment since we will constrain the mean of Y^r to match the mean of Y^e (Magdalou and Nock, 2011).

We will now turn to the construction of a norm vector Y^r that accords with the principles of EOp and FfP.

1.3.3 Baseline Measure

Norm Vector. Let $\mathcal{D} \subseteq \mathbb{R}^N_+$ be the set containing all possible income distributions d. In the following we will define subsets of eligible distributions $\mathcal{D}^h \in \mathcal{D}$ that are consistent with the normative intuitions embodied in the principles of EOp and FfP.

First, since we are concerned with the fair distribution of available resources in a given society, we follow the inequality measurement literature and rule out creatio ex nihilo:

$$\mathcal{D}^1 = \left\{ d \in \mathcal{D} \,\middle|\, \sum_{i \in \mathcal{N}} y_i^r = \sum_{i \in \mathcal{N}} y_i^e \right\}. \tag{8}$$

Thus, $\mathcal{D}^1$ is the subset of distributions in which the total amount of available resources match their empirical counterpart. By fixing the total amount of resources we let the distribution of these resources be the only margin of difference between the observed and the benchmark situation.[14] This assumption is standard in the literature on inequality measurement but highlights an important difference to the quest for optimal tax design. The latter is concerned with trading off equity and efficiency concerns. In such a framework, restriction (8) would rule out behavioral responses to taxation and only makes sense in a first-best setting.[15]

Second, we characterize the EOp principle by reference to the principle of ex-ante compensation (Fleurbaey and Peragine, 2013; Ramos and Van de gaer, 2016) which states that the expected income of an individual should not be correlated to her circumstance type. Hence, we are infinitely inequality averse with respect to inequalities *between* circumstance types and the ideal of an equal-opportunity society is realized if there is equality across average

[14] Cappelen and Tungodden (2017) call restriction $\mathcal{D}^1$ the "no-waste-condition".

[15] The efficiency costs of reaching the norm distribution are never part of inequality measurement. Accounting for efficiency costs, however, could be part of further analysis. Assuming the joint minimization of EOp and FfP to be a goal of public policy, our framework could be integrated into models of fair taxation (Fleurbaey and Maniquet, 2006; Ooghe and Peichl, 2015; Saez and Stantcheva, 2016; Weinzierl, 2014) in which the planner seeks to realize a specific notion of fairness while taking behavioral responses to taxation into account. See also Fleurbaey and Maniquet (2018) for a recent overview on the integration of fairness principles into the standard Mirrleesian optimal tax framework.

type incomes $\mu^e_{\mathcal{T}(\omega)}$. $\mathcal{D}^2$ is the subset of distributions for which this criterion is satisfied:

$$\mathcal{D}^2 = \left\{ d \in \mathcal{D} \,\middle|\, \mu^r_{\mathcal{T}(\omega)} = \frac{1}{N_{\mathcal{T}(\omega)}} \sum_{i \in \mathcal{T}(\omega)} y^r_i = \frac{1}{N} \sum_{j \in \mathcal{N}} y^e_j = \mu^e,\ \forall\, \omega \in \Omega \right\}. \tag{9}$$

Note that in this specification we implicitly treat the correlation between Ω and Θ as morally objectionable. This assumption is in line with the normative account of Roemer (1998). However, we provide sensitivity checks to this assumption in section 1.5.

Third, we maintain that people have a claim for a minimum level of resources y_{min} even if their outcomes can be ascribed to factors within their realm of control. Opportunity equalization alone does not achieve this objective. Next to compensating circumstances Ω, opportunity-egalitarians want to preserve income differences due to effort exertion. Consistent with this idea, we impose that within types $\mathcal{T}(\omega)$ efforts should be respected by distributing incomes proportionally to empirical incomes y^e_i:

$$\frac{y^r_i}{y^r_j} = \frac{y^e_i}{y^e_j},\ \forall\, i, j \in \mathcal{T}(\omega),\ \forall\, \omega \in \Omega.^{16} \tag{10}$$

However, while such ex-post proportionality within $\mathcal{T}(\omega)$ maintains relative differences in effort exertion, it may keep (push) some $i \in \mathcal{P}$ ($i \in \mathcal{R}$) below y_{min}. To realize FfP we therefore want to identify those who are poor due to a lack of effort exertion instead of exogenous circumstances and compensate them so that they are able to make ends meet. In line with this insight we define a partition according to which people are labeled (non-)poor after considering their counterfactual gains from opportunity equalization while holding them responsible for their individual efforts θ_i:

$$\mathcal{P}(\omega) = \left\{ i \in \mathcal{T}(\omega) \,\middle|\, y^e_i \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \leq y_{\text{min}} \right\},\ \forall\, \omega \in \Omega, \tag{11}$$

$$\mathcal{R}(\omega) = \left\{ i \in \mathcal{T}(\omega) \,\middle|\, y^e_i \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} > y_{\text{min}} \right\},\ \forall\, \omega \in \Omega. \tag{12}$$

Based on the definition of $\mathcal{P}(\omega)$ and $\mathcal{R}(\omega)$, we formulate the FfP requirement:

$$\mathcal{D}^3 = \left\{ d \in \mathcal{D} \,\middle|\, y^r_i = y_{\text{min}},\ \forall\, i \in \mathcal{P}(\omega),\ \forall\, \omega \in \Omega \right\}. \tag{13}$$

[16] This condition is a relative version of the "equal-transfer-for-equal-[circumstance]" condition laid out in Bossert and Fleurbaey (1995).

The FfP requirement can be broken down into two parts: $y_i^r = \frac{1}{N_{\mathcal{P}(\omega)}} \sum_{j \in \mathcal{P}(\omega)} y_j^r = \mu^r_{\mathcal{P}(\omega)}$ and $\mu^r_{\mathcal{P}(\omega)} = y_{\text{min}}$. The first component states infinite inequality aversion with respect to income differences among the poor – they all have an *equal* claim to a certain level of resources. The second component states infinite inequality aversion with respect to the average shortfall of the poor population from the poverty line. Within an equal-opportunity society, they all have an equal claim to nothing less (but also nothing more) than exactly the *minimum subsistence level* y_{min}.

Fourth, combining the proportionality requirement (10) with the FfP condition (13), there is zero inequality aversion with respect to the share of income that exceeds the poverty line. Hence, $\mathcal{D}^4$ denotes the subset of distributions in which within-type inequality of excess income above the poverty line remains unaltered in comparison to the counterfactual equal-opportunity income distribution:

$$\mathcal{D}^4 = \left\{ d \in \mathcal{D} \,\middle|\, \frac{y_i^r - y_{\text{min}}}{y_j^r - y_{\text{min}}} = \frac{y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}, \; \forall \, i, j \in \mathcal{R}(\omega), \; \forall \, \omega \in \Omega \right\}. \tag{14}$$

The intersection $\cap_{h=1}^4 \mathcal{D}^h$ characterizes our baseline norm distribution which is summarized in Proposition 1:

Proposition 1. *Suppose $\mu^e > y_{min}$. Then, the intersection $\cap_{h=1}^4 \mathcal{D}^h$ yields a singleton which defines the norm distribution Y^r:*

$$y_i^r = \begin{cases} y_{min}, & \forall \, i \in \mathcal{P}(\omega), \; \forall \, \omega \in \Omega, \\ y_{min} + \underbrace{(y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{min})}_{=\tilde{y}_i} \underbrace{\frac{(\mu^e - y_{min})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}} (\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{min})}}_{=\delta_{\mathcal{T}(\omega)}}, & \forall \, i \in \mathcal{R}(\omega), \; \forall \, \omega \in \Omega. \end{cases} \tag{15}$$

Conversely, if $\mu \leq y_{min}$, then $\cap_{h=1}^4 \mathcal{D}^h = \emptyset$. The proof for this proposition is given in Appendix A.1.

Individuals in $\mathcal{P}(\omega)$ receive a norm income of y_{min}. This prescription directly follows from the FfP requirement specified in (13): Those who are poor due to factors other than exogenous circumstances are owed compensation to make ends meet but nothing more.

Norm incomes for individuals in $\mathcal{R}(\omega)$ are determined by the individual share of (counterfactual) income above the poverty threshold, $\tilde{y}_i \in (0,\infty)$, and a type-specific scaling factor, $\delta_{\mathcal{T}(\omega)} \in (0,\infty)$. First, conditional on the individual circumstance type, y_i^r increases with $\tilde{y}_i$. This relation follows from the proportionality condition in (14): In absence of additional normative grounds for income inequality aversion, relative income differences among people with similar circumstance characteristics that are able to make ends meet need to be preserved. Second, the type-specific scaling factor $\delta_{\mathcal{T}(\omega)}$ increases with the total amount of resources that are available relative to the poverty line ($\mu^e - y_{\text{min}}$). This relation follows from the constant resource requirement specified in (8) and from fixing incomes of the poor population $\mathcal{P}(\omega)$ at the minimum threshold y_{min} (13): The higher the total amount of available resources, the smaller the share of resources that needs to be given up by the members of $\mathcal{R}(\omega)$ in order to realize FfP. Lastly, the type-specific scaling factor $\delta_{\mathcal{T}(\omega)}$ decreases with the share of non-poor individuals in a type ($N_{\mathcal{R}(\omega)}/N_{\mathcal{T}(\omega)}$) and their average (counterfactual) income in excess of the minimum threshold ($\mu^e_{\mathcal{R}(\omega)}\frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}$). This relation follows from combining the EOp requirement given in (9) with the FfP requirement given in (13) while observing the proportionality requirement given in (14): EOp requires equal average outcomes across types. The higher the total volume of transfers to the poor members of a type, the higher the proportional charge levied on the non-poor members of the same type in order to maintain the EOp requirement.

Equation (15) shows that the fair distribution of incomes Y^r is a function of simple summary statistics of the empirical income distribution Y^e. Some may argue that the normatively desirable distribution of incomes should be independent of the actual distribution of incomes. However, we note that such a dependence is not particular to our measurement approach but characterizes many standard measures of inequality, poverty and inequality of opportunity.[17]

Furthermore, we note that the extent of such dependence can be strengthened or loosened in several ways. In fact, whether and to what extent an insulation of Y^r from Y^e is desirable, depends on the normative intuitions one strives to capture. For illustrative purposes we will give two examples in the following. First, y_{min} can be set i) in absolute terms or ii) in relative terms as some functional of Y^r. Option i) is preferable if one thinks that the poverty concept

[17] For example, the standard approach to inequality measurement can be characterized as finding a suitable distance measure between the actual distribution and the norm distribution where every individual has the mean of the distribution. The properties of the distance measure can be further specified (for example, the Pigou-Dalton property, the scale independence property, decomposability, etc.). But as the empirical vector changes, the norm vector also changes. For instance, for the conventional Gini coefficient it holds that $y_i^r = \mu^e, \forall\, i \in \mathcal{N}$, implying that Y^r changes with μ^e.

applies to basic human needs, while option ii) is preferable if one aims to capture aspects of social deprivation as well (Foster, 1998). In our baseline, we choose an absolute poverty threshold and therefore insulate y_{min} from changes in Y^e but provide sensitivity checks to this choice in section 1.5.3. Second, $\mathcal{D}^4$ proposes to honor within-type income differences since we interpret them as indicators of differential effort exertion. In line with this normative interpretation, our baseline Y^r is dependent on changes in the intra-type variance of incomes and therefore Y^e. However, in section 1.3.4 we show how the dependence between Y^r and Y^e can be loosened by harmonizing intra-type variances in Y^r across circumstance types. More generally: While the construction of Y^r may depend on Y^e to varying degrees, the underlying principles that inform the construction of Y^r are always independent of the observed distribution of incomes.

Measure and Comparative Statics. Substituting the norm distribution given in (15) into the divergence measure given in (7), we obtain our baseline measure of unfair inequality:

$$\begin{aligned} D(Y^e||Y^r) =& \frac{1}{N}\sum_{i\in\mathcal{P}(\omega)}\left\{\ln\frac{y_{\text{min}}}{y_i^e}-\left(\frac{y_{\text{min}}-y_i^e}{y_{\text{min}}}\right)\right\} \\ &+\frac{1}{N}\sum_{i\in\mathcal{R}(\omega)}\left\{\ln\left(\frac{y_{\text{min}}+\tilde{y}_i\delta_{\mathcal{T}(\omega)}}{y_i^e}\right)-\left(\frac{(y_{\text{min}}+\tilde{y}_i\delta_{\mathcal{T}(\omega)})-y_i^e}{y_{\text{min}}+\tilde{y}_i\delta_{\mathcal{T}(\omega)}}\right)\right\}, \end{aligned} \tag{16}$$

where $\delta_{\mathcal{T}(\omega)}$ represents the type-specific scaling factor that is applied to $\tilde{y}_i$ – the share of counterfactual income above y_{min}. To further illustrate the properties of this measure, we provide some comparative statics in the following.

(1) Assume $y_{\text{min}} \to 0$. The limiting case of $y_{\text{min}} = 0$ is equivalent to abstracting from the concern for FfP altogether, whereas EOp remains the only normative foundation for inequality aversion. In the limit, this leads to $\mathcal{P}(\omega) = \emptyset$, $\mu^e_{\mathcal{R}(\omega)} = \mu^e_{\mathcal{T}(\omega)}$, and $N_{\mathcal{R}(\omega)} = N_{\mathcal{T}(\omega)}$. As a consequence, $\delta_{\mathcal{T}(\omega)} = 1,\ \forall\,\omega\in\Omega$. The resulting norm vector as well as the ensuing measure of unfair inequality read as follows:

$$y_i^r = y_i^e\frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}},\ \forall\, i\in\mathcal{T}(\omega),\ \forall\,\omega\in\Omega, \tag{17}$$

$$D(Y^e||Y^r) = \frac{1}{N}\sum_{i\in\mathcal{N}}\ln\frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}}. \tag{18}$$

With $y_{\text{min}} = 0$, unfair inequality collapses to inequality in the distribution of average outcomes of circumstance types. Hence, as $y_{\text{min}} \to 0$, the measure converges to the standard ex-ante utilitarian measure of inequality of opportunity in which the MLD is applied to a smoothed distribution of type-specific mean incomes.

(2) Assume $N_{\mathcal{P}(\omega)} \to 0,\ \forall\, \omega \in \Omega$. Note the difference to our previous thought experiment in which we abstracted from the concern for FfP altogether. The limiting case of $N_{\mathcal{P}(\omega)} = 0$ corresponds to a society that values FfP below the threshold of y_{min} but happens to be in the fortunate position of having zero poverty incidence once incomes are corrected for the unequal opportunities faced by people with different circumstances. At the limit, $\mathcal{P}(\omega) = \emptyset$, $\mu^e_{\mathcal{R}(\omega)} = \mu^e_{\mathcal{T}(\omega)}$ and $N_{\mathcal{R}(\omega)} = N_{\mathcal{T}(\omega)}$. As a consequence, $\delta_{\mathcal{T}(\omega)} = 1,\ \forall\, \omega \in \Omega$ and the resulting norm vector as well as the ensuing measure of unfair inequality read as follows:

$$y^r_i = y^e_i \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}},\ \forall\, i \in \mathcal{T}(\omega),\ \forall\, \omega \in \Omega, \tag{19}$$

$$D(Y^e||Y^r) = \frac{1}{N} \sum_{i \in \mathcal{N}} \ln \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}}. \tag{20}$$

In spite of the fact that the concern for FfP remains intact, opportunity equalization is sufficient to satisfy the criteria of both EOp and FfP if $N_{\mathcal{P}(\omega)} = 0,\ \forall\, \omega \in \Omega$. Hence, the measure of unfair inequality again converges to the standard ex-ante utilitarian measure of inequality of opportunity. The limiting case of $N_{\mathcal{P}(\omega)} = 0,\ \forall\, \omega \in \Omega$ thus illustrates that the measure continues to detect unfairness through violations of EOp even if FfP is perfectly satisfied.

(3) Assume we reduce the number of criteria that constitute unfair outcome determinants from an opportunity-egalitarian perspective. This can be represented by letting the number of circumstance types go to one, i.e. $T \to 1$. At the limit, the entire population would be considered as a single circumstance type and FfP remains the only normative foundation for inequality aversion. $T = 1$ leads to $\mathcal{T}(\omega) = \mathcal{N}$, $\mathcal{P}(\omega) = \mathcal{P}$, and $\mathcal{R}(\omega) = \mathcal{R}$. Furthermore, $N_{\mathcal{P}(\omega)} = N_{\mathcal{P}}$, $\mu^e_{\mathcal{T}(\omega)} = \mu^e$, and $\mu^e_{\mathcal{P}(\omega)} = \mu^e_{\mathcal{P}}$. As a consequence, $\tilde{y}_i = y^e_i - y_{\text{min}}$ and $\delta_{\mathcal{T}(\omega)} = \frac{(\mu^e - y_{\text{min}})}{N_{\mathcal{R}}/N(\mu^e_{\mathcal{R}} - y_{\text{min}})} = \left(1 - \frac{N_{\mathcal{P}}/N(y_{\text{min}} - \mu^e_{\mathcal{P}})}{N_{\mathcal{R}}/N(\mu^e_{\mathcal{R}} - y_{\text{min}})}\right) = \delta$. Thus, the corresponding norm vector as well as the resulting measure of unfair inequality read as follows:

$$y_i^r = \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{P}, \\ y_{\text{min}} + (y_i^e - y_{\text{min}}) \underbrace{\left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\text{min}} - \mu_{\mathcal{P}}^e)}{\frac{N_{\mathcal{R}}}{N}(\mu_{\mathcal{R}}^e - y_{\text{min}})}\right)}_{=\delta}, & \forall\, i \in \mathcal{R}, \end{cases} \tag{21}$$

$$\begin{aligned} D(Y^e \| Y^r) = & \underbrace{\frac{1}{N}\sum_{i\in\mathcal{P}} \ln\left(\frac{y_{\text{min}}}{y_i^e}\right)}_{\text{Watts Index}} - \underbrace{\frac{1}{N}\sum_{i\in\mathcal{P}}\left(\frac{y_{\text{min}} - y_i^e}{y_{\text{min}}}\right)}_{\text{Poverty Gap}} \\ & + \frac{1}{N}\sum_{i\in\mathcal{R}}\left\{\ln\left(\frac{y_{\text{min}} + (y_i^e - y_{\text{min}})\delta}{y_i^e}\right) - \left(\frac{(y_i^e - y_{\text{min}})(\delta - 1)}{y_{\text{min}} + (y_i^e - y_{\text{min}})\delta}\right)\right\}. \end{aligned} \tag{22}$$

Abstracting from the concern for EOp, leads to a scaling factor δ that is uniform across all $i \in \mathcal{R}$. δ is determined by the ratio of the poverty gap to the amount of excess income above the poverty line. This contrasts with the baseline case in which the transfer rate $\delta_{\mathcal{T}(\omega)}$ is decreasing with the type-specific share of non-poor individuals and their average excess income above the poverty threshold.

The decomposability property of the MN-measures allows us evaluate unfairness in the truncated distribution $Y_{\mathcal{P}}^e = [y_1^e, y_2^e, ..., y_{\text{min}}]$. Up to y_{min}, unfair inequality is characterized by the difference between the Watts index (Zheng, 1993) and the poverty gap ratio. Individually, these are well-known measures of poverty. However, also their combination bears a number of desirable properties that have been identified in the literature on poverty measurement (e.g. Ravallion and Chen, 2003). These include *monotonicity* (as opposed for example to the headcount ratio), the *principle of transfers* (as opposed for example to the poverty gap taken as a stand-alone measure) and *additive decomposability*. Note that we do not obtain a measure of poverty that satisfies the *focus axiom*. Our approach frames poverty as an aspect of inequality and thus imposes requirements on how the funds to eradicate poverty should be raised – see condition (14). Therefore, it is not indifferent to transfers between individuals with incomes above the poverty line y_{min} (the third term in equation (22)) and thus violates the *focus axiom*.

(4) Let $\mu_{\mathcal{T}(\omega)}^e \to \mu^e$, $\forall\, \omega \in \Omega$. Note the difference to our previous thought experiment, in which we let $T \to 1$ and abstracted from the concern for EOp altogether. In contrast to the previous case, the normative concern for EOp remains intact, however, the EOp principle is increasingly satisfied as $\mu^e_{\mathcal{T}(\omega)} \to \mu^e,\ \forall\, \omega \in \Omega$. The limiting case corresponds to an equal-opportunity society without disparities in average outcomes across circumstance types. At the limit, $\tilde{y}_i = y^e_i - y_{\text{min}}$, $\delta_{\mathcal{T}(\omega)} = \frac{(\mu^e - y_{\text{min}})}{N_{\mathcal{R}(\omega)}/N_{\mathcal{T}(\omega)}(\mu^e_{\mathcal{R}(\omega)} - y_{\text{min}})} = \left(1 - \frac{N_{\mathcal{P}(\omega)}/N_{\mathcal{T}(\omega)}(y_{\text{min}} - \mu^e_{\mathcal{P}(\omega)})}{N_{\mathcal{R}(\omega)}/N_{\mathcal{T}(\omega)}(\mu^e_{\mathcal{R}(\omega)} - y_{\text{min}})}\right)$. The resulting norm vector and the corresponding measure of unfair inequality read as follows:

$$y^r_i = \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{P},\ \forall\, \omega \in \Omega, \\ y_{\text{min}} + (y^e_i - y_{\text{min}}) \underbrace{\left(1 - \dfrac{\frac{N_{\mathcal{P}(\omega)}}{N_{\mathcal{T}(\omega)}}(y_{\text{min}} - \mu^e_{\mathcal{P}(\omega)})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}(\mu^e_{\mathcal{R}(\omega)} - y_{\text{min}})}\right)}_{=\delta_{\mathcal{T}(\omega)}}, & \forall\, i \in \mathcal{R},\ \forall\, \omega \in \Omega. \end{cases} \tag{23}$$

$$\begin{aligned} D(Y^e \| Y^r) = & \underbrace{\frac{1}{N} \sum_{i \in \mathcal{P}} \ln\left(\frac{y_{\text{min}}}{y^e_i}\right)}_{\text{Watts Index}} - \underbrace{\frac{1}{N} \sum_{i \in \mathcal{P}} \left(\frac{y_{\text{min}} - y^e_i}{y_{\text{min}}}\right)}_{\text{Poverty Gap}} \\ & + \frac{1}{N} \sum_{i \in \mathcal{R}} \left\{ \ln\left(\frac{y_{\text{min}} + (y^e_i - y_{\text{min}})\delta_{\mathcal{T}(\omega)}}{y^e_i}\right) - \left(\frac{(y^e_i - y_{\text{min}})(\delta_{\mathcal{T}(\omega)} - 1)}{y_{\text{min}} + (y^e_i - y_{\text{min}})\delta_{\mathcal{T}(\omega)}}\right) \right\}. \end{aligned} \tag{24}$$

Since our concern for EOp remains intact we calculate poverty-eradicating transfers across types by reference to the *type-specific* poverty gap and the *type-specific* income share that exceeds y_{min}. The limiting case of $\mu^e_{\mathcal{T}(\omega)} = \mu^e,\ \forall\, \omega \in \Omega$ shows that our measure continues to detect unfairness through violations of FfP even if EOp is perfectly satisfied.

The previous comparative statics illustrate some particular advantages of our measure of unfair inequality. First, it is easily interpretable since it nests well-known measures of both EOp and FfP. If we abstract from the concern for FfP ($y_{\text{min}} = 0$), we obtain a standard measure for inequality of opportunity. If we abstract from the concern for EOp ($T = 1$), we obtain a combination of the Watts index and the poverty gap ratio, both of which are well-established measures of poverty.

Second, the proposed measure treats EOp and FfP as co-equal principles and therefore detects unfair inequality even if either of the two principles is perfectly satisfied.[18] If there is zero poverty incidence ($N_{\mathcal{P}(\omega)} = 0,\ \forall\, \omega \in \Omega$), it still detects unfair inequality based on aver-

[18] In contrast, the "opportunity-sensitive poverty" measures proposed by Brunori et al. (2013) do not have this property. Since the EOp principle is a mere weighting factor for incomes below the poverty line, the measure does not detect any violations of the EOp principle once the FfP principle is satisfied.

age outcome differences across circumstance types. If the income distribution is perfectly opportunity-egalitarian ($\mu^e_{\mathcal{T}(\omega)} = \mu^e, \,\forall\, \omega \in \Omega$), it still requires type-specific transfers from rich to poor in order to assure the satisfaction of both FfP and EOp.

1.3.4 Alternative Conceptualizations

Our baseline measure provides one way of reconciling the principles of EOp and FfP. However, the extensive literature on the measurement of EOp shows that there are different ways of conceptualizing its underlying normative ideas (Roemer and Trannoy, 2016). In this section we discuss two alternations to the EOp concept and show how these impact the reconciliation of EOp with FfP.[19]

First, the baseline norm demands the equalization of average incomes across circumstance types. This is a *weak criterion of equality of opportunity* since it only requires the expectation of outcomes to be identically distributed across circumstance types (Lefranc et al., 2009). To the contrary, a *strong criterion of equality of opportunity* requires equality of outcomes conditional on exerting similar levels of effort. For the purpose of formulating a stronger version of the EOp requirement, we follow Roemer (1998) and identify effort tranches by the quantiles of the type-specific income distributions. Hence, i and j are part of the same effort tranche if they both sit at the q-th quantile of their respective type income distribution.[20] Compensation requires to equalize outcomes in each effort tranche, and hence to equalize all moments of the type-specific income distributions. As such, the strong conceptualization of EOp contrasts with the weak conceptualization embodied in our baseline measure since the latter required equalizing one moment of the type income distributions only. Furthermore, note that the satisfaction of strong EOp implies the satisfaction of weak EOp.

[19] In addition to varying the conceptualizations of EOp and FfP, our measurement approach allows us to introduce other normative foundations for inequality aversion. These may include affluence aversion due to concerns about political capture by income elites (Piketty, 2014) and the emergence of concentrated market structures in which massive returns accrue to an increasingly small number of "superstar" agents (Autor et al., 2020; König, 2019). While a precise formulation of these normative concerns is beyond the scope of this paper, we briefly illustrate in Supplementary Material A.2.4 how additional inequality aversion may be introduced into our framework. Furthermore, we show in Supplementary Material A.2.5 how the heterogeneity in individual needs could be integrated based on individual-specific deprivation thresholds.

[20] This "Roemerian Identification Assumption" relies on a relative conception of effort. The distribution of absolute effort like the propensity to study or to work long hours may vary across circumstance types. However, the focus on type-specific quantile distributions forces the type-specific effort distributions to be equal. Hence, the absolute effort exertion of individuals is evaluated relative to the distribution of efforts within their circumstance type.

TABLE 1.1 – Overview Alternative Conceptualizations

	Weak/Strong	Separability	Norm Distribution
Baseline	Weak	No	$y_i^r = \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{P}(\omega),\ \forall\, \omega \in \Omega \\ y_{\text{min}} + (y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}) \frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}(\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}})}, & \forall\, i \in \mathcal{R}(\omega),\ \forall\, \omega \in \Omega \end{cases}$
Alternative (a)	Strong	No	$y_i^r = \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{P}(\theta),\ \forall\, \theta \in \Theta \\ y_{\text{min}} + (\mu^e_{\mathcal{S}(\theta)} - y_{\text{min}}) \frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}(\Theta)}}{N}(\mu^e_{\mathcal{R}(\Theta)} - y_{\text{min}})}, & \forall\, i \in \mathcal{R}(\theta),\ \forall\, \theta \in \Theta \end{cases}$
Alternative (b)	Weak	Yes	$y_i^r = \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{T}(\omega) \cap \mathcal{P},\ \forall\, \omega \in \Omega \\ y_{\text{min}} + (y_i^e - y_{\text{min}}) \frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} - y_{\text{min}})}, & \forall\, i \in \mathcal{T}(\omega) \cap \mathcal{R},\ \forall\, \omega \in \Omega \end{cases}$
Alternative (c)	Strong	Yes	$y_i^r = \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{S}(\theta) \cap \mathcal{P},\ \forall\, \theta \in \Theta \\ y_{\text{min}} + (\mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}} - y_{\text{min}}) \frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\text{min}})}, & \forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R},\ \forall\, \theta \in \Theta \end{cases}$

Second, the baseline norm evaluates type-specific opportunity sets by reference to the average incomes of all $i \in \mathcal{T}(\omega)$. Moreover, the (non-)poor fraction of the population is identified by evaluating incomes in a counterfactual income distribution that corrects for unequal opportunities across circumstance types. The baseline norm thus treats EOp and FfP as *non-separable* in their scope of application: The assessment of type advantages (EOp) depends on both poor and non-poor individuals, whereas the identification of poverty (FfP) depends on the counterfactual income an individual would obtain in an opportunity-egalitarian world. In contrast to this conceptualization, one may claim that the requirements of EOp and FfP operate on separate supports of the income distribution Y^e. While FfP characterizes the normative requirement for $\mathcal{P}$, i.e. for people with incomes below y_{min}, the distributional ideal of EOp only applies to $\mathcal{R}$, i.e. to those individuals whose basic needs are satisfied. According to such an argument the normative principles of EOp and FfP are *separable* in their scope of application.

While our baseline measure adheres to *weak EOp* and *non-separability*, we can construct alternative measures by invoking either *strong EOp* or *separability*, or both. These three alternatives are presented in Table 1.1. Detailed expositions of their construction are provided in Supplementary Material A.2 and comparative statics are shown in Supplementary Material A.3.

The main features of the alternatives are as follows: First, alternatives (a) and (c) are based on strong EOp. Hence, under the assumption of non-separability (separability) the proportionality requirement for raising funds in the non-deprivation set refers to average tranche incomes in excess of the deprivation threshold, $\mu^e_{\mathcal{S}(\theta)} - y_{\text{min}}$ ($\mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}} - y_{\text{min}}$), instead of individual incomes

$y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\min}$ $(y_i^e - y_{\min})$. All else equal, one would expect the measures based on strong EOp to yield higher levels of unfair inequality. Second, alternatives (b) and (c) operate on the assumption of *separability*. Therefore, individuals are not assigned to the deprived and non-deprived fractions of society based on the counterfactual income distributions of a weakly (strongly) opportunity-egalitarian society – indicated by $\mathcal{P}(\omega)$ and $\mathcal{R}(\omega)$ ($\mathcal{P}(\theta)$ and $\mathcal{R}(\theta)$) – but based on the actual income distribution – indicated by $\mathcal{P}$ and $\mathcal{R}$. All else equal, one would expect the measures based on the separability assumption to yield lower levels of unfair inequality.

1.4 Empirical Application

To illustrate the proposed measure of unfair inequality we provide two empirical applications. First, we use the Panel Study of Income Dynamics (PSID) to analyze the development of unfair inequality in the US over the time period 1969-2014. Second, we combine the PSID and the EU Statistics on Income and Living Conditions (EU-SILC) to conduct a cross-sectional analysis in which we benchmark unfair inequality in the US against unfair inequality in 31 European countries in 2010.[21]

1.4.1 Unfair Inequality in the US over Time

Data Source. The PSID is a main resource for the study of inequality, poverty and intergenerational transmission processes in the US (see Johnson et al., 2018; Smeeding, 2018, and the overview articles in the same issue). At its inception in 1968 the PSID consisted of a nationally representative sample of 2,930 families and an oversample of 1,872 low-income families that are tracked until the present day. All individuals who leave their original households automatically become independent units in the PSID sampling frame. To match compositional changes of the US population through post-1968 immigration flows, the PSID added a Latino sample and an immigrant sample in its 1990 and 1997 waves, respectively.[22] Starting in 1997

[21] Note that much of the recent literature on inequality trends draws on administrative data sources (Burkhauser et al., 2012). However, in the context of this study survey data such as the PSID or EU-SILC provide important advantages since the operationalization of EOp and FfP requires detailed information on individual background characteristics and an accurate representation of the lower tail of the income distribution. Administrative data is restricted in both dimensions since tax returns collect only basic demographic information and because the bottom half of the distribution pays little personal income tax.

[22] We exclude the Latino sample from our investigation as it was dropped in 1995 and did not reflect the full range of post-1968 immigrants.

it has switched from an annual to a biennial survey rhythm. In its most recent waves, the PSID covers the members of more than 9,000 families and provides rich information on their incomes, family background characteristics and living practices.

In this study we focus on individuals aged 25-60 over the survey (income reference) periods 1970-2015 (1969-2014).[23] We will now briefly outline the construction of the inputs to our inequality measure: Y^e, Ω, Θ, and y_{min}. Further detail on the construction of all relevant variables as well as descriptive statistics are disclosed in Supplementary Materials A.4 and A.6.

Outcome Variable. To assess the distribution of economic resources from a fairness perspective, we use the income components created by the PSID Cross-National Equivalence File (CNEF) to define annual disposable household income as the sum of total household income from labor, asset flows, windfall gains, private transfers, public transfers, private retirement income and social security pensions. We deduct total household taxes as calculated by the NBER TAXSIM calculator (Butrica and Burkhauser, 1997).

Our measure of unfairness puts a strong emphasis on the lower end of the income distribution. It is well-known that poverty estimates based on survey data tend to be upward biased due to the under-reporting of government benefit receipts (Meyer and Mok, 2019; Mittag, 2019). Furthermore, it has been shown that households with extremely low reported incomes tend to misreport their income from earnings (Brewer et al., 2017; Meyer et al., 2019). To mitigate distortions from benefit under-reporting we use the time series provided in Meyer et al. (2015) to scale reported public transfers by a year-specific under-reporting factor that is calculated based on a comparison between the aggregate level of benefits receipts reported in the PSID and the aggregate expenditure levels from administrative program data. To cushion distortions from the under-reporting of labor incomes we identify individuals that report zero earnings but non-zero working hours in the reporting period. We then replace their reported earnings level by a prediction from a Mincer wage regression, and adjust household labor

[23] We employ cross-sectional sample weights for all calculations. However, one may worry that infrequent PSID updates for compositional changes in the US population distort comparisons over time. To address such concerns, we calculate population weights for 48 age-sex-race-cells ($8 \times 2 \times 3$) in the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC) and rescale the provided PSID individual weights to match their CPS-ASEC counterparts. This rescaling has a negligible effect on our results suggesting that the standard PSID weights do a good job in representing the underlying US population.

income by the sum of these correction values over all household members. In total only about 1% of our person-year observations are affected by this imputation procedure.

To account for differences in need and standard of living by household composition we scale all household incomes by the modified OECD equivalence scale. For the sake of inter-period and between-country comparisons we deflate all income figures with the purchasing power parity (PPP) adjustment factors for household consumption provided by the Penn World Tables (Feenstra et al., 2015). Lastly, we curb the influence of outliers by winsorizing at the 1st and the 99.5th-percentile of the year-specific income distribution.

Circumstance Types and Effort Tranches. In an equal-opportunity society there are no differences in outcomes across individuals with different circumstance characteristics but comparable levels of effort. Our measure of unfairness therefore requires to partition the population into circumstance types. Thereby a tension arises. On the one hand, the more parsimonious the type partition, the more we underestimate the influence of individual circumstances on life outcomes (Ferreira and Gignoux, 2011). On the other hand, limited degrees of freedom suggest restrictions on the granularity of the type partition to avoid noisy estimates of the relevant type parameters. In this work we use four circumstance variables to partition the population into a maximum of 36 circumstance types.[24] First, we include the biological sex of the respondent. Second, we include a binary indicator differentiating among non-Hispanic white individuals and the remaining population. Third, we construct a categorical variable based on whether the highest educated parent (i) dropped out of secondary education, (ii) attained a secondary school degree, or (iii) acquired at least some tertiary education. Lastly, we proxy the occupational status of parents by grouping them in (i) elementary occupations, (ii) semi-skilled occupations, or (iii) skilled occupations. These are standard circumstances used in the empirical literature on inequality of opportunity. However, we present sensitivity analyses based on alternative type partitions in section 1.5.2.

Replacing our baseline notion of weak EOp with strong EOp additionally requires the identification of effort tranches. To this end, we further partition each type-specific income distribution into 20 quantiles and replace individual incomes with the within-type average of their re-

[24] Brunori et al. (2020) use machine learning techniques to find the optimal type partition for the same set of European countries that are used for our second empirical application, see section 3.5.2. Their results suggest that type partitions with more than 40 types tend to overfit the data. We therefore adhere to a threshold of 36 types.

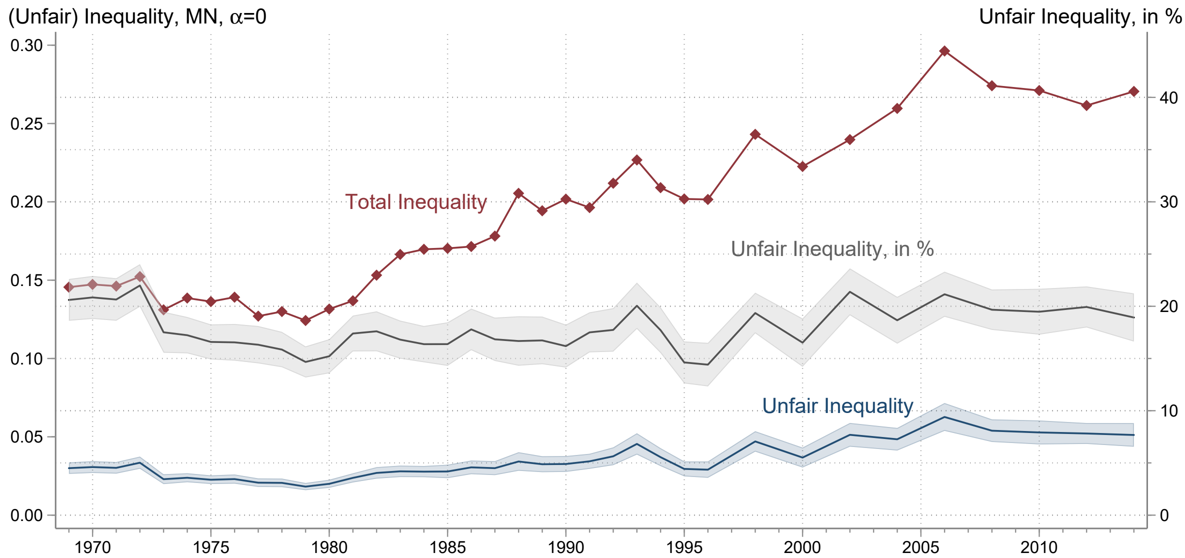
spective effort tranche. Hence, for each year we perform our calculations on a maximum population of 36×20 cells, where each cell represents a particular circumstance-effort combination. In Figure A.3 we show that this standardization of income distributions has a negligible impact on conventional inequality and poverty measures in the time period of interest.

Minimum Threshold. The specification of poverty thresholds that allow for meaningful comparisons over time and across countries is a topic of widespread academic debate. For example, the official US poverty line is based on expenditure data from the 1950s to reflect three times the cost of a well-balanced diet. Since then it has been updated only by inflation adjustments without taking account of potential changes in the needs of different family types (Meyer and Sullivan, 2012). The international poverty line of the World Bank is currently set at $1.90 per capita and day in PPP-adjusted dollars. In view of its low value it is criticized for being irrelevant in countries outside of the developing world (Allen, 2017). Lastly, both EU and OECD define relative poverty lines as a fraction of median equivalized disposable household income. Poverty measurement based on relative lines, however, may react to changes in the upper percentiles of the distribution irrespective of changes in the shortfall of those in need from what is required to make ends meet (Foster, 1998).

For our baseline estimates we rely on a revised set of international poverty lines as calculated by Jolliffe and Prydz (2016) in a two-step procedure. First, they match official national poverty headcounts to the PovcalNet expenditure data of the World Bank and calculate the implied poverty thresholds. Second, they group the resulting range of national poverty lines according to indicators of economic development and take the group median as an internationally comparable poverty line for the respective class of countries. Their procedure recovers the $1.90 line for the least developed economies but yields more relevant poverty thresholds for economically advanced countries. In our baseline estimate, we take their set of national poverty lines and group countries in quintiles of PPP-adjusted household final consumption expenditure per capita. For single households in the US, this procedure yields a PPP-adjusted poverty line of $12,874 annually that we hold constant (in real terms) over the period of our analysis. Sensitivity analyses based on alternative poverty thresholds are presented in section 1.5.3.

Baseline Results. Figure 1.1 displays the development of (unfair) inequality in the US over the time period 1969-2014. The upper line shows the development of total inequality as measured by the divergence of the empirical income distribution from a perfectly outcome egalitarian distribution in which $y_i^r = \mu^e, \forall\, i \in \mathcal{N}$. The time series replicates the well-documented pattern of inequality development in the US (among others Burkhauser et al., 2012; Heathcote et al., 2010a; Piketty et al., 2018): Slight inequality decreases throughout the 1970s are followed by strong inequality increases in the 1980s. This trend continues until the present day, most notably interrupted by the economic crises following the burst of the dot-com bubble at the turn of the century and the global financial crisis in the late 2000s.

FIGURE 1.1 – Unfair Inequality in the US, 1969-2014, Baseline Results



Data: PSID.
Note: Own calculations. This figure displays the development of (unfair) inequality in the US over the period 1969-2014. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The shaded areas show bootstrapped 95-% confidence intervals based on 500 draws.

The lower blue line displays the development of unfair inequality as measured by the divergence of the empirical income distribution from a norm distribution in which the ideals of EOp and FfP are realized to their full extent (see equation 16). It is unsurprising that unfair inequality remains at a much lower level than total inequality as the latter provides an upper bound for the former in any given country at any given point in time. However, it is noteworthy that unfair inequality seems to follow a similar time trend as total inequality. Starting with decreases of unfair inequality until 1980, we observe a steady increase of unfairness until the present day and downward movements that are largely coincidental with economic downturns.

The intermediate black line shows the share of total inequality that is in violation of EOp and FfP. It is calculated as the ratio between unfair inequality and total inequality and converted into percentage terms. Starting from a level of approximately 20% in 1969, unfair inequality drops to a share of 15% until 1980. This development suggests that the observed decreases in inequality over the 1970s were accompanied by an even stronger reduction of unfair inequality. In spite of an inequality increase by approximately 50% in the 1980s, the share of inequality attributable to violations of EOp and FfP remained roughly stable at this level until 1990. While the subsequent two decades are characterized by a more erratic pattern, we also note that unfair inequality follows a steeper growth curve than total inequality. Starting at a level of around 16% in 1990, the unfair share of inequality climbs to levels of close to 21% in the mid 2000s and stalls at a level of approximately 19% in the latest period of observation. Some may be surprised by the low relative share of unfair inequality. However, we emphasize that our measures are based on disposable household income and therefore evaluate the remaining unfairness after taking transfers through existing welfare state institutions as well as redistribution within households into account.[25]

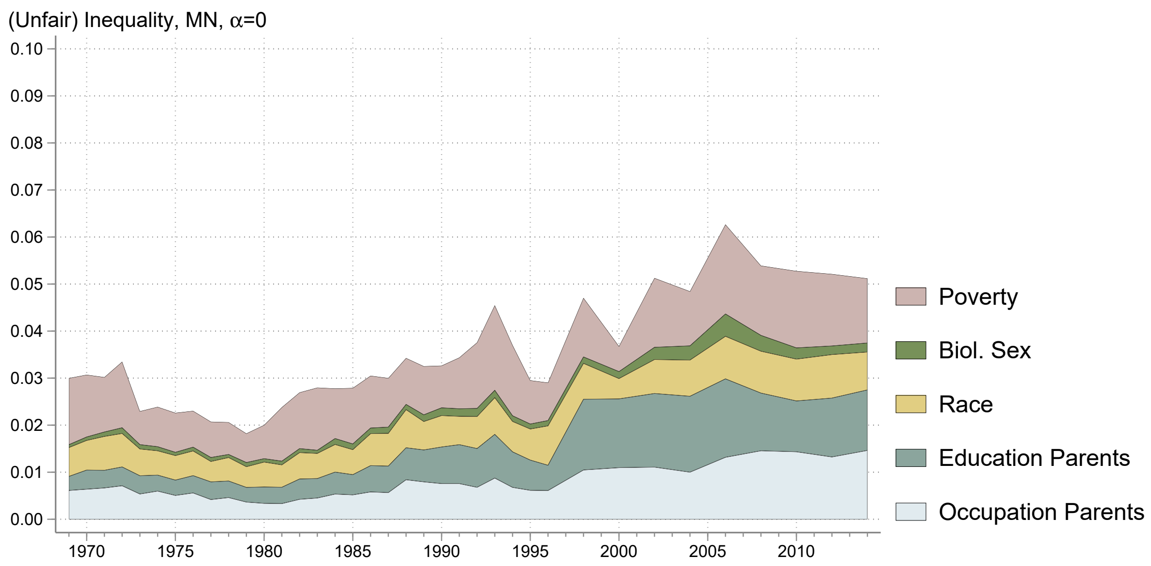
Decomposition. To develop a better understanding for the observed inequality trends, we conduct a Shapley value decomposition (Shorrocks, 2012) to identify the contributions of the different components that underpin our normative principles. That is, we quantify the contributions of FfP and EOp, respectively. Furthermore, we decompose the latter into the contributions from the circumstance characteristics biological sex, race, parental education, and parental occupation. This decomposition furthermore allows us to embed our measure of unfairness into the larger literature branches on US trends in poverty, gender income gaps, racial disparities, and social mobility.

The Shapley value procedure quantifies the contribution of each of the aforementioned factors by calculating the average marginal decline in unfair inequality once we eliminate it from our calculations. For example, one could quantify the marginal impact of FfP on unfair inequality by decreasing y_{min} from our baseline threshold of \$12,874 to \$0. Analogously, one could

[25] Moreover, it is well understood in the empirical literature that standard estimates of inequality of opportunity provide only lower bounds of their true value (Ferreira and Gignoux, 2011; Hufe et al., 2017).

quantify the marginal impact of biological sex by excluding it from the list of variables that define our type partition. However, in both steps the estimate of the marginal impact depends on the specification of the remaining normative criteria. To avoid such path-dependencies, we estimate the individual contribution of each factor by averaging their marginal impacts on unfair inequality across all possible elimination sequences (Shorrocks, 2012). The results of this decomposition are shown in Figure 1.2.

FIGURE 1.2 – Unfair Inequality in the US, 1969-2014, Decomposition



Data: PSID.
Note: Own calculations. This figure displays a decomposition of unfair inequality in the US over the period 1969-2014. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. The decomposition is based on the Shapley value procedure proposed in Shorrocks (2012).

At the end of the 1960s, approximately half of unfair inequality, that is 10% of total inequality, could be attributed to violations of the FfP principle. The previously described attenuation of relative unfairness in the 1970s can be almost exclusively attributed to decreased violations of the FfP principle. While EOp shows only a slightly decreasing trend over the 1970s, the contribution of FfP to total inequality is halved, dropping from 0.014 points (10%) to 0.007 points (5%). Following the sharp decreases of the 1970s, the contribution of FfP bounces back to its initial levels in the 1980s and subsequently follows a by and large flat time trend that persists until the present day.[26] In 2014, violations of FfP contribute 0.014 points to our measure of unfairness and explain roughly 5% of total inequality.

At first glance, our results on poverty are in line with official statistics that also show a flat time trend in poverty rates across the period of investigation (U.S. Bureau of the Census, 2019). However, the official poverty concept in the US differs from ours in important aspects such that this analogy only holds superficially. Official poverty statistics rely on the poverty headcount ratio applied to an annually adjusted poverty line that is based on the pre-government income of families. To the contrary, we apply a time-constant absolute poverty threshold to disposable household income after taxes and transfers and measure poverty as a linear combination of the poverty gap ratio and the Watts index (Section 3.2). In fact, applying the headcount ratio to our income concept and the time-constant poverty line, we find that the share of poor individuals drops by more than 40% over time (Figure A.7 and Table A.4).[27] However, while the share of poor households has constantly decreased over time the intensity of poverty as measured by the poverty gap ratio and the Watts index has first decreased in the 1970s and then rebounded since the mid-1990s. As a consequence, we also find a relatively constant poverty trend over time, but for different reasons than the official US government statistics.

The stable poverty trend, however, is superseded by marked increases in the violations of EOp. After slight decreases in the 1970s, the EOp contribution to total inequality increases from 10% in 1980, over 12% (14%) in 1990 (2000) to 14% in the latest period of observation.

Analyzing the EOp component in further detail, we note that the contribution of biological sex to overall inequality is negligible and hovers around the 1%-mark in relative terms. Hence, our measure does not reflect the well-documented decrease in earnings differences between males and females (Blau and Kahn, 2017). This deviation is not unexpected and follows from our focus on disposable household income. Accounting for resource sharing at the household level evens out any intra-household inequality among males and females. As such, all our results on biological sex are driven by single-headed households. Within this group the flat time trend in the contribution of sex-based differences to total inequality can be rationalized by two countervailing forces that are displayed in Figure A.8. First, income differences among

[26] Note that while the absolute contribution of FfP is rather stable between 1969 and 2014, its relative contribution is halved from 10% to 5%. However, this decrease in the relative contribution follows mechanically from the increase in total inequality. For further illustration, see also Figure A.6 in which we fit locally smoothed time trends for the relative contributions of both EOp and FfP.

[27] See Wimer et al. (2016) for similar results.

male and female-headed single households have been decreasing over the time period 1969-2014. Second, the prevalence of single-headed households has been steadily increasing for both males and females. While the first trend depresses the contribution of sex-based differences to total inequality, the second trend magnifies the remaining differences leading to relatively time-constant contributions of this component to unfairness in the US.[28]

In analogy to biological sex, the contribution of race to unfairness in the US is largely stagnant at approximately 0.007 points across the time period of observation. In relative terms the contribution of race slightly decreases from 4% to 3%, again reflecting the marked increase of total inequality. This flat trend echoes previous findings that there has been little progress in closing the black-white earnings gap since the 1970s (Bayer and Charles, 2018; Derenoncourt and Montialoux, 2019).[29]

With the contributions from sex- and race-based differences rather constant over time, the witnessed increase of the EOp component is entirely driven by the increased importance of parental background variables – namely parental education and occupation. While these factors jointly contributed 0.009 points (6%) in 1969, their importance has tripled to 0.028 points (10%) in 2014. Interpreting the covariances between parental education and occupation and individual income as a proxy for social mobility, our findings suggest that the US has become increasingly immobile in the time period from 1969 to 2014. This finding is in line with Aaronson and Mazumder (2008) and Davis and Mazumder (2019) who find that the intergenerational elasticity of income has declined for cohorts entering the labor market after 1980 as well as Hilger (2019) who documents a similar time trend for educational mobility. However, we note that the assessment of intergenerational mobility trends in the US is contentious. In contrast to the previously cited works, Chetty et al. (2014c), Lee and Solon (2009), and Song et al. (2019) conclude that intergenerational mobility has stayed constant over the time period of investigation. The disparity of results is explained by various drawbacks of the underlying data sources as well as different measurement choices. While our measurement approach is not strictly comparable to either of these papers, our results are in line with the first set of works.[30]

[28] See also S. Lundberg et al. (2016) on the interaction between changing gender gaps, family structures and the intergenerational transmission of advantages.

[29] See also Figure A.9 for complementary evidence on the stability of non-white disposable income gaps in our data.

[30] Mobility measures can be decomposed into i) the copula of parental background characteristics and child outcomes, and ii) the marginal distributions of child outcomes and parental background characteristics, respectively (Chetty et al., 2014c). Rank-mobility measures such as intergenerational correlations (IGC) and rank-rank

To summarize: In terms of its trend, unfair inequality largely replicates the development of total inequality in the US. However, due to marked decreases in poverty, unfairness showed an even stronger decrease than total inequality in the 1970s. To the contrary, the steeper growth of unfair inequality since the 1990s is almost exclusively attributable to increased violations of the EOp principle and the growing importance of parental background variables in particular.

1.4.2 Cross-Country Differences in Unfair Inequality

Data Source. For the purpose of an international comparison we combine the PSID with the 2011 wave of EU-SILC. EU-SILC serves as the official database for monitoring inequality, poverty and social exclusion in the EU (see for example Atkinson et al. (2017) and the references cited therein) and covers a total of 31 countries.[31] We use the 2011 cross-sectional wave as it contains a special survey module on parental background information that allows us to construct types from a broad range of circumstance variables.[32] As in the PSID, incomes are reported for the year preceding the survey leading to 2010 as the year of our cross-sectional comparison. The data preparation closely follows the procedures outlined for the PSID. Further detail on the variable construction as well as descriptive statistics are provided in Supplementary Materials A.4 and A.6.

correlations depend on i) while holding ii) constant. To the contrary, mobility measures like the intergenerational elasticity (IGE) allow for changes in ii). Clearly, our measurement approach is closer to the second class as we compare different marginal distributions in the parent and the child generation that we allow to change over time. However, our measure differs from a typical IGE estimate in at least three important dimensions. i) We model child income as a function of parental education and occupation instead of parental income. ii) We summarize persistence by calculating inequality in a predicted distribution instead of interpreting regression parameters. iii) Child outcomes refer to annual incomes at various points of the life-cycle instead of modeling them so as to mimic lifetime income (Nybom and Stuhler, 2016). To provide a closer analogy to standard IGE estimates we re-estimate our measure of unfairness for different age groups at different points in time while excluding all determinants of unfairness except for parental background characteristics. The results, displayed in Figure A.10, suggest that relative mobility has decreased at all points of the individual life-cycle with more pronounced changes at older ages. This pattern is consistent with earnings profiles that fan out over the life-cycle.

[31] The sample consists of Austria (AT), Belgium (BE), Bulgaria (BG), Switzerland (CH), Cyprus (CY), Czech Republic (CZ), Germany (DE), Denmark (DK), Estonia (EE), Greece (EL), Spain (ES), Finland (FI), France (FR), Croatia (HR), Hungary (HU), Ireland (IE), Iceland (IS), Italy (IT), Malta (MT), Lithuania (LT), Luxembourg (LU), Latvia (LV), Netherlands (NL), Norway (NO), Poland (PL), Portugal (PT), Romania (RO), Sweden (SE), Slovenia (SI), Slovakia (SK), and the United Kingdom (UK).

[32] In contrast to the PSID, EU-SILC consists of rotating panels and each household stays in the data for only 4 years. Hence, one cannot use the panel dimension to construct circumstance variables.

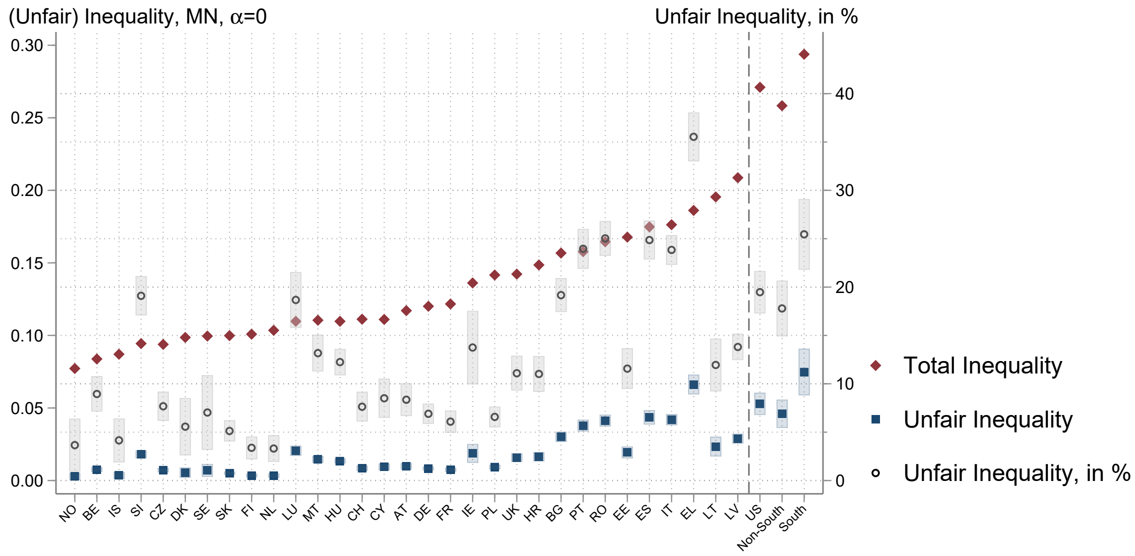
Outcome Variable. We construct disposable household income as the sum of total household income from labor, asset flows, private transfers, public transfers, private retirement income and social security pensions, and deduct taxes on wealth (if applicable), income and social security contributions. In analogy to the PSID, we scale reported public transfers by a country-specific under-reporting factor and adjust labor incomes by imputing individual labor incomes of respondents with zero labor incomes but non-zero working hours. Only about 1% of respondents are affected by the latter imputation. Furthermore, we deflate household incomes by the modified OECD equivalence scale, adjust for purchasing power parities and winsorize country-specific income distributions at the 1st and 99.5th percentiles.

Circumstance Types and Effort Tranches. For each country we partition the population based on the following circumstance characteristics: i) biological sex, ii) migration background, iii) educational achievement of the highest educated parent, and iv) the highest occupation category of either parent. While circumstances i), iii), and iv) mirror the PSID specification, we replace the binary race variable of the PSID with a binary indicator for whether respondents were born in their current country of residence. In total we partition the population into 36 circumstance types which we again subdivide into 20 quantiles to identify effort tranches. As evidenced in Figure A.3 this transformation is innocuous with respect to cross-country comparisons of inequality and poverty statistics.

Minimum Threshold. Internationally comparable absolute poverty thresholds are again constructed based on the procedure suggested by Jolliffe and Prydz (2016). 21 out of the 31 European countries belong to the highest quintile of countries in terms of PPP-adjusted household final consumption expenditures per capita and are hence characterized by the same poverty threshold as the US: $12,874 per annum (PPP-adj.). However, 10 Eastern European countries only belong to the second highest quintile and are therefore characterized by a lower poverty threshold of $3,957 per annum (PPP-adj.).

Baseline Results. Figure 1.3 replicates Figure 1.1 for the cross-country comparison. The red diamonds indicate total inequality, the blue squares unfair inequality. The black hollow circles show the relative share of unfair inequality. Countries are ordered from left to right by their level of total inequality. The dashed vertical line separates the European countries from the US sample. Acknowledging the special role of the Southern states in terms of intergenerational transmission processes (Bratberg et al., 2017; Chetty et al., 2014b) and poverty prevalence (Ziliak, 2006), we also provide results separating the South of the US from the rest of the country (Northeast, Midwest, West) based on the census region groupings of the US Census Bureau.

FIGURE 1.3 – Unfair Inequality across Countries, 2010, Baseline Results

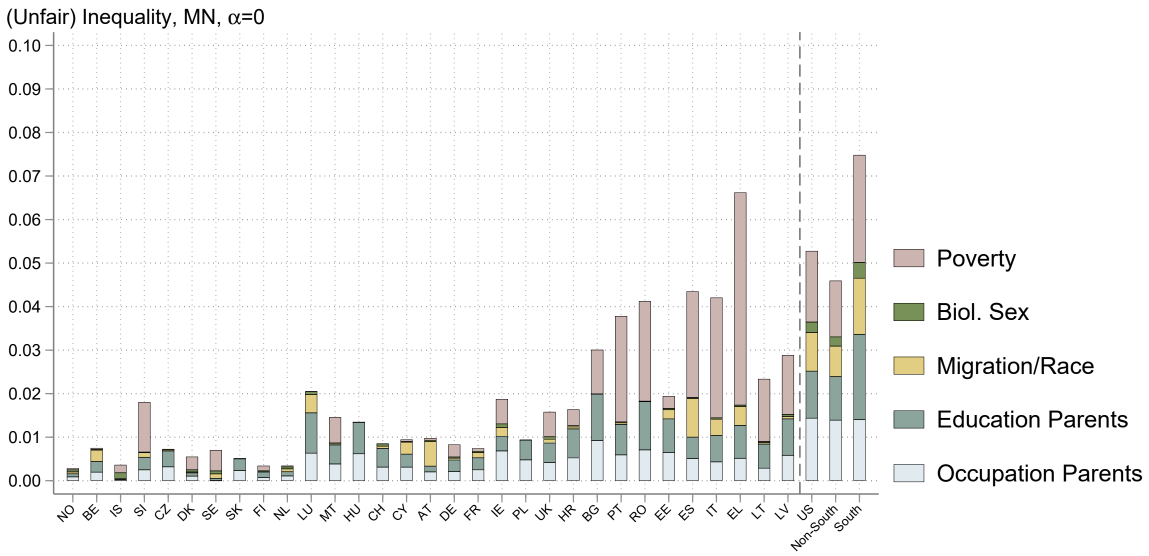


Data: PSID and EU-SILC.
Note: Own calculations. This figure displays cross-country differences in (unfair) inequality in 2010. Data points to the left of the vertical dashed line refer to the European country sample. Data points to the right of the vertical dashed line refer to the US and its census regions. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The shaded areas show bootstrapped 95-% confidence intervals based on 500 draws.

The US are by far the most unequal society in our country sample with inequality figures about 25% higher than the most unequal European societies. At the other end of the spectrum we find Norway, Iceland and Belgium. The most unfair societies in 2010 are Greece, the US, Spain, Italy, and Romania closely followed by Portugal. Treating the South of the US as a separate country, it would attain the highest level of unfairness of all countries. In relative terms, EOp and FfP explain roughly 25% of total inequality in the European countries of this group – even 35% in Greece. The US attains an unfair share of approximately 19%. The lower unfairness share of the US follows mechanically from its higher levels of total inequality. The group of countries with the least extent of unfair inequality consists of Scandinavian countries plus the Netherlands. It is important to emphasize that country rankings differ depending on whether we analyze total inequality or unfairness. While for example Belgium is among the top three countries of least total inequality, it is not among the top ten countries of least unfair inequality.

Decomposition. The US differs markedly from its European counterparts in terms the processes that determine unfair inequality. Figure 1.4 shows the results of a Shapley value decomposition of unfair inequality into its different components.

FIGURE 1.4 – Unfair Inequality across Countries, 2010, Decomposition



Data: PSID and EU-SILC.
Note: Own calculations. This figure displays a decomposition of cross-country differences in unfair inequality in 2010. Data points to the left of the vertical dashed line refer to the European country sample. Data points to the right of the vertical dashed line refer to the US and its census regions. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. The decomposition is based on the Shapley value procedure proposed in Shorrocks (2012).

In the European group of countries with the highest unfairness (Greece, Portugal, Romania, Spain, Italy), violations of the FfP principle consistently explain more than half of the detected unfair inequality. 2010 marks a peak year of the European sovereign debt crisis, and Greece, Portugal, Spain and Italy were among the countries most affected by it. To highlight the differential impact of the economic crisis on unfairness in Europe and the US, we calculate the difference between the Watts index and the poverty gap ratio for the six most unfair societies in our country sample (Greece, US, Spain, Italy, Romania, Portugal) from 2006 to 2014. Since the FfP component nests the difference between these two poverty measures, it can be interpreted as a proxy statistic for the longitudinal development of FfP in these countries. The results are displayed in Figure A.11. Romania is the least economically developed country in the considered country group. In Romania the financial crisis ended a trend of decreasing poverty and led to increased violations of the FfP principle in its aftermath. Similarly, in the group of Southern European countries the FfP proxy increases markedly after 2008. This evidence suggests that the high levels of unfair inequality among the European countries in 2010 followed from the economic downturn that accompanied the financial crisis and which in turn led to increased violations of the FfP principle.

In contrast to the European group, the difference between Watts index and poverty gap ratio is completely flat in the US over the crisis years. Instead, Figure 1.4 shows that unfairness in the US is strongly driven by the EOp component. This difference is not caused by the differential importance of biological sex. Due to our focus on disposable household income, income differences across the sexes have a negligible impact on unfair inequality in Europe and the US alike. Neither is this difference a mere consequence of replacing the race indicator with the immigration background indicator. Even abstracting from the migration/race circumstance, the US would be characterized by the highest degree of unequal opportunities in our country sample. It is the contributions of parental education and occupation that are the highest among all countries under consideration and place the US among the most unfair societies in our country sample. In line with the findings of Chetty et al. (2014b) and Hilger (2019) the lack of social mobility is particularly pronounced in the Southern states of the US. However, even when focusing on the non-Southern states only, the US ranks among the countries with the highest intergenerational persistence in our country sample (see also Corak, 2013).

1.5 Sensitivity Analysis

In this section, we investigate the sensitivity of our baseline results to alternative normative assumptions. First, we provide empirical results for all three alternative conceptualizations laid out in Table 1.1.

Second, in principle the measurement approach adopted in this paper takes a neutral stance on the specification of the model primitives Ω, Θ, and y_{min}. Hence, it may accommodate a wide array of different views on the responsibility cut as well as the appropriate minimum threshold y_{min}. Yet, we acknowledge that the precise cut between circumstances and effort, as well as the choice of y_{min} are normatively contentious. While it is not the ambition of this paper to resolve such disagreement, we provide results for alternative choices of Ω, Θ, and y_{min} in sections 1.5.2 and 1.5.3, respectively.

Third, differences between Y^r and Y^e may be aggregated by different divergence measures that put different weights on positive and negative divergences from norm incomes, respectively. We therefore provide robustness analyses with respect to the use of different divergence measures in section 1.5.4.

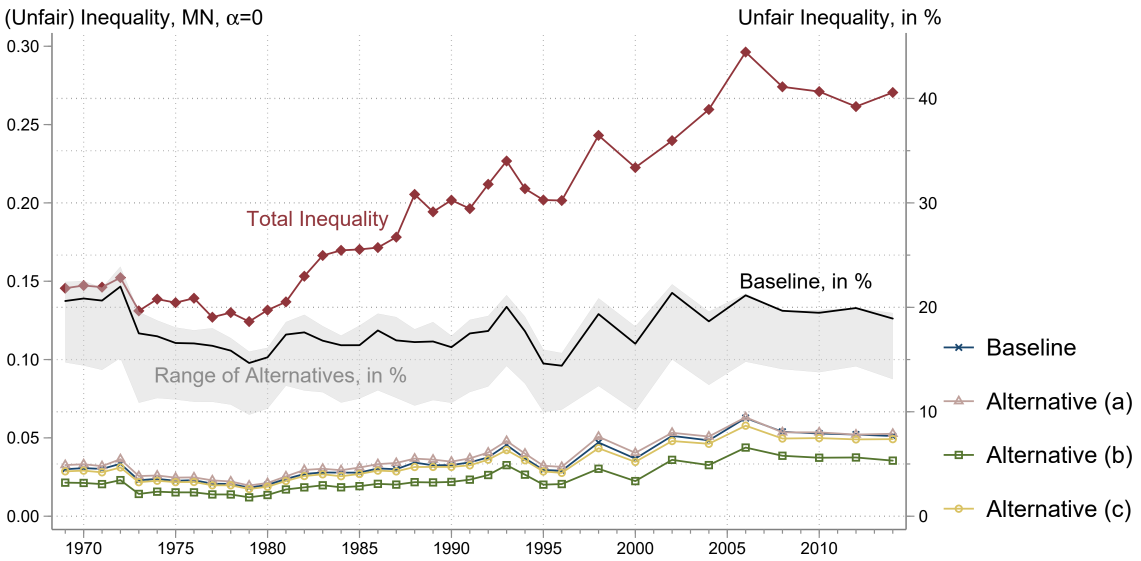
For brevity, we only present robustness checks for the longitudinal analysis of the US in the main body of this paper. However, every sensitivity check is conducted in an analogous way for the cross-country comparison – see Figures A.12-A.15 and Table A.6 in the Supplementary Material.

1.5.1 Alternative Norm Distributions

Our baseline estimates of unfair inequality rely on a measure that is based on a weak conceptualization of EOp and reconciles EOp and FfP in a non-separable way. In Table 1.1 we have presented alternative norm distributions that divert from the baseline by operating on a strong notion of EOp (Alternatives (a) and (c)) and/or assume separability between EOp and FfP (Alternatives (b) and (c)). Figure 1.5 presents the development of (unfair) inequality in the US with the upper line again marking the development of total inequality and the lower lines marking unfair inequality under each of these different conceptualizations. The black line marks the relative share of unfair inequality from our baseline estimate. The gray area shows the range between the lower and the upper envelope of the relative share of unfairness according to the alternative measurement specifications.

We note that our conclusions with respect to the time trend of unfair inequality in the US is robust to the different conceptualizations: A decrease in the relative share of unfair inequality until 1980 is followed by a stagnation throughout the following decade and increases throughout the 1990s until the present day. However, level differences exist. While Alternatives (a) and (c) yield results that are largely congruent to our baseline, Alternative (b) consistently detects lower levels of unfair inequality than the remaining measures. This result directly follows from the separability assumption according to which (i) opportunity sets of circumstance types are evaluated by excess incomes above y_{min} only, and ii) excluding empirically poor individuals

FIGURE 1.5 – Unfair Inequality in the US, 1969-2014, Alternative Norm Distributions



Data: PSID.
Note: Own calculations. This figure displays the development of (unfair) inequality in the US over the period 1969-2014 according to the alternative norm distributions outlined in Table 1.1. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The gray area shows the range of unfair inequality in percent (in %) of total inequality depending on the alternative measurement specifications.

from compensation through opportunity-equalizing transfers beyond the poverty line. Both features make the distribution of type-specific advantages more homogeneous and therefore require less transfers across types to attain the normatively desirable distribution of incomes. If one prefers the conceptualization of Alternative (b) over our baseline measure, one would conclude that unfairness amounts to 13% instead of 19% of total inequality in 2014.

1.5.2 Alternative Responsibility Cuts

Any measurement of a responsibility-sensitive version of egalitarianism requires a stance on the features of life for which people should be held responsible. In our baseline estimates we assume that people should not be held responsible for i) their biological sex, ii) their race, iii) the occupation of their parents, and iv) the education of their parents. However, there may be further characteristics beyond individual control that evoke normative concern. Examples could be the quality of neighborhoods in which people grew up (Chetty et al., 2016a), parenting practices (Doepke et al., 2019) or even genetic endowments (Papageorge and Thom, 2020).

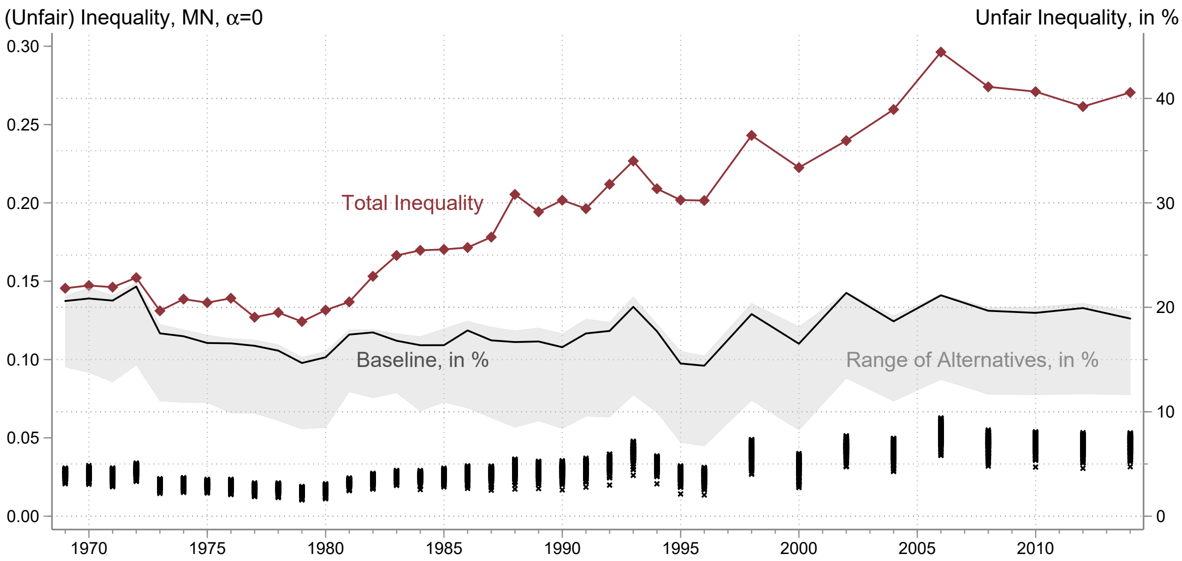
To be sure, the PSID puts strong constraints on testing the influence of different circumstance characteristics.[33] We therefore proceed as follows: First, we extract two additional circumstances that are consistently measured across the period of our analysis: i) the census region in which respondents grew up, and ii) the migration background of parents. We convert both variables into a vector of binary indicators and add them to our set of circumstances. Second, we repeat our analysis for all circumstance combinations that yield the same number of types as in our baseline analysis (36 types).[34] Hence, we repeat our analysis for 210 different specifications of Ω. The results are presented in Figure 1.6, where each black cross represents a different specification of Ω in any given year. The black line again marks the relative share of unfair inequality from our baseline estimate while the gray area shows the range between the lower and the upper envelope of the relative share of unfairness according to the alternative measurement specifications.

Our conclusions with respect to the time trend of unfair inequality in the US remains unaffected by the specification of Ω. However, we again register level differences depending on the factors for which we hold people responsible. According to the most conservative specification of Ω, unfair inequality in the US amounts to roughly 12% of total inequality in 2014 (with the upper range being 20%). We acknowledge that the alternative circumstance information in the PSID remains limited to geographical and migration background information. EU-SILC avails a broader range of circumstance characteristics from different domains that are consistently elicited across all sample countries. These include i) the relationship status of parents, ii) the number of siblings, iii) the financial situation of the parental household, as well as iv) property ownership of parents. We again test 210 different specifications of Ω for the EU-SILC countries holding the maximum number of types constant at 36. However, Figure A.13 reveals that in spite of level differences the general conclusions from our cross-country comparison remain robust to this broader set of alternative circumstance characteristics.

[33] The PSID has introduced the Child and Development Supplement (CDS) in 1997 with follow-up waves in 2002/03 and 2007/08. The CDS provides very detailed information on the living environments of 3,563 children aged 0-12 in the initial wave. However, even the oldest children from the 1997 CDS cohort are only now in their early 30s – an age that is commonly believed to be the minimum threshold to approximate long-term earnings potential. Respecting sensible age thresholds and due to sample attrition over time, the CDS sample is too small to exploit its richer circumstance information for the income decompositions that underlie our empirical analysis – see also our discussion in section 1.4.1.

[34] We keep the granularity of the type partition constant to ensure the comparability to our baseline results and to balance the concerns for underestimating the influence of circumstances and noisy estimates of the relevant type parameters – see also our discussion in section 1.4.1.

FIGURE 1.6 – Unfair Inequality in the US, 1969-2014, Alternative Circumstance Sets

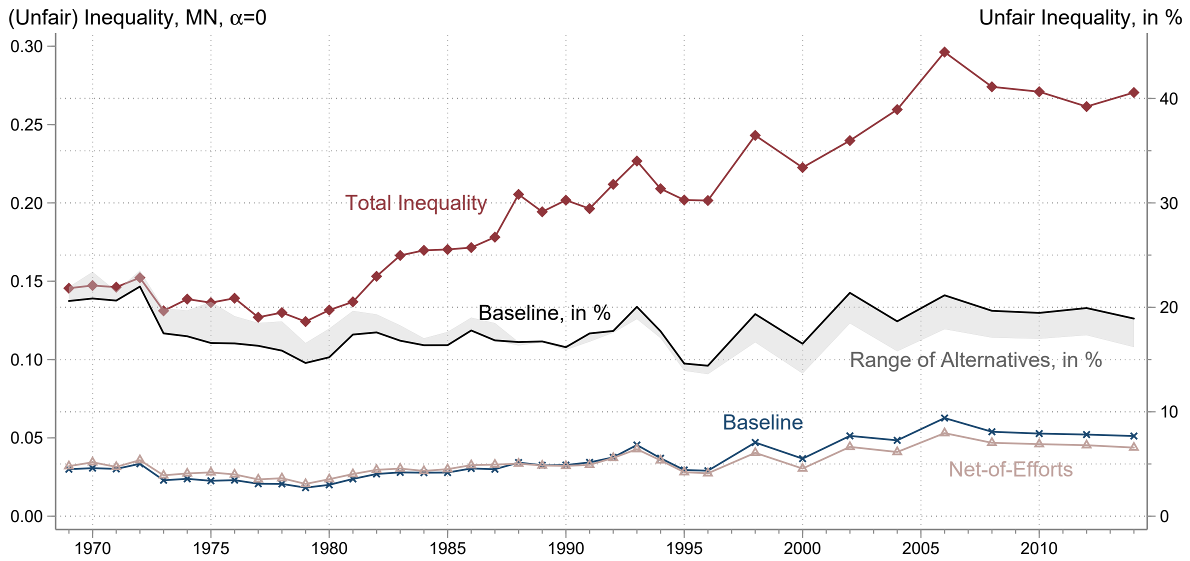


Data: PSID.
Note: Own calculations. This figure displays the development of (unfair) inequality in the US over the period 1969-2014 according to alternative specifications of the circumstance set Ω. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The gray area shows the range of unfair inequality in percent (in %) of total inequality depending on the alternative measurement specifications.

Another normative assumption relates to the correlation between circumstances Ω and efforts Θ. In our baseline measure we treat the correlation between both components as morally objectionable. For example, part of the income gap between whites and non-whites can be explained by differences in educational attainment (Gelbach, 2016) which itself is at least partially under the control of individuals. Circumstances thus exert a direct and an indirect effect on life outcomes. While in our baseline we follow Roemer (1998) and consider both effects as normatively objectionable, others have suggested to hold people responsible for effort and preference variables regardless of how they are formed (Barry, 2005). To test the sensitivity of our baseline results to this alternative normative stance, we repeat our analysis while partialling out the indirect effect that circumstances exert through individual efforts. To this end, we consider two variables that are partially under the control of individuals and highly predictive of incomes – i) educational attainment, and ii) annual working hours – and clean circumstances from their correlation with these effort variables before repeating our analysis.[35] If circumstances had no impact independent of the considered efforts, we would see a sharp drop of unfair inequality in comparison to our baseline results.

[35] We describe the exact steps of this procedure in Supplementary Material A.7.

Figure 1.7 – Unfair Inequality in the US, 1969-2014, Accounting for Preferences



Data: PSID.
Note: Own calculations. This figure displays the development of (unfair) inequality in the US over the period 1969-2014 according to alternative treatments of the correlation between the effort set Θ and the circumstance set Ω. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The gray area shows the range of unfair inequality in percent (in %) of total inequality depending on the alternative measurement specifications.

Figure 1.7 shows the differences between our baseline and the alternative responsibility cut. We note a moderation of the previously described time trend when holding people responsible for the correlation between circumstances Ω and efforts Θ. In contrast to our baseline, unfair inequality starts at higher levels in 1969 and increases much more moderately in the 1990s. Combining this moderation of the time trend in absolute unfair inequality with the increasing slope of total inequality, the relative share of unfairness decreases over time and remains slightly above the 15%-mark in 2014. The differential development of our baseline and the alternative measure is consistent with evidence on increasing college wage premia (Heathcote et al., 2010b), longer working hours among the highly educated (Fuentes and Leamer, 2019) and the increasing stratification of college completion by parental background characteristics (Davis and Mazumder, 2019; Hilger, 2019). Once we shut down educational attainment and working hours as channels of circumstance influence, unfairness does no longer reflect the growing importance of these factors for the determination of incomes over time.

1.5.3 Alternative Minimum Thresholds

There is no clear consensus on how to set an income threshold that captures the material requirements of what it takes to make ends meet. Acknowledging the arbitrariness of any threshold, Foster (1998) suggests to move beyond normative and empirical disagreements on the correct value of y_{min} and to show the robustness of the main conclusions based on different plausible specifications of y_{min} instead. In this spirit we provide alternative measures of unfair inequality based on four different poverty lines. First, Allen (2017) uses a linear programming approach to calculate the PPP-adjusted minimal cost of a basic needs consumption basket containing food to satisfy nutritional requirements, as well as fuel for heating, clothing and shelter for different climatic regions of the world. For the four countries overlapping with our sample (US, Lithuania, UK, France) he calculates an average basic needs poverty (BNP) line of \$3.96 (PPP-adj.) per capita and day which we apply to all countries and years in our sample. Second, we repeat our analysis by using the official country-year-specific national poverty lines of the US Census Bureau and EUROSTAT. Third, we calculate relative poverty lines based on the suggestions of the OECD and EUROSTAT. While the OECD proposes a poverty line at 50% of the median equivalized disposable household income, EUROSTAT proposes an at-risk-of-poverty (AROP) line at 60% of the median of the same distribution.[36] The results for these different poverty thresholds are shown in Figure 1.8.

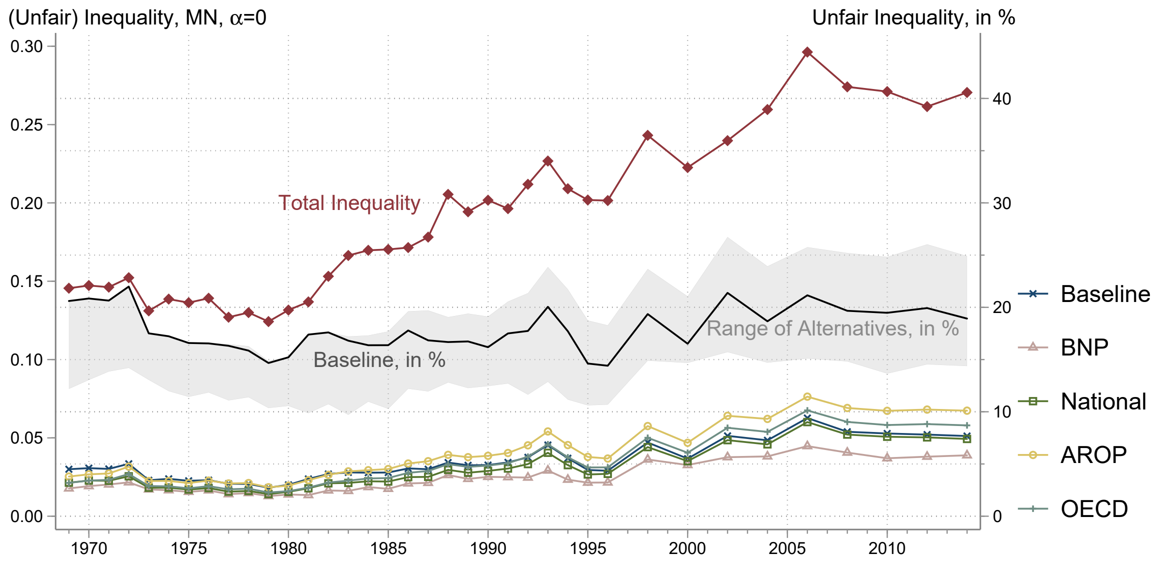
We note that our general conclusions with respect to the trend of unfairness in the US are insensitive to the specification of the poverty threshold. If anything, the relative poverty thresholds of the OECD and AROP tend to magnify the relative increase of unfairness since the 1990s. However, unsurprisingly we observe sharp level differences in unfair inequality depending on the stringency of the poverty threshold. Proponents of the AROP threshold (\$18,737) would conclude that unfairness explained 25% of total inequality in the US in 2014, while proponents of the BNP (\$1,445) threshold would detect a relative share of 14%.

1.5.4 Alternative Divergence Measures

Our baseline measure of unfair inequality employs the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$. In addition to alternations in the weighting parameter α, we now present results based on the measures put forward by Cowell (1985) and Almås et al. (2011). The family put forward by Cowell (1985) is another generalization of the entropy class of inequality indexes that varies with an inequality aversion parameter α. The Cowell-family and the MN-family coincide exactly for $\alpha = 1$. Moreover, we employ the unfairness Gini

[36] Note that the official poverty statistics of EUROSTAT are also calculated by reference to the AROP threshold. The AROP lines presented in this work differ nevertheless from the national poverty lines provided by EUROSTAT since we calculate them by observing the sample restrictions and variable definitions used in this paper.

FIGURE 1.8 – Unfair Inequality in the US, 1969-2014, Alternative Minimum Thresholds



Data: PSID.
Note: Own calculations. This figure displays the development of (unfair) inequality in the US over the period 1969-2014 according to alternative specifications of the poverty threshold $y_{\min}$. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The gray area shows the range of unfair inequality in percent (in %) of total inequality depending on the alternative measurement specifications. The construction of the alternative minimum thresholds is discussed in Supplementary Material A.4.

proposed by Almås et al. (2011) which tends to put relatively less weight on large negative divergences from the reference distribution.

In spite of their differences, all measures yield highly comparable results in terms of cross-period comparisons of unfair inequality. Table 1.2 shows rank-correlations for the different measures and their parameterizations for the US sample. All correlation coefficients are at a level of at least 0.96. Hence, we conclude that our results are robust to alternations in the way in which divergences between Y^e and Y^r are aggregated.

1.6 Conclusion

In this paper we have provided a new measure of unfair inequality that reconciles the ideals of equality of opportunity (EOp) and freedom from poverty (FfP). In fact, we provide the first work that combines these widely-endorsed principles of justice into a joint measure of unfair inequality by treating both as co-equal grounds for compensation.

Next to illustrating our measurement approach and showcasing its flexibility to various normative alternations, we provide two empirical applications. First, we analyze the development

Table 1.2 – Rank Correlation across Years, US

	Magdalou and Nock			Cowell			Almås et al.
	$\alpha = 0$	$\alpha = 1$	$\alpha = 2$	$\alpha = 0$	$\alpha = 1$	$\alpha = 2$	
Magdalou and Nock							
$\alpha = 0$	1.00	.	.	.	.	.	.
$\alpha = 1$	0.99	1.00	.	.	.	.	.
$\alpha = 2$	0.97	0.99	1.00	.	.	.	.
Cowell							
$\alpha = 0$	0.99	1.00	0.98	1.00	.	.	.
$\alpha = 1$	0.99	1.00	0.99	1.00	1.00	.	.
$\alpha = 2$	0.98	1.00	0.99	1.00	1.00	1.00	.
Almås et al.							
	0.96	0.98	1.00	0.98	0.98	0.98	1.00

Data: PSID.
Note: Own calculations. This table displays rank correlations for unfair inequality across years based on different divergence measures. Unfair inequality is calculated based on the divergence measures proposed by Magdalou and Nock (2011), Cowell (1985), and Almås et al. (2011).

of inequality in the US over the time period 1969-2014 from the normative perspective of our unfairness measure. Second, we provide a corresponding international comparison between the US and 31 European countries in 2010. In combination, both analyses yield important implications for current debates on inequality. First, the US trend in unfair inequality has largely traced the marked increase of total inequality since the beginning of the 1980s. Second, since the 1990s unfair inequality follows a steeper growth curve than total inequality. Third, this trend is mainly driven by a less equal distribution of opportunities across people that face different circumstances beyond their individual control. Fourth, unfairness in the US shows a remarkably different structure than in comparable European societies. While unfairness in Europe in 2010 seems to be largely driven by the consequences of European debt crisis, unfairness in the US is driven by the intergenerational transmission of disadvantages. The underlying determinants of the latter are arguably much more persistent than income shortfalls due to economic downturns which illustrates the enormous challenge presented to policymakers willing to address unfairness in the US.

While we provide comprehensive robustness checks for our findings, there are shortcomings which suggest a wide avenue for further research. At the empirical level, it includes addressing the well-known drawbacks of survey data by the use of suitable administrative datasets.

Furthermore, we have shown in this work that our measurement approach lends itself to various refinements and extensions with respect to the conceptualization of unfairness. While we were careful to choose our guiding principles to broadly match the fairness perceptions of a larger public, we look forward to tailor our approach even stronger to forthcoming empirical evidence on the normative preferences upheld by individuals.

Appendix A.1 Proof

Proof of Proposition 1. The proposition is trivially true for the (counterfactually) poor population $\mathcal{P}(\omega)$ as their norm incomes are prescribed by the FfP condition (13). Furthermore, for each type $\mathcal{T}(\omega)$ we can use (14) to rewrite y_i^r for the non-poor population $\mathcal{R}(\omega)$:

$$y_i^r = y_{\text{min}} + \frac{y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}(y_j^r - y_{\text{min}}). \tag{25}$$

We use this expression together with the FfP condition ($y_i^r = y_{\text{min}},\ \forall\, i \in \mathcal{P}(\omega)$) in the EOp condition (9):

$$\underbrace{\frac{1}{N_{\mathcal{T}(\omega)}}\left[\sum_{i\in\mathcal{P}(\omega)} y_{\text{min}} + \sum_{i\in\mathcal{R}(\omega)}\left(y_{\text{min}} + \frac{y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}(y_j^r - y_{\text{min}})\right)\right]}_{=\mu^r_{\mathcal{T}(\omega)}} = \mu^e. \tag{26}$$

We simplify (26) as follows:

$$\begin{aligned}
y_{\text{min}} + \frac{1}{N_{\mathcal{T}(\omega)}}\sum_{i\in\mathcal{R}(\omega)} \frac{y_j^r - y_{\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}(y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}) &= \mu^e \\
y_{\text{min}} + \frac{y_j^r - y_{\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}\frac{1}{N_{\mathcal{T}(\omega)}}\sum_{i\in\mathcal{R}(\omega)} (y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}) &= \mu^e \\
y_{\text{min}} + \frac{y_j^r - y_{\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}}\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}(\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}}) &= \mu^e.
\end{aligned}$$

We solve for y_j^r to obtain the norm income of any $j \in \mathcal{R}(\omega)$:

$$y_j^r = y_{\text{min}} + (y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}})\frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}(\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}})}. \tag{27}$$

As evidenced by (26), $\mu^r_{\mathcal{T}(\omega)}$ is a continuous and monotonically increasing function of y_j^r and we know that $y_j^r \in (y_{\text{min}}, \infty)$. It is straightforward that $\mu^r_{\mathcal{T}(\omega)} \to y_{\text{min}}$ for $y_j^r \to y_{\text{min}}$ and $\mu^r_{\mathcal{T}(\omega)} \to \infty$ for $y_j^r \to \infty$. Under the assumption that $\mu^e > y_{\text{min}}$ and invoking the intermediate value theorem, it follows that there is a unique value of y_j^r for which (26) holds. Since the choice of $i, j \in \mathcal{R}(\omega)$ was arbitrary, expressions (27) and (15) hold for all $i, j \in \mathcal{R}(\omega),\ \forall\, \omega \in \Omega$.

However, such a unique value only exists if $\mu^e > y_{\text{min}}$. Assume this was not true, i.e. $\mu^e \leq y_{\text{min}}$. Then, it would still hold that $\mu^r_{\mathcal{T}(\omega)} \to y_{\text{min}}$ for $y^r_j \to y_{\text{min}}$ and $\mu^r_{\mathcal{T}(\omega)} \to \infty$ for $y^r_j \to \infty$. Hence, $\mu^r_{\mathcal{T}(\omega)} \in (y_{\text{min}}, \infty)$. However, from the EOp requirement (9) we also know that $\mu^r_{\mathcal{T}(\omega)} = \mu^e$. If $\mu^e \leq y_{\text{min}}$, either of these statements must be false and hence $\cap^4_{h=1}\mathcal{D}^h = \emptyset$. Intuitively, if $\mu^e \leq y_{\text{min}}$ one cannot lift all people above the minimum threshold ($\mathcal{D}^3$), without drawing non-poor people below the minimum threshold ($\mathcal{D}^4$), while maintaining the equal resource requirement ($\mathcal{D}^1$).

■

Appendix A.2 Alternative Conceptualizations

In this appendix we provide the formal derivations of the alternative norm distributions discussed in section 1.3.4 and displayed in Table 1.1. Furthermore, we show how additional inequality aversion may be introduced into our framework and how to operationalize the FfP concept based on individual-specific deprivation thresholds.

A.2.1 Alternative (a).

For this alternative measure we divert from the baseline by replacing weak EOp with strong EOp. The satisfaction of strong EOp requires the equalization of all moments of the type-specific income distribution. We therefore reformulate (9) as follows:

$$\mathcal{D}^{2a} = \left\{ d \in \mathcal{D} \,\middle|\, y_i^r = \tfrac{1}{N_{\mathcal{S}(\theta)}} \sum_{j \in \mathcal{S}(\theta)} y_j^r = \mu^r_{\mathcal{S}(\theta)},\ \forall\, i \in \mathcal{S}(\theta),\ \forall\, \theta \in \Theta \right\}. \tag{28}$$

Since we adhere to non-separability, invoking strong EOp requires a subsequent redefinition of the poor and the non-poor fraction of the population. As in the baseline, we construct a counterfactual income distribution that complies with the EOp principle in order to identify those below the poverty threshold y_{min}:

$$\mathcal{P}(\theta) = \left\{ i \in \mathcal{S}(\theta) \,\middle|\, y_i^e \frac{\mu^e_{\mathcal{S}(\theta)}}{y_i^e} \leq y_{\text{min}} \right\},\ \forall\, \theta \in \Theta, \tag{29}$$

$$\mathcal{R}(\theta) = \left\{ i \in \mathcal{S}(\theta) \,\middle|\, y_i^e \frac{\mu^e_{\mathcal{S}(\theta)}}{y_i^e} > y_{\text{min}} \right\},\ \forall\, \theta \in \Theta. \tag{30}$$

Furthermore, we define $\mathcal{R}(\Theta) = \cup_h \mathcal{R}(\theta)$.

As a consequence, the FfP and the proportionality requirement are formulated with respect to the counterfactual distribution in which strong EOp is realized:

$$\mathcal{D}^{3a} = \left\{ d \in \mathcal{D} \,\middle|\, y_i^r = y_{\text{min}},\ \forall\, i \in \mathcal{P}(\theta),\ \forall\, \theta \in \Theta \right\}, \tag{31}$$

$$\mathcal{D}^{4a} = \left\{ d \in \mathcal{D} \,\middle|\, \frac{y_i^r - y_{\text{min}}}{y_j^r - y_{\text{min}}} = \frac{\mu^e_{\mathcal{S}(\theta)} - y_{\text{min}}}{\mu^e_{\mathcal{S}(\theta')} - y_{\text{min}}},\ \forall\, i \in \mathcal{R}(\theta),\ \forall\, j \in \mathcal{R}(\theta'),\ \forall\, \theta \in \Theta \right\}. \tag{32}$$

Invoking strong EOp leads to the following proposition:

Proposition 2. *Suppose $\mu^e > y_{min}$. Then, the intersection $\mathcal{D}^1 \cap \mathcal{D}^{2a} \cap \mathcal{D}^{3a} \cap \mathcal{D}^{4a}$ yields a singleton which defines the norm distribution Y^r:*

$$y_i^r = \begin{cases} y_{min}, & \forall\, i \in \mathcal{P}(\theta),\ \forall\, \theta \in \Theta, \\ y_{min} + (\mu^e_{\mathcal{S}(\theta)} - y_{min}) \frac{(\mu^e - y_{min})}{\frac{N_{\mathcal{R}(\Theta)}}{N}(\mu^e_{\mathcal{R}(\Theta)} - y_{min})}, & \forall\, i \in \mathcal{R}(\theta),\ \forall\, \theta \in \Theta. \end{cases} \tag{33}$$

Conversely, if $\mu^e \leq y_{min}$, then $\mathcal{D}^1 \cap \mathcal{D}^{2a} \cap \mathcal{D}^{3a} \cap \mathcal{D}^{4a} = \emptyset$.

Proof of Proposition 2. The proof proceeds in analogy to the proof of Proposition 1. The proposition is trivially true for the (counterfactually) poor population $\mathcal{P}(\theta)$ as their norm incomes are prescribed by (31). We can use (32) to rewrite y_i^r for the members of tranches that are non-poor on average and use this expression in the constant resources constraint (8):

$$\underbrace{\frac{1}{N}\sum_{\theta\in\Theta}\left[\sum_{i\in\mathcal{P}(\theta)} y_{\text{min}} + \sum_{i\in\mathcal{R}(\theta)}\left(y_{\text{min}} + \frac{\mu^e_{\mathcal{S}(\theta)} - y_{\text{min}}}{\mu^e_{\mathcal{S}(\theta')} - y_{\text{min}}}(y_j^r - y_{\text{min}})\right)\right]}_{=\mu^r} = \mu^e. \tag{34}$$

Solving for y_j^r we obtain:

$$y_j^r = y_{\text{min}} + (\mu^e_{\mathcal{S}(\theta')} - y_{\text{min}})\frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}(\Theta)}}{N}(\mu^e_{\mathcal{R}(\Theta)} - y_{\text{min}})}. \tag{35}$$

As evidenced by (34), μ^r is a continuous and monotonically increasing function of y_j^r and we know that $y_j^r \in (y_{\text{min}}, \infty)$. Under the assumption that $\mu^e > y_{\text{min}}$ and invoking the intermediate value theorem, it follows that there is a unique value of y_j^r for which (8) holds. Since the choice of $i \in \mathcal{R}(\theta)$ and $j \in \mathcal{R}(\theta')$ was arbitrary, expression (33) holds $i \in \mathcal{R}(\theta), j \in \mathcal{R}(\theta'),\ \forall\, \theta \in \Theta$.

However, such a unique value only exists if $\mu^e > y_{\text{min}}$. Assume this was not true, i.e. $\mu^e \leq y_{\text{min}}$. Then, it would still hold that $\mu^r \to y_{\text{min}}$ for $y_j^r \to y_{\text{min}}$ and $\mu^r \to \infty$ for $y_j^r \to \infty$. Hence, $\mu^r \in (y_{\text{min}}, \infty)$. However, from the constant resources requirement (8) we also know that $\mu^r = \mu^e$. If $\mu^e < y_{\text{min}}$, either of these statements must be false and hence $\mathcal{D}^1 \cap \mathcal{D}^{2a} \cap \mathcal{D}^{3a} \cap \mathcal{D}^{4a} = \emptyset$. Intuitively, if $\mu^e < y_{\text{min}}$ one cannot lift all people above the minimum threshold ($\mathcal{D}^{3a}$), without drawing non-poor people below the minimum threshold ($\mathcal{D}^{4a}$), while maintaining the equal resource requirement ($\mathcal{D}^1$). ■

A.2.2 Alternative (b).

For this alternative measure we divert from the baseline by replacing non-separability with separability. In line with this normative assumption we reformulate the EOp requirement as follows:

$$\mathcal{D}^{2b} = \left\{ d \in \mathcal{D} \,\middle|\, \underbrace{\frac{1}{N_{\mathcal{T}(\omega)\cap\mathcal{R}}} \sum_{i \in \mathcal{T}(\omega)\cap\mathcal{R}} y_i^r}_{=\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}}} = \underbrace{\frac{1}{N_{\mathcal{R}}} \sum_{j \in \mathcal{R}} y_j^e \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\text{min}} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\text{min}})}\right)}_{=\mu^r_{\mathcal{R}}}, \, \forall\, \omega \in \Omega \right\}. \tag{36}$$

Instead of rating type-specific advantages by $\mu^e_{\mathcal{T}(\omega)}$, (36) draws on the average excess income above the poverty line, $\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}}$, to evaluate opportunity sets. Note that the type-specific average income above y_{min} must be equalized with respect to the norm (not the empirical) income of the rich population. This is a direct consequence of the constant resource restriction given in (8): Maintaining constant resources it is impossible to satisfy FfP without reducing the resources of the non-poor fraction $\mathcal{R}$ accordingly.

Separability of EOp and FfP entails that the incomes of $i \in \mathcal{P}$ are compared to a norm income of y_{min}, while the gains from opportunity equalization only accrue to $i \in \mathcal{R}$. As a consequence, the FfP and the proportionality requirement are formulated with respect to the sets $\mathcal{P}$ and $\mathcal{R}$ instead of their counterfactual analogues $\mathcal{P}(\omega)$ and $\mathcal{R}(\omega)$:

$$\mathcal{D}^{3b} = \left\{ d \in \mathcal{D} \,\middle|\, y_i^r = y_{\text{min}}, \, \forall\, i \in \mathcal{P} \right\}, \tag{37}$$

$$\mathcal{D}^{4b} = \left\{ d \in \mathcal{D} \,\middle|\, \frac{y_i^r - y_{\text{min}}}{y_j^r - y_{\text{min}}} = \frac{y_i^e - y_{\text{min}}}{y_j^e - y_{\text{min}}}, \, \forall\, i, j \in \mathcal{T}(\omega) \cap \mathcal{R}, \, \forall\, \omega \in \Omega \right\}. \tag{38}$$

Invoking the separability assumption leads to the following proposition:

Proposition 3. *Suppose $\mu^e > y_{min}$. Then, the intersection $\mathcal{D}^1 \cap \mathcal{D}^{2b} \cap \mathcal{D}^{3b} \cap \mathcal{D}^{4b}$ yields a singleton which defines the norm distribution Y^r:*

$$y_i^r = \begin{cases} y_{min}, & \forall\, i \in \mathcal{T}(\omega) \cap \mathcal{P},\ \forall\, \omega \in \Omega, \\ y_{min} + (y_i^e - y_{min}) \frac{(\mu^e - y_{min})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} - y_{min})}, & \forall\, i \in \mathcal{T}(\omega) \cap \mathcal{R},\ \forall\, \omega \in \Omega. \end{cases} \tag{39}$$

Conversely, if $\mu^e \leq y_{min}$, then $\mathcal{D}^1 \cap \mathcal{D}^{2b} \cap \mathcal{D}^{3b} \cap \mathcal{D}^{4b} = \emptyset$.

Proof of Proposition 3. The proof proceeds in analogy to the proof of Proposition 1. The proposition is trivially true for the poor population $\mathcal{P}$ as their norm incomes are prescribed by (37). For each type $\mathcal{T}(\omega)$ we can use (38) to rewrite y_i^r for the non-poor population and use this expression in the reformulated EOp condition (36):

$$\underbrace{\frac{1}{N_{\mathcal{T}(\omega)\cap\mathcal{R}}} \left[\sum_{i \in \mathcal{T}(\omega)\cap\mathcal{R}} \left(y_{\text{min}} + \frac{y_i^e - y_{\text{min}}}{y_j^e - y_{\text{min}}} (y_j^r - y_{\text{min}}) \right) \right]}_{=\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}}} = \mu^r_{\mathcal{R}}. \tag{40}$$

We use the constant resource condition (8) to express $\mu^r_{\mathcal{R}}$ in terms of observable quantities:

$$\frac{1}{N_{\mathcal{T}(\omega)\cap\mathcal{R}}} \left[\sum_{i \in \mathcal{T}(\omega)\cap\mathcal{R}} \left(y_{\text{min}} + \frac{y_i^e - y_{\text{min}}}{y_j^e - y_{\text{min}}} (y_j^r - y_{\text{min}}) \right) \right] = \frac{\mu^e - \frac{N_{\mathcal{P}}}{N} y_{\text{min}}}{\frac{N_{\mathcal{R}}}{N}}. \tag{41}$$

Solving for y_j^r we obtain:

$$y_j^r = y_{\text{min}} + (y_j^e - y_{\text{min}}) \frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} - y_{\text{min}})}. \tag{42}$$

As evidenced by (41), $\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}}$ is a continuous and monotonically increasing function of y_j^r and we know that $y_j^r \in (y_{\text{min}}, \infty)$. Invoking the proportionality condition (38) it must also be that $\mu^r_{\mathcal{R}} > y_{\text{min}}$. Under the assumption that $\mu^e > y_{\text{min}}$ and invoking the intermediate value theorem, it follows that there is a unique value of y_j^r for which (36) holds. Since the choice of $i, j \in \mathcal{T}(\omega) \cap \mathcal{R}$ was arbitrary, expression (39) holds $\forall\, i, j \in \mathcal{T}(\omega) \cap \mathcal{R},\ \forall\, \omega \in \Omega$.

However, such a unique value only exists if $\mu^e > y_{\text{min}}$. Assume this was not true, i.e. $\mu^e \leq y_{\text{min}}$. Then, it would still hold that $\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}} \to y_{\text{min}}$ for $y_j^r \to y_{\text{min}}$ and $\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}} \to \infty$ for $y_j^r \to \infty$. Hence, $\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}} \in (y_{\text{min}}, \infty)$. However, from the reformulated EOp requirement (36) we also know that $\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^r_{\mathcal{R}} \left(= (\mu^e - \frac{N_{\mathcal{P}}}{N} y_{\text{min}})/\frac{N_{\mathcal{R}}}{N}\right)$. If $\mu^e \leq y_{\text{min}}$, either of these statements must be false and hence $\mathcal{D}^1 \cap \mathcal{D}^{2b} \cap \mathcal{D}^{3b} \cap \mathcal{D}^{4b} = \emptyset$. Intuitively, if $\mu^e \leq y_{\text{min}}$ one cannot lift all people above the minimum threshold ($\mathcal{D}^{3b}$), without drawing non-poor people below the minimum threshold ($\mathcal{D}^{4b}$), while maintaining the equal resource requirement ($\mathcal{D}^1$). ■

A.2.3 Alternative (c).

For this alternative measure we divert from the baseline by adhering to both strong EOp and separability. We therefore reformulate (9) as follows:

$$\mathcal{D}^{2c} = \left\{ d \in \mathcal{D} \,\middle|\, y^r_i = \frac{1}{N_{\mathcal{S}(\theta)\cap\mathcal{R}}} \sum_{j\in\mathcal{S}(\theta)\cap\mathcal{R}} y^r_j = \mu^r_{\mathcal{S}(\theta)\cap\mathcal{R}},\ \forall\, i \in \mathcal{S}(\theta)\cap\mathcal{R},\ \forall\, \theta \in \Theta \right\}. \tag{43}$$

Separability of EOp and FfP entails that the incomes of $i \in \mathcal{P}$ are compared to a norm income of y_{min}, while the gains from opportunity equalization only accrue to $i \in \mathcal{R}$. As a consequence, the FfP and the proportionality requirement are formulated with respect to the sets $\mathcal{P}$ and $\mathcal{R}$ instead of a counterfactual analogue:

$$\mathcal{D}^{3c} = \left\{ d \in \mathcal{D} \,\middle|\, y^r_i = y_{\text{min}},\ \forall\, i \in \mathcal{P} \right\}, \tag{44}$$

$$\mathcal{D}^{4c} = \left\{ d \in \mathcal{D} \,\middle|\, \frac{y^r_i - y_{\text{min}}}{y^r_j - y_{\text{min}}} = \frac{\mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}} - y_{\text{min}}}{\mu^e_{\mathcal{S}(\theta')\cap\mathcal{R}} - y_{\text{min}}},\ \forall\, i \in \mathcal{R}(\theta),\ \forall\, j \in \mathcal{R}(\theta'),\ \forall\, \theta \in \Theta \right\}. \tag{45}$$

Invoking strong EOp and the separability assumption leads to the following proposition:

Proposition 4. *Suppose $\mu^e > y_{\text{min}}$. Then, the intersection $\mathcal{D}^1 \cap \mathcal{D}^{2c} \cap \mathcal{D}^{3c} \cap \mathcal{D}^{4c}$ yields a singleton which defines the norm distribution Y^r:*

$$y^r_i = \begin{cases} y_{min}, & \forall\, i \in \mathcal{S}(\theta)\cap\mathcal{P},\ \forall\, \theta \in \Theta, \\ y_{min} + (\mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}} - y_{min}) \frac{(\mu^e - y_{min})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{min})}, & \forall\, i \in \mathcal{S}(\theta)\cap\mathcal{R},\ \forall\, \theta \in \Theta. \end{cases} \tag{46}$$

Conversely, if $\mu^e \leq y_{min}$, then $\mathcal{D}^1 \cap \mathcal{D}^{2c} \cap \mathcal{D}^{3c} \cap \mathcal{D}^{4c} = \emptyset$.

Proof of Proposition 4. The proof proceeds in analogy to the proof of Proposition 1. The proposition is trivially true for the poor population $\mathcal{P}$ as their norm incomes are prescribed by (44). We can use (45) to rewrite y_i^r for the non-poor members of each effort tranche and use this expression in the constant resources constraint (8):

$$\underbrace{\frac{1}{N}\sum_{\theta\in\Theta}\left[\sum_{i\in\mathcal{S}(\theta)\cap\mathcal{P}} y_{\text{min}} + \sum_{i\in\mathcal{S}(\theta)\cap\mathcal{R}}\left(y_{\text{min}} + \frac{\mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}} - y_{\text{min}}}{\mu^e_{\mathcal{S}(\theta')\cap\mathcal{R}} - y_{\text{min}}}(y_j^r - y_{\text{min}})\right)\right]}_{=\mu^r} = \mu^e. \tag{47}$$

Solving for y_j^r we obtain:

$$y_j^r = y_{\text{min}} + (\mu^e_{\mathcal{S}(\theta')\cap\mathcal{R}} - y_{\text{min}})\frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\text{min}})}. \tag{48}$$

As evidenced by (47), μ^r is a continuous and monotonically increasing function of y_j^r and we know that $y_j^r \in (y_{\text{min}}, \infty)$. Under the assumption that $\mu^e > y_{\text{min}}$ and invoking the intermediate value theorem, it follows that there is a unique value of y_j^r for which (8) holds. Since the choice of $i \in \mathcal{S}(\theta)\cap\mathcal{R}$ and $j \in \mathcal{S}(\theta')\cap\mathcal{R}$ was arbitrary, expression (46) holds for all $i \in \mathcal{S}(\theta)\cap\mathcal{R}$, $\forall\, j \in \mathcal{S}(\theta')\cap\mathcal{R}$, $\forall\,\theta\in\Theta$.

However, such a unique value only exists if $\mu^e > y_{\text{min}}$. Assume this was not true, i.e. $\mu^e \leq y_{\text{min}}$. Then, it would still hold that $\mu^r \to y_{\text{min}}$ for $y_j^r \to y_{\text{min}}$ and $\mu^r \to \infty$ for $y_j^r \to \infty$. Hence, $\mu^r \in (y_{\text{min}}, \infty)$. However, from the constant resources requirement (8) we also know that $\mu^r = \mu^e$. If $\mu^e \leq y_{\text{min}}$, either of these statements must be false and hence $\mathcal{D}^1 \cap \mathcal{D}^{2c} \cap \mathcal{D}^{3c} \cap \mathcal{D}^{4c} = \emptyset$. Intuitively, if $\mu^e \leq y_{\text{min}}$ one cannot lift all people above the minimum threshold ($\mathcal{D}^{3c}$), without drawing non-poor people below the minimum threshold ($\mathcal{D}^{4c}$), while maintaining the equal resource requirement ($\mathcal{D}^1$). ∎

A.2.4 Additional Progressiveness

We are able to accommodate additional inequality aversion by relaxing the proportionality assumption and allowing for additional progressiveness in the intra-type distribution of excess income above y_{min}. To this end, let us reformulate the proportionality restriction as follows:

$$\mathcal{D}^{4d} = \left\{ d\in\mathcal{D} \,\middle|\, \frac{y_i^r - y_{\text{min}}}{y_j^r - y_{\text{min}}} = \frac{y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \mathcal{W}_i(\sigma) - y_{\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \mathcal{W}_j(\sigma) - y_{\text{min}}},\ \forall\, i,j\in\mathcal{R}(\omega),\ \forall\,\omega\in\Omega \right\}, \tag{49}$$

where $\mathcal{W}_i(\sigma)$ is an income weight subject to the parameter $\sigma \in [0,1]: \mathcal{W}_i(\sigma) = \left(1 - \sigma \frac{y_i^e - \mu^e_{\mathcal{R}(\omega)}}{y_i^e}\right)$.

Accounting for additional inequality aversion in the upper end of the income distribution leads to the following proposition:

Proposition 5. *Suppose $\mu^e > y_{min}$. Then, the intersection $\mathcal{D}^1 \cap \mathcal{D}^2 \cap \mathcal{D}^3 \cap \mathcal{D}^{4d}$ yields a singleton which uniquely defines the norm distribution Y^r:*

$$y_i^r = \begin{cases} y_{min}, & \forall\, i \in \mathcal{P}(\omega),\ \forall\, \omega \in \Omega, \\ y_{min} + (y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \mathcal{W}_i(\sigma) - y_{min}) \frac{(\mu^e - y_{min})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}(\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{min})}, & \forall\, i \in \mathcal{R}(\omega),\ \forall\, \omega \in \Omega. \end{cases} \tag{50}$$

Conversely, if $\mu \leq y_{min}$, then $\mathcal{D}^1 \cap \mathcal{D}^2 \cap \mathcal{D}^3 \cap \mathcal{D}^{4d} = \emptyset$.

Proof of Proposition 5. The proof proceeds in analogy to the proof of Proposition 1. The proposition is trivially true for the (counterfactually) poor population $\mathcal{P}(\omega)$ as their norm incomes are prescribed by (13). For each type $\mathcal{T}(\omega)$ we can use (49) to rewrite y_i^r for the non-poor population and use this expression together with the FfP condition (13) in the EOp condition (9):

$$\underbrace{\frac{1}{N_{\mathcal{T}(\omega)}} \left[\sum_{i \in \mathcal{P}(\omega)} y_{\text{min}} + \sum_{i \in \mathcal{R}(\omega)} \left(y_{\text{min}} + \frac{y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \mathcal{W}_i(\sigma) - y_{\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \mathcal{W}_j(\sigma) - y_{\text{min}}} (y_j^r - y_{\text{min}}) \right) \right]}_{=\mu^r_{\mathcal{T}(\omega)}} = \mu^e. \tag{51}$$

Solving for y_j^r we obtain:

$$y_j^r = y_{\text{min}} + (y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \mathcal{W}_j(\sigma) - y_{\text{min}}) \frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}} (\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}})}. \tag{52}$$

As evidenced by (51), $\mu^r_{\mathcal{T}(\omega)}$ is a continuous and monotonically increasing function of y_j^r and we know that $y_j^r \in (y_{\text{min}}, \infty)$. Under the assumption that $\mu^e > y_{\text{min}}$ and invoking the intermediate value theorem, it follows that there is a unique value of y_j^r for which (9) holds. Since the choice of $i, j \in \mathcal{R}(\omega)$ was arbitrary, expression (50) holds $\forall\, i, j \in \mathcal{R}(\omega),\ \forall\, \omega \in \Omega$.

However, such a unique value only exists if $\mu^e > y_{\text{min}}$. Assume this was not true, i.e. $\mu^e \leq y_{\text{min}}$. Then, it would still hold that $\mu^r_{\mathcal{T}(\omega)} \to y_{\text{min}}$ for $y_j^r \to y_{\text{min}}$ and $\mu^r_{\mathcal{T}(\omega)} \to \infty$ for $y_j^r \to \infty$. Hence,

$\mu^r_{\mathcal{T}(\omega)} \in (y_{\text{min}}, \infty)$. However, from the EOp requirement (9) we also know that $\mu^r_{\mathcal{T}(\omega)} = \mu^e$. If $\mu^e \leq y_{\text{min}}$, either of these statements must be false and hence $\mathcal{D}^1 \cap \mathcal{D}^2 \cap \mathcal{D}^3 \cap \mathcal{D}^{4d} = \emptyset$. Intuitively, if $\mu^e \leq y_{\text{min}}$ one cannot lift all people above the minimum threshold ($\mathcal{D}^3$), without drawing non-poor people below the minimum threshold ($\mathcal{D}^{4d}$), while maintaining the equal resource requirement ($\mathcal{D}^1$). ■

Note that σ can be interpreted as an inequality aversion parameter with respect to excess income above y_{min}.[1] To see this, note that $\frac{\partial y^r_i}{\partial \sigma} > 0$ ($\frac{\partial y^r_i}{\partial \sigma} < 0$) if $y^e_i < \mu^e_{\mathcal{R}(\omega)}$ ($y^e_i > \mu^e_{\mathcal{R}(\omega)}$) and $\frac{\partial^2 y^r_i}{\partial \sigma \partial y^e_i} < 0$. Hence, increasing σ leads to higher norm incomes for those below the type-specific mean of excess income. The positive effect monotonically decreases for increasing y^e_i until it turns negative for incomes above the type-specific mean of excess income.

Letting σ travel to one, $\mathcal{W}_i(\sigma) \to \mu^e_{\mathcal{R}(\omega)}/y^e_i$ and the norm distribution collapses to the following expression:

$$\lim_{\sigma \to 1} y^r_i = \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{P}(\omega),\ \forall\, \omega \in \Omega, \\ y_{\text{min}} + (y^e_i \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \mathcal{W}_i(\sigma) - y_{\text{min}}) \frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}} (\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\text{min}})}, & \forall\, i \in \mathcal{R}(\omega),\ \forall\, \omega \in \Omega, \end{cases} \tag{53}$$

$$= \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{P}(\omega),\ \forall\, \omega \in \Omega, \\ y_{\text{min}} + \frac{(\mu^e - y_{\text{min}})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}}, & \forall\, i \in \mathcal{R}(\omega),\ \forall\, \omega \in \Omega, \end{cases} \tag{54}$$

$$= \begin{cases} y_{\text{min}}, & \forall\, i \in \mathcal{P}(\omega),\ \forall\, \omega \in \Omega, \\ \mu^r_{\mathcal{R}(\omega)}, & \forall\, i \in \mathcal{R}(\omega),\ \forall\, \omega \in \Omega. \end{cases} \tag{55}$$

Hence, increasing σ indicates increasing inequality aversion with respect to income disparities among the non-poor population of a particular type. With $\sigma = 1$, the norm income of each non-poor type member is given by the average norm income of the non-poor constituents of its respective type. As a consequence, average income differences between the poor and the non-poor members of each type remain as the sole justifiable source of inequality. Reversely, taking limits towards zero inequality aversion, $\mathcal{W}_i(\sigma) \to 1$, and we obtain the baseline norm (see equation (15)) according to which excess norm incomes above y_{min} are distributed proportionally to their empirical analogues.

[1] For the sake of illustration we treat σ as a uniform parameter for all $\omega \in \Omega$. However, it is easy to allow for heterogeneity in σ across types.

A.2.5 Individual Minimum Thresholds

In our baseline analysis we account for differential needs across individuals by applying an equivalence scale. Alternatively, one could also account for differential needs by replacing the population-wide minimum threshold y_{min} with individual-specific minimum thresholds $y_{i,\text{min}}$. As a consequence, one would have to redefine the set of poor and non-poor individuals as follows:

$$\mathcal{P}(\omega) = \left\{ i \in \mathcal{N} \;\middle|\; y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} \leq y_{i,\text{min}} \right\}, \; \forall\, \omega \in \Omega \tag{56}$$

$$\mathcal{R}(\omega) = \left\{ i \in \mathcal{N} \;\middle|\; y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} > y_{i,\text{min}} \right\}, \; \forall\, \omega \in \Omega. \tag{57}$$

Similarly, the FfP principle and the proportionality requirement would have to be redefined in terms of the individual-specific minimum thresholds $y_{i,\text{min}}$:

$$\mathcal{D}^{3e} = \left\{ d \in \mathcal{D} \;\middle|\; y_i^r = y_{i,\text{min}}, \; \forall\, i \in \mathcal{P}(\omega), \; \forall\, \omega \in \Omega \right\}. \tag{58}$$

$$\mathcal{D}^{4e} = \left\{ d \in \mathcal{D} \;\middle|\; \frac{y_i^r - y_{i,\text{min}}}{y_j^r - y_{j,\text{min}}} = \frac{y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{i,\text{min}}}{y_j^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{j,\text{min}}}, \; \forall\, i, j \in \mathcal{R}(\omega), \; \forall\, \omega \in \Omega \right\}. \tag{59}$$

In addition let us define the type-specific average of poverty thresholds $\mu^{\text{min}}_{\mathcal{T}(\omega)} = \frac{1}{N_{\mathcal{T}(\omega)}} \sum_{i \in \mathcal{T}(\omega)} y_{i,\text{min}}$ and the type-specific average of poverty thresholds among its non-poor constituents $\mu^{\text{min}}_{\mathcal{R}(\omega)} = \frac{1}{N_{\mathcal{R}(\omega)}} \sum_{i \in \mathcal{R}(\omega)} y_{i,\text{min}}$.

These reformulations lead to the following proposition:

Proposition 6. *Suppose $\mu^e > \mu^{min}_{\mathcal{T}(\omega)}$, $\forall\, \omega \in \Omega$. Then, the intersection $\mathcal{D}^1 \cap \mathcal{D}^2 \cap \mathcal{D}^{3e} \cap \mathcal{D}^{4e}$ yields a singleton which uniquely defines the norm distribution Y^r:*

$$y_i^r = \begin{cases} y_{i,min}, & \forall\, i \in \mathcal{P}(\omega), \; \forall\, \omega \in \Omega, \\ y_{i,min} + \left(y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{i,min}\right) \frac{(\mu^e - \mu^{min}_{\mathcal{T}(\omega)})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}} \left(\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - \mu^{min}_{\mathcal{R}(\omega)}\right)}, & \forall\, i \in \mathcal{R}(\omega), \; \forall\, \omega \in \Omega. \end{cases} \tag{60}$$

Conversely, if $\exists\, \omega \in \Omega \colon \mu^e \leq \mu^{min}_{\mathcal{T}(\omega)}$, then $\mathcal{D}^1 \cap \mathcal{D}^2 \cap \mathcal{D}^{3e} \cap \mathcal{D}^{4e} = \emptyset$.

Proof of Proposition 6. The proof proceeds in analogy to the proof of Proposition 1. The proposition is trivially true for the (counterfactually) poor population $\mathcal{P}(\omega)$ as their norm incomes are prescribed by (58). For each type $\mathcal{T}(\omega)$ we can use (59) to rewrite y_i^r for the non-poor population and use this expression together with the FfP condition (58) in the EOp condition (9):

$$\underbrace{\frac{1}{N_{\mathcal{T}(\omega)}}\left[\sum_{i\in\mathcal{P}(\omega)} y_{i,\text{min}} + \sum_{i\in\mathcal{R}(\omega)}\left(y_{i,\text{min}} + \frac{y_i^e\frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{i,\text{min}}}{y_j^e\frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{j,\text{min}}}(y_j^r - y_{j,\text{min}})\right)\right]}_{=\mu^r_{\mathcal{T}(\omega)}} = \mu^e. \tag{61}$$

Solving for y_j^r we obtain:

$$y_j^r = y_{j,\text{min}} + (y_j^e\frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{j,\text{min}})\frac{(\mu^e - \mu^{\text{min}}_{\mathcal{T}(\omega)})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}(\mu^e_{\mathcal{R}(\omega)}\frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - \mu^{\text{min}}_{\mathcal{R}(\omega)})}. \tag{62}$$

As evidenced by (61), $\mu^r_{\mathcal{T}(\omega)}$ is a continuous and monotonically increasing function of y_j^r and we know that $y_j^r \in (y_{j,\text{min}}, \infty)$. Under the assumption that $\mu^e > \mu^{\text{min}}_{\mathcal{T}(\omega)},\ \forall\, \omega \in \Omega$ and invoking the intermediate value theorem, it follows that there is a unique value of y_j^r for which (9) holds. Since the choice of $i, j \in \mathcal{R}(\omega)$ was arbitrary, expression (60) holds $\forall\, i, j \in \mathcal{R}(\omega),\ \forall\, \omega \in \Omega$.

However, such a unique value only exists if $\mu^e > \mu^{\text{min}}_{\mathcal{T}(\omega)},\ \forall\, \omega \in \Omega$. Assume this was not true, i.e. $\exists\, \omega \in \Omega\colon\ \mu^e \leq \mu^{\text{min}}_{\mathcal{T}(\omega)}$. Then, it would still hold that $\mu^r_{\mathcal{T}(\omega)} \to \mu^{\text{min}}_{\mathcal{T}(\omega)}$ for $y_j^r \to y_{j,\text{min}}$ and $\mu^r_{\mathcal{T}(\omega)} \to \infty$ for $y_j^r \to \infty$. Hence, $\mu^r_{\mathcal{T}(\omega)} \in \left(\mu^{\text{min}}_{\mathcal{T}(\omega)}, \infty\right)$. However, from the EOp requirement (9) we also know that $\mu^r_{\mathcal{T}(\omega)} = \mu^e$. If $\mu^e \leq \mu^{\text{min}}_{\mathcal{T}(\omega)}$, either of these statements must be false and hence $\mathcal{D}^1 \cap \mathcal{D}^2 \cap \mathcal{D}^{3e} \cap \mathcal{D}^{4e} = \emptyset$. Intuitively, if $\exists\, \omega \in \Omega\colon\ \mu^e \leq \mu^{\text{min}}_{\mathcal{T}(\omega)}$ one cannot lift all people above the minimum threshold ($\mathcal{D}^{3e}$), without drawing non-poor people below the minimum threshold ($\mathcal{D}^{4e}$), while maintaining the equal resource requirement ($\mathcal{D}^1$). ■

Appendix A.3 Comparative Statics

In this appendix we give a comprehensive overview over the comparative statics of all norms listed in Table 1.1. A general overview can be found in Table A.1. Each of the illustrated comparative static scenarios is discussed verbally in the following.

(a) EOp or FfP Only

(1) Assume $y_{\text{min}} \to 0$. The limit case with $y_{\text{min}} = 0$ is equivalent to abstracting from the concern for FfP altogether.

- **Baseline**: Leads to $\mathcal{P}(\omega) = \emptyset$, $\mu^e_{\mathcal{R}(\omega)} = \mu^e_{\mathcal{T}(\omega)}$, $N_{\mathcal{R}(\omega)} = N_{\mathcal{T}(\omega)}$, $\forall\, \omega \in \Omega$. As a consequence, realizing weak EOp remains the only normative concern.
- **Alternative (a)**: Leads to $\mathcal{P}(\theta) = \emptyset$, $\mu^e_{\mathcal{R}(\theta)} = \mu^e$, $N_{\mathcal{R}(\theta)} = N$, $\forall\, \theta \in \Theta$. As a consequence, realizing strong EOp remains the only normative concern.
- **Alternative (b)**: Leads to $\mathcal{P} = \emptyset$, $N_{\mathcal{R}} = N$, and $\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^e_{\mathcal{T}(\omega)}$, $\forall\, \omega \in \Omega$. As a consequence, realizing weak EOp remains the only normative concern.
- **Alternative (c)**: Leads to $\mathcal{P} = \emptyset$, $N_{\mathcal{R}} = N$, and $\mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}} = \mu^e_{\mathcal{S}(\theta)}$, $\forall\, \theta \in \Theta$. As a consequence, realizing strong EOp remains the only normative concern.

(2) Assume $T \to 1$. The limit case with $T = 1$ is equivalent to abstracting from the concern for EOp altogether. It also leads to $\mathcal{P}(\omega) = \mathcal{P} = \mathcal{P}(\theta)$.

- **Baseline**: Leads to $N_{\mathcal{R}(\omega)} = N_{\mathcal{R}}$, $N_{\mathcal{T}(\omega)} = N$, $\mu^e_{\mathcal{R}(\omega)} = \mu^e_{\mathcal{R}}$, $\mu^e_{\mathcal{T}(\omega)} = \mu^e$, $\forall\, \omega \in \Omega$. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.
- **Alternative (a)**: Leads to $N_{\mathcal{R}(\Theta)} = N_{\mathcal{R}}$, $\mu^e_{\mathcal{R}(\Theta)} = \mu^e_{\mathcal{R}}$, $\mu^e_{\mathcal{S}(\theta)} = y^e_i$, $\forall\, i \in \mathcal{S}(\theta)$, $\forall\, \theta \in \Theta$. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.

- **Alternative (b)**: Leads to $\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^e_{\mathcal{R}}, \; \forall\, \omega \in \Omega$. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.

- **Alternative (c)**: Leads to $\mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}} = y^e_i, \; \forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R}, \; \forall\, \theta \in \Theta$. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.

(b) Freedom from Poverty

(3) Assume $N_{\mathcal{P}(\omega)} \to 0, \; \forall \omega \in \Omega$. The limit case with $N_{\mathcal{P}(\omega)} = 0$ is equivalent to zero poverty incidence if resources were distributed in accordance with weak EOp.

- **Baseline**: Leads to $\mathcal{P}(\omega) = \emptyset$, $\mu^e_{\mathcal{R}(\omega)} = \mu^e_{\mathcal{T}(\omega)}$, $N_{\mathcal{R}(\omega)} = N_{\mathcal{T}(\omega)}, \; \forall\, \omega \in \Omega$. As a consequence, realizing weak EOp remains the only normative concern.

- **Alternative (a)**: $\cup_k \mathcal{P}(\omega) = \emptyset$ implies $\cup_l \mathcal{P}(\theta) = \emptyset$. Hence, $\mu^e_{\mathcal{R}(\Theta)} = \mu^e$, and $N_{\mathcal{R}(\Theta)} = N$. As a consequence, realizing strong EOp remains the only normative concern.

- **Alternative (b)**: No difference. The poor are identified and tied to a norm income of y_{min} based on $\mathcal{P}$. Since $\cup_k \mathcal{P}(\omega) = \emptyset$ does not imply $\mathcal{P} = \emptyset$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (c)**: No difference. The poor are identified and tied to a norm income of y_{min} based on $\mathcal{P}$. Since $\cup_k \mathcal{P}(\omega) = \emptyset$ does not imply $\mathcal{P} = \emptyset$ the calculation of the norm remains unaffected even in the limit case

(4) Assume $N_{\mathcal{P}} \to 0$. The limit case with $N_{\mathcal{P}} = 0$ is equivalent to zero poverty incidence in the empirical income distribution.

- **Baseline**: No difference. The poor are identified and tied to a norm income of y_{min} based on $\mathcal{P}(\omega)$. Since $\mathcal{P} = \emptyset$ does not imply $\cup_k \mathcal{P}(\omega) = \emptyset$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (a)**: $\mathcal{P} = \emptyset$ implies $\cup_l \mathcal{P}(\theta) = \emptyset$. Hence, $\mu^e_{\mathcal{R}(\Theta)} = \mu^e$, and $N_{\mathcal{R}(\Theta)} = N$. As a consequence, realizing strong EOp remains the only normative concern.

- **Alternative (b)**: Leads to $N_{\mathcal{R}} = N$, and $\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^e_{\mathcal{T}(\omega)}, \; \forall\, \omega \in \Omega$. As a consequence, realizing weak EOp through a type-specific linear transfer rate on excess income above y_{min} remains the only normative concern.

- **Alternative (c)**: Leads to $N_{\mathcal{R}} = N$, and $\mu^e_{R} = \mu^e$, $\mu^e_{\mathcal{S}(\omega)\cap\mathcal{R}} = \mu^e_{\mathcal{S}(\omega)}, \; \forall\, \theta \in \Theta$. As a consequence, realizing strong EOp remains the only normative concern.

(5) Assume $N_{\mathcal{P}(\theta)} \to 0, \; \forall\, \theta \in \Theta$. The limit case with $N_{\mathcal{P}(\theta)} = 0$ is equivalent to zero poverty incidence if resources were distributed in accordance with strong EOp.

- **Baseline**: No difference. The poor are identified and tied to a norm income of y_{min} based on $\mathcal{P}(\omega)$. Since $\cup_l \mathcal{P}(\theta) = \emptyset$ does not imply $\cup_k \mathcal{P}(\omega) = \emptyset$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (a)**: Leads to $\mu^e_{\mathcal{R}(\Theta)} = \mu^e$, $N_{\mathcal{R}(\Theta)} = N$. As a consequence, realizing strong EOp remains the only normative concern.

- **Alternative (b)**: No difference. The poor are identified and tied to a norm income of y_{min} based on $\mathcal{P}$. Since $\cup_l \mathcal{P}(\theta) = \emptyset$ does not imply $\mathcal{P} = \emptyset$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (c)**: No difference. The poor are identified and tied to a norm income of y_{min} based on $\mathcal{P}$. Since $\cup_l \mathcal{P}(\theta) = \emptyset$ does not imply $\mathcal{P} = \emptyset$ the calculation of the norm remains unaffected even in the limit case.

(c) Equality of Opportunity

(6) Assume $\mu^e_{\mathcal{T}(\omega)} \to \mu^e, \; \forall\, \omega \in \Omega$. The limit case with $\mu^e_{\mathcal{T}(\omega)} = \mu^e, \; \forall\, \omega \in \Omega$ corresponds to a society in which weak EOp is realized. It also leads to $\cup_k \mathcal{P}(\omega) = \mathcal{P}$.

- **Baseline**: Weak EOp under the non-separability assumption is realized by assumption. As a consequence, poverty eradication through a type-specific linear transfer rate on excess income above y_{min} remains the only normative concern.

- **Alternative (a)**: No difference. Strong EOp under the non-separability assumption requires equalizing all moments of the type distribution. Since $\mu^e_{\mathcal{T}(\omega)} = \mu^e, \; \forall\, \omega \in \Omega$

does not imply $y_i^e = y_j^e = \mu^e_{\mathcal{S}(\theta)}, \forall\, i, j \in \mathcal{S}(\theta), \forall\, \theta \in \Theta$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (b)**: No difference. Weak EOp under the separability assumption requires equalizing type mean incomes above y_{min} only. Since $\mu^e_{\mathcal{T}(\omega)} = \mu^e, \forall\, \omega \in \Omega$ does not imply $\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^r_{\mathcal{R}}, \forall\, \omega \in \Omega$ the calculation of the norm remains unaffected even in the limit case.
- **Alternative (c)**: No difference. Strong EOp under the separability assumption requires equalizing all incomes of the non-poor tranche members. Since $\mu^e_{\mathcal{T}(\omega)} = \mu^e, \forall\, \omega \in \Omega$ does not imply $y_i^e = y_j^e = \mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}}, \forall\, i, j \in \mathcal{S}(\theta) \cap \mathcal{R}, \forall\, \theta \in \Theta$ the calculation of the norm remains unaffected even in the limit case.

(7) Assume $y_i^e \to \mu^e_{\mathcal{S}(\theta)}, \forall\, i \in \mathcal{S}(\theta), \forall\, \theta \in \Theta$. The limit case with $y_i^e = \mu^e_{\mathcal{S}(\theta)}, \forall\, i \in \mathcal{S}(\theta), \forall\, \theta \in \Theta$ corresponds to a society in which strong EOp is realized. It also leads to $\cup_k \mathcal{P}(\omega) = \mathcal{P} = \cup_l \mathcal{P}(\theta)$.

- **Baseline**: Weak EOp under the non-separability assumption requires equalizing type mean incomes. $y_i^e = \mu^e_{\mathcal{S}(\theta)}, \forall\, i \in \mathcal{S}(\theta), \forall\, \theta \in \Theta$ implies $\mu^e_{\mathcal{T}(\omega)} = \mu^e, \forall\, \omega \in \Omega$. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.
- **Alternative (a)**: Strong EOp under the non-separability assumption is realized by assumption. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.
- **Alternative (b)**: Weak EOp under the separability assumption requires equalizing type mean incomes above y_{min} only. $y_i^e = y_j^e = \mu^e_{\mathcal{S}(\theta)}, \forall\, i, j \in \mathcal{S}(\theta), \forall\, \theta \in \Theta$ implies $\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^e_{\mathcal{R}}, \forall\, \omega \in \Omega$. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.
- **Alternative (c)**: Strong EOp under the separability assumption requires equalizing all incomes of the non-poor tranche members. $y_i^e = y_j^e = \mu^e_{\mathcal{S}(\theta)}, \forall\, i, j \in \mathcal{S}(\theta), \forall\, \theta \in \Theta$ implies $y_i^e = y_j^e = \mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}}, \forall\, i, j \in \mathcal{S}(\theta) \cap \mathcal{R}, \forall\, \theta \in \Theta$. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.

(8) Assume $\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} \rightarrow \mu^e_{\mathcal{R}}$, $\forall\, \omega \in \Omega$. The limit case corresponds to a society in which weak EOp is realized under the separability assumption.

- **Baseline**: No difference. Weak EOp under the non-separability assumption requires equalizing type mean incomes. Since $\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^e_{\mathcal{R}}$, $\forall\, \omega \in \Omega$ does not imply $\mu^e_{\mathcal{T}(\omega)} = \mu^e$, $\forall\, \omega \in \Omega$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (a)**: No difference. Strong EOp under the non-separability assumption requires equalizing all moments of the type distribution. Since $\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^e_{\mathcal{R}}$, $\forall\, \omega \in \Omega$ does not imply $y^e_i = y^e_j = \mu^e_{\mathcal{S}(\theta)}$, $\forall\, i, j \in \mathcal{S}(\theta)$, $\forall\, \theta \in \Theta$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (b)**: Weak EOp under the separability assumption is realized by assumption. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.

- **Alternative (c)**: No difference. Strong EOp under the separability assumption requires equalizing all incomes of the non-poor tranche members. Since $\mu^e_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^e_{\mathcal{R}}$, $\forall\, \omega \in \Omega$ does not imply $y^e_i = \mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}}$, $\forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R}$, $\forall\, \theta \in \Theta$ the calculation of the norm remains unaffected even in the limit case.

(9) Assume $y^e_i \rightarrow \mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}}$, $\forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R}$, $\forall\, \theta \in \Theta$. The limit case corresponds to a society in which strong EOp is realized under the separability assumption.

- **Baseline**: No difference. Weak EOp under the non-separability assumption requires equalizing type mean incomes. Since $y^e_i = \mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}}$, $\forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R}$, $\forall\, \theta \in \Theta$ does not imply $\mu^e_{\mathcal{T}(\omega)} = \mu^e$, $\forall\, \omega \in \Omega$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (a)**: No difference. Strong EOp under the non-separability assumption requires equalizing all moments of the type distribution. Since $y^e_i = \mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}}$, $\forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R}$, $\forall\, \theta \in \Theta$ does not imply $y^e_i = y^e_j = \mu^e_{\mathcal{S}(\theta)}$, $\forall\, i, j \in \mathcal{S}(\theta)$, $\forall\, \theta \in \Theta$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (b)**: No difference. Weak EOp under the separability assumption requires equalizing type mean incomes above y_{min} only. Since $y^e_i = \mu^e_{\mathcal{S}(\theta)\cap\mathcal{R}}$, $\forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R}$, $\forall\, \theta \in$

Θ does not imply $\mu^r_{\mathcal{T}(\omega)\cap\mathcal{R}} = \mu^r_{\mathcal{R}}, \forall \omega \in \Omega$ the calculation of the norm remains unaffected even in the limit case.

- **Alternative (c)**: Strong EOp under the separability assumption is realized by assumption. As a consequence, poverty eradication through a linear transfer rate on excess income above y_{min} remains the only normative concern.

TABLE A.1 – Overview Comparative Statics

	Baseline	Alternative (a)	Alternative (b)	Alternative (c)
Norm Distribution Y^r				
	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P}(\omega),\, \forall\, \omega \in \Omega \\ y_{\min} + (y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\min}) \frac{(\mu^e - y_{\min})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}(\mu^e_{\mathcal{R}(\omega)} \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}} - y_{\min})}, & \forall\, i \in \mathcal{R}(\omega),\, \forall\, \omega \in \Omega \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P}(\theta),\, \forall\, \theta \in \Theta \\ y_{\min} + (\mu^e_{\mathcal{S}(\theta)} - y_{\min}) \frac{(\mu^e - y_{\min})}{\frac{N_{\mathcal{R}(\Theta)}}{N}(\mu^e_{\mathcal{R}(\Theta)} - y_{\min})}, & \forall\, i \in \mathcal{R}(\theta),\, \forall\, \theta \in \Theta \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{T}(\omega) \cap \mathcal{P},\, \forall\, \omega \in \Omega \\ y_{\min} + (y_i^e - y_{\min}) \frac{(\mu^e - y_{\min})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{T}(\omega) \cap \mathcal{R}} - y_{\min})}, & \forall\, i \in \mathcal{T}(\omega) \cap \mathcal{R},\, \forall\, \omega \in \Omega \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{S}(\theta) \cap \mathcal{P},\, \forall\, \theta \in \Theta, \\ y_{\min} + (\mu^e_{\mathcal{S}(i) \cap \mathcal{R}} - y_{\min}) \frac{(\mu^e - y_{\min})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}, & \forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R},\, \forall\, \theta \in \Theta \end{cases}$
(a) EOp or FfP Only				
(1) $y_{\min} = 0$	$y_i^r = y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}},\, \forall\, i \in \mathcal{T}(\omega),\, \forall\, \omega \in \Omega$	$y_i^r = y_i^e \frac{\mu^e_{\mathcal{S}(\theta)}}{y_i^e},\, \forall\, i \in \mathcal{S}(\theta),\, \forall\, \theta \in \Theta$	$y_i^r = y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}},\, \forall\, i \in \mathcal{T}(\omega),\, \forall\, \omega \in \Omega$	$y_i^r = y_i^e \frac{\mu^e_{\mathcal{S}(\theta)}}{y_i^e},\, \forall\, i \in \mathcal{S}(\theta),\, \forall\, \theta \in \Theta$
(2) $T = 1$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$
(b) Freedom from Poverty				
(3) $\mathcal{P}(\omega) = 0,\, \forall\, \omega \in \Omega$	$y_i^r = y_i^e \frac{\mu^e}{\mu^e_{\mathcal{T}(\omega)}},\, \forall\, i \in \mathcal{T}(\omega),\, \forall\, \omega \in \Omega$	$y_i^r = y_i^e \frac{\mu^e_{\mathcal{S}(\theta)}}{y_i^e},\, \forall\, i \in \mathcal{S}(\theta),\, \forall\, \theta \in \Theta$	No difference	No difference
(4) $\mathcal{P} = 0$	No difference	$y_i^r = y_i^e \frac{\mu^e_{\mathcal{S}(\theta)}}{y_i^e},\, \forall\, i \in \mathcal{S}(\theta),\, \forall\, \theta \in \Theta$	$y_i^r = y_{\min} + (y_i^e - y_{\min}) \frac{(\mu^e - y_{\min})}{(\mu^e_{\mathcal{T}(\omega)} - y_{\min})},\, \forall\, i \in \mathcal{P}(\omega),\, \forall\, \omega \in \Omega$	$y_i^r = y_i^e \frac{\mu^e_{\mathcal{S}(\theta)}}{y_i^e},\, \forall\, i \in \mathcal{S}(\theta),\, \forall\, \theta \in \Theta$
(5) $\mathcal{P}(\theta) = 0,\, \forall\, \theta \in \Theta$	No difference	$y_i^r = y_i^e \frac{\mu^e_{\mathcal{S}(\theta)}}{y_i^e},\, \forall\, i \in \mathcal{S}(\theta),\, \forall\, \theta \in \Theta$	No difference	No difference
(c) Equality of Opportunity				
(6) $\mu^e_{\mathcal{T}(\omega)} = \mu^e,\, \forall\, \omega \in \Omega$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P}(\omega),\, \forall\, \omega \in \Omega \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}(\omega)}}{N_{\mathcal{T}(\omega)}}(y_{\min} - \mu^e_{\mathcal{P}(\omega)})}{\frac{N_{\mathcal{R}(\omega)}}{N_{\mathcal{T}(\omega)}}(\mu^e_{\mathcal{R}(\omega)} - y_{\min})}\right), & \forall\, i \in \mathcal{R}(\omega),\, \forall\, \omega \in \Omega \end{cases}$	No difference	No difference	No difference
(7) $y_i^e = \mu^e_{\mathcal{S}(\theta)},\, \forall\, i \in \mathcal{S}(\theta),\, \forall\, \theta \in \Theta$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$
(8) $\mu^e_{\mathcal{T}(\omega) \cap \mathcal{R}} = \mu^e_{\mathcal{R}},\, \forall\, \omega \in \Omega$	No difference	No difference	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$	No difference
(9) $y_i^e = \mu^e_{\mathcal{S}(\theta) \cap \mathcal{R}},\, \forall\, i \in \mathcal{S}(\theta) \cap \mathcal{R},\, \forall\, \theta \in \Theta$	No difference	No difference	No difference	$y_i^r = \begin{cases} y_{\min}, & \forall\, i \in \mathcal{P} \\ y_{\min} + (y_i^e - y_{\min}) \left(1 - \frac{\frac{N_{\mathcal{P}}}{N}(y_{\min} - \mu^e_{\mathcal{P}})}{\frac{N_{\mathcal{R}}}{N}(\mu^e_{\mathcal{R}} - y_{\min})}\right), & \forall\, i \in \mathcal{R} \end{cases}$

Appendix A.4 Data Appendix

A.4.1 Disposable Household Income

PSID. We construct disposable household income as the sum of household labor income, household asset income, household private transfers, household private pensions, other household income, household public pensions, household public cash assistance minus total household taxes. These income aggregates are calculated and provided by PSID CNEF.

In view of changes in the handling of negative incomes across waves, we consistently set household asset income and household private transfers to zero if they are negative or missing.

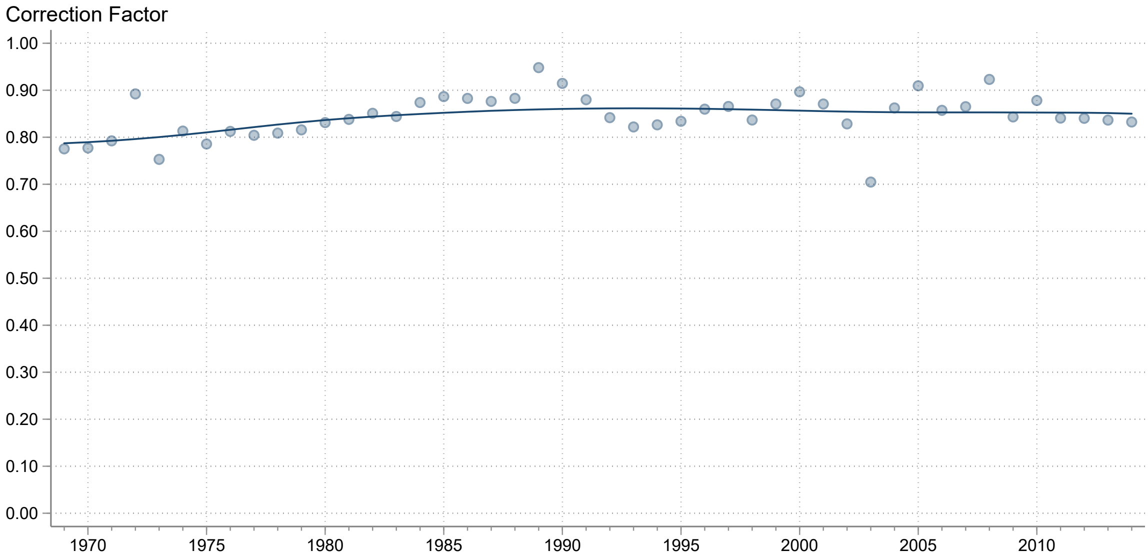
We account for the under-reporting of government transfer income by scaling up household public cash assistance of each recipient household in year t by the inverse of the following adjustment factor:

$$UR_t = \frac{V_{pt}}{\sum_p V_{pt}} * UR_{pt}^{PSID}, \tag{63}$$

where UR_{pt}^{PSID} is the share of transfer income from government program p in year t reported by PSID households when comparing their cumulated reports to government statistics on annual spending in the respective program. V_{pt} indicates the total volume of government spending on program p in year t. UR_{pt}^{PSID} and V_{pt} are taken from the time series provided in Meyer et al. (2015). The government programs p include Unemployment Insurance (UI), Workers' Compensation (WC), Social Security Retirement and Survivors Insurance (OASI), Social Security Disability Insurance (SSDI), Supplemental Security Income (SSI), the Food Stamp Program (SNAP), and Aid to Families with Dependent Children/Temporary Assistance for Needy Families (AFDC/TANF). Since their time series end in 2010 we fit UR_{pt}^{PSID} to a second-order polynomial of the year-variable and impute UR_{pt}^{PSID} for 2012 and 2014 with the predicted values. The time series for UR_t^{PSID} is displayed in Figure A.1.

We account for the under-reporting of labor income by imputing individual labor incomes according to the following procedure. First, we identify individuals with zero or missing labor income information but non-zero working hours. Second we run the following Mincer

FIGURE A.1 – Correction Factor for Under-reporting of Transfer Income (US), 1969-2014



Data: Meyer et al. (2015).
Note: Own calculations. This figure displays the correction factor for under-reported transfer incomes in the PSID over the time period 1969-2014. The correction factor is calculated based on equation (63) and the time series presented in Meyer et al. (2015). The solid lines display Lowess smoothed time trends where each data point is constructed using 80% of all data points (Bandwidth 0.8).

regression on the pooled PSID sample:[2]

$$\ln y_{ict} = \beta_0 + \beta_1 Hours_{ict} + \beta_2 Hours^2_{ict} + \beta_3 Age_{ict} + \beta_4 Age^2_{ict} + \beta_5 Race_{ict} + \beta_6 Male_{ict} + \beta_7 Education_{ict} + \gamma_t + \epsilon_{ict}. \tag{64}$$

Third, we impute individual labor incomes of the identified individuals with the income predictions from the Mincer regression. Fourth, we aggregate the volume of imputed incomes across all members of a household and add the imputed incomes to the household labor income provided by PSID CNEF.

The resulting variable for disposable household income is converted to equivalized disposable household income using the modified OECD equivalence scale, winsorized at the 1st and 99.5th percentiles, and converted into PPP-adjusted US Dollar using the conversion factors provided by the Penn World Tables (Feenstra et al., 2015).

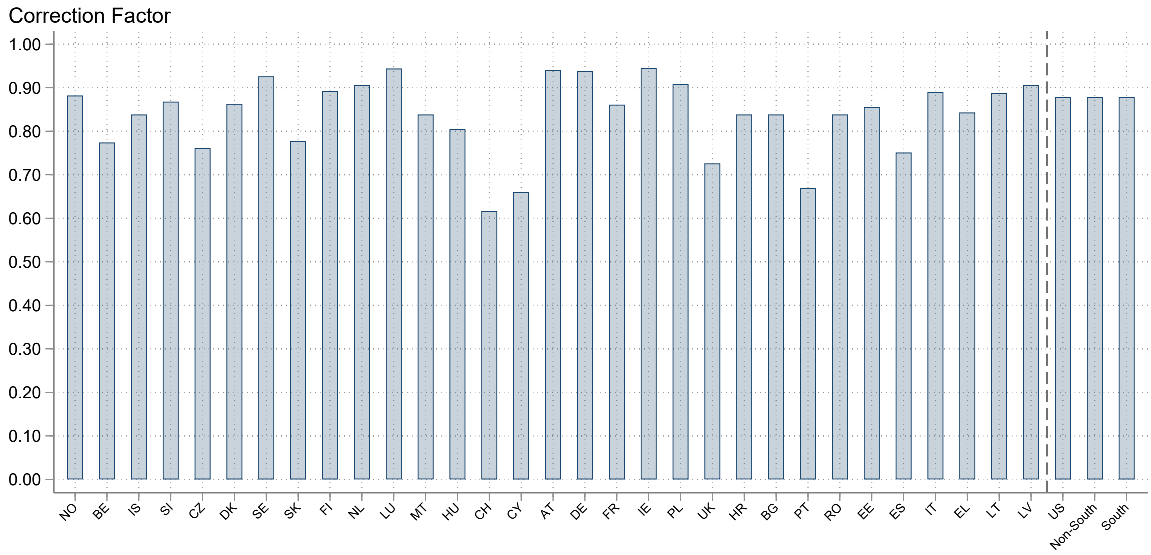
EU-SILC. We construct household disposable income as the sum of household labor income, household asset income, household private transfers, household private pensions, other

[2] The underlying variables are constructed according to the details provided in this Data Appendix. Regression results are available upon request.

household income, household public pensions, household public cash assistance minus total household taxes.

For consistency with the PSID, we set household asset income and household private transfers to zero if they are negative or missing. We account for the under-reporting of government transfer income by scaling up household public cash assistance of each recipient household in country c by the inverse of the adjustment factor UR_c^{SILC}. UR_c^{SILC} is extracted from EUROSTAT (2013) – a report in which EUROSTAT compares various income sources from EU-SILC with the corresponding national accounts aggregates. Specifically, UR_c^{SILC} contains family/children-related allowances, unemployment benefits, old-age benefits, survivors' benefits, sickness benefits, disability benefits, education-related allowances, and social exclusion benefits not elsewhere classified. This exercise is conducted for the income reference period 2008 and we write the calculated values forward to 2010. Furthermore, five of our sample countries were excluded from the analysis due to a lack of information from either of the two data sources (Bulgaria, Malta, Romania, Iceland and Croatia). For these countries we impute UR_c^{SILC} with the European cross-country sample mean. The values for UR_c^{SILC} are displayed in Figure A.2.

Figure A.2 – Correction Factor for Under-reporting of Transfer Income (Cross-Country Sample), 2010



Data: EUROSTAT (2013) and Meyer et al. (2015).
Note: Own calculations. This figure displays the correction factor for under-reported transfer incomes in the the cross-country sample in 2010. The correction factor is calculated based on equation (63) and the time series presented in Meyer et al. (2015) as well as the under-reporting factors reported in EUROSTAT (2013). Data points to the left of the vertical dashed line refer to the European country sample. Data points to the right of the vertical dashed line refer to the US and its census regions.

We account for the under-reporting of labor income by imputing individual labor incomes in the same way as in the PSID. To this end we construct a EU-SILC country-panel spanning the time period 2006-2014. In contrast to the PSID we run the underlying Mincer regression separately for each country in the EU-SILC sample and replace the race indicator with the migration background indicator:[3]

$$\begin{aligned}\ln y_{ict} = \beta_0 + \beta_1 Hours_{ict} + \beta_2 Hours^2_{ict} + \beta_3 Age_{ict} + \beta_4 Age^2_{ict} \\ + \beta_5 Mig.\ Background_{ict} + \beta_6 Male_{ict} + \beta_7 Education_{ict} + \gamma_t + \epsilon_{ict}.\end{aligned} \tag{65}$$

Again, the resulting variable for disposable household income is converted to equivalized disposable household income using the modified OECD equivalence scale, winsorized at the 1st and 99.5th percentiles, and converted into PPP-adjusted US Dollar using the conversion factors provided by the Penn World Tables (Feenstra et al., 2015).

A.4.2 Biological Sex

PSID. We use the binary biological sex variable provided by PSID CNEF. Using the panel dimension of the PSID we replace the few missing values with the mode of all records for the respective individual.

EU-SILC. We use the binary biological sex variable provided by EU-SILC. Respondents with missing information are dropped through list-wise deletion.

A.4.3 Race/Migration Background

PSID. We use the 6-category race indicator (White, Black, Am. Indian-Inuit, Asian-Pacific Islander, Black, Hispanic, Other) provided by PSID CNEF and transform it into a binary indicator for non-Hispanic whites and others. Using the panel dimension of the PSID we replace missing values with the mode of all records for the respective individual.

EU-SILC. We use the 3-category migration background indicator (born in country of residence, born in other European country, born elsewhere) provided by EU-SILC and transform it

[3] The underlying variables are constructed according to the details provided in this Data Appendix. Regression results are available upon request.

Table A.2 – Harmonization of Education Codes

	PSID	EU-SILC
High	(1) College BA and no advanced degree mentioned (2) College and advanced or professional degree (3) College but no degree	(1) At least first stage of tertiary education (2) – (3) –
Middle	(4) 12 grades (5) 12 grades plus non-academic training	(4) Upper secondary education (5) –
Low	(6) 0-5 grades (7) 6-8 grades (8) 9-11 grades (9) Could not read or write	(6) Pre-primary, primary education, lower secondary education (7) Father (mother) could neither read nor write (8) Don't know (9) –

into a binary indicator for whether the respondent was born in her current country of residence or not. Respondents with missing information are dropped through list-wise deletion.

A.4.4 Parental Education

PSID. We use the 9-category indicator for paternal and maternal education provided by the PSID and transform them into a 3-category indicator for high, medium, and low education according to the classification scheme outlined in Table A.2. We retain the highest information of either parent. We replace missing information by the highest recorded education level from previous years. Since educational attainment cannot be downgraded we also replace lower educational attainments by the highest recorded education level from previous years.

EU-SILC. We use the 5-category indicator for paternal and maternal education provided by EU-SILC and transform them into a 3-category indicator for high, medium, and low education according to the classification scheme outlined in Table A.2. We then retain the highest information of either parent. Respondents with missing information are dropped through list-wise deletion.

A.4.5 Parental Occupation

PSID. In the PSID, waves 1970-2001 report occupation codes with reference to 1970 census codes. Waves 2003-2015 report occupation codes with reference to 2000 census codes. If available on 3-digit level, we use the cross-walk routine provided by Autor and Dorn (2013) to standardize codes based on the 1990 census classification. 1 (28) of the 1970 (2000) 3-digit occupational codes available in the PSID are not included in the cross-walks provided by Autor

TABLE A.3 – Harmonization of Occupation Codes

	Census 1970	Census 1990	ISCO-08
High	(1) Professional, Technical and Kindred workers	(1) Managerial and Professional Specialty Occ.	(1) Managers
	(2) Managers, Officials and Proprietors	(2) Technical and Sales Op.	(2) Professionals
	(3) Self-Employed Businessmen	–	(3) Technicians and Associate Professionals
Middle	(4) Clerical and Sales Workers	(3) Administrative Support Occ., Including Clerical	(4) Clerical Support Workers
	(5) Craftsmen, Foremen and Kindred Workers	(5) Precision Production, Craft, and Repair Occ.	(5) Service and Sales workers
	(6) Operatives and Kindred Workers	(7) Machine Op., Assemblers, and Inspectors	(7) Craft and Related Trade Workers
	–	(6) Extractive and Precision Production Occ.	(8) Plant and Machine Op.s and Assemblers
Low	(7) Laborers, Service Workers and Farm Laborers	(4) Service, Farming, Forestry, and Fishing Occ.	(6) Skilled Agric., Forestry and Fishery Workers
	(8) Farmers and Farm Managers	(8) Transportation and Material Moving Occ., Handlers, Equipment Cleaners, Helpers, and Laborers	(9) Elementary Occ.
	(9) Miscellaneous (incl. Armed Services, Protective Workers etc.)	(9) Military Occ.	(0) Armed Forces Occ.
	(-) Not in Labor Force	(-) Not in Labor Force	(-) Not in Labor Force

and Dorn (2013). These categories are matched to their 1990 census classification analogues by the authors of this paper. This classification is available on request. We then aggregate all codes to the 1-digit level and apply the classification scheme outlined in Table A.3.

Additionally, wives of household heads report parental occupation codes in terms of 1970 codes at the 2-digit level in the 1976 wave. We aggregate them to the 1-digit level and apply the classification scheme outlined in Table A.3. Using the panel dimension of the PSID we replace missing values with the mode of all records for the respective individual.

EU-SILC. In EU-SILC, the 2011 wave reports occupation codes with reference to the ISCO-08 classification. We aggregate all codes to the 1-digit level and apply the classification scheme outlined in Table A.3. Respondents with missing information are dropped through list-wise deletion.

A.4.6 Other Circumstances

PSID. For the robustness checks presented in section 1.5 we construct two additional circumstance variables. First, the PSID collects the census region of upbringing for all individuals. Furthermore, we transform the resulting 4-category variable into three binary indicators. Second, the PSID reports the state of upbringing of both mother and father of individual respondents. We transform this variable into a binary variable indicating whether either the mother or the father had been raised in a foreign country. Using the panel dimension of the PSID we replace missing values in both variables with the mode of all records for the respective individual.

EU-SILC. For the robustness checks presented in section 1.5 we construct four additional circumstance variables. First, EU-SILC provides a 5-category variable indicating whether respondents at the age of 14 lived with i) both parents (or persons considered as parents), ii) father only (or person considered as a father), iii) mother only (or person considered as a mother), iv) in a private household without any parent, or v) in a collective household or institution. We transform this variable into a binary variable indicating whether individuals lived with both parents at the age of 14. Second, EU-SILC provides a categorical variable indicating the number of children in the household in which they lived at age 14. We transform this variable into a binary variable indicating whether individuals lived with less than 3 siblings at age 14. Third, EU-SILC provides a 6-category variable indicating whether the financial situation of the household in which respondents lived at the age of 14 was i) very bad, ii) bad, iii) moderately bad, iv) moderately good, v) good or vi) very good. We transform this variable into a binary variable indicating whether individuals lived in a household in which the situation was at least moderately good. Fourth, EU-SILC provides a 3-category variable indicating whether respondents at the age of 14 lived in i) owner-occupied housing, ii) as tenants or iii) in a household to which accommodation was provided for free. We transform this variable into a binary variable indicating whether individuals lived in owner-occupied housing. Respondents with missing information in any of these variables are dropped through list-wise deletion.

A.4.7 Individual Working Hours

PSID. PSID CNEF reports the total annual working hours of individuals. We replace missing hours information with zero if the respondent reports to be unemployed. In each year, we winsorize the resulting distribution from above at the 99th percentile.

EU-SILC. EU-SILC reports weekly working hours of individuals in their main and side jobs. We set hours to zero if the respondent reports to be unemployed, retired or otherwise inactive in the labor market. We add hours in the main and the side jobs to obtain total weekly working hours and multiply by 52 to obtain total annual working hours. In each year, we winsorize the resulting distribution from above at the 99th percentile.

A.4.8 Individual Education

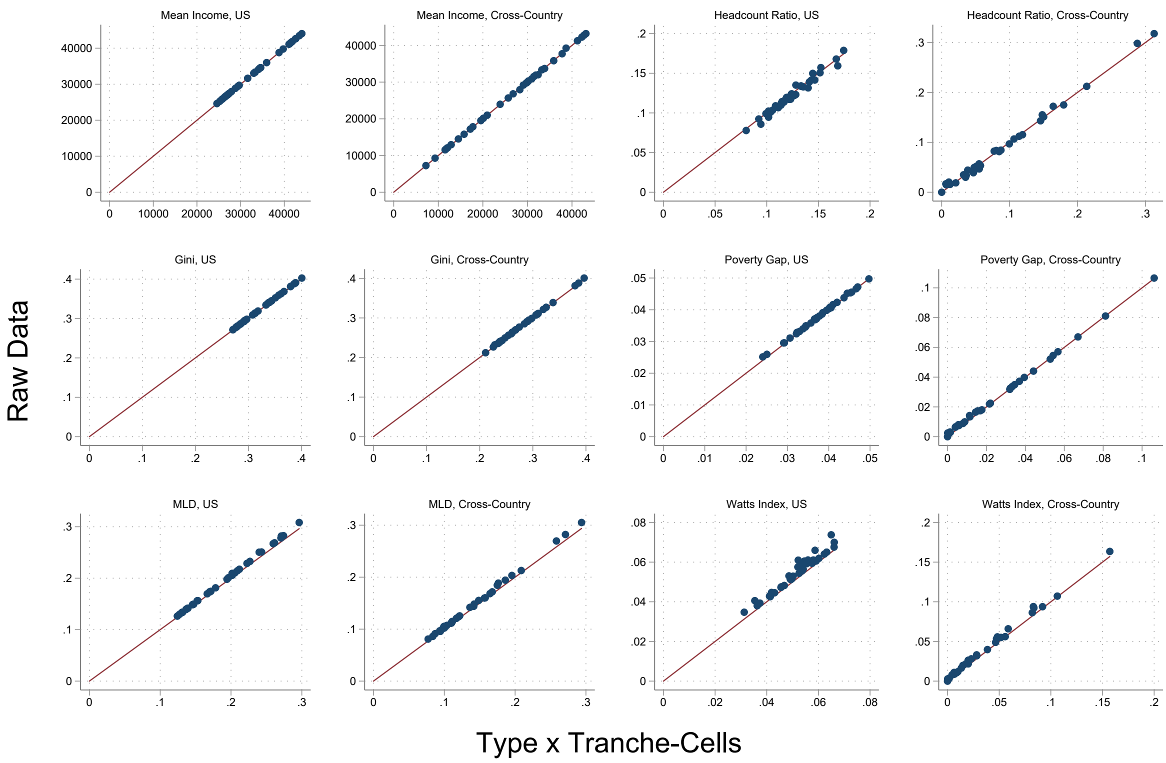
PSID. PSID CNEF reports individual educational attainment by total years of education. We map years of education into a 5-point categorical variable that corresponds to the ISCED-11 classification: (Pre-)Primary (1-6 years), Lower Secondary (7-11 years), Upper Secondary (12 years), Post-Secondary Non-Tertiary (13-14 years), Tertiary ($>$14 years). We replace missing information by the highest recorded education level from previous years. Since educational attainment cannot be downgraded we also replace lower educational attainments by the highest recorded education level from previous years.

EU-SILC. EU-SILC reports individual educational attainment in terms of the ISCED-11 classification. In view of small cell sizes we reduce the scale from 7 categories to 5 categories by merging Pre-Primary and Primary Education and First Stage Tertiary and Second Stage Tertiary Education. This merger corresponds to the 5-point categorical variable that we have coded for the PSID. Respondents with missing information are dropped through list-wise deletion.

A.4.9 Transformation to Type-Tranche Cells

In each country-year cell of our data we partition the population into a maximum of 36 circumstance types. These types are divided into 20 quantiles ordered by increasing incomes that identify Roemerian effort tranches. Since we use population weights, individual observations with high weights may span more than one effort tranche. To assure the existence of all effort tranches in every type, we duplicate the respective individual observations and divide their weight by two. We repeat this procedure until all type-effort cells are populated. We then collapse the data to the type-tranche level by replacing individual incomes and effort variables (individual education, individual working hours) by their respective cell average. Hence, each country-year cell of our data contains a maximum of 36 x 20 observations. In Figure A.3 we plot summary statistics of the raw distribution of our outcome variable against the same statistics calculated on the collapsed data. These statistics include the mean, the Gini coefficient, the mean log deviation, the poverty headcount ratio, the poverty gap ratio, as well as the Watts index. Results are presented separately for the US sample over time and the cross-country comparison sample. The closer the data points align to the 45 degree line, the smaller the information loss from collapsing the raw data to the type-tranche level.

FIGURE A.3 – Raw Data vs. Type x Tranche-Cells

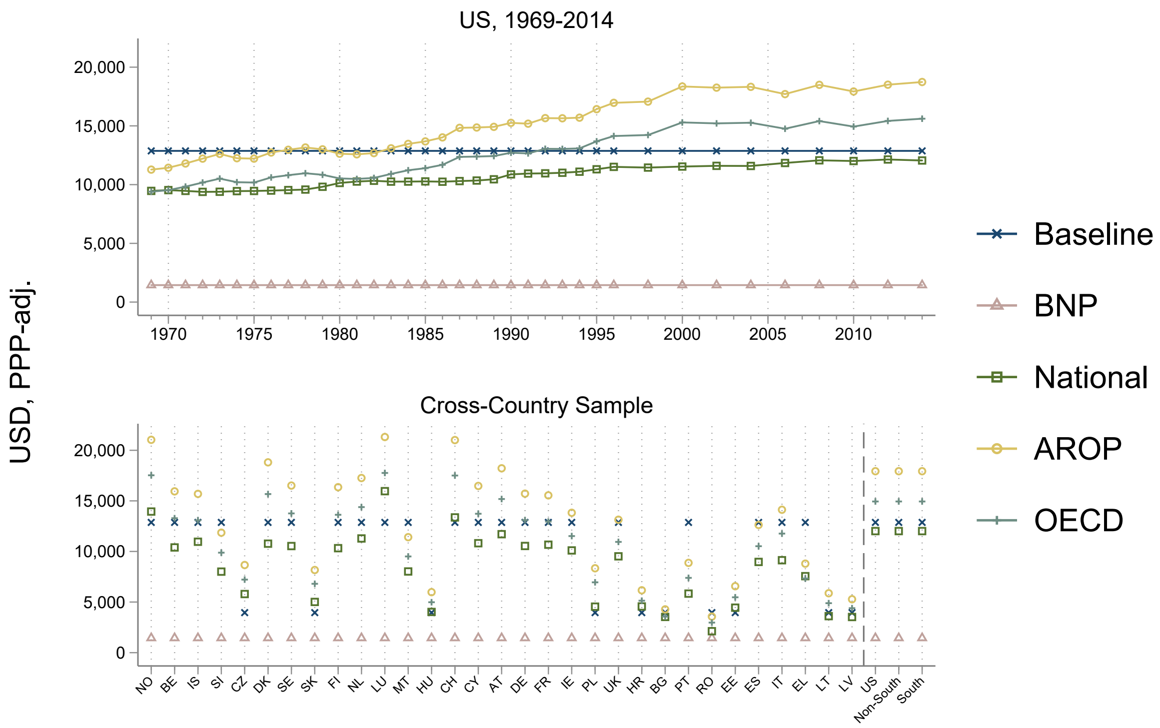


Data: PSID and EU-SILC.
Note: Own calculations. This figure plots standard measures of inequality and poverty estimated on the raw data against the corresponding estimates on data that is collapsed to type-tranche cells. The maroon line displays the 45 degree line. If inequality and poverty estimates on the raw data and the collapsed data were perfectly identical, all data points would align on the 45 degree line.

A.4.10 Poverty Lines

The PPP-adjusted US Dollar values of all poverty lines are displayed in Figure A.4.

FIGURE A.4 – Alternative Poverty Lines



Data: PSID, EU-SILC, EUROSTAT, US Census Bureau, and Allen (2017).
Note: Own calculations. This figure displays the value of alternative poverty thresholds $y_{\min}$ for each country-year cell in our data samples. The upper panel refers to the longitudinal US sample. The lower panel refers to the cross-country sample. All poverty lines are expressed in PPP-adjusted US Dollar (USD). Data points to the left of the vertical dashed line refer to the European country sample. Data points to the right of the vertical dashed line refer to the US and its census regions.

Baseline. Jolliffe and Prydz (2016) provide national poverty lines and average consumption expenditures per capita in PPP-adjusted US Dollar per day for a sample of 126 countries. With the exception of Malta and Cyprus all countries of our sample are covered in their data base. Based on average per capita consumption expenditures we divide the data sample into quintiles. We assign the median poverty line of each consumption expenditure quintile to the respective countries. The resulting five poverty lines are multiplied by 365 to obtain national poverty lines in terms of PPP-adjusted US Dollar per capita and year. Following the suggestion of van den Boom et al. (2015) we divide each poverty line by 0.7 to convert the poverty lines from per capita into adult-equivalent terms. In view of their high-income status we assign Malta and Cyprus the same poverty line as the countries from the highest consumption expenditure quintile.

National Poverty Line. For the US we retrieve the time series of the official poverty line for unrelated individuals under the age of 65 from the US Census Bureau and convert it into PPP-adjusted US Dollar using the conversion factors provided by the Penn World Tables (Feenstra et al., 2015). Similarly, we retrieve the official poverty lines for all European countries in 2010 from EUROSTAT. The poverty lines are provided in PPP-adjusted units already, requiring no further adjustment.

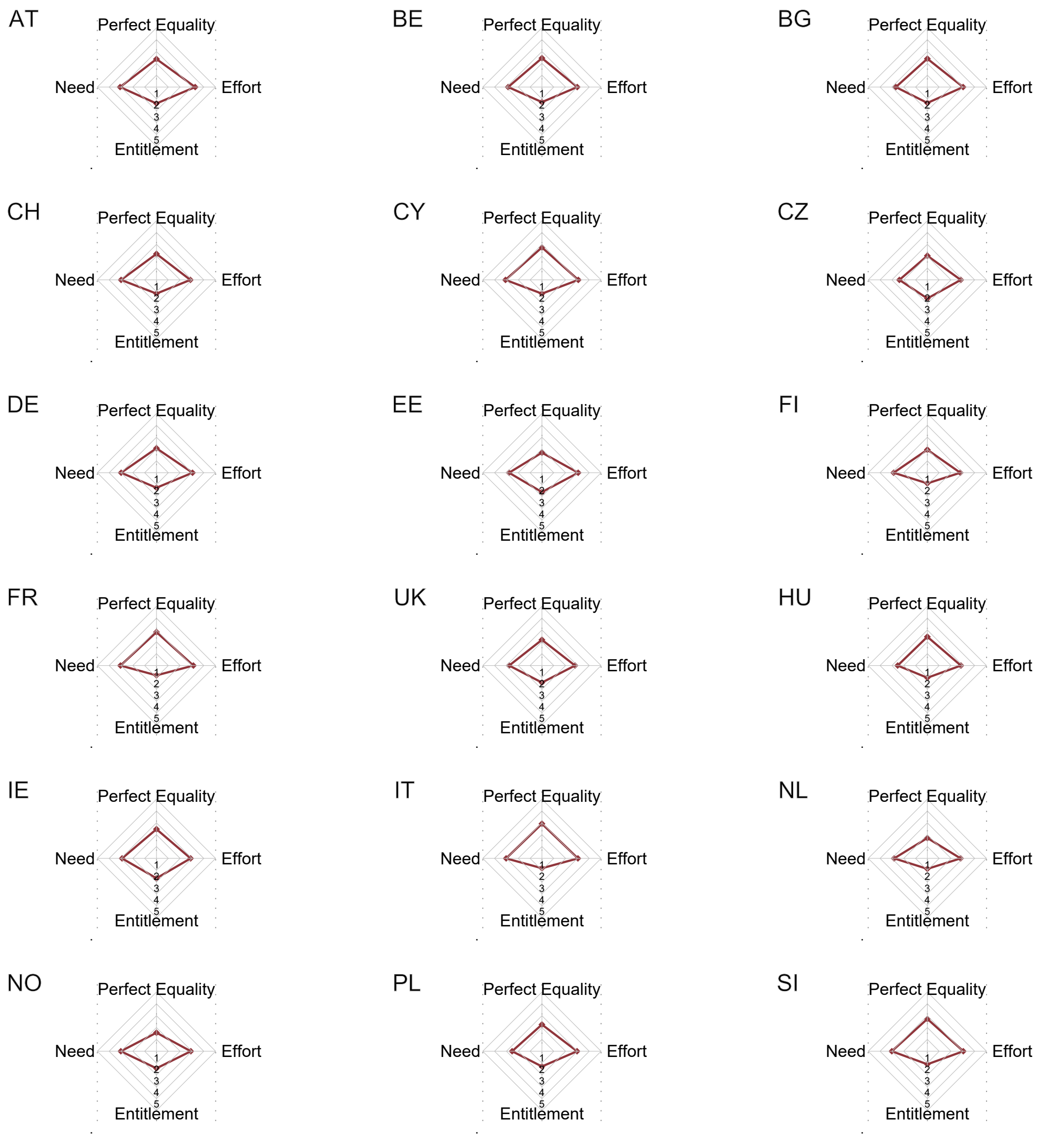
Basic Needs Poverty (BNP) Line. Allen (2017) provides basic needs adjusted poverty lines in PPP-adjusted US Dollar per day for four countries in our sample: Lithuania ($4.62), United Kingdom ($3.49), United States ($3.72) and France ($4.02). Taking the unweighted average across these poverty lines yields a value of $3.96 which we multiply by 365 to obtain the annual BNP line. We apply this BNP line to all countries and years in our sample.

At-Risk-of-Poverty (AROP) Line. In each country-year cell we calculate the median of the distribution of disposable household income (see above). The AROP line is then drawn at 60% of the respective country-year-specific median.

OECD Poverty Line. The OECD poverty line is calculated as the AROP line. However, the OECD line is drawn at 50% of the respective country-year-specific median.

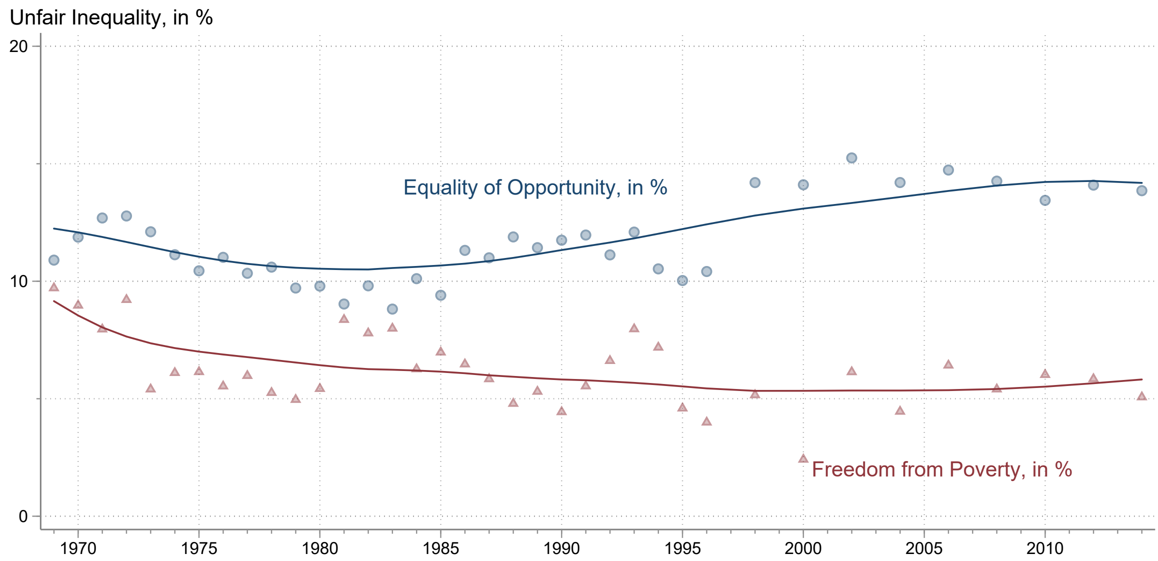
Appendix A.5 Supplementary Figures

FIGURE A.5 – Normative Preferences



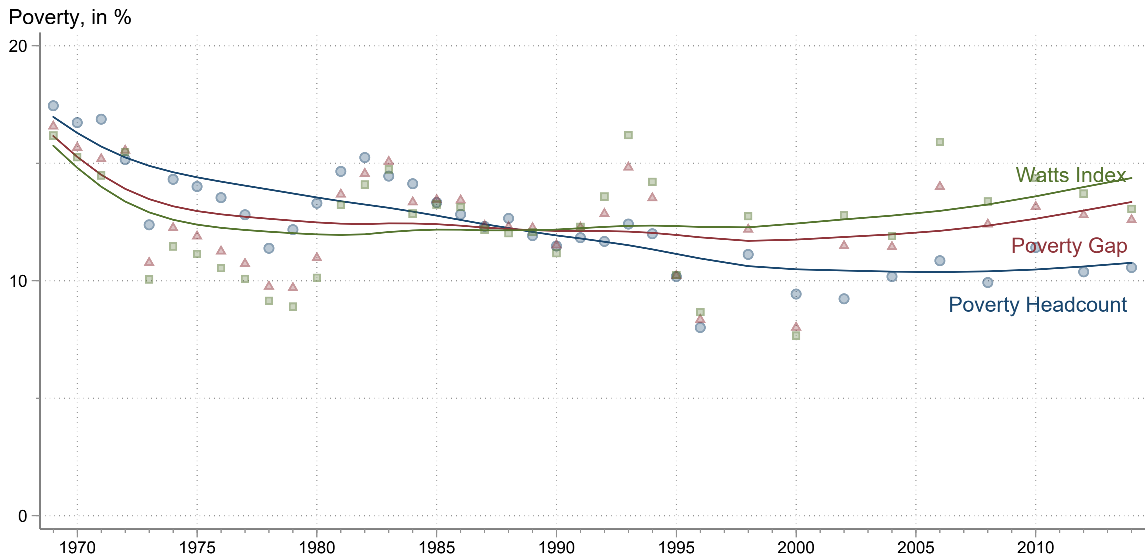
Data: European Social Survey (2018).
Note: Own calculations. This figure displays the average support for four different principles of justice in 18 of our sample countries. Answers are given on a 5-point Likert scale ranging from 1 (Agree Strongly) to 5 (Disagree Strongly). We invert the scale such that higher values indicate stronger support. The questions for the different dimensions are based on Hülle et al. (2018) and read as as follows. i) Perfect Equality: A society is fair when income and wealth are equally distributed among all people. ii) Effort: A society is fair when hard-working people earn more than others. iii) Need: A society is fair when it takes care of those who are poor and in need regardless of what they give back to society. iv) Entitlement: A society is fair when people from families with high social status enjoy privileges in their lives.

Figure A.6 – Decomposition by Principle (US), 1969-2014



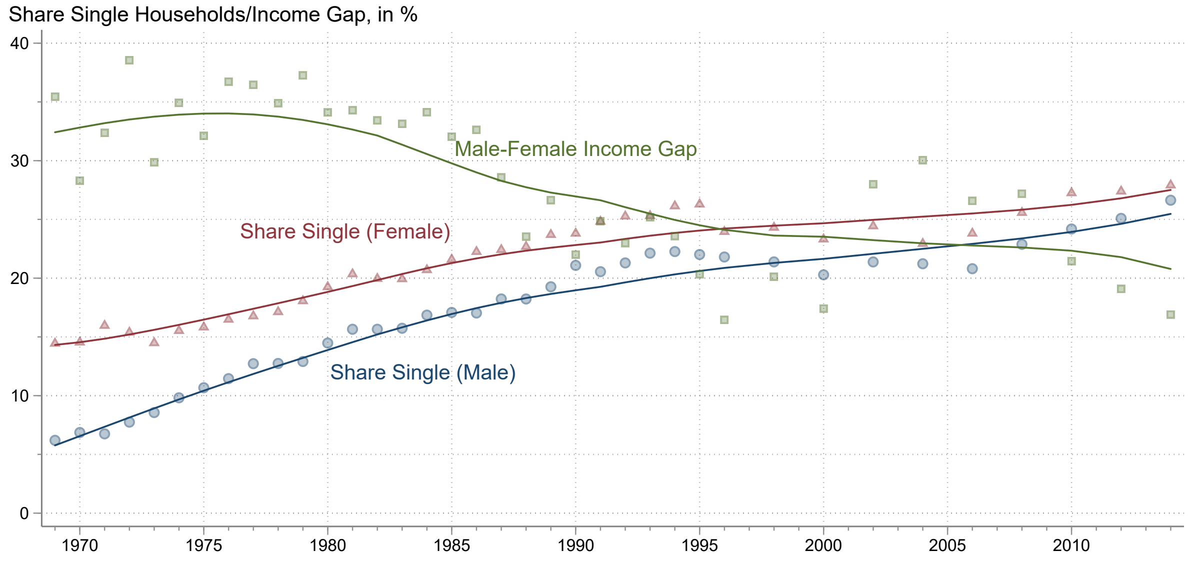
Data: PSID.
Note: Own calculations. This figure displays the contribution of EOp and FfP to total inequality in the US over the period 1969-2014. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The decomposition is based on the Shapley value procedure proposed in Shorrocks (2012). The solid lines display Lowess smoothed time trends where each data point is constructed using 80% of all data points (Bandwidth 0.8).

Figure A.7 – Poverty in the US, 1969-2014



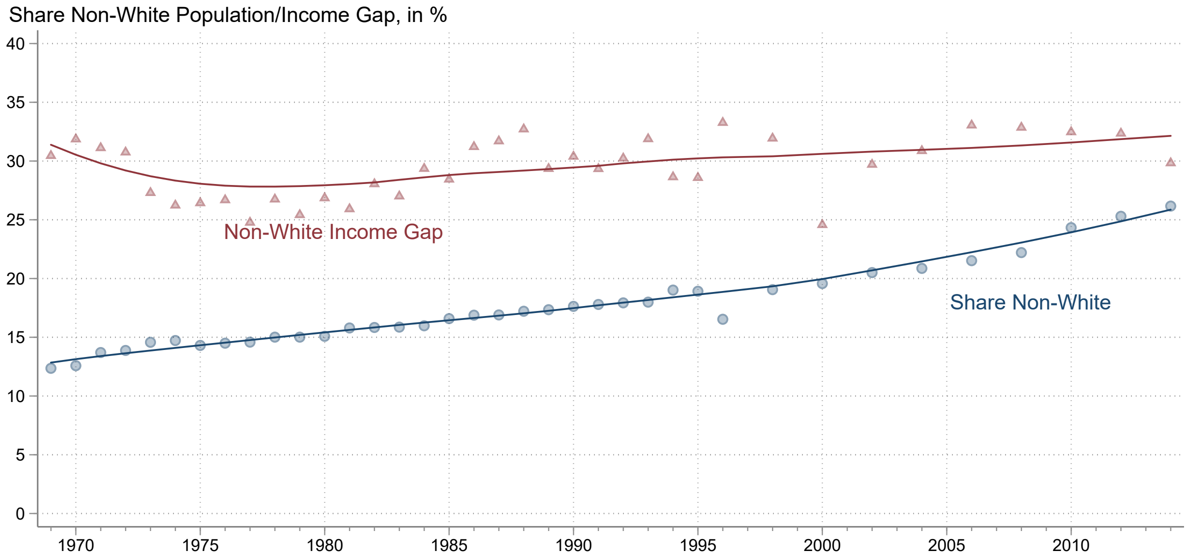
Data: PSID.
Note: Own calculations. This figure displays the development of poverty in the US over the period 1969-2014 according to different poverty measures. Poverty statistics are displayed in units of the poverty headcount ratio (in %): All data points are rescaled by multiplying with the cross-year mean of the poverty headcount ratio and dividing by the cross-year mean of the respective poverty measure. The solid lines display Lowess smoothed time trends where each data point is constructed using 80% of all data points (Bandwidth 0.8).

Figure A.8 – Income Gaps and Population Shares of Single Households in the US, 1969-2014



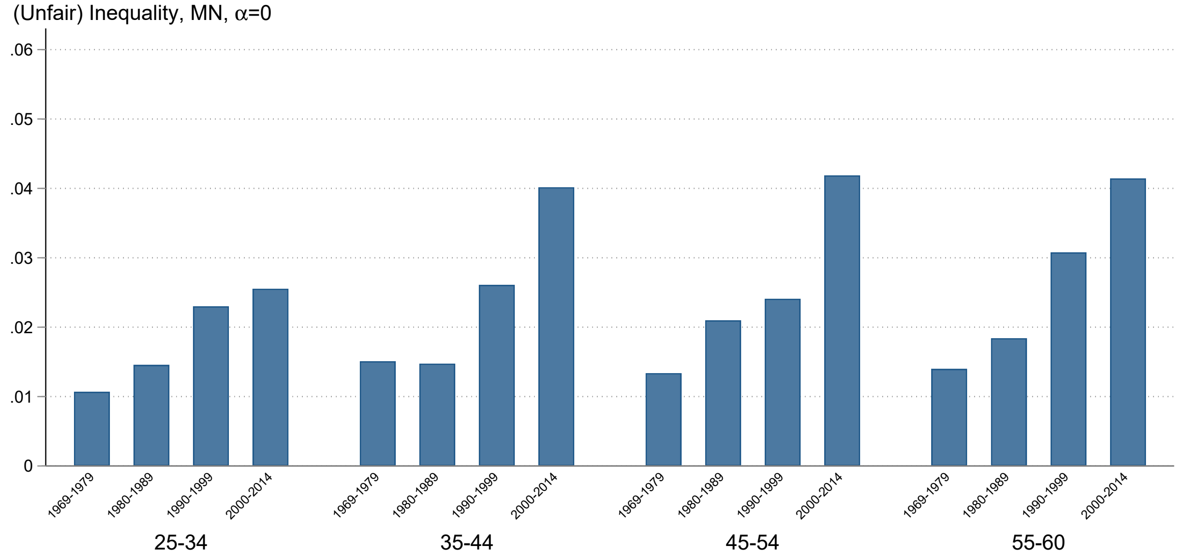
Data: PSID.
Note: Own calculations. This figure displays the share of females (males) living in households with only one adult present (in %) and the female-male income gap among those households. The female-male income gap is calculated as $\left(1-\frac{\mu_{ft}}{\mu_{mt}}\right)*100$ where μ_{ft} (μ_{mt}) is the average disposable household income of females (males) living in households with only one adult present in year t. The solid lines display Lowess smoothed time trends where each data point is constructed using 80% of all data points (Bandwidth 0.8).

Figure A.9 – (Non-)White Income Gaps and Population Shares in the US, 1969-2014



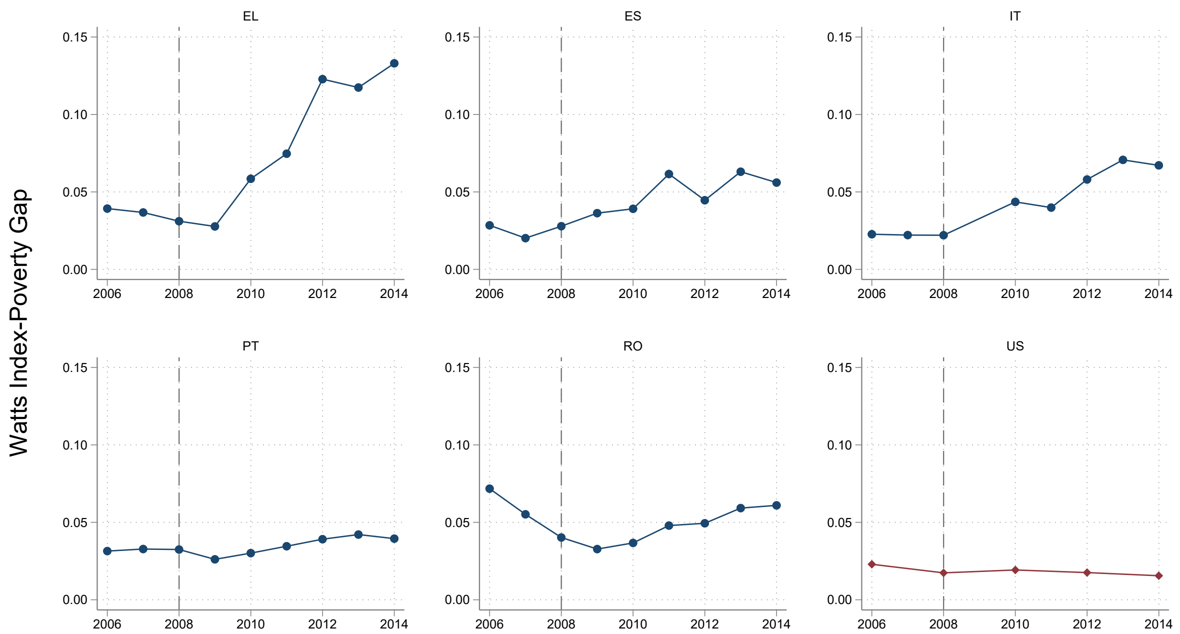
Data: PSID.
Note: Own calculations. This figure displays the share of individuals classified as non-white/Hispanic (in %) and the average income gap in comparison to individuals classified as white/non-Hispanic. The income gap is calculated as $\left(1-\frac{\mu_{nt}}{\mu_{wt}}\right)*100$ where μ_{nt} (μ_{wt}) is the average disposable household income of the non-white/Hispanic (white/non-Hispanic) population in year t. The solid lines display Lowess smoothed time trends where each data point is constructed using 80% of all data points (Bandwidth 0.8).

Figure A.10 – Social Mobility in the US, 1969-2014



Data: PSID.
Note: Own calculations. This figure displays estimates of unfair inequality considering parental education and parental occupation as the only relevant circumstance characteristics while abstracting from the concern for FfP. The calculation is conducted for each age bin-year-cell and then aggregated to the indicated year bins by taking unweighted averages. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality.

Figure A.11 – Poverty, 2006-2014



Data: PSID and EU-SILC.
Note: Own calculations. This figure displays the development of FfP as measured by the difference between the Watts index and the poverty gap ratio over the period 2006-2014. The selected countries represent the six most unfair societies of our cross-country sample in 2010. The vertical dashed line marks the starting year of the global financial crisis.

Appendix A.6 Supplementary Tables

Table A.4 – Descriptive Statistics US, 1969-2014

	Income	Circumstances				Efforts		Poverty
		Male	Race	Educ.	Occ.	Hours	Educ.	
1969	24,636	0.53	0.88	1.37	1.73	1,575	2.98	0.18
1970	25,254	0.53	0.87	1.39	1.74	1,551	3.01	0.17
1971	25,718	0.52	0.86	1.40	1.75	1,537	3.03	0.16
1972	26,597	0.52	0.86	1.42	1.75	1,557	3.06	0.15
1973	27,110	0.48	0.85	1.61	1.77	1,519	3.10	0.12
1974	26,689	0.48	0.85	1.64	1.78	1,485	3.26	0.14
1975	26,342	0.48	0.86	1.70	1.79	1,459	3.30	0.13
1976	27,392	0.48	0.86	1.72	1.80	1,478	3.33	0.13
1977	27,093	0.48	0.85	1.74	1.81	1,507	3.36	0.12
1978	27,481	0.48	0.85	1.75	1.82	1,548	3.37	0.11
1979	27,105	0.48	0.85	1.77	1.83	1,552	3.39	0.12
1980	26,668	0.48	0.85	1.78	1.83	1,553	3.41	0.13
1981	25,934	0.48	0.84	1.81	1.85	1,553	3.43	0.14
1982	26,854	0.48	0.84	1.83	1.86	1,531	3.45	0.16
1983	27,968	0.48	0.84	1.85	1.87	1,551	3.47	0.15
1984	28,854	0.48	0.84	1.90	1.89	1,642	3.62	0.14
1985	29,413	0.48	0.83	1.92	1.91	1,647	3.65	0.13
1986	29,704	0.47	0.83	1.94	1.92	1,647	3.66	0.14
1987	31,644	0.48	0.83	1.96	1.93	1,669	3.68	0.12
1988	33,380	0.48	0.83	1.99	1.94	1,689	3.70	0.12
1989	33,061	0.48	0.83	2.00	1.95	1,704	3.71	0.12
1990	34,134	0.48	0.82	2.02	1.96	1,719	3.72	0.11
1991	33,301	0.48	0.82	2.03	1.97	1,693	3.73	0.12
1992	34,607	0.48	0.82	2.06	1.98	1,662	3.74	0.11
1993	34,567	0.48	0.82	2.08	2.00	1,671	3.75	0.12
1994	34,478	0.49	0.81	2.10	2.02	1,699	3.73	0.12
1995	36,012	0.49	0.81	2.12	2.03	1,748	3.75	0.09
1996	38,791	0.49	0.83	2.25	2.09	1,780	3.80	0.08
1998	39,776	0.49	0.81	2.21	2.11	1,808	3.76	0.11
2000	41,579	0.49	0.80	2.23	2.13	1,791	3.75	0.09
2002	41,104	0.49	0.79	2.23	2.15	1,755	3.75	0.09
2004	42,586	0.49	0.79	2.22	2.14	1,750	3.82	0.10
2006	44,061	0.48	0.78	2.23	2.16	1,735	3.83	0.11
2008	43,496	0.48	0.78	2.24	2.17	1,681	3.86	0.10
2010	41,268	0.48	0.76	2.25	2.19	1,606	4.00	0.11
2012	41,874	0.48	0.75	2.27	2.21	1,659	4.03	0.10
2014	42,675	0.48	0.74	2.29	2.22	1,703	4.05	0.10

Data: PSID.
Note: Own calculations. This table displays descriptive statistics for the longitudinal US sample. *Male* displays the share of males. *Race* displays the share of white/non-Hispanics. The circumstance variables *Educ.* (*Occ.*) show the average education (occupation) level of the parent with the highest education (occupation) status measured on a 3-point scale. *Hours* show the average working hours per year. The effort variable *Educ.* shows the average education level measured on a 6-point scale. *Poverty* shows the share of people below the baseline poverty line. Further detail on the construction of all variables is disclosed in Supplementary Material A.4.

Table A.5 – Descriptive Statistics Cross-Country Sample, 2010

	Income	Circumstances				Efforts		Poverty
		Male	Mig./ Race	Educ.	Occ.	Hours	Educ.	
AT	35,829	0.50	0.79	1.80	1.94	1,599	3.37	0.03
BE	31,917	0.50	0.84	1.79	2.29	1,574	3.65	0.03
BG	9,295	0.50	1.00	1.70	1.94	1,631	3.27	0.12
CH	42,784	0.49	0.69	1.89	2.27	1,710	3.60	0.02
CY	33,336	0.48	0.78	1.49	1.92	1,671	3.38	0.04
CZ	17,836	0.44	0.96	1.56	2.24	1,695	3.32	0.00
DE	30,311	0.50	0.87	2.05	2.22	1,597	3.50	0.08
DK	33,699	0.52	0.94	2.04	2.31	1,623	3.66	0.04
EE	14,526	0.48	0.87	2.03	2.30	1,596	3.69	0.05
EL	19,526	0.50	0.89	1.38	1.81	1,311	3.26	0.32
ES	25,679	0.51	0.84	1.32	1.92	1,392	3.13	0.17
FI	30,887	0.52	0.97	1.85	1.85	1,549	3.76	0.05
FR	31,520	0.49	0.90	1.40	2.01	1,616	3.40	0.05
HR	12,952	0.50	0.89	1.61	1.95	1,299	3.15	0.05
HU	12,098	0.48	0.99	1.55	2.00	1,425	3.30	0.02
IE	29,921	0.42	0.79	1.74	1.97	1,167	3.70	0.08
IS	27,941	0.51	0.89	1.90	2.26	1,828	3.53	0.05
IT	26,813	0.50	0.88	1.32	1.99	1,435	2.91	0.16
LT	11,848	0.48	0.94	1.66	1.94	1,528	3.83	0.08
LU	43,214	0.50	0.49	1.66	2.18	1,595	3.07	0.02
LV	11,545	0.47	0.88	1.83	2.14	1,480	3.51	0.11
MT	23,952	0.50	0.95	1.37	1.99	1,420	2.68	0.15
NL	32,002	0.50	0.88	1.91	2.34	1,450	3.61	0.03
NO	37,728	0.54	0.93	2.15	2.38	1,718	3.62	0.02
PL	17,200	0.47	1.00	1.70	1.90	1,622	3.35	0.02
PT	20,140	0.48	0.91	1.15	1.93	1,574	2.31	0.30
RO	7,264	0.50	1.00	1.25	1.57	1,602	3.22	0.21
SE	29,750	0.55	0.91	1.99	1.00	1,526	3.73	0.06
SI	20,999	0.51	0.88	1.54	2.05	1,598	3.37	0.17
SK	15,795	0.49	0.99	1.78	2.13	1,667	3.42	0.02
UK	29,198	0.47	0.87	1.71	2.34	1,596	3.81	0.08
US	41,268	0.48	0.76	2.25	2.19	1,606	4.00	0.11
US (Non-South)	42,268	0.48	0.80	2.28	2.20	1,615	4.00	0.10
US (South)	39,261	0.48	0.68	2.21	2.16	1,589	4.00	0.14

Data: PSID and EU-SILC.
Note: Own calculations. This table displays descriptive statistics for the cross-country sample. *Male* displays the share of males. *Mig./Race* displays the share of people born in their current country of residence (white/non-Hispanics) in the European (US) sample. The circumstance variables *Educ.* (*Occ.*) show the average education (occupation) level of the parent with the highest education (occupation) status measured on a 3-point scale. *Hours* show the average working hours per year. The effort variable *Educ.* shows the average education level measured on a 6-point scale. *Poverty* shows the share of people below the baseline poverty line. Further detail on the construction of all variables is disclosed in Supplementary Material A.4.

Appendix A.7 Decorrelating Ω and Θ

First, we regress the outcome of interest (y_i^e) on a vector of type fixed effects ($\delta_{\mathcal{T}(\omega)}$), a categorical variable for educational attainment (educ$_i$) and annual working hours (hours$_i$):

$$y_i^e = \delta_{\mathcal{T}(\omega)} + \beta_1 \text{hours}_i + \beta_2 \text{educ}_i + \epsilon_i. \tag{66}$$

Second, we construct a counterfactual distribution $\tilde{Y}^e$ by adding residuals to the estimated type averages net of their correlation with the considered effort variables:

$$\tilde{y}_i^e = \hat{\delta}_{\mathcal{T}(\omega)} + \hat{\epsilon}_i. \tag{67}$$

Third, we use $\tilde{Y}^e$ as an input to the construction of the reference distribution Y^r (see equation 15) and repeat our analysis according to the usual steps.

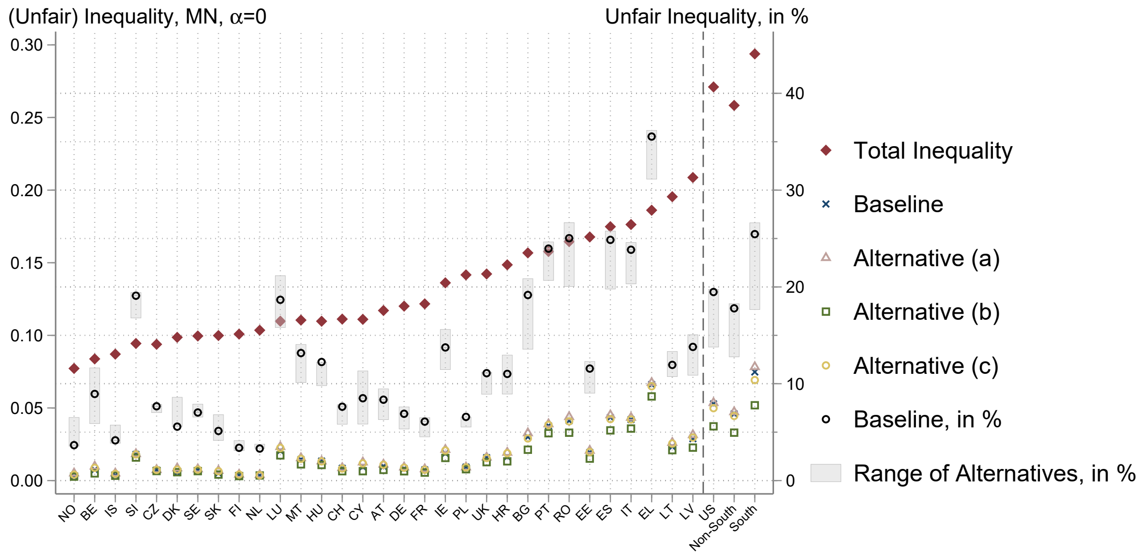
To develop an intuition for this procedure consider the polar case in which circumstances influenced outcomes only indirectly through their impact on education and working hours. Then $\hat{\delta}_{\mathcal{T}(\omega)} = \mu^e,\ \forall\, \omega \in \Omega$ and our measure of unfairness collapses to the case in which we abstracted from the concern for EOp altogether (see equations (21) and (22)). This is precisely what the normative stance of Barry (2005) requires.

Reversely, consider the polar case in which there is zero correlation between circumstances on the one hand, and education and working hours on the other hand. In this case circumstances influence outcomes only directly without affecting intermediate outcomes that are partially under the control of individuals. Then $\hat{\delta}_{\mathcal{T}(\omega)} = \mu^e_{\mathcal{T}(\omega)},\ \forall\, \omega \in \Omega$, and we would recover exactly our baseline measure of unfair inequality (see equations (15) and (16)).[4]

[4] Another way to think about this procedure is that the alternative normative stance of Barry (2005) does not require perfect equalization of type means tout court, but perfect equalization of type means once they are cleaned from effort influence.

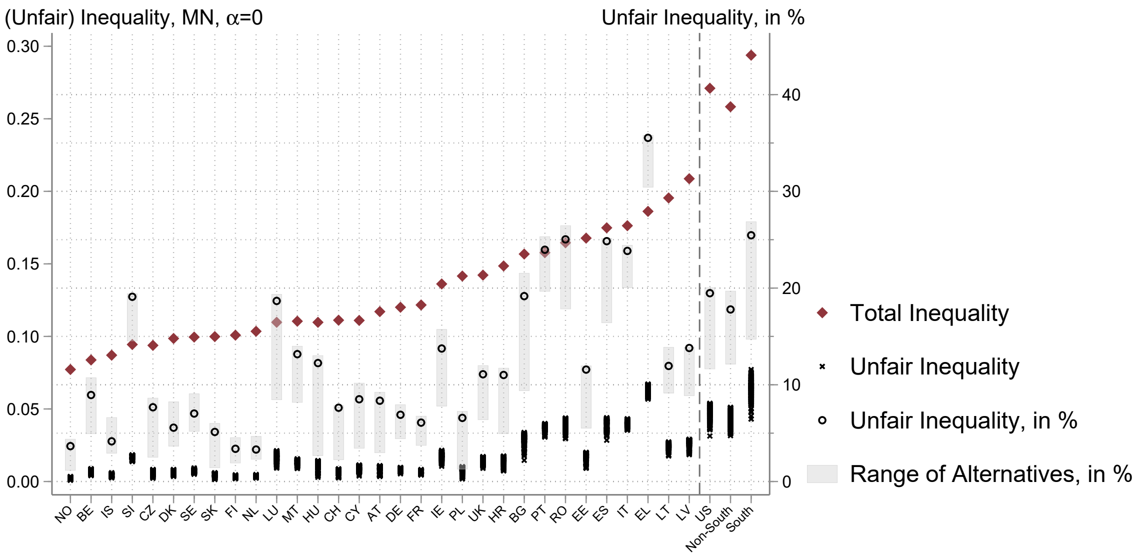
Appendix A.8 Sensitivity Analysis Cross-Country Comparison

FIGURE A.12 – Unfair Inequality across Countries, 2010, Alternative Norm Distributions



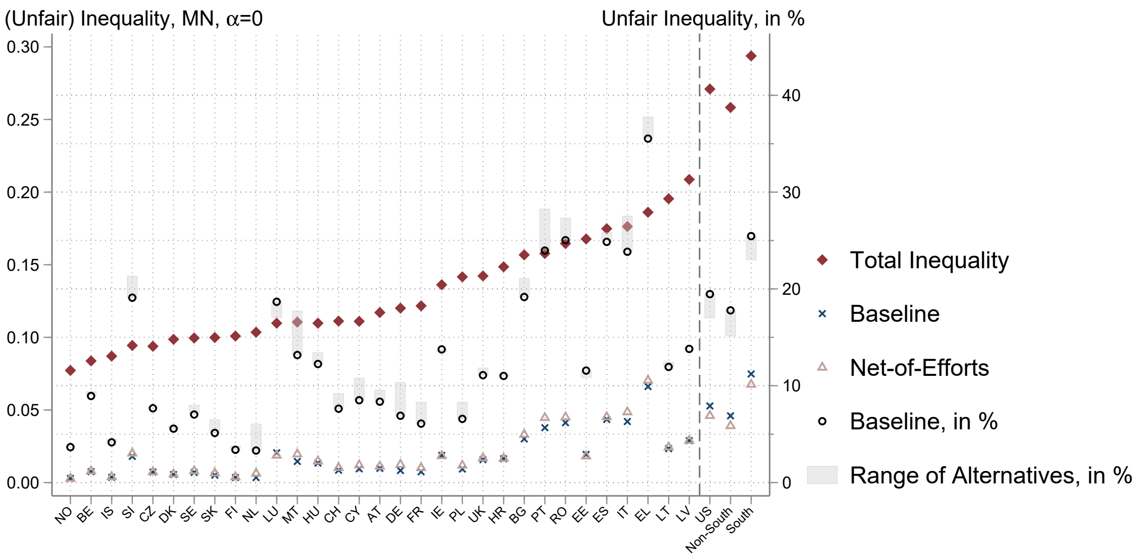
Data: PSID and EU-SILC.
Note: Own calculations. This figure displays cross-country differences in (unfair) inequality in 2010 according to the alternative norm distributions outlined in Table 1.1. Data points to the left of the vertical dashed line refer to the European country sample. Data points to the right of the vertical dashed line refer to the US and its census regions. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The gray area shows the range of unfair inequality in percent (in %) of total inequality depending on the alternative measurement specifications.

Figure A.13 – Unfair Inequality across Countries, 2010, Alternative Circumstance Sets



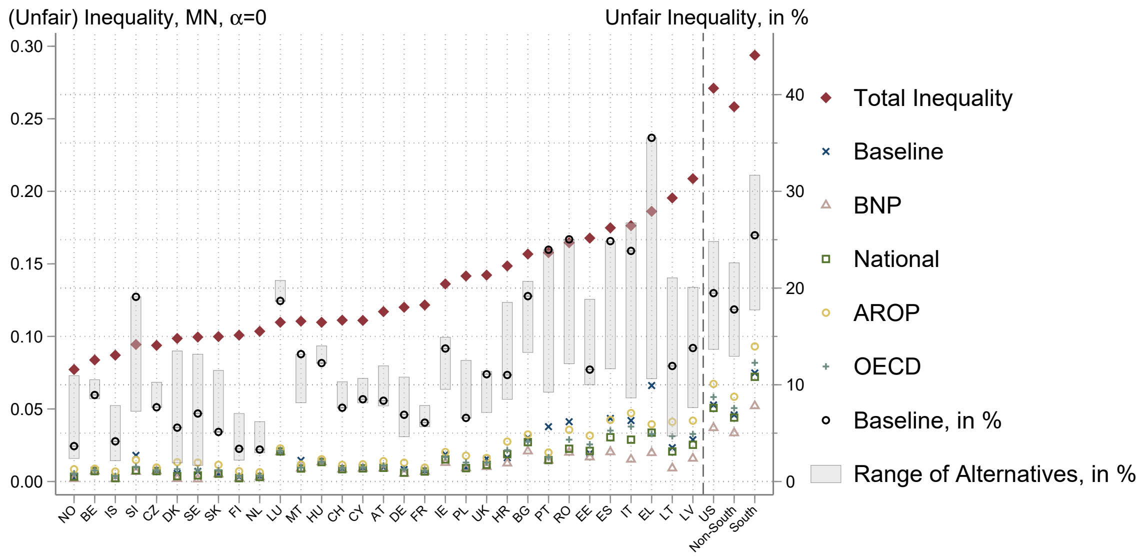
Data: PSID and EU-SILC.
Note: Own calculations. This figure displays cross-country differences in (unfair) inequality in 2010 according to alternative specifications of the circumstance set Ω. Data points to the left of the vertical dashed line refer to the European country sample. Data points to the right of the vertical dashed line refer to the US and its census regions. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The gray area shows the range of unfair inequality in percent (in %) of total inequality depending on the alternative measurement specifications.

Figure A.14 – Unfair Inequality across Countries, 2010, Accounting for Preferences



Data: PSID and EU-SILC.
Note: Own calculations. This figure displays cross-country differences in (unfair) inequality in 2010 according to alternative treatments of the correlation between the effort set Θ and the circumstance set Ω. Data points to the left of the vertical dashed line refer to the European country sample. Data points to the right of the vertical dashed line refer to the US and its census regions. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The gray area shows the range of unfair inequality in percent (in %) of total inequality depending on the alternative measurement specifications.

FIGURE A.15 – Unfair Inequality across Countries, 2010, Alternative Minimum Thresholds



Data: PSID and EU-SILC.
Note: Own calculations. This figure displays cross-country differences in (unfair) inequality in 2010 according to alternative specifications of the poverty threshold y_{min}. Data points to the left of the vertical dashed line refer to the European country sample. Data points to the right of the vertical dashed line refer to the US and its census regions. (Unfair) inequality is calculated based on the divergence measure proposed by Magdalou and Nock (2011) with $\alpha = 0$ (MN, $\alpha = 0$) which corresponds to the MLD for total inequality. Relative measures of unfair inequality are expressed in percent (in %) of total inequality. The gray area shows the range of unfair inequality in percent (in %) of total inequality depending on the alternative measurement specifications. The construction of the alternative minimum thresholds is discussed in Supplementary Material A.4.

TABLE A.6 – Rank Correlation across Countries, 2010

	Magdalou and Nock			Cowell			Almås et al.
	$\alpha = 0$	$\alpha = 1$	$\alpha = 2$	$\alpha = 0$	$\alpha = 1$	$\alpha = 2$	
Magdalou and Nock							
$\alpha = 0$	1.00	.	.	.	.	.	.
$\alpha = 1$	0.98	1.00	.	.	.	.	.
$\alpha = 2$	0.95	0.99	1.00	.	.	.	.
Cowell							
$\alpha = 0$	0.99	0.99	0.98	1.00	.	.	.
$\alpha = 1$	0.98	1.00	0.99	0.99	1.00	.	.
$\alpha = 2$	0.97	1.00	0.99	0.99	1.00	1.00	.
Almås et al.							
	0.95	0.98	0.98	0.97	0.98	0.98	1.00

Data: PSID and EU-SILC.
Note: Own calculations. This table displays rank correlations for unfair inequality across countries based on different divergence measures. Unfair inequality is calculated based on the divergence measures proposed by Magdalou and Nock (2011), Cowell (1985), and Almås et al. (2011).

2 The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees and Forests

This chapter is based on the paper "The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees and Forests" and has been co-authored with Paolo Brunori and Daniel Gerszon Mahler.

2.1 Introduction

Equality of opportunity is an important ideal of distributive justice. It has widespread support in the general public (Alesina et al., 2018; Cappelen et al., 2007) and its realization has been identified as an important goal of public policy intervention (Chetty et al., 2016b; Corak, 2013). In spite of its popularity, providing empirical estimates of equality of opportunity is notoriously difficult. Next to normative dissent about the precise factors that should be viewed as contributing to unequal opportunities, current approaches to estimate inequality of opportunity are encumbered by ad-hoc model selection that lead researchers to over- or underestimate inequality of opportunity.

In this paper we propose the use of machine learning methods to overcome the issue of ad-hoc model selection. Machine learning methods allow for flexible models of how unequal opportunities come about while imposing statistical discipline through criteria of out-of-sample replicability. These features serve to establish inequality of opportunity estimates that are less prone to upward or downward bias. For example, in comparison to our preferred method, current estimation approaches overestimate inequality of opportunity in Scandinavian countries by close to 300%. While these figures may inform policy debates about inclusive institutions, they are the result of overfitted estimation models that fail to replicate in independent samples of the same underlying population. This example illustrates that the choice of appropriate model specifications is of great importance for the analysis of institutional configurations and the ensuing policy debate.

The empirical literature on the measurement of unequal opportunities has been flourishing since John Roemer's (1998) seminal contribution, *Equality of Opportunity*. At the heart of Roemer's formulation is the idea that individual outcomes are determined by two sorts of factors: those factors over which individuals have control, which he calls *effort*, and those factors for which individuals cannot be held responsible, which he calls *circumstances*. While outcome differences due to effort exertion are morally permissible, differences due to circumstances are inequitable and call for compensation.[1] Grounded on this distinction, inequality of opportunity measures quantify the extent to which individual outcomes are determined by circumstance characteristics. In particular, inequality of opportunity is frequently measured by using a set of circumstances to predict an outcome of interest and calculating inequality in the predicted outcomes: the more predicted outcomes diverge, the more circumstances beyond individual control influence outcomes, and the more inequality of opportunity there is.

In spite of their policy relevance, current approaches to estimate inequality of opportunity suffer from biases that are the consequence of critical choices in model selection. First, researchers have to decide which circumstance variables to consider for estimation.[2] The challenge of this task grows with the increasing availability of high-quality datasets that provide very detailed information with respect to individual circumstances (Björklund et al., 2012a; Hufe et al., 2017). On the one hand, discarding relevant circumstances from the estimation model limits the explanatory scope of circumstances and leads to downward biased estimates of inequality of opportunity (Ferreira and Gignoux, 2011). On the other hand, including too many circumstances overfits the data and leads to upward biased estimates of inequality of opportunity (Brunori et al., 2019). Second, researchers must choose the functional form according to which circumstances co-produce the outcome of interest. For example, it is a well-established finding that the influence of similar child care arrangements on various life outcomes varies strongly by biological sex (Felfe and Lalive, 2018; García et al., 2018). In contrast to such evidence, many empirical applications presume that the effect of circumstances on individual outcomes is log-linear and additive while abstracting from

[1] The distinction between circumstances and efforts underpins many prominent literature branches in economics such as the ones on intergenerational mobility (Chetty et al., 2014b,c), the gender pay gap (Blau and Kahn, 2017) and racial differences (Kreisman and Rangel, 2015). For different notions of equality of opportunity, see Arneson (2018).

[2] Roemer does not provide a fixed list of circumstance variables. Instead he suggests that the set of circumstances should evolve from a political process (Roemer and Trannoy, 2015). In empirical implementations typical circumstances include biological sex, socioeconomic background and race.

possible interaction effects (Bourguignon et al., 2007; Ferreira and Gignoux, 2011). On the one hand, restrictive functional form assumptions limit the ability of circumstances to explain variation in the outcome of interest and thus force another downward bias on inequality of opportunity estimates. On the other hand, limitations in the available degrees of freedom may prove a statistically meaningful estimation of complex models with many parameters infeasible.

This discussion highlights the non-trivial challenge of selecting the appropriate model for estimating inequality of opportunity. Researchers must balance different sources of bias while avoiding ad-hoc solutions. While this task is daunting for the individual researcher, it is a standard application for machine learning algorithms that are designed to make out-of-sample predictions of a dependent variable based on a number of observable predictors. In this paper, we use conditional inference regression trees and forests to estimate inequality of opportunity (Hothorn et al., 2006). Introduced and popularized by Breiman et al. (1984), Breiman (2001), and Morgan and Sonquist (1963), regression trees and forests belong to a set of machine learning methods that is increasingly integrated into the statistical toolkit of economists (Athey, 2018; Mullainathan and Spiess, 2017; Varian, 2014). By drawing on a clear-cut algorithm, they obtain predictions without assumptions about which and how circumstances interact in shaping individual opportunities. Hence, the model specification is no longer a judgment call of the researcher but an outcome of data analysis. As a consequence they cushion downward bias by flexibly accommodating different ways of how circumstance characteristics shape the distribution of outcomes. Moreover, the conditional inference algorithm branches trees (and constructs forests) by a sequence of hypothesis tests that prevents the inclusion of noisy circumstance parameters. This reduces the potential for upward biased estimates of inequality of opportunity through model overfitting. Hence, regression trees and forests address the detrimental consequences of ad-hoc model selection in a way that is sensitive to both upward and downward bias.

To showcase the advantages of regression trees and forests we compare them to existing estimation approaches in a cross-sectional dataset of 31 European countries. We demonstrate that current estimation approaches overfit (underfit) the data which in turn leads to upward (downward) biased estimates of inequality of opportunity. These biases are sizable. For example, some standard methods overestimate inequality of opportunity in Scandinavian countries by close to 300%, whereas they underestimate the extent of inequality of opportunity in Germany by more than 40%. Hence, cross-country comparisons based on standard estimation approaches yield misleading recommendations with respect to the need for policy intervention in different societies. We illustrate how regression trees and forests can be used to analyze opportunity structures in different societies. We find that mothers' education and occupation are the most important predictors of children's income in Eastern Europe, while in Western/Southern Europe fathers' occupation and education are most important, and in Northern Europe area of birth is most important. Although we are careful to highlight the non-causal nature of our estimates, such analyses provide useful starting points for policymakers to target areas for opportunity equalizing reforms.

In a parallel paper, J. Blundell and Risa (2019) apply machine learning methods to the estimation of intergenerational mobility – a literature in which similar issues of model selection arise.[3] In particular, they use machine learning methods to validate rank-rank estimates of intergenerational mobility against an extended set of child circumstances to assess the completeness of the prevalent intergenerational mobility approach as a measure of equal opportunities. In contrast to their work, we directly estimate inequality of opportunity statistics. As a consequence, our focus is less on the downward bias that follows from focusing on one circumstance characteristic only (e.g. parental income) but on balancing both downward and upward bias if the set of available circumstances is large in relation to a given sample size.

The remainder of this paper is organized as follows: section 2.2 gives a brief introduction to current empirical approaches in the literature on inequality of opportunity. Section 2.3 introduces conditional inference regression trees and forests, and illustrates how to use them in the context of inequality of opportunity estimations. An empirical illustration based on the EU Survey of Income and Living Conditions is contained in section 2.4. In this section we also highlight the particular advantages of tree and forest-based estimation methods by comparing them to the prevalent estimation approaches in the literature. Lastly, section 3.6 concludes.

[3] These issues include the influence of non-linearities along the parental distribution (Björklund et al., 2012b; Corak and Piraino, 2011) and the question of whether intergenerational persistence is sufficiently characterized by focusing on the parent-child link only (Braun and Stuhler, 2018; Mare, 2011). Furthermore, recent works in this branch of the literature go beyond single indicator models and use many proxy variables to construct comprehensive indicators for the underlying parental social status (Vosters, 2018; Vosters and Nybom, 2017).

2.2 Empirical Approaches to Equality of Opportunity

Theoretical Set-up and Notation. Consider a population $\mathbf{N} := \{1, ..., N\}$ and an associated vector of non-negative outcomes $y = (y_1, ..., y_N)$. Outcomes are the result of two sets of factors: First, a set of *circumstances* beyond individual control: $\mathbf{\Omega} := C^1 \times ... \times C^P$. Second, a set of *efforts* $\mathbf{\Theta} := E^1 \times ... \times E^Q$. In what follows, $\mathbf{\Omega}$ and $\mathbf{\Theta}$ will be referred to as the circumstance and effort space, spanned by the dimensions ($C^p,\ p = 1, ..., P$) and ($E^q,\ q = 1, ..., Q$), respectively. We define the $(P \times 1)$-vector $\omega_i \in \mathbf{\Omega}$ as a comprehensive description of the circumstances with which $i \in \mathbf{N}$ is endowed. Analogously we define the $(Q \times 1)$-vector $\theta_i \in \mathbf{\Theta}$ as a comprehensive description of the efforts that are exerted by $i \in \mathbf{N}$.

The outcome generating function can be defined as follows:

$$\mathbf{\Omega} \times \mathbf{\Theta} \ni (\omega, \theta) \mapsto d(\omega, \theta) =: y,\ y \in \mathbb{R}_+, \tag{68}$$

such that for every $i \in \mathbf{N}$, the individual outcome y_i is a function of her circumstances ω_i and the effort θ_i she exerts. Individual effort exertion is plausibly co-determined by circumstance characteristics. We follow Roemer (1998) in adopting a relative conception of effort. Normatively, this assumption entails a stance according to which outcome differences due to a correlation between circumstances and effort constitute a violation of the opportunity egalitarian ideal. For example, if individuals work shorter hours due to wage discrimination in the labor market we would deem the ensuing income differences worth of compensation. Econometrically, this assumption entails that θ is purged of its correlation with circumstance characteristics ω such that effort is independently distributed of circumstance characteristics (see Lefranc et al., 2009; Roemer and Trannoy, 2015, for discussions). While such a conception is in line with the majority of the literature, our estimation approach is not dependent on it and can be easily extended to alternative cuts between ω and θ (Jusot et al., 2013).

Based on the realizations of individual circumstances ω_i the population can be partitioned into a set of *types*. We define the type partition $\mathbf{T} = \{t_1, ..., t_M\}$, such that individuals are member of one type if they share the same set of circumstances: $i, j \in t_m \Leftrightarrow \omega_i = \omega_j,\ \forall t_m \in \mathbf{T},\ \forall i, j \in \mathbf{N}$. Hence, types define one particular way of partitioning the population into groups, where group membership indicates uniformity in circumstances.

Measurement. Opportunity egalitarians are averse to inequality to the extent that it is rooted in circumstance factors that are beyond individual control. They are agnostic towards inequalities that originate from differences in effort exertion. In spite of the intuitive appeal of this idea, the literature has suggested a variety of formulations that differ in their precise normative content. Each of these different formulations is pinned down by combining a *principle of compensation* with a *principle of reward* (Aaberge et al., 2011; Almås et al., 2011; Fleurbaey, 1995; Ramos and Van de gaer, 2016). The former specifies how differences due to circumstances should be compensated. The latter specifies to what extent differences due to effort should be respected. In this work we exclusively focus on the principles of *ex-ante compensation* and *utilitarian reward*. Measures satisfying these two principles were first proposed in Checchi and Peragine (2010) and Van de gaer (1993). They are the most widely applied formulations in empirical works on equality of opportunity. To keep our analysis tractable we restrict ourselves to this particular conception of inequality of opportunity. However, our estimation approach is not dependent on it and can be easily extended to alternative measures of inequality of opportunity.

The ex-ante view of compensation focuses on between-type differences in the value of opportunity sets without paying attention to the specific effort realizations of individual type members. That means, we always prefer a distribution y' over y if the former is obtained from the latter by making a transfer from a more advantaged type to a less advantaged type. Utilitarian reward specifies zero inequality aversion with respect to income differences within a type. As a consequence, the value of the opportunity set of a type is pinned down by the expected value of its outcomes, $\mathbb{E}[y|\omega]$. Thus, the distribution of opportunities in a population can be expressed by the following counterfactual distribution y^C:

$$y^C = (y_1^C, ..., y_i^C, ..., y_N^C) = (\mathbb{E}[y_1|\omega_1], ..., \mathbb{E}[y_i|\omega_i], ..., \mathbb{E}[y_N|\omega_N]). \tag{69}$$

From this distribution one can construct ex-ante utilitarian measures of inequality of opportunity by choosing any functional $I()$ that satisfies the following two properties:

1. $I(y^C)$ decreases (increases) through transfers from i to j if i is from a circumstance type with a higher (lower) expected value of outcomes than the recipient j.

2. $I(y^C)$ remains unaffected by transfers from i to j if they are members of the same type.

In most empirical applications $I()$ represents an inequality index satisfying the standard properties of anonymity, the principle of transfers, population replication, and scale invariance (FCowell, 2016).[4] Examples of the latter are the Gini index or any member of the generalized entropy class. Note that the choice of $I()$ is normative in itself as it specifies the extent of inequality aversion at different points of the counterfactual distribution y^C. For example, the mean logarithmic deviation (MLD) would value compensating transfers to the most disadvantaged types more than the Gini index. In this work we are agnostic towards the normatively correct choice of $I()$. While we will present our main results in terms of the Gini index, we provide robustness checks based on other inequality indexes in Supplementary Material B.6.

Note that the measurement of inequality of opportunity can also be understood as a decomposition exercise where total inequality is split into a between- and a within-group component. It thereby relates to the broad literature on distributional decompositions in labor economics (Fortin et al., 2011). However, it is important to highlight that opportunity egalitarians view differences among circumstance groups as normatively objectionable regardless of whether these differences are the result of compositional differences in (un)observed characteristics (e.g. educational achievement and occupational choices) or the return to such characteristics. While distinctions among these different explanations are important for the design of appropriate policy responses, they are of indifference for the measurement of inequality of opportunity in the ex-ante utilitarian sense.

Given the measurement decisions described above, we require an estimate of the conditional outcome distribution y^C. The data generating process described in equation 68 can be rewritten as follows:

$$y = d(\omega, \theta) = f(\omega) + \epsilon = \mathbb{E}(y|\omega) + \epsilon = y^C + \epsilon, \tag{70}$$

where $\mathbb{E}(y|\omega)$ captures variation due to observed circumstances. The iid error term ϵ captures variation due to unobserved circumstances and individual effort. The fact that ϵ representes both fair (individual effort) and unfair (unobserved circumstances) determinants of individual

[4] The β coefficient from intergenerational mobility regressions can also be interpreted as an ex-ante utilitarian measure of inequality of opportunity. In the intergenerational mobility framework, $\beta = \frac{E(y_{ic}|y_{ip})}{y_{ip}}$, where y_{ip} equals parental income as the sole circumstance. Hence, the functional applied to the distribution of conditional expectations can be written as $I() = \frac{1}{y_{ip}}$. Note that β decreases (increases) through transfers from children from advantaged (disadvantaged) backgrounds to children from less (more) advantaged backgrounds. However, β remains unaffected by transfers between children from parental households of equal income.

outcomes illustrates that the resulting measures of inequality of opportunity have a lower bound interpretation.

Estimating y^C is a prediction task in which the researcher tries to answer the following question: What outcome y_i do we expect for an individual that faces circumstances ω_i? This task is complicated by the fact that the precise form of $f()$ is a priori unknown. In the vast majority of empirical applications, researchers address this lack of knowledge by invoking strong functional form assumptions. For example, they perform a log-linear regression of the outcome of interest on the set of observed circumstances and construct an estimate of y^C from the predicted values:

$$\ln(y_i) = \beta_0 + \sum_{p=1}^{P} \beta_p \omega_i^p + \epsilon_i, \tag{71}$$

$$\hat{y}_i^C = \exp\left[\beta_0 + \sum_{p=1}^{P} \hat{\beta}_p \omega_i^p\right], \tag{72}$$

where $\omega_i^p \in \mathbf{\Omega}$. The literature refers to this estimation procedure as the *parametric approach* (Bourguignon et al., 2007; Ferreira and Gignoux, 2011).

Another common estimator of y^C comes from an approach where the researcher partitions the sample into mutually exclusive types based on the realizations of all circumstances under consideration. An estimate of y^C is then constructed from the average outcome values within types:

$$\hat{y}_i^C = \mu_{m(i)} = \frac{1}{N_m} \sum_{j=1}^{N_m} y_j, \ \forall j \in t_m, \ \forall t_m \in \mathbf{T}. \tag{73}$$

The literature refers to this estimation procedure as the *non-parametric approach* (Checchi and Peragine, 2010).

Both approaches face empirical challenges which are typically resolved by discretionary decisions of the researcher. For example, the parametric approach assumes a log-linear impact of all circumstances and therefore neglects the existence of interdependencies between circumstances and other non-linearities. To alleviate this shortcoming the researcher may integrate interaction terms and higher order polynomials into equation (71). However, such extensions remain at her discretion. Reversely, the non-parametric approach does not restrict the interdependent impact of circumstances. However, if the data is rich enough in information on circumstances, the researcher may be forced to reduce the observed circumstance space to obtain statistically meaningful estimates of the relevant parameters. Assume for example, that the researcher observes ten circumstance variables with three expressions each – a quantity easily observed in many household surveys. Implementing the non-parametric approach would require the estimation of $3^{10} = 59,049$ group means which is hardly feasible given the sample sizes of most household surveys. The necessary process of restricting the circumstance space again remains at the researcher's discretion.

The previous discussion illustrates that common approaches leave the researcher to her own devices when it comes to selecting the best model for estimating the distribution y^C. In this paper, we provide an automated solution to this problem. Similarly, Li Donni et al. (2015) propose the use of latent class modeling to obtain type partitions that allow for estimates of y^C according to the non-parametric procedure outlined in equation (73). In their approach, observable circumstances are considered indicators of membership in an unobservable latent type, t_m. For each possible number of latent types, M, individuals are assigned to types so as to minimize the within-type correlation of observable circumstances. Then the optimal number of types, M^*, is selected by minimizing an appropriate model selection criterion such as Schwarz's Bayesian Information Criterion (BIC). The latent class approach therefore partly solves the issue of arbitrary model selection. However, it cannot solve the problem of model selection once the potential number of type characteristics exceeds the available degrees of freedom. In such cases, the latent class approach replicates the limitations of the parametric and the non-parametric approach: the researcher must pre-select the relevant set of circumstances, their subpartition, and the respective interactions. Furthermore, latent classes are obtained by minimizing the within-type correlation of circumstances while ignoring the correlation of circumstance variables with the outcome variable. As a consequence, they are not well-suited for capturing the dependence between circumstances and a particular outcome of interest.

In the following we will show how the outlined shortcomings of existing approaches can be addressed by regression trees and forests.

2.3 Estimating Inequality of Opportunity from Regression Trees and Forests

Regression trees and forests belong to the class of supervised learning methods that were developed to make out-of-sample predictions of a dependent variable based on a number of observable predictors. As we will outline in the following, they can be straightforwardly applied to inequality of opportunity estimations and solve the issue of model selection.

While there are many supervised learning methods to solve prediction problems, trees and forests are particularly attractive in our setting since they are very flexible in accounting for non-linearities and effective in excluding features that are unrelated to the outcome of interest (Athey and Imbens, 2019). Moreover, in the context of inequality of opportunity estimations they strike a balance between prediction accuracy and interpretability. More complex ensemble methods that obtain predictions as a weighted average from hundreds of models will tend to make smaller prediction errors, but often be harder to interpret. In many applications, as exemplified by the Netflix prize challenge (Bell and Koren, 2007), there are good reasons to neglect interpretability and focus exclusively on predictive performance. However, inequality of opportunity estimates are policy-relevant statistics designed to inform debates on potential policy interventions. Therefore, interpretability of the output is of great importance, making approaches based on trees and forests an attractive option in the context of inequality of opportunity estimations.

First, we will introduce conditional inference regression trees. By providing predictions based on identifiable groups, they closely connect to Roemer's theoretical formulation of inequality of opportunity. Furthermore, their simple graphical illustration is particularly instructive for longitudinal or cross-sectional comparisons of opportunity structures. Second, we will introduce conditional inference forests, which are – loosely speaking – a collection of many conditional inference trees. While forests do not have the intuitive appeal of regression trees, they perform better in terms of out-of-sample prediction accuracy and hence provide better estimates of the counterfactual distribution y^C.

Conditional Inference Trees

Tree-based methods obtain predictions for outcome y as a function of the input variables $x = (x^1, ..., x^k)$. Specifically, they use the sample $\mathcal{S} = \{(y_i, x_i)\}_{i=1}^S$ to divide the population into non-overlapping groups, $\mathbf{G} = \{g_1, ..., g_m, ..., g_M\}$, where each group g_m is homogeneous in the expression of some input variables. These groups are called *terminal nodes* or *leafs* in a regression tree context. The conditional expectation for observation i is estimated from the mean outcome $\hat{\mu}_m$ of the group g_m to which the i^{th} observation is assigned. Hence, in addition to the observed outcome vector $y = (y_1, ..., y_i, ..., y_N)$ one obtains a vector of predicted values $\hat{y} = (\hat{f}(x_1), ..., \hat{f}(x_i), ..., \hat{f}(x_N))$, where

$$\hat{f}(x_i) = \hat{\mu}_{m(i)} = \frac{1}{N_m} \sum_{j \in g_m} y_j, \tag{74}$$

and N_m is the size of each group.

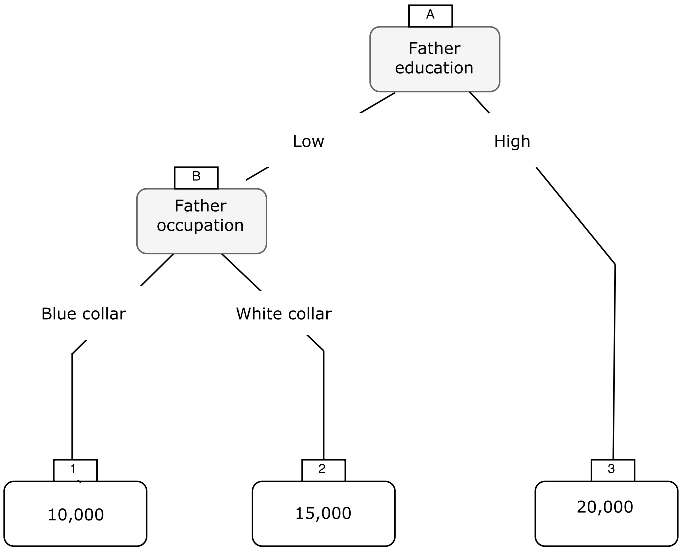
The mapping from regression trees to equality of opportunity estimation is straightforward. Conditional on the input variables being circumstances only, each resulting group $g_m \in \mathbf{G}$ can be interpreted as a circumstance type $t_m \in \mathbf{T}$. Furthermore, $\hat{y}$ is analogous to an estimate of the counterfactual distribution y^C which in turn can be used for the construction of ex-ante utilitarian measures of inequality of opportunity.

Tree Construction. Regression trees partition the sample into M types by *recursive binary splitting*. Recursive binary splitting starts by dividing the full sample into two distinct groups according to the value they take in one input variable $\omega^p \in \mathbf{\Omega}$. If ω^p is a continuous or ordered variable, then $i \in t_l$ if $\omega_i^p < \tilde{\omega}^p$ and $i \in t_m$ if $\omega_i^p \geq \tilde{\omega}^p$, where $\tilde{\omega}^p$ is a splitting value chosen by the algorithm. If ω^p is a categorical variable then the categories can be split into any two arbitrary groups. The process is continued such that one of the two groups is divided into further subgroups (potentially based on another $\omega^q \in \mathbf{\Omega}$), and so on. Graphically, this division into groups can be presented like an upside-down tree (Figure 2.1).

The exact manner in which the split is conducted depends on the type of regression tree that is used. In this paper, we follow the conditional inference methodology proposed by Hothorn et al. (2006). Conditional inference trees are grown by a series of permutation tests according to the following 4-step procedure:

0. Choose a significance level α^*.

FIGURE 2.1 – Exemplary Tree Representation



Note: Artificial example of a regression tree. The gray boxes indicate splitting points, while the white boxes indicate terminal nodes. The values inside the terminal nodes show estimates for the conditional expectation y^C.

1. Test the null hypothesis of density function independence: $H_0^{\omega^p} : D(Y|\omega^p) = D(Y)$, for all $\omega^p \in \mathbf{\Omega}$, and obtain a p-value associated with each test, p^{ω^p}.

   $\Rightarrow$ Adjust the p-values for multiple hypothesis testing, such that $p_{adj.}^{\omega^p} = 1-(1-p^{\omega^p})^P$ (Bonferroni Correction).

2. Select the variable ω^* with the lowest p-value, i.e.

$$\omega^* = \underset{\omega^p}{\text{argmin}}\{p_{adj.}^{\omega^p} : \omega^p \in \mathbf{\Omega},\ p = 1, ..., P\}.$$

   $\Rightarrow$ If $p_{adj.}^{\omega^*} > \alpha^*$: Exit the algorithm.

   $\Rightarrow$ If $p_{adj.}^{\omega^*} \leq \alpha^*$: Continue, and select ω^* as the splitting variable.

3. Test the null hypothesis of density function independence between the subsamples for each possible binary partition splitting point s based on ω^* and obtain a p-value associated with each test, $p^{\omega_s^*}$.

   $\Rightarrow$ Split the sample based on ω^*, by choosing the splitting point s that yields the lowest p-value, i.e. $\tilde{\omega}^* = \underset{\omega_s^*}{\text{argmin}}\{p^{\omega_s^*} : \omega_s^* \in \mathbf{\Omega}\}$.

4. Repeat steps 1.–3. for each of the resulting subsamples.

In words, conditional inference start by a series of univariate hypothesis tests that test the relationship between the outcome and each circumstance variable. The circumstance that is most related to the outcome is chosen as the potential splitting variable. If the dependence between the outcome and the splitting variable is sufficiently strong, then a split is made. If not, no split is made. Whenever a circumstance can be split in several ways, the sample is split into two subsamples such that the dependence with the outcome variable is maximized. This procedure is repeated in each of the two subsamples until no circumstance in any subsample is sufficiently related to the outcome variable. Note that the structure and depth of the resulting opportunity tree hinges crucially on the level of α^*. The less stringent the α^*-requirement, the more we allow for false positives, i.e. the more splits will be detected as significant and the deeper the tree will be grown. In our empirical application we fix $\alpha^* = 0.01$, which is in line with the disciplinary convention for hypothesis tests. To illustrate the robustness of this choice we show comparisons to setting $\alpha^* = 0.05$ and choosing α^* through cross-validation in Appendix Figure B.1.

Conditional Inference Forests

Regression trees solve the model selection problem outlined in section 2.2 and provide a simple and standardized way of dividing the population into types. However, constructing estimates for the counterfactual distribution y^C from conditional inference trees suffers from three shortcomings: first, the structure of trees – and therefore the estimate of the relevant distribution y^C – is fairly sensitive to alternations in the respective data samples. This issue is particularly pronounced if there are various circumstances that are close competitors for defining the first splits (Friedman et al., 2009). Second, trees assume a non-linear data generating process that imposes interactions while ruling out the linear influence of circumstances. On the one hand, this is fully consistent with Roemer's theory by which circumstances partition the population into types. On the other hand, the best model for constructing $\hat{y}$ may in fact be linear in some circumstances. Third, trees make only limited use of the information inherent in the set of observed circumstances since some of the circumstances $\omega^p \in \mathbf{\Omega}$ are not used for the construction of the tree. However, circumstances may possess informational content that can increase predictive power even if they are not significantly associated with y at level α^*. In analogy to the problem of multicollinear regressors in regression analysis, this is a particular issue if two or more important circumstances are highly correlated. Once a split is done using either of the two, the other will unlikely yield enough information to cause another split.

In what follows we will introduce conditional inference forests (Biau and Scornet, 2016; Breiman, 2001) which address all three of these shortcomings.

Forest Construction. Random forests create many trees and average over all of these when making predictions. Trees are constructed according to the same 4-step procedure outlined in the previous subsection. However, two tweaks are made. First, given the sample $\mathcal{S} = \{(y_i, \omega_i)\}_{i=1}^{S}$ each tree is estimated on a random subsample $\mathcal{S}' \subset \mathcal{S}$. In our case, we randomly select half of the observations for each tree, and estimate B^* such trees in total. Second, only a random subset of circumstances $\{\omega^p \in \mathbf{\Omega} : p \in \bar{\mathbf{P}} \subset \{1, ..., P\}\}$ of size $\bar{P}^*$ is allowed to be used at each splitting point. Together these two tweaks remedy the shortcomings of single conditional inference trees. First, averaging over the B^* predictions cushions the variance in the estimates of y^C and smoothes the non-linear impact of circumstance characteristics. Second, drawing only on subsets of all circumstance variables increases the likelihood that all observed circumstances with informational content will be identified as the splitting variable ω^* at some point.

Predictions are formed as follows:

$$\hat{f}(\omega; \alpha^*, \bar{P}^*, B^*) = \frac{1}{B^*} \sum_{b=1}^{B^*} \hat{f}^b(\omega; \alpha^*, \bar{P}^*). \tag{75}$$

Equation (75) illustrates that individual predictions are a function of α^* – the significance level governing the implementation of splits, $\bar{P}^*$ – the number of circumstances to be considered at each splitting point, and B^* – the number of subsamples to be drawn from the data. In our empirical illustration we fix $B^* = 200$ and determine α^* and $\bar{P}^*$ by minimizing the *out-of-bag* error (MSE^{OBB}). Details on these choices and the empirical procedures are disclosed in Appendix B.1.

2.4 Empirical Application

In this section we provide an illustration of the machine learning approach using harmonized survey data from 31 European countries. We will compare the results from trees and forests with results from the prevalent estimation approaches in the extant literature; namely parametric, non-parametric and latent class models. Comparisons will be made along two dimensions.

First, we evaluate the different estimation approaches by comparing their out-of-sample mean squared error (MSE). The MSE provides a standard statistic to evaluate the prediction quality of different models by representing the variance-bias trade-off. In the context of constructing an estimate of the conditional income distribution y^C, this property is equivalent to trading-off upward and downward biases in inequality of opportunity estimates: The more parsimonious the model, the higher the prediction bias (underfitting) and the stronger the downward bias in inequality of opportunity estimates. The more complex the model, the higher the prediction variance (overfitting) and the stronger the upward bias of inequality of opportunity estimates. A thorough illustration of this mapping is provided in Appendix B.2.

Second, we compare the inequality of opportunity estimates emanating from the set of benchmark methods to the ones from regression trees and forests.

In a last step, we illustrate how regression trees and forests can be used to analyze opportunity structures in the population of interest.

Data

We base our empirical illustration on the 2011 wave of the European Union Statistics on Income and Living Conditions (EU-SILC). EU-SILC provides harmonized survey data with respect to income, poverty, and living conditions on an annual basis and covers a cross-section of 31 European countries in the 2011 wave.[5] For each country, EU-SILC provides a random sample of all resident private households. The data is collected by the various national statistical agencies following common variable definitions and data collection procedures. It provides the official reference source for comparative statistics on income distribution and social inclusion in the European Union (EU) and therefore provides a degree of harmonization that makes it particularly suitable for methodological comparisons. We draw on the 2011 wave since it contains an ad-hoc module about the intergenerational transmission of (dis)advantages which

[5] The sample consists of Austria (AT), Belgium (BE), Bulgaria (BG), Switzerland (CH), Cyprus (CY), Czech Republic (CZ), Germany (DE), Denmark (DK), Estonia (EE), Greece (EL), Spain (ES), Finland (FI), France (FR), Croatia (HR), Hungary (HU), Ireland (IE), Iceland (IS), Italy (IT), Malta (MT), Lithuania (LT), Luxembourg (LU), Latvia (LV), Netherlands (NL), Norway (NO), Poland (PL), Portugal (PT), Romania (RO), Sweden (SE), Slovenia (SI), Slovak Republic (SK), and Great Britain (UK).

allows us to construct finely-grained circumstance type partitions. The space of observed circumstances Ω and their respective expressions are listed in Table 2.1. The list includes all variables of EU-SILC containing information about the respondent's characteristics at birth and their living conditions during childhood. Descriptive statistics concerning circumstances are reported in Supplementary Material B.5.

The unit of observation is the individual and the outcome of interest is equivalized disposable household income. The latter is obtained by dividing household disposable income by the square root of household size. Reported incomes refer to the year preceding the survey wave, i.e. 2010 in the case of our empirical application. In line with the literature we focus on equivalized household income as it provides the closest income analogue to consumption possibilities and general economic well-being. Aware that inequality statistics tend to be heavily influenced by outliers (Cowell and Victoria-Feser, 1996) we adopt a standard winsorization method according to which we set all non-positive incomes to 1 and scale back all incomes exceeding the 99.5th percentile of the country-specific income distribution to this lower threshold. Our analysis is focused on the working age population. Therefore, we restrict the sample to respondents aged between 30 and 59. To assure the representativeness of our country samples all results are calculated by using appropriate individual cross-sectional weights.

Table 2.2 shows considerable heterogeneity in the income distributions of the European country sample. While the average households in Norway (NO) and Switzerland (CH) obtained incomes above €40,000 in 2010, the average household income in Bulgaria (BG), Romania (RO) and Lithuania (LT) did not exceed the €5,000 mark. The lowest inequality prevails in the Nordic countries of Norway (NO), Sweden (SE) and Iceland (IS), all of which have Gini coefficients below 0.220. At the other end of the spectrum we find the Eastern European countries of Latvia (LV), Lithuania (LT) and Romania (RO) with Gini coefficients well above 0.330.

Table 2.2 also shows the sample size for each country. These figures include observations with missing values in one or more of the circumstances we use. The parametric approach, the non-parametric approach, and latent class analysis handle missing values by listwise deletion. In contrast, conditional inference trees and forests make use of the full sample by allowing for surrogate splits. For each splitting point $\tilde{\omega}^*$, the algorithm searches for an alternative splitting point $\tilde{\omega}^+$ that mimicks the sample partition of $\tilde{\omega}^*$ to the greatest extent.

Table 2.1 – List of Circumstances

1. Respondent's sex:
   - Male
   - Female
2. Respondent's country of birth:
   - Respondent's present country of residence
   - European country
   - Non-European country
3. Presence of parents at home*:
   - Both present
   - Only mother
   - Only father
   - Without parents
   - Lived in a private household without any parent
4. Number of adults (aged 18 or more) in respondent's household*
5. Number of working adults (aged 18 or more) in respondent's household*
6. Number of children (under 18) in respondent's household*
7. Father's/mother's country of birth and citizenship:
   - Born/citizen of the respondent's present country of residence
   - Born/citizen of another EU-27 country
   - Born/citizen of another European country
   - Born/citizen of a country outside Europe
8. Father's/mother's education (based on the International Standard Classification of Education 1997 (ISCED-97))*:
   - Unknown father/mother
   - Illiterate
   - Low (0-2 ISCED-97)
   - Medium (3-4 ISCED-97)
   - High (5-6 ISCED-97)
9. Father's/mother's occupational status*:
   - Unknown or dead father/mother
   - Employed
   - Self-employed
   - Unemployed
   - Retired
   - House worker
   - Other inactive
10. Father's/mother's main occupation (based on the International Standard Classification of Occupations, published by the International Labour Office ISCO-08)*:
    - Managers (I-01)
    - Professionals (I-02)
    - Technicians (I-03)
    - Clerical support workers (I-04)
    - Service and sales workers (including also armed force) (I-05 and 10)
    - Skilled agricultural, forestry and fishery workers (I-06)
    - Craft and related trades workers (I-07)
    - Plant and machine operators, and assemblers (I-08)
    - Elementary occupations (I-09)
    - Armed forces occupation (I-00)
    - Father/mother did not work, was unknown or was dead
11. Managerial position of the father/mother*:
    - Supervisory
    - Non-supervisory
12. Tenancy status of the house in which the respondent was living*:
    - Owned
    - Not owned

Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Questions marked with * refer to the period when the respondent was approximately 14 years old. Item 11 is missing for Finland. We exclude subjective questions about the financial situation and the level of deprivation of the household of origin from the list of circumstances.

All observations that lack information on $\tilde{\omega}^*$ are then allocated to subbranches based on $\tilde{\omega}^+$. As a consequence, there are differences in the actual sample sizes available for the different methods. When comparing inequality of opportunity estimates across methods, we tolerate these differences in sample sizes since we want to compare inequality of opportunity estimates

Table 2.2 – Summary Statistics

Country	N	Equivalized Disposable Household Income μ	σ	Gini
AT	6,220	25,451	13,971	0.268
BE	6,011	23,291	10,948	0.249
BG	7,154	3,714	2,491	0.333
CH	7,583	42,208	24,486	0.279
CY	4,589	21,058	11,454	0.279
CZ	8,711	9,006	4,320	0.250
DE	12,683	22,221	12,273	0.276
DK	5,897	32,027	13,836	0.232
EE	5,338	6,922	3,912	0.330
EL	6,184	13,184	8,651	0.334
ES	15,481	17,088	10,597	0.329
FI	9,743	27,517	13,891	0.246
FR	11,078	24,299	14,583	0.288
HR	6,969	6,627	3,819	0.306
HU	13,330	5,327	2,863	0.276
IE	4,318	24,867	14,307	0.296
IS	3,684	22,190	9,232	0.210
IT	21,070	18,786	11,730	0.309
LT	5,403	4,774	3,150	0.344
LU	6,765	37,911	19,977	0.271
LV	6,423	5,334	3,618	0.363
MT	4,701	13,006	6,747	0.277
NL	11,411	25,210	11,414	0.235
NO	5,026	43,260	16,971	0.202
PL	15,545	6,103	3,690	0.316
PT	5,899	10,781	7,296	0.334
RO	7,867	2,562	1,646	0.337
SE	6,599	26,346	10,700	0.215
SI	13,183	13,772	5,994	0.225
SK	6,779	7,304	3,416	0.257
UK	7,391	25,936	16,815	0.320

Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: N indicates the total number of observations in the respective country sample. The last three columns refer to the country-specific distribution of equivalized disposable household incomes measured in €. μ indicates the mean, σ the standard deviation, and the last column shows inequality as measured by the Gini coefficient.

by respecting all methods to the greatest extent. To the contrary, when comparing the out-of-sample performance we use the smallest sample size across methods for all calculations, such that the relative out-of-sample performance cannot be driven by sample size differences or non-random attrition through listwise deletion. A thorough discussion of the sensitivity of all methods to different sample sizes is provided in Appendix B.3.

Benchmark Methods

We compare our estimates from trees and forests against three benchmark methods that have been proposed in the extant literature.

First, we draw on the parametric approach as proposed by Bourguignon et al. (2007) and Ferreira and Gignoux (2011). In line with equation (71), estimates are obtained by a Mincerian regression of equivalent household income on the following circumstances: father's occupation (10 categories), father's and mother's education (5 categories), area of birth (3 categories), and tenancy status of the household (2 categories). The model specification therefore includes 20 binary variables and resembles the specification used in Palomino et al. (2019).

Second, we draw on the non-parametric approach as proposed by Checchi and Peragine (2010). In line with equation (73), non-parametric estimates are obtained by calculating average outcomes in non-overlapping circumstance types. In this application we construct 40 such types. Individuals in type t_m are homogeneous with respect to the educational achievement of their highest educated parent (5 categories) as well as their migration status (2 categories). The latter is indicated by a binary variable for whether the respondent is a first or second generation immigrant. Furthermore, they have fathers working in the same occupation (4 categories). To minimize the frequency of sparsely populated types we divert from the occupational list given in Table 2.1 by re-coding occupations into the following categories: high-skilled non-manual (I-01–I-03), low-skilled non-manual (I-04–I-05 and I-10), skilled manual and elementary occupation (I-06–I-09), and unemployed/unknown/dead. This partition is similar but more parsimonious than the one used in Checchi et al. (2016) who base their analysis on a total of 96 types. Notably, in contrast to Checchi et al. (2016) we exclude age from the list of circumstances since it is fairly controversial whether age qualifies as a circumstance characteristic in the relevant sense.

Lastly, we compare our estimates against the latent class approach as proposed by Li Donni et al. (2015). The eligible set of circumstances is the full set of observable circumstances. For the latent class analysis, we follow Li Donni et al. (2015) and select the number of latent types by minimizing BIC.

Model Performance

In order to assess the prediction accuracy of different models, we follow the machine learning practice of splitting our sample into a *training set* with $i^{-H} \in \{1, ..., N^{-H}\}$ and a *test set* with $i^{H} \in \{1, ..., N^{H}\}$. For each country in our sample, $N^{-H} = \frac{2}{3}N$ while $N^{H} = \frac{1}{3}N$. We fit our models on the training set and compare their performance on the test set according to the following procedure:

1. Run the chosen models on the training data (for the specific estimation procedures, see section 2.3 for trees and forests, and section 2.4 for our benchmark methods).

2. Store the prediction functions $\hat{f}^{-H}()$.

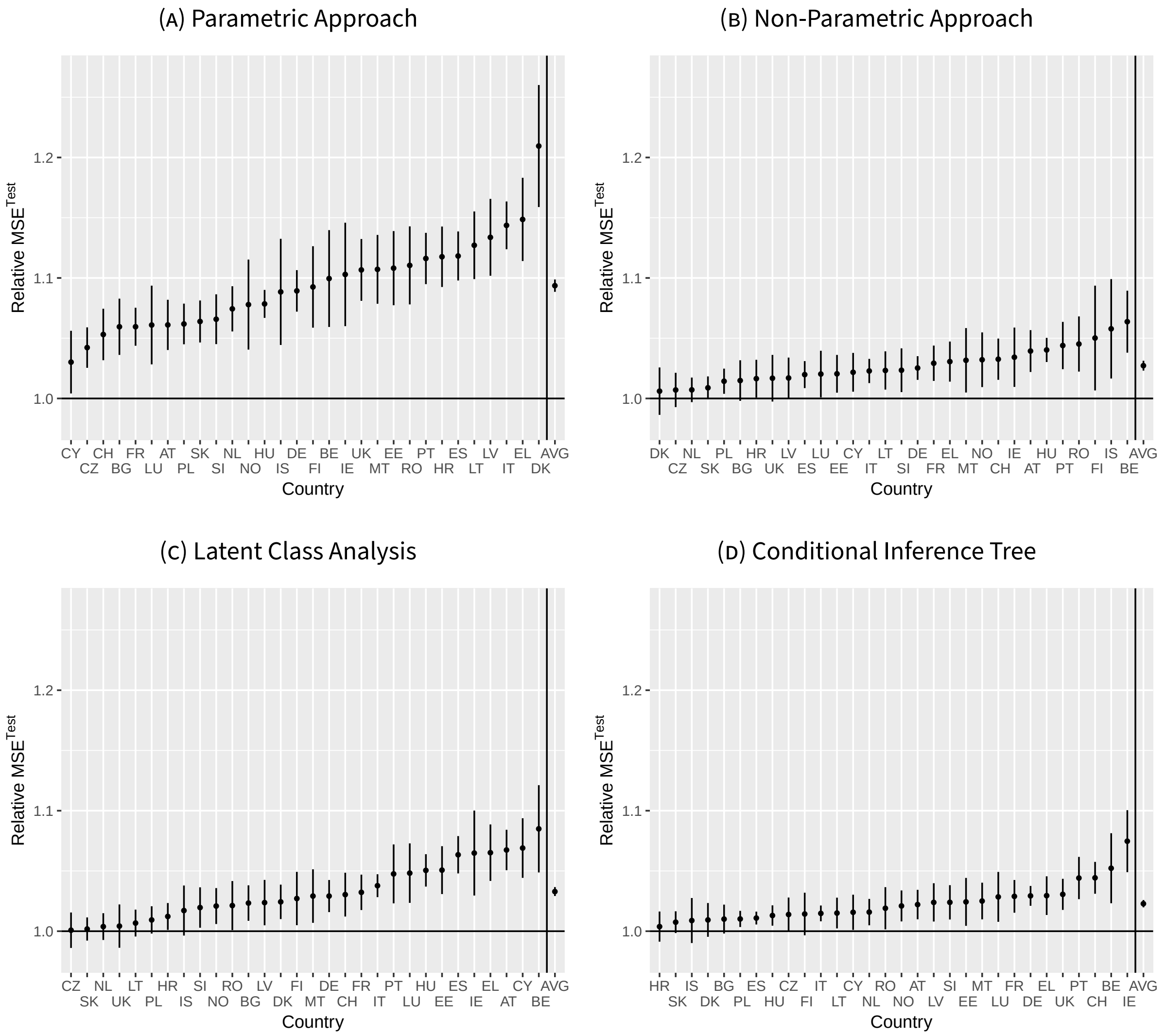
3. Calculate the mean squared error in the test sample:
   $\text{MSE}^{\text{Test}} = \frac{1}{N^H} \sum_{i \in H} [y_i - \hat{f}^{-H}(\omega_i)]^2$.

Figure 2.2 compares the resulting MSE^{Test} of the different models. For each country, MSE^{Test} of random forests is standardized to equal 1, such that an MSE^{Test} larger than 1 represents a worse out-of-sample fit. This implies that the respective method performs worse than forests in trading off upward and downward bias – either by making poor use of circumstance information or overfitting the data. We derive 95% confidence intervals based on 200 bootstrapped re-samples of the test data using the normal approximation method (DiCiccio and Efron, 1996).

Random forests outperform all other methods in all cases. On average, the parametric approach gives a fit that is 9.4% worse than forests (Figure 2.2, Panel (a)). This average, however, masks considerable heterogeneity. While the relative test error for Cyprus only slightly exceeds the 3%-mark, the test error of the parametric model for Denmark and Sweden exceed the benchmark method by more than 20%. For all countries, the benchmark MSE lies outside the 95% confidence band of the parametric approach.

With average shortfalls of around 3%, out-of-sample prediction errors are less pronounced for the non-parametric (Figure 2.2, Panel (b)) and latent class models (Figure 2.2, Panel (c)). Yet, as in the case of the parametric approach, MSE^{Test} statistics of conditional inference forests lie outside the 95% confidence band of the respective method for the vast majority of the country cases in our sample. Hence, relative to random forests, the benchmark methods either underutilize or overutilize the information contained in $\boldsymbol{\Omega}$. As we will see in section 2.4, the parametric and the non-parametric models are overfitting the data and are therefore upward biased. To the contrary, the type partition delivered by latent class analysis tends to be too coarse and therefore downward biased. The relatively good performance of the non-parametric approach could suggest that it is a sustainable alternative to forests. However, since the model specification remains under the discretion of the researcher, this performance is a luck of the draw rather than a property inherent to the estimation approach. In this particular case, had we followed the specification of Checchi et al. (2016) exactly by incorporating age as

Figure 2.2 – Comparison of Models' Test Error



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The y-axis shows the MSE^{Test} of the different estimation approaches relative to the benchmark of random forests. MSE^{Test} for random forests is standardized to 1, such that a relative test error > 1 indicates worse fit than random forests. 95% confidence intervals are derived based on 200 bootstrapped re-samples of the test data using the normal approximation method. For better result visibility Sweden is excluded from the figure since it is an outlier. The test errors for Sweden are 1.43 $[1.21, 1.66]$ for the parametric approach, 1.11 $[1.01, 1.21]$ for the non-parametric approach, 1.06 $[1.02, 1.11]$ for latent class analysis, and 1.06 $[1.01, 1.11]$ for conditional inference trees.

a circumstance characteristic, the type partition would more than double and be accompanied with a significant deterioration in the out-of-sample performance (see section 3.3).

On average, conditional inference trees are closest to the test error rate of forests. With the exception of two country cases, the test error of trees exceeds the test error of forests by less than 5%. Yet, as outlined in section 2.3, they also fall short of the performance of forests due to their poorer utilization of the information given in Ω.

We conclude that among all considered methods, conditional inference forests deliver the highest out-of-sample prediction accuracy. Hence, they perform best in trading off upward and downward bias in inequality of opportunity estimations. One may suspect that other machine learning algorithms perform even better in predicting outcomes out-of-sample. However, we note that in social science applications the gain in prediction accuracy is typically small when alternating between algorithms that allow for sufficient model flexibility. For example, in the context of intergenerational mobility estimations J. Blundell and Risa (2019) show that there is no difference in the performance of random forests, neural nets and gradient boosted trees.[6] To demonstrate the substantive relevance of this property, we now turn to a comparison of the equality of opportunity estimates emanating from the considered set of estimation approaches.

Estimates of Inequality of Opportunity

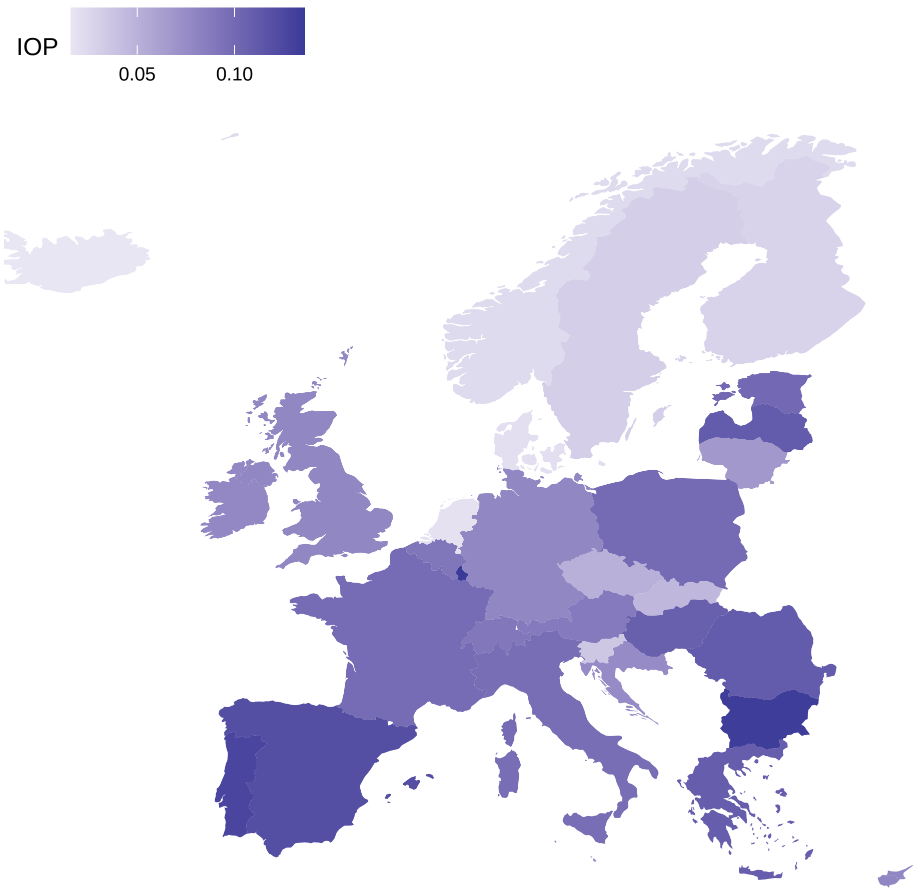
Figure 2.3 plots inequality of opportunity estimates based on random forests for our European country sample in 2010. We observe a clear North-South gradient with the Scandinavian countries being characterized by the lowest level on inequality of opportunity. Similarly, we observe a slight East-West gradient with many countries from the former Warsaw pact being characterized by higher levels of inequality of opportunity. Notable exceptions are Czech Republic and Slovakia.

It is important to emphasize that the results of the random forests cannot be interpreted as recovering the truth. However, they provide a benchmark estimate since forests have the lowest test error for all countries, therefore perform best in balancing concerns about upward and downward bias, and hence provide the best *approximation of the truth* among all methods we consider. Following this insight, Figure 2.4 plots inequality of opportunity estimates based on each method relative to the estimates from conditional inference forests on a logarithmic scale.[7] For all methods, inequality of opportunity estimates are obtained by calculating the Gini index in the estimated counterfactual distribution $\hat{y}^C$. As discussed in section 2.2, there is a class of functionals that can be used to summarize the distribution of $\hat{y}^C$. We therefore

[6] Although it is not explicitly part of our methodological comparison, we provide the exact time necessary to run a single iteration for all countries for each method in the following. (i) Non-parametric approach: 2.45 seconds, (ii) Parametric approach: 1.55 seconds, (iii) Latent class analysis: 1.02 hours, (iv) conditional inference trees: 39.14 seconds, (v) conditional inference forests: 2.06 hours. The run times are measured for a computer with a of 2.3 GHz Intel Core i5 central processor.

[7] The results from Figure 2.2 and 2.4 are not directly comparable since they use different samples. See the text following Table 2.2 for details.

FIGURE 2.3 – Inequality of Opportunity in Europe, 2010

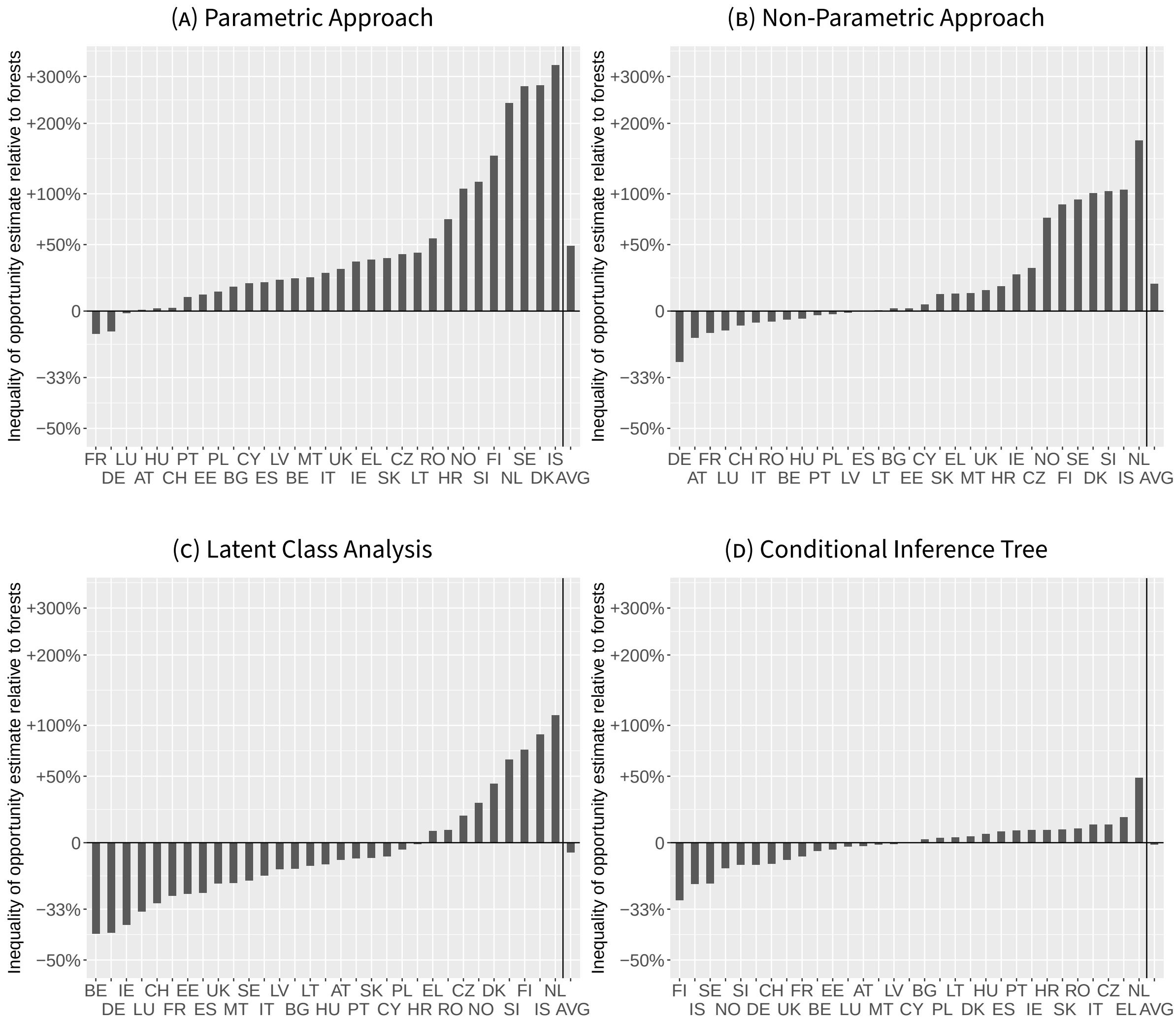


Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Inequality of opportunity is measured by the Gini coefficient in the estimated counterfactual distribution $\hat{y}^C$. $\hat{y}^C$ is constructed based on the predictions from conditional inference forests. Darker shaded colors indicate higher levels of inequality of opportunity. The displayed inequality of opportunity estimates are reported in the last column of Table B.1.

provide estimates for alternative inequality indexes in Supplementary Material B.6. For each country- and method-specific estimate we divide by the estimate from random forests to obtain the relative divergence between the respective benchmark and our preferred method. This implies that, for a given country, inequality of opportunity estimates larger than those obtained from forests overfit the data and vice versa. An overview table of the underlying point estimates including 95% confidence bands is disclosed in Appendix B.4.

Panel (a) plots the estimates from the parametric approach relative to the forest estimates. For 28 out of 31 countries the inequality of opportunity estimates are higher than the results from conditional inference forests. The given specification of the parametric approach inflates inequality of opportunity statistics by 47% on average. The most pronounced overstatement is observed for Iceland where the parametric approach yields an estimate more than four

FIGURE 2.4 – Comparison of Estimates by Method



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: In each panel, the y-axis shows the inequality of opportunity estimate from the method in question divided by the inequality of opportunity estimate from forests, displayed on a logarithmic scale. Country-estimates above the black line indicate an overestimation of inequality of opportunity relative to the random forest benchmark. Reversely, country-estimates below the black line indicate an underestimation of inequality of opportunity relative to the random forest benchmark. For all methods inequality of opportunity is measured by the Gini coefficient in the estimated counterfactual distribution $\hat{y}^C$.

times higher than the forest analogue. Similarly, the figures of Sweden and Denmark are inflated by a factor of 3.8. Also in terms of country rankings, the parametric approach delivers markedly different results in comparison to our preferred method. While the parametric approach identifies Romania (RO), Bulgaria (BG) and Greece (EL) as the countries in which opportunities are most unequally distributed, these countries rank 6th, 2nd and 7th in the case of forests.

Panel (b) illustrates that the benchmark specification of the non-parametric approach takes a middle-ground between the parametric approach and our preferred method. For 19 out of 31 countries the non-parametric estimate exceeds its forest-based analogue. The non-parametric specification inflates inequality of opportunity statistics at a rate of 18% on average. Also in terms of country rankings the non-parametric approach shows a much closer resemblance to our preferred method than the parametric approach. For example, it identifies Bulgaria (BG), Portugal (PT) and Luxembourg (LU) as the countries in which opportunities are most unequally distributed. This ranking is congruent with the top three countries identified by forests. However, the resemblance should be interpreted as a luck of the draw rather than a property inherent to the estimation approach. Under alternative type partitions the estimates from the non-parametric approach may diverge much more strongly than under the partition adopted in this work.

As shown in Panel (c), the latent class model tends to provide lower estimates than the previous methods. For 22 out of 31 countries the latent class estimate falls short of the forest-based estimate. Given the set of observed circumstances latent class analysis understates inequality of opportunity by 6% on average. The most pronounced understatement of inequality of opportunity is observed for Belgium and Germany. For these countries the latent class model provides estimates more than 40% lower than the forest-based analogues. However, in spite of the tendency to underestimate, there remain four countries for which latent class analysis overstates inequality of opportunity by more than 50% relative to the forest benchmark. Also in terms of country rankings the latent class approach differs markedly from our preferred method. It identifies Romania (RO), Greece (EL) and Portugal (PT) as the countries in which opportunities are most unequally distributed, whereas these countries rank 6th, 7th and 3rd in the case of forests.

Finally, Panel (d) shows that trees and forests tend to produce similar results. The correlation between estimates is high (0.98) and in contrast to all other approaches there is no general tendency to over- or underestimate inequality of opportunity relative to random forests. In view of the discussed shortcomings of trees, it is unsurprising that some estimates divert from their forest-based analogues. However, even the most notable outliers – Finland at the lower end, and the Netherlands at the upper end – remain well below the extrema of the benchmark methods considered previously.

To summarize: according to our benchmark specifications the parametric and the non-parametric approach tend to overestimate inequality of opportunity. To the contrary, estimates based on latent class analysis tends to underestimate inequality of opportunity. The poor out-of-sample replicability of standard estimation approaches in conjunction with the large divergences of their inequality of opportunity estimates from approaches that perform better in the first dimension, illustrate the importance of appropriate model specifications when comparing societies with respect to their need for opportunity equalizing policy interventions.

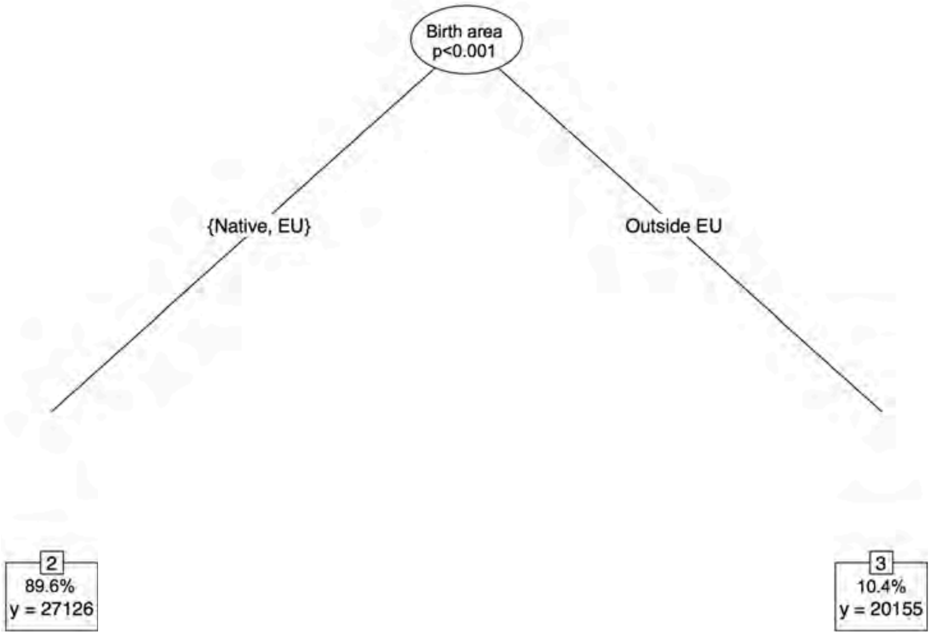
Opportunity Structure

Endowed with an estimate of inequality of opportunity, adequate policy responses must be informed by the opportunity structure of a society. Policymakers want to learn about the particular circumstance characteristics which drive the existence of inequality of opportunity. In this section we illustrate such analyses for both trees and forests. To keep the analysis intelligible we restrict ourselves to two interesting country cases: Sweden and Germany. Readers interested in the opportunity structures of the remaining 29 countries are referred to Supplementary Material B.7.

We are careful to emphasize that one cannot ascribe any causality to our estimates. However, in spite of the correlative nature of the displayed opportunity structures, they may provide useful starting points for decisionmakers to locate policy areas for opportunity equalizing reforms or to stimulate further academic investigation by means of detailed decomposition or causal analyses (Fortin et al., 2011). In the case of trees, it is also worthwhile to keep in mind that their structure remains rather sensitive to small perturbations of the data. In this application, however, tree structures are affirmed by variable importance calculations based on forests which are less sensitive to such perturbations. This validation is a tentative confirmation that graphical tree representations can serve as useful starting points for the analysis of opportunity structures.

Trees. Figure 2.5 illustrates that the opportunity structure of Sweden can be summarized by a tree with two terminal nodes. Inequality of opportunity in Sweden is due to marked differences between first-generation immigrants born outside of Europe and the collective group of native residents and European immigrants. The former group accounts for about 10% of the population and on average obtains an equivalent household income that is 26% lower than the corresponding income of the latter group. The between-type Gini is 0.025 or about 12% of total inequality. We note that our estimates differ from Björklund et al. (2012a)

FIGURE 2.5 – Opportunity Tree (Sweden)



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

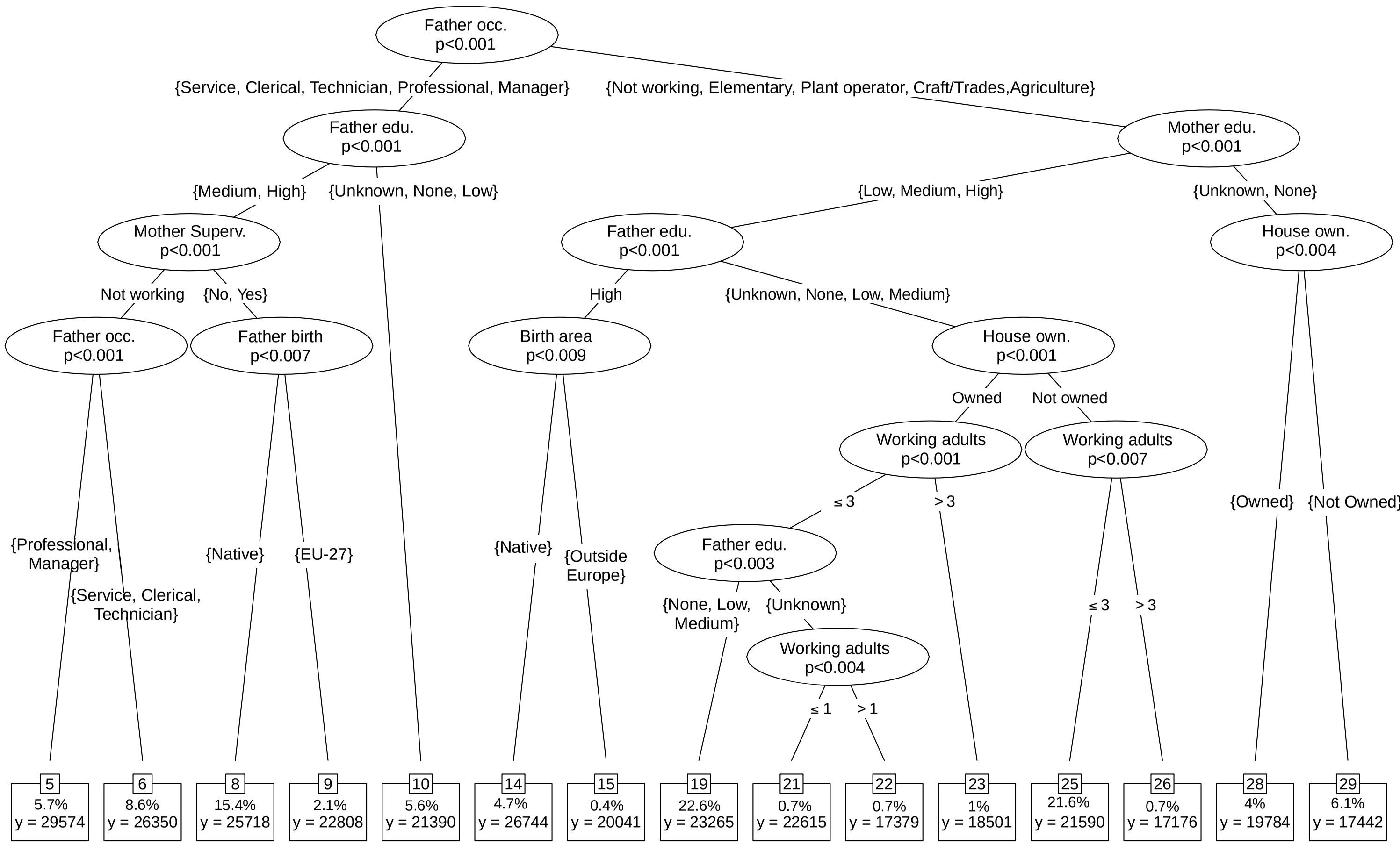
who use Swedish registry data to estimate inequality of opportunity at about 28% of total inequality. These estimates, however, are not strictly comparable to ours since Björklund et al. (2012a) focus on a younger (32-38) male-only sample and market income instead of disposable household income.

A different picture arises when considering Germany (Figure 2.6). Parental occupation, parental education, migration status, the number of working adults in the household, and parental tenancy status interact in creating a complex tree made of 14 splits and 15 terminal nodes. The null hypothesis of equality of opportunity is most firmly rejected for individuals whose fathers work in different occupations. If a respondent's father worked in one of the higher ranked occupations (I-01–I-05), the individual belongs to a more advantaged circumstance type than otherwise (Terminal nodes 5-10). These types together account for 37.4% of the population and have an average outcome of €26,380 – far above the population average of €22,221. However, the advantage of this circumstance characteristic is contingent on the educational status of the father. If a respondent's father had no or low education, the offspring earned less (€21,390) than the country average in spite of the fact that fathers made a career in a high-rank occupation. Conditional on the father both being highly educated and working in a high-rank occupation, the intra-household division of labor plays an important role. On the one hand, those individuals coming from single-earner households in which the mother stayed at home are the most advantaged circumstance types of Germany in 2010 – especially so if their father worked as a manager or professional (Terminal nodes 5 and 6). On the other hand, offspring of double-earner households tend to be differentiated by their migration status. Comparing terminal nodes 8 and 9 we learn that the advantage of coming from a highly-educated double-earner household is substantially diminished from €25,718 to €22,808 if the respondent's father was born outside of Germany. A similar distinction based on migration status can be observed on the right-hand side of the tree, in which individuals were born to fathers with a lower occupational status (I-05–I-00). Individuals in this group lived in above average income households if both of their parents were fairly educated *and* their father had no migration background (Terminal node 14). This advantage again vanishes substantially if the respondent's father was born outside of Europe (Terminal node 15).

There is marked heterogeneity in tree structures across countries (Supplementary Material B.7). For the remaining countries in our sample, terminal nodes range from three (Denmark, Iceland and Norway) to 27 (Italy). It is noteworthy that the rank-rank correlation between the number of terminal nodes and the inequality of opportunity estimates presented in section 2.4 is positive but not perfect. Whether a split is conducted is a function of the average income difference and the sample size of the ensuing types. Hence, if the sample size is large enough, the statistical tests underlying the splitting algorithm have sufficient power to detect even minor differences in average incomes across groups. Such small differences, however, have little impact on inequality in the estimated counterfactual distribution $\hat{y}^C$.

Forests. Forests cannot be analyzed in the straightforward graphical manner of trees. However, we can use variable importance measures to assess the impact of circumstance variables for the construction of opportunity forests. One measure of variable importance, as proposed by Strobl et al. (2007), is obtained by permuting input variable ω^p such that its dependence with y is lost. After this, the out-of-bag error rate MSE^{OOB} is re-computed. The increase of MSE^{OOB} in comparison to the baseline out-of-bag error indicates the importance of the input variable for prediction accuracy. Repeating this procedure for all $\omega^p \in \boldsymbol{\Omega}$ affords a relative comparison of the importance of all circumstances.

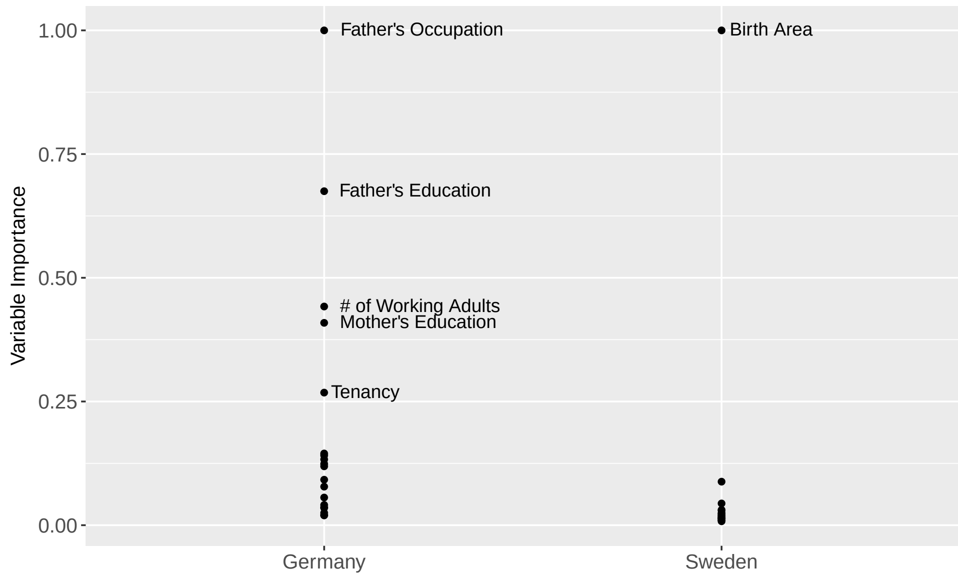
Figure 2.6 – Opportunity Tree (Germany)



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the type mean ($\hat{\mu}_m$).

Figure 2.7 shows the results from this procedure for our example cases of Germany and Sweden. Each black dot is the importance of one of the variables in the set of observed circumstances Ω. We standardize the ensuing results such that the variable importance measure for the circumstance with the greatest impact in each country equals one. For the case of Sweden birth area is the only circumstance that has a meaningful predictive value. In Germany, father's occupation and father's education are most important, followed by the number of working adults in the household and mother's education.

FIGURE 2.7 – Variable Importance for Germany and Sweden



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Each dot shows the importance of a particular circumstance variable ω^P. Variable importance is measured by the decrease in MSE^{OOB} after permuting ω^P such that it is orthogonal to y. The importance measure is standardized such that the circumstance with the greatest importance in each country equals 1. The forests are constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1.

It is reassuring that these findings are in line with the graphical analysis of opportunity trees. In Supplementary Material B.7 we show variable importance plots for all countries in our sample. Broadly, we can divide our country sample into three groups according to the circumstances that determine their opportunity structure. First, there is a handful of primarily Nordic countries where the respondent's birth area is the most important circumstance. Second, there is a large group of primarily Western and Southern European countries where father's occupation and father's education are most important. Third, there is a group of Eastern European countries where mother's education and occupation are most important.

These results have important implications for analyses that use inequality of opportunity estimates as left-hand side variables. Researchers have become increasingly interested in the opportunity equalizing properties of specific policy reforms (e.g. Andreoli et al., 2018). Our results suggest that a one-size-fits-all approach is insufficient to capture the underlying opportunity structures in different societies. Hence, one should be cautious in comparing equality of opportunity estimates based on the same model within a particular country before and after a change in institutional configurations due to a policy reform. While the coefficients on particular circumstance characteristics may change in the course of a reform, the *relevant* model $f()$ may change as well. Therefore, to the extent that researchers are interested in the aggregate opportunity-equalizing effect of a particular reform, they need to take both of these channels into account.

2.5 Conclusion

In this paper we propose the use of conditional inference trees and forests to estimate inequality of opportunity. Both estimation approaches minimize arbitrary model selection by the researcher while trading off downward and upward biases in inequality of opportunity estimates. Conditional inference forests outperform all methods considered in this paper in terms of their out-of-sample performance. Hence, they deliver the best estimates of inequality of opportunity. Conditional inference trees, on the other hand, are econometrically less complex and provide a handy graphical illustration that can be used to analyze opportunity structures. The fact that trees are very close to forests in terms of their out-of-sample performance, their inequality of opportunity estimates, and the importance they assign to specific circumstances makes us confident that they are a useful tool for communicating issues related to inequality of opportunity to a larger audience.

To be sure, the development of machine learning algorithms and their integration into the analytical toolkit of economists is a highly dynamic process. We are well aware that finding the best machine learning algorithm for inequality of opportunity estimations is a methodological horse race with frequent entry of new competitors that eventually will lead to some method outperforming the ones employed in this work. Therefore, the main contribution of this work should be understood as paving the way for new methods that are able to handle the intricacies of model selection for inequality of opportunity estimations. A particularly interesting extension may be the application of local linear forests (Friedberg et al., 2018) that outperform more traditional forest algorithms in their ability to capture the linear impact of particular predictor variables.

Finally, while we restricted ourselves to ex-ante utilitarian measures of inequality of opportunity, the exploration of these algorithms for other methods in the inequality of opportunity literature, such as ex-post measures à la Pistolesi (2009) or ex-ante and ex-post tests à la Kanbur and Snell (2018) and Lefranc et al. (2009), provides another interesting avenue for future research.

Appendix B.1 Empirical Choices

Tuning of Trees. Alternatively to specifying α^* a priori, it can be chosen by K-fold cross-validation (CV), which – under some minimal assumptions (Friedman et al., 2009) – provides unbiased estimates of the out-of-sample MSE. To perform cross-validation, one starts by splitting the sample into K roughly equal-sized folds. Then, one implements the conditional inference algorithm on the union of $K-1$ folds for varying levels of α, while leaving out the kth subsample. This makes it possible to compare the predictions emanating from the $K-1$ folds with the unused data points observed in the kth fold. One then calculates the out-of-sample MSE as a function of α:

$$\text{MSE}_k^{\text{CV}}(\alpha) = \frac{1}{N^k}\sum_{i\in k}(y_i^k - \hat{f}^{-k}(\omega_i;\alpha))^2,\ \omega_i \in \mathbf{\Omega},\ i \in \mathbf{N}, \tag{76}$$

where $\hat{f}^{-k}()$ denotes the estimation function $\hat{f}()$ constructed while leaving out the kth fold. Note that every fold may render a new $\hat{f}()$. This exercise is repeated for all K folds, so that $\text{MSE}^{\text{CV}}(\alpha) = \frac{1}{K}\sum_k \text{MSE}_k^{\text{CV}}(\alpha)$. One then chooses α^* such that

$$\alpha^* = \underset{\alpha}{\text{argmin}}\{\text{MSE}^{\text{CV}}(\alpha) : \alpha \in (0,1)\}. \tag{77}$$

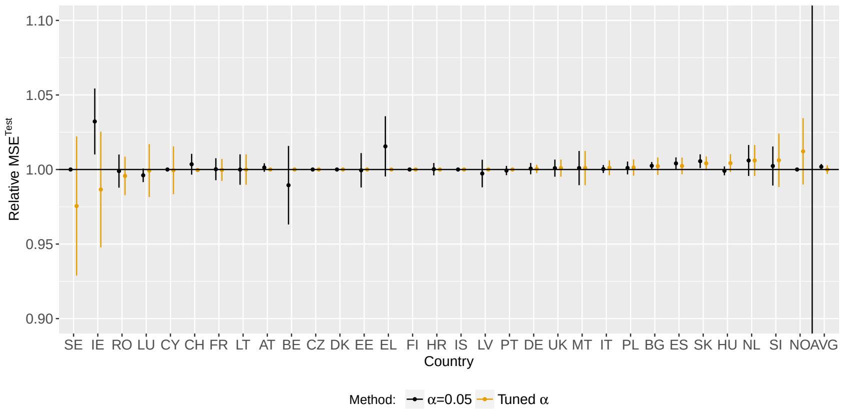
Figure B.1 reveals that selecting α^* based on cross-validation or setting $\alpha^* = 0.05$ has little bearing on our results.

Tuning of Forests. The grid of parameters $(\alpha, \bar{P}, B)$ can be imposed a priori by the researcher or tuned to optimize the out-of-sample fit of the model. In our empirical illustration we proceed as follows. First, to reduce computational costs we fix B^* at a level at which the marginal gain of drawing an additional subsample in terms of out-of-sample prediction accuracy becomes negligible. Empirical tests show that this is the case with $B^* = 200$ for most countries in our sample (Figure B.2).

Second, we determine α^* and $\bar{P}^*$ by minimizing the *out-of-bag* error. This entails the following three steps for a grid of values of α and $\bar{P}$:

1. Run a random forest with B^* subsamples, where $\bar{P}$ circumstances are randomly chosen to be considered at each splitting point, and α is used as the critical value for the hypothesis tests.

FIGURE B.1 – Tuning Conditional Inference Trees



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The y-axis shows the MSE$^{\text{Test}}$ for different specifications of α^* relative to the baseline specification of $\alpha^* = 0.01$. The MSE$^{\text{Test}}$ for the baseline specification of $\alpha^* = 0.01$ is standardized to 1, such that a relative test error > 1 indicates worse fit than the baseline specification. 95% confidence intervals are derived based on 200 bootstrapped re-samples of the test data using the normal approximation method. When no confidence intervals are shown, the methods give the same MSE$^{\text{Test}}$. Trees are constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference trees is detailed in Table 2.1. For the construction of MSE$^{\text{Test}}$, see section 2.4. Black dots and the associated confidence bands show results for $\alpha^* = 0.05$. Orange dots and the associated confidence bands show results for cross-validated α^* using $K = 5$ folds.

2. Calculate the average predicted value of observation i using each of the prediction functions estimated in the subsamples $\mathcal{B}_{-i} := \{\mathcal{S}' \subset \mathcal{S} : \mathcal{S}' \cap \{(y_i, \omega_i)\} = \emptyset\}$ (the so called *bags*) in which i does not enter: $\hat{f}^{\text{OOB}}(\omega_i; \alpha, \bar{P}) = \frac{1}{N_{\mathcal{B}_{-i}}} \sum_{\mathcal{S}' \in \mathcal{B}_{-i}} \hat{f}^{\mathcal{S}'}(\omega_i; \alpha, \bar{P})$.

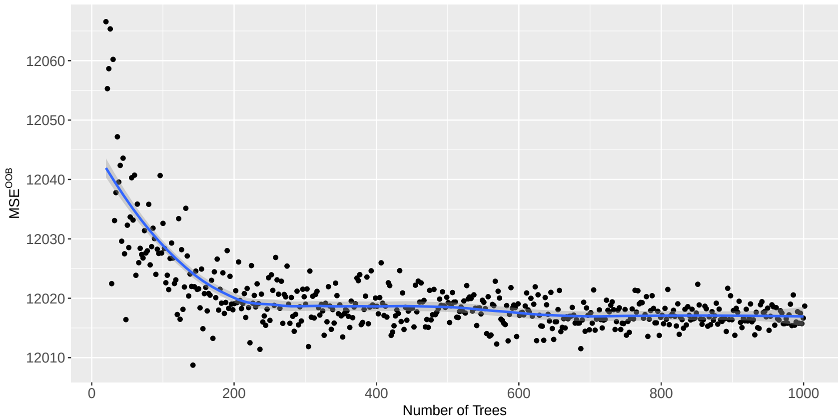
3. Calculate the out-of-bag mean squared error:
MSE$^{\text{OOB}}(\alpha, \bar{P}) = \frac{1}{N} \sum_i [y_i - \hat{f}^{\text{OOB}}(\omega_i; \alpha, \bar{P})]^2$.

One then chooses the combination of parameter values that delivers the lowest MSE$^{\text{OOB}}$:

$$(\alpha^*,\ \bar{P}^*) = \underset{\alpha, \bar{P}}{\text{argmin}}\{\text{MSE}^{\text{OOB}} : (\alpha, \bar{P}) \in (0,1) \times \bar{\mathbf{P}}\}. \tag{78}$$

The logic behind this tuning exercise is similar to cross-validation. However, instead of leaving out the kth fraction of the dataset to make out-of-sample predictions, we leverage the fact that each tree of a forest is grown on a subsample $\mathcal{S}' \subset \mathcal{S}$ that excludes some observations $i \in \{1, ..., S\}$. Hence, for each tree we can use the out-of-bag data points to evaluate the predictive accuracy of the respective model. Using this out-of-bag procedure, the optimal level of α is often very high, meaning that the trees are grown very deep. At the limit, α may be set to 1, in which case splits are made as long as each end node has at least a minimum number of observations. If we were to extract a single tree from such a forest, then this tree

Figure B.2 – Optimal Size of Forests



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The x-axis shows the parameter value for B^*, i.e. the number of trees per forest. The dots show the MSE$^{\text{OOB}}$ obtained from estimating a random forest with the given number of trees for the case of Germany. We allow for 6 circumstances to be considered at each splitting point ($\bar{P}^* = 6$). Due to the randomness in the observations selected for each tree and the randomness in the circumstances allowed at each splitting point, even when estimating multiple forests with the same number of trees, the associated MSE$^{\text{OOB}}$ will differ. The blue line is a non-parametric fitted line of the MSE$^{\text{OOB}}$ estimates and the shaded area the 95% confidence interval of this line. Evidently, as the tree size approaches 200, on expectation, the MSE$^{\text{OOB}}$ stops improving much.

would surely overfit the data and perform poorly out-of-sample. However, when averaging over many overfitted trees this drawbacks gets remedied. Hence, setting $\alpha^* = 1$ a priori is a sensible strategy to reduce the computational cost of random forests.

Appendix B.2 Upward Bias, Downward Bias and the MSE

A standard statistic adopted to assess prediction accuracy is the mean squared error (MSE):

$$\text{MSE} = \mathbb{E}_{\mathcal{S}}[(y - \hat{f}(\omega))^2], \tag{79}$$

where y is the observed outcome and $\hat{f}(\omega)$ the estimator of the individual's conditional expectation $\mathbb{E}(y|\omega)$ in a random sample $\mathcal{S}$. The MSE can be decomposed into three components (Friedman et al., 2009):

$$\begin{aligned} \text{MSE} &= \text{Var}(\hat{f}(\omega)) + \mathbb{E}_{\mathcal{S}}[f(\omega) - \hat{f}(\omega)]^2 + \text{Var}(\epsilon), && (80)\\ &= \underbrace{\text{Var}(f(\omega) - \hat{f}(\omega))}_{(1)} + \underbrace{(f(\omega) - \mathbb{E}_{\mathcal{S}}[\hat{f}(\omega)])^2}_{(2)} + \underbrace{\text{Var}(\epsilon)}_{(3)}. && (81)\end{aligned}$$

In the literature on statistical learning this is referred to as the bias-variance decomposition. Components (1) and (2) can be directly linked to concerns of upward and downward biases in inequality of opportunity estimates. To illustrate this point, notice that we minimize (1) by imposing the following model specification $y = \hat{f}(\omega) + \epsilon = \beta_0 + \epsilon$. This model specification assumes that individual outcomes are best predicted by the sample mean $\mu^{\mathcal{S}}$. For the sake of illustration, furthermore assume that each population sample $\mathcal{S}$ is large enough such that its mean corresponds to the underlying population truth: $\mu^{\mathcal{S}} = \mu$. Obviously, this is a stark assumption which we only make for illustration purposes. In reality, there will always be some variance in the sample means as long as one does not capture the entire underlying population. As a consequence, (1) drops out and the MSE is entirely captured by components (2) and (3):

$$\begin{aligned} \text{MSE} &= \text{Var}(f(\omega) - \hat{f}(\omega)) + (f(\omega) - \mathbb{E}_{\mathcal{S}}[\hat{f}(\omega)])^2 + \text{Var}(\epsilon)\\ &= (f(\omega) - \mu)^2 + \text{Var}(\epsilon). \end{aligned}$$

This shows that the variance-minimizing model is established by the assumption that everybody in the population faces exactly the same circumstance set, i.e. $\omega_i = 1 \, \forall i \in \mathbf{N}$ – and hence that the value of every individual opportunity set is best estimated by the sample mean $\mu^{\mathcal{S}}$. Under the given assumptions, the model cannot give an upward biased estimate of inequality of opportunity since it is restricted in a way that does not allow for any role of circumstance characteristics in the explanation of individual outcome differences. In fact, for any functional $I()$ that satisfies the measurement criteria outlined in section 2.2, $I(\hat{y}^C) = 0$.

Reversely, one could ask which model would minimize component (2) of the MSE. To this end, we would have to specify a complex model that allows for the full set of relevant circumstances, their mutual interactions and non-linearities such that *in expectation* we would obtain unbiased estimates of $f(\omega)$. In this case the MSE would be entirely captured by components (1) and (3):

$$\begin{aligned}\text{MSE} &= \text{Var}(f(\omega) - \hat{f}(\omega)) + (f(\omega) - \mathbb{E}_{\mathcal{S}}[\hat{f}(\omega)])^2 + \text{Var}(\epsilon),\\ &= \text{Var}(f(\omega) - \hat{f}(\omega)) + \text{Var}(\epsilon).\end{aligned}$$

While such a model in expectation provides unbiased estimates of y^C, the conditional expectations within a particular sample $\mathcal{S}$ may be estimated with error: $\hat{y}_i^C$=$y_i^C + u_i$. u_i is an iid error the importance of which tends to increase with model complexity (Friedman et al., 2009). Hence, model complexity leads to measurement error which in turn inflates the variance of $\hat{y}^C$ in comparison to the underlying truth: Var$(\hat{y}^C) =$ Var$(y^C) +$ Var(u). As shown in Brunori et al. (2019), applying any functional $I()$ that satisfies the measurement criteria outlined in section 2.2 to the variance inflated distribution $\hat{y}^C$ results in upward biased estimates of inequality of opportunity.

Appendix B.3 Sensitivity to Sample Size

Table B.1 shows the country sample sizes from our empirical application (Section 2.4) by estimation method. Note that the parametric and the non-parametric approach tend to have smaller sample sizes since they rely on list-wise deletion of observations in case of item non-response.

Table B.1 – Sample Size by Method

Country	Parametric Approach	Non-parametric Approach	CI Forests and Trees
AT	6,042	6,107	6,220
BE	4,528	5,375	6,011
BG	5,952	6,210	7,154
CH	6,420	6,754	7,583
CY	4,483	4,525	4,589
CZ	6,438	6,524	8,711
DE	10,539	11,139	12,683
DK	2,107	2,223	5,897
EE	4,857	5,004	5,338
EL	5,743	5,862	6,184
ES	14,640	14,816	15,481
FI	2,900	3,207	9,743
FR	10,104	10,391	11,078
HR	5,945	6,159	6,969
HU	12,139	12,525	13,330
IE	3,080	3,138	4,318
IS	1,447	1,492	3,684
IT	20,238	20,800	21,070
LT	4,539	4,703	5,403
LU	6,528	6,654	6,765
LV	6,046	6,192	6,423
MT	4,048	4,117	4,701
NL	5,414	5,518	11,411
NO	2,329	2,400	5,026
PL	12,676	13,182	15,545
PT	5,689	5,795	5,899
RO	5,701	6,145	7,867
SE	467	561	6,599
SI	4,691	4,747	13,183
SK	6,170	6,401	6,779
UK	5,756	5,922	7,391
Minimum (N^{min})	467	561	3,684

Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Each column refers to the number of observations used in the estimation of inequality of opportunity for the particular approach.

To analyze the extent to which inequality of opportunity estimates are a function of sample sizes we rely on the following procedure.

1. For each country-method cell we make 10 random draws from the full country sample. The size of each subsample is determined by the smallest method-specific country sample: $N^{\text{min}}_{\text{trees}} = N^{\text{min}}_{\text{forests}} = 3,684$, $N^{\text{min}}_{\text{non-parametric}} = 561$ and $N^{\text{min}}_{\text{parametric}} = 467$.

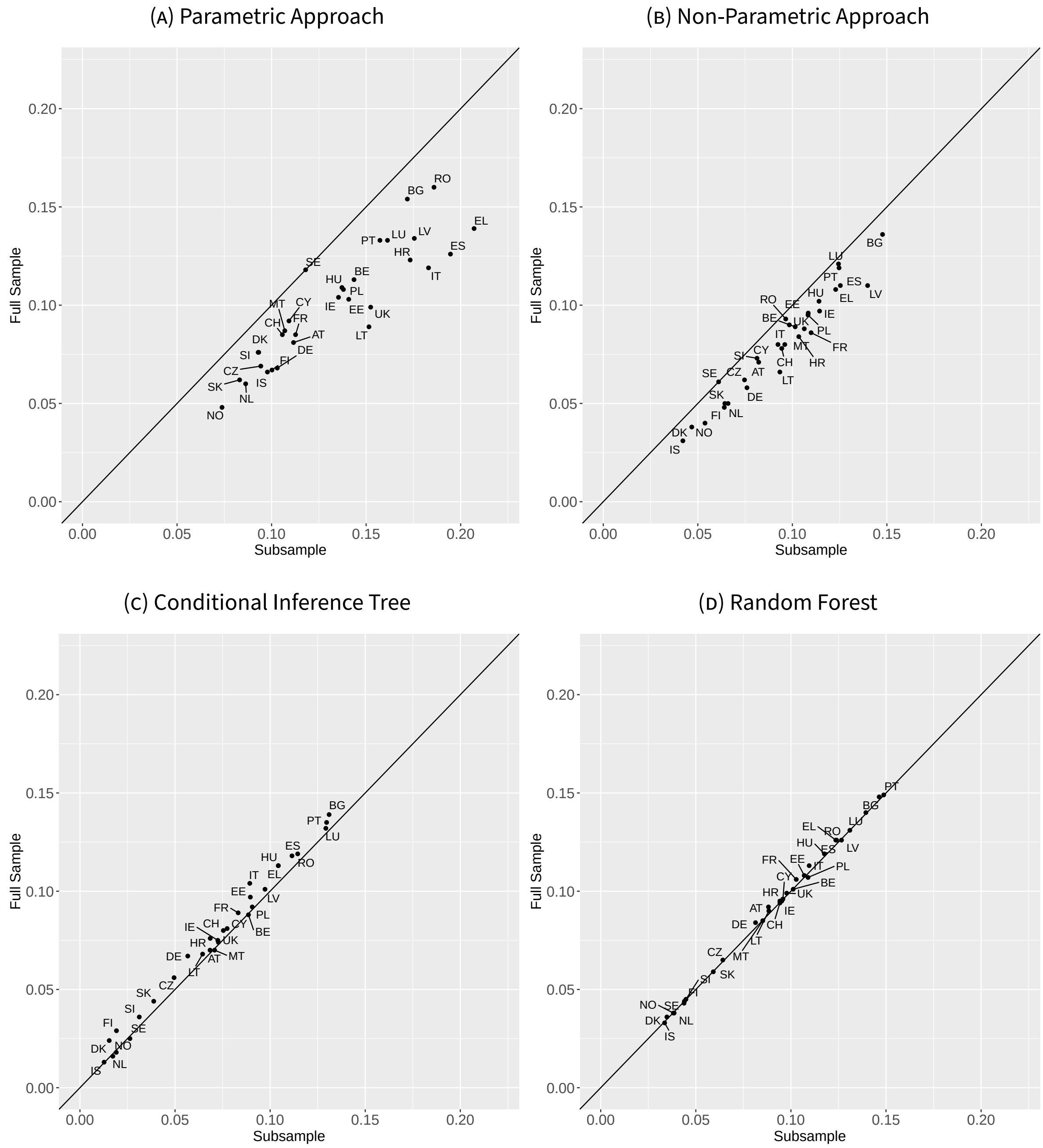
2. We estimate inequality of opportunity on each of the 10 subsamples and average over these 10 iterations to obtain an estimate for each country-method cell.

Figure B.1 plots the estimates based on the full sample against the estimates from the subsamples as derived from the procedure outlined above.

Panels (a) and (b) illustrate that the parametric and non-parametric approaches tend to overestimate inequality of opportunity as sample sizes decline. This is a direct consequence of fixing the number of model parameters a priori. As the available degrees of freedom decline, the model tends to overfit, which translates into upward biased estimates of inequality of opportunity (see Appendix Section B.2). To the contrary, in the case of trees and forests (Panels (c) and (d)) country estimates align closely along the 45-degree line. This pattern illustrates that trees and forests are less sensitive to variations in the sample size. On the one hand, if the sample size is small, the p-values of the hypothesis tests will increase and less splits will be conducted. This prevents the models from overfitting and safeguards the inequality of opportunity estimate against overestimation. On the other hand, both methods allow for extremely flexible functional forms in the construction of $\hat{y}^C$. Hence, even when reducing the sample size by a factor of 5.8 (Italy) the estimate from the subsample closely aligns with the inequality of opportunity estimate from the full country sample.

We conclude that inequality of opportunity comparisons across countries based on trees and forests are more robust to sample size variations than alternative estimation approaches.

FIGURE B.1 – Inequality of opportunity with full sample and random subsamples



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: In each panel the x-axis indicates the country-specific inequality of opportunity estimate of the respective estimation approach on the subsample $N^{\min}$, which differs by method. Analogously, the y-axis indicates the country-specific inequality of opportunity estimate of the respective estimation approach on the full sample N. For all methods inequality of opportunity is measured by the Gini coefficient in the estimated counterfactual distribution $\hat{y}^C$. The black solid line is the 45-degree line. Country-estimates above the 45-degree line indicate an underestimation of inequality of opportunity relative to the full sample benchmark. Reversely, country-estimates below the 45-degree line indicate an overestimation of inequality of opportunity relative to the full sample benchmark. To avoid the intricacies of tuning forests, we set $\bar{P}^* = 8$ and $\alpha^* = 1$. As a consequence, our benchmark estimates may slightly differ from the ones reported in the main text.

Appendix B.4 Point Estimates and Confidence Intervals

Table B.1 – Inequality of Opportunity Estimates

	Benchmark Methods			Conditional Inference	
Country	Parametric	Non-Parametric	Latent Class	Tree	Random Forest
AT	0.089 [0.081;0.097]	0.075 [0.067;0.083]	0.080 [0.070;0.090]	0.087 [0.076;0.097]	0.088 [0.080;0.096]
BE	0.111 [0.100;0.121]	0.087 [0.080;0.094]	0.053 [0.036;0.071]	0.087 [0.078;0.096]	0.091 [0.084;0.098]
BG	0.154 [0.145;0.163]	0.136 [0.126;0.145]	0.115 [0.106;0.124]	0.136 [0.127;0.146]	0.134 [0.124;0.144]
CH	0.092 [0.081;0.103]	0.083 [0.075;0.091]	0.063 [0.047;0.079]	0.080 [0.068;0.091]	0.090 [0.082;0.098]
CY	0.094 [0.085;0.103]	0.083 [0.073;0.094]	0.074 [0.058;0.090]	0.080 [0.066;0.094]	0.080 [0.065;0.095]
CZ	0.072 [0.065;0.079]	0.066 [0.059;0.073]	0.060 [0.051;0.069]	0.057 [0.048;0.066]	0.051 [0.044;0.058]
DE	0.070 [0.063;0.078]	0.059 [0.053;0.064]	0.047 [0.039;0.054]	0.070 [0.062;0.077]	0.079 [0.074;0.085]
DK	0.077 [0.046;0.108]	0.041 [0.030;0.052]	0.029 [0.018;0.040]	0.021 [0.011;0.031]	0.020 [0.015;0.026]
EE	0.111 [0.098;0.124]	0.102 [0.091;0.113]	0.074 [0.059;0.090]	0.097 [0.084;0.110]	0.101 [0.088;0.113]
EL	0.148 [0.130;0.165]	0.121 [0.110;0.132]	0.117 [0.099;0.134]	0.126 [0.111;0.142]	0.109 [0.094;0.124]
ES	0.142 [0.132;0.152]	0.120 [0.114;0.126]	0.089 [0.069;0.109]	0.128 [0.122;0.135]	0.120 [0.105;0.135]
FI	0.069 [0.049;0.088]	0.052 [0.041;0.062]	0.048 [0.032;0.063]	0.020 [0.009;0.031]	0.028 [0.021;0.034]
FR	0.086 [0.080;0.092]	0.086 [0.080;0.093]	0.072 [0.062;0.081]	0.090 [0.082;0.099]	0.098 [0.092;0.104]
HR	0.131 [0.117;0.146]	0.088 [0.080;0.097]	0.076 [0.064;0.088]	0.082 [0.070;0.095]	0.076 [0.066;0.087]
HU	0.110 [0.104;0.116]	0.103 [0.098;0.109]	0.095 [0.087;0.104]	0.113 [0.108;0.119]	0.108 [0.102;0.114]
IE	0.105 [0.092;0.118]	0.097 [0.087;0.108]	0.048 [0.029;0.068]	0.084 [0.070;0.099]	0.078 [0.069;0.087]
IS	0.067 [0.029;0.104]	0.032 [0.021;0.043]	0.030 [0.017;0.042]	0.012 [0.004;0.021]	0.016 [0.010;0.022]
IT	0.121 [0.113;0.130]	0.091 [0.086;0.095]	0.080 [0.068;0.091]	0.108 [0.102;0.113]	0.097 [0.090;0.104]
LT	0.095 [0.079;0.110]	0.067 [0.058;0.077]	0.059 [0.048;0.070]	0.069 [0.053;0.085]	0.067 [0.055;0.080]
LU	0.134 [0.125;0.143]	0.121 [0.114;0.127]	0.090 [0.072;0.109]	0.133 [0.125;0.140]	0.136 [0.130;0.142]
LV	0.134 [0.119;0.148]	0.110 [0.100;0.120]	0.095 [0.079;0.112]	0.110 [0.097;0.124]	0.111 [0.100;0.122]
MT	0.087 [0.075;0.099]	0.080 [0.071;0.089]	0.057 [0.047;0.067]	0.071 [0.059;0.083]	0.072 [0.062;0.082]
NL	0.066 [0.050;0.082]	0.053 [0.047;0.059]	0.041 [0.029;0.053]	0.028 [0.020;0.037]	0.019 [0.015;0.024]
NO	0.048 [0.032;0.064]	0.041 [0.031;0.050]	0.030 [0.019;0.041]	0.020 [0.012;0.028]	0.023 [0.018;0.029]
PL	0.111 [0.104;0.118]	0.097 [0.091;0.104]	0.095 [0.088;0.102]	0.102 [0.095;0.109]	0.099 [0.092;0.106]
PT	0.138 [0.128;0.148]	0.124 [0.113;0.134]	0.116 [0.102;0.129]	0.136 [0.124;0.149]	0.127 [0.114;0.140]
RO	0.170 [0.158;0.182]	0.104 [0.094;0.114]	0.119 [0.105;0.134]	0.120 [0.109;0.132]	0.111 [0.100;0.122]
SE	0.118 [0.037;0.199]	0.060 [0.043;0.078]	0.025 [0.007;0.043]	0.025 [0.016;0.033]	0.031 [0.025;0.038]
SI	0.077 [0.069;0.085]	0.073 [0.066;0.080]	0.059 [0.051;0.067]	0.032 [0.024;0.039]	0.036 [0.032;0.040]
SK	0.063 [0.055;0.071]	0.051 [0.045;0.057]	0.042 [0.033;0.051]	0.050 [0.041;0.058]	0.046 [0.039;0.053]
UK	0.101 [0.087;0.115]	0.090 [0.080;0.099]	0.062 [0.042;0.082]	0.071 [0.056;0.087]	0.079 [0.071;0.087]

Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Inequality of opportunity is measured by the Gini coefficient in the estimated counterfactual distribution $\hat{y}^C$. $\hat{y}^C$ is constructed by the respective estimation approach indicated in the table header. 95% confidence intervals are derived based on 200 bootstrapped re-samples using the normal approximation method.

Appendix B.5 Descriptive Statistics

Table B.2 – Descriptive Statistics (Individual and Household)

		Birth Area		Parents in HH		HH Composition			
Country	Male	Native	EU	Both	None	Adults	Working Ad.	Children	Home Owner
AT	0.501	0.790	0.070	0.856	0.017	2.730	1.760	2.600	0.585
BE	0.498	0.824	0.076	0.855	0.019	2.380	1.590	2.780	0.750
BG	0.500	0.994	0.001	0.904	0.012	2.440	2.010	2.070	0.910
CH	0.505	0.684	0.197	0.837	0.017	2.550	1.900	2.530	0.546
CY	0.525	0.787	0.096	0.900	0.015	2.640	1.670	2.700	0.784
CZ	0.508	0.964	0.026	0.851	0.013	2.090	1.920	2.240	0.597
DE	0.496	0.868	0.000	0.830	0.020	2.240	1.680	2.320	0.499
DK	0.505	0.923	0.026	0.809	0.027	2.220	2.310	2.240	0.736
EE	0.525	0.847	0.000	0.756	0.011	2.100	1.800	2.090	0.859
EL	0.498	0.890	0.025	0.931	0.019	2.310	1.560	2.330	0.834
ES	0.495	0.834	0.051	0.893	0.012	2.880	2.110	2.430	0.819
FI	0.499	0.954	0.018	0.829	0.016	2.360	1.750	2.300	0.772
FR	0.509	0.885	0.036	0.820	0.022	2.470	1.660	1.750	0.630
HR	0.501	0.875	0.017	0.874	0.020	2.560	1.350	2.310	0.902
HU	0.517	0.988	0.008	0.844	0.041	2.140	1.750	2.270	0.830
IE	0.524	0.783	0.149	0.893	0.078	3.170	3.200	3.200	0.727
IS	0.507	0.920	0.042	0.899	0.012	2.420	1.900	2.630	0.893
IT	0.502	0.880	0.040	0.901	0.011	2.590	1.620	2.410	0.685
LT	0.521	0.939	0.004	0.846	0.016	2.320	2.020	2.460	0.698
LU	0.499	0.480	0.401	0.868	0.020	2.530	1.640	2.710	0.734
LV	0.520	0.865	0.000	0.763	0.012	1.970	1.760	2.280	0.455
MT	0.497	0.944	0.000	0.932	0.020	3.020	1.840	2.680	0.576
NL	0.509	0.903	0.020	0.882	0.016	2.100	1.540	3.250	0.575
NO	0.511	0.907	0.041	0.913	0.014	2.020	1.760	1.870	0.922
PL	0.504	0.999	0.000	0.889	0.015	2.700	1.960	2.440	0.644
PT	0.506	0.906	0.022	0.854	0.017	2.680	2.230	2.680	0.544
RO	0.506	0.999	0.000	0.919	0.009	2.770	1.900	2.270	0.861
SE	0.493	0.846	0.050	0.820	0.035	2.070	1.780	2.350	0.757
SI	0.496	0.876	0.000	0.855	0.019	2.530	1.770	2.200	0.746
SK	0.519	0.987	0.010	0.920	0.010	2.520	2.080	2.340	0.694
UK	0.507	0.848	0.042	0.825	0.024	2.340	2.240	2.410	0.649

Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Omitted circumstance expressions listed in order of the circumstance categories are: “Female”; “Non-EU”; “Only Mother/Only Father/Collective House”; “House Not Owned”.

Table B.3 – Descriptive Statistics (Fathers)

	Birth Area		Citizenship		Education			Activity				Occupation (ISCO-08 1-Digit)									
Country	Native	EU	Resid.	EU	Low	Med.	High	Empl.	Self-Empl.	Un-empl.	House Work	1	2	3	4	5	6	7	8	9	Superv.
AT	0.743	0.093	0.777	0.068	0.398	0.421	0.135	0.714	0.215	0.003	0.001	0.043	0.046	0.064	0.051	0.138	0.145	0.284	0.063	0.085	0.338
BE	0.748	0.100	0.762	0.093	0.491	0.199	0.178	0.699	0.179	0.007	0.002	0.068	0.126	0.104	0.084	0.054	0.057	0.209	0.127	0.041	0.278
BG	0.933	0.004	0.936	0.001	0.466	0.333	0.081	0.899	0.028	0.005	0.000	0.022	0.065	0.047	0.029	0.035	0.135	0.216	0.207	0.142	0.093
CH	0.588	0.286	0.603	0.280	0.227	0.487	0.151	0.653	0.292	0.001	0.000	0.086	0.131	0.140	0.057	0.060	0.111	0.223	0.077	0.054	0.397
CY	0.803	0.082	0.808	0.094	0.667	0.178	0.091	0.566	0.381	0.004	0.000	0.011	0.071	0.074	0.029	0.104	0.161	0.245	0.122	0.125	0.229
CZ	0.878	0.065	0.910	0.036	0.602	0.216	0.090	0.891	0.017	0.001	0.000	0.033	0.070	0.125	0.036	0.035	0.039	0.305	0.195	0.053	0.233
DE	0.800	0.200	0.855	0.145	0.125	0.496	0.213	0.819	0.123	0.008	0.001	0.046	0.104	0.158	0.051	0.061	0.059	0.266	0.154	0.040	0.299
DK	0.935	0.025	0.970	0.020	0.368	0.418	0.214	0.708	0.272	0.004	0.001	0.111	0.122	0.070	0.043	0.103	0.160	0.288	0.072	0.009	0.447
EE	0.603	0.270	0.637	0.233	0.300	0.338	0.165	0.823	0.006	0.003	0.002	0.076	0.092	0.053	0.014	0.013	0.034	0.221	0.253	0.053	0.153
EL	0.887	0.016	0.911	0.015	0.587	0.135	0.084	0.449	0.517	0.002	0.000	0.073	0.047	0.026	0.087	0.046	0.308	0.210	0.099	0.055	0.182
ES	0.836	0.047	0.846	0.046	0.762	0.064	0.081	0.702	0.219	0.006	0.001	0.056	0.045	0.076	0.055	0.087	0.145	0.191	0.113	0.137	0.191
FI	0.827	0.007	0.827	0.007	0.491	0.182	0.162	0.592	0.209	0.016	0.001	0.041	0.089	0.085	0.016	0.046	0.135	0.146	0.138	0.044	
FR	0.789	0.078	0.857	0.057	0.695	0.073	0.095	0.753	0.170	0.003	0.001	0.084	0.068	0.111	0.072	0.038	0.103	0.155	0.055	0.223	0.335
HR	0.822	0.006	0.834	0.004	0.464	0.312	0.063	0.763	0.103	0.037	0.016	0.025	0.041	0.088	0.036	0.072	0.049	0.214	0.103	0.228	0.129
HU	0.962	0.017	0.969	0.012	0.599	0.241	0.087	0.892	0.043	0.001	0.001	0.037	0.060	0.052	0.017	0.053	0.094	0.279	0.193	0.137	0.117
IE	0.792	0.107	0.758	0.094	0.574	0.258	0.112	0.659	0.221	0.049	0.002	0.104	0.092	0.042	0.022	0.072	0.155	0.149	0.065	0.158	0.344
IS	0.918	0.050	0.923	0.044	0.334	0.486	0.139	0.638	0.332	0.001	0.000	0.115	0.121	0.076	0.024	0.094	0.180	0.220	0.094	0.042	0.570
IT	0.823	0.022	0.827	0.020	0.708	0.136	0.038	0.614	0.244	0.016	0.004	0.054	0.040	0.074	0.057	0.068	0.099	0.227	0.105	0.118	0.199
LT	0.899	0.004	0.926	0.004	0.538	0.228	0.085	0.916	0.011	0.000	0.001	0.049	0.074	0.038	0.017	0.023	0.080	0.241	0.179	0.214	0.110
LU	0.387	0.467	0.400	0.466	0.484	0.316	0.120	0.757	0.174	0.001	0.001	0.063	0.093	0.118	0.048	0.035	0.112	0.228	0.183	0.039	0.251
LV	0.572	0.248	0.642	0.165	0.381	0.297	0.098	0.767	0.005	0.002	0.003	0.036	0.083	0.037	0.010	0.019	0.069	0.199	0.218	0.083	0.070
MT	0.952	0.041	0.953	0.040	0.561	0.180	0.059	0.717	0.214	0.013	0.001	0.062	0.046	0.106	0.045	0.141	0.050	0.244	0.099	0.106	0.225
NL	0.829	0.028	0.888	0.022	0.376	0.285	0.198	0.726	0.173	0.006	0.006	0.087	0.124	0.155	0.051	0.069	0.086	0.200	0.079	0.031	0.310
NO	0.897	0.046	0.908	0.041	0.328	0.390	0.278	0.712	0.255	0.002	0.001	0.116	0.110	0.167	0.029	0.057	0.111	0.227	0.100	0.032	0.285
PL	0.955	0.012	0.980	0.003	0.462	0.448	0.070	0.701	0.238	0.002	0.001	0.036	0.044	0.053	0.025	0.042	0.237	0.254	0.157	0.078	0.111
PT	0.932	0.006	0.945	0.006	0.700	0.031	0.031	0.650	0.248	0.002	0.001	0.047	0.032	0.060	0.038	0.082	0.185	0.264	0.114	0.077	0.190
RO	0.938	0.001	0.939	0.001	0.726	0.088	0.030	0.642	0.237	0.004	0.013	0.004	0.040	0.034	0.016	0.018	0.253	0.249	0.121	0.104	0.045
SE	0.945	0.022	0.851	0.061	0.422	0.350	0.182	0.745	0.211	0.002	0.001	0.043	0.118	0.067	0.031	0.092	0.086	0.230	0.108	0.019	0.337
SI	0.769	0.200	0.000	0.000	0.684	0.166	0.085	0.773	0.099	0.013	0.011	0.024	0.052	0.100	0.037	0.052	0.089	0.257	0.080	0.173	0.242
SK	0.935	0.020	0.945	0.011	0.362	0.497	0.075	0.921	0.011	0.002	0.001	0.042	0.060	0.095	0.028	0.043	0.030	0.285	0.209	0.128	0.145
UK	0.800	0.064	0.869	0.039	0.508	0.228	0.150	0.795	0.147	0.025	0.002	0.095	0.142	0.085	0.040	0.075	0.036	0.236	0.133	0.083	0.398

Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Omitted circumstance expressions listed in order of the circumstance categories are: “Non-EU”; “Not Europe”; “Dead/Unknown/Illiterate”; “Dead/Unknown/Retired/Other Inactive”; "Dead/Unknown/Not Working/Armed Forces”; “Dead/Unknown/Not Working/Non-Supervisory”. Compare also to Table 2.1.

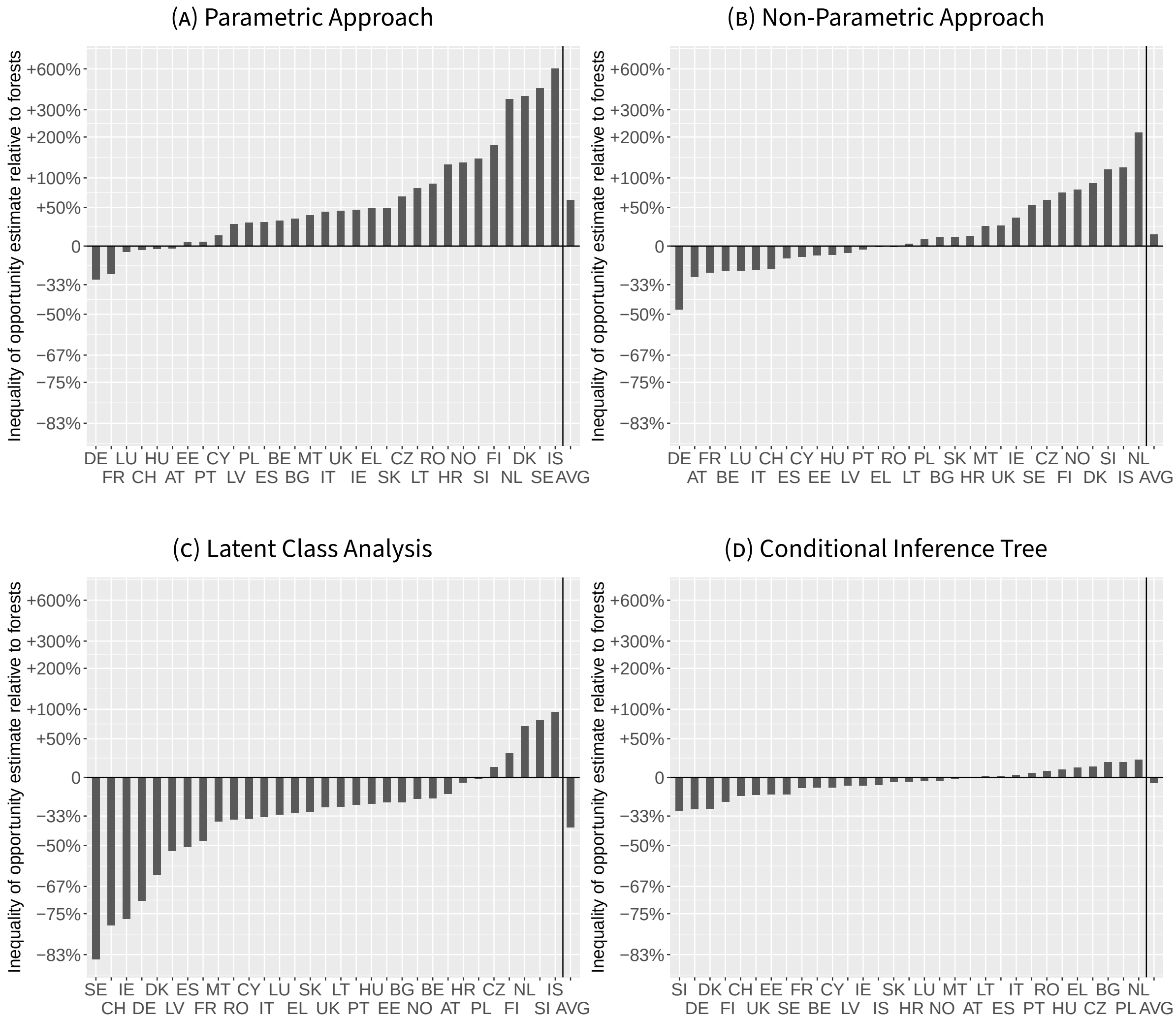
Table B.4 – Descriptive Statistics (Mothers)

	Birth Area		Citizenship		Education			Activity				Occupation (ISCO-08 1-Digit)									
Country	Native	EU	Resid.	EU.	Low	Med.	High.	Empl.	Self-Empl.	Un-empl.	House Work	1	2	3	4	5	6	7	8	9	Superv.
AT	0.740	0.096	0.789	0.065	0.587	0.328	0.041	0.369	0.169	0.002	0.435	0.009	0.024	0.009	0.071	0.154	0.128	0.045	0.010	0.087	0.092
BE	0.755	0.097	0.790	0.092	0.564	0.201	0.126	0.320	0.117	0.006	0.508	0.009	0.081	0.045	0.058	0.046	0.002	0.016	0.024	0.069	0.034
BG	0.931	0.003	0.981	0.002	0.464	0.357	0.099	0.878	0.026	0.007	0.058	0.008	0.123	0.040	0.092	0.140	0.181	0.099	0.064	0.152	0.030
CH	0.567	0.307	0.599	0.286	0.410	0.399	0.057	0.382	0.152	0.001	0.429	0.027	0.056	0.069	0.069	0.125	0.055	0.039	0.025	0.068	0.064
CY	0.804	0.080	0.812	0.091	0.684	0.162	0.057	0.325	0.166	0.001	0.492	0.001	0.045	0.020	0.037	0.067	0.036	0.022	0.042	0.220	0.048
CZ	0.882	0.061	0.946	0.037	0.670	0.261	0.043	0.898	0.007	0.003	0.064	0.011	0.080	0.104	0.149	0.160	0.074	0.105	0.080	0.139	0.088
DE	0.811	0.189	0.862	0.138	0.284	0.475	0.081	0.482	0.050	0.009	0.438	0.012	0.051	0.079	0.089	0.116	0.025	0.015	0.087	0.033	0.059
DK	0.922	0.029	0.935	0.023	0.531	0.283	0.186	0.630	0.069	0.006	0.274	0.020	0.103	0.095	0.123	0.225	0.035	0.052	0.026	0.001	0.122
EE	0.601	0.272	0.726	0.250	0.334	0.391	0.208	0.906	0.004	0.001	0.051	0.049	0.169	0.109	0.097	0.110	0.084	0.051	0.124	0.113	0.085
EL	0.888	0.016	0.916	0.016	0.592	0.133	0.044	0.193	0.277	0.001	0.517	0.023	0.027	0.004	0.039	0.048	0.223	0.034	0.021	0.049	0.026
ES	0.836	0.046	0.849	0.046	0.802	0.048	0.040	0.186	0.069	0.001	0.718	0.010	0.025	0.010	0.021	0.059	0.028	0.021	0.009	0.071	0.029
FI	0.826	0.007	0.933	0.006	0.559	0.238	0.145	0.658	0.204	0.019	0.055	0.012	0.126	0.091	0.122	0.145	0.048	0.046	0.057	0.202	
FR	0.806	0.067	0.880	0.047	0.724	0.079	0.072	0.454	0.085	0.001	0.424	0.011	0.036	0.050	0.111	0.107	0.059	0.049	0.005	0.109	0.072
HR	0.823	0.008	0.848	0.003	0.634	0.189	0.043	0.352	0.053	0.027	0.538	0.003	0.058	0.036	0.046	0.070	0.022	0.034	0.013	0.122	0.033
HU	0.964	0.016	0.980	0.012	0.655	0.243	0.053	0.729	0.022	0.001	0.216	0.014	0.049	0.063	0.113	0.118	0.061	0.075	0.087	0.167	0.044
IE	0.787	0.114	0.761	0.103	0.546	0.324	0.097	0.253	0.048	0.007	0.673	0.022	0.061	0.007	0.052	0.059	0.017	0.014	0.007	0.060	0.082
IS	0.905	0.059	0.924	0.046	0.626	0.275	0.075	0.598	0.102	0.001	0.278	0.030	0.095	0.045	0.109	0.180	0.064	0.028	0.013	0.130	0.149
IT	0.820	0.024	0.862	0.024	0.779	0.112	0.023	0.224	0.080	0.005	0.637	0.011	0.038	0.022	0.029	0.051	0.035	0.031	0.022	0.062	0.041
LT	0.902	0.002	0.959	0.003	0.519	0.316	0.106	0.867	0.014	0.001	0.095	0.035	0.129	0.046	0.049	0.109	0.067	0.112	0.034	0.293	0.068
LU	0.374	0.483	0.393	0.485	0.587	0.245	0.071	0.318	0.106	0.000	0.538	0.028	0.049	0.046	0.036	0.061	0.054	0.015	0.024	0.108	0.047
LV	0.585	0.234	0.793	0.182	0.414	0.399	0.125	0.891	0.003	0.002	0.064	0.031	0.138	0.084	0.098	0.121	0.085	0.093	0.023	0.221	0.074
MT	0.950	0.043	0.957	0.038	0.652	0.145	0.026	0.073	0.015	0.001	0.894	0.003	0.019	0.007	0.009	0.018	0.002	0.004	0.009	0.010	0.011
NL	0.829	0.027	0.907	0.023	0.532	0.288	0.087	0.282	0.056	0.003	0.628	0.010	0.050	0.038	0.052	0.089	0.016	0.011	0.008	0.060	0.037
NO	0.877	0.048	0.891	0.043	0.368	0.437	0.181	0.623	0.106	0.008	0.239	0.031	0.041	0.142	0.114	0.209	0.053	0.017	0.026	0.091	0.065
PL	0.957	0.010	0.990	0.004	0.524	0.410	0.057	0.518	0.261	0.008	0.174	0.018	0.057	0.053	0.071	0.096	0.262	0.080	0.018	0.118	0.050
PT	0.928	0.008	0.950	0.007	0.631	0.029	0.028	0.359	0.197	0.003	0.383	0.016	0.031	0.017	0.025	0.075	0.158	0.059	0.032	0.145	0.048
RO	0.936	0.001	0.939	0.001	0.728	0.112	0.020	0.370	0.219	0.005	0.300	0.001	0.034	0.024	0.026	0.050	0.218	0.076	0.040	0.080	0.010
SE	0.942	0.024	0.855	0.058	0.409	0.369	0.201	0.731	0.058	0.002	0.189	0.006	0.087	0.033	0.057	0.152	0.016	0.009	0.021	0.035	0.095
SI	0.791	0.178	0.000	0.000	0.752	0.148	0.058	0.578	0.071	0.005	0.296	0.006	0.047	0.093	0.085	0.090	0.061	0.066	0.006	0.193	0.089
SK	0.932	0.023	0.980	0.010	0.451	0.482	0.039	0.846	0.006	0.004	0.111	0.010	0.075	0.110	0.107	0.161	0.034	0.096	0.052	0.203	0.048
UK	0.808	0.064	0.877	0.036	0.679	0.099	0.124	0.577	0.051	0.087	0.271	0.026	0.097	0.068	0.078	0.152	0.005	0.028	0.044	0.127	0.104

Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Omitted circumstance expressions listed in order of the circumstance categories are: "Non-EU"; "Not Europe"; "Dead/Unknown/Illiterate"; "Dead/Unknown/Retired/Other Inactive"; "Dead/Unknown/Not Working/Armed Forces"; "Dead/Unknown/Not Working/Non-Supervisory". Compare also to Table 2.1.

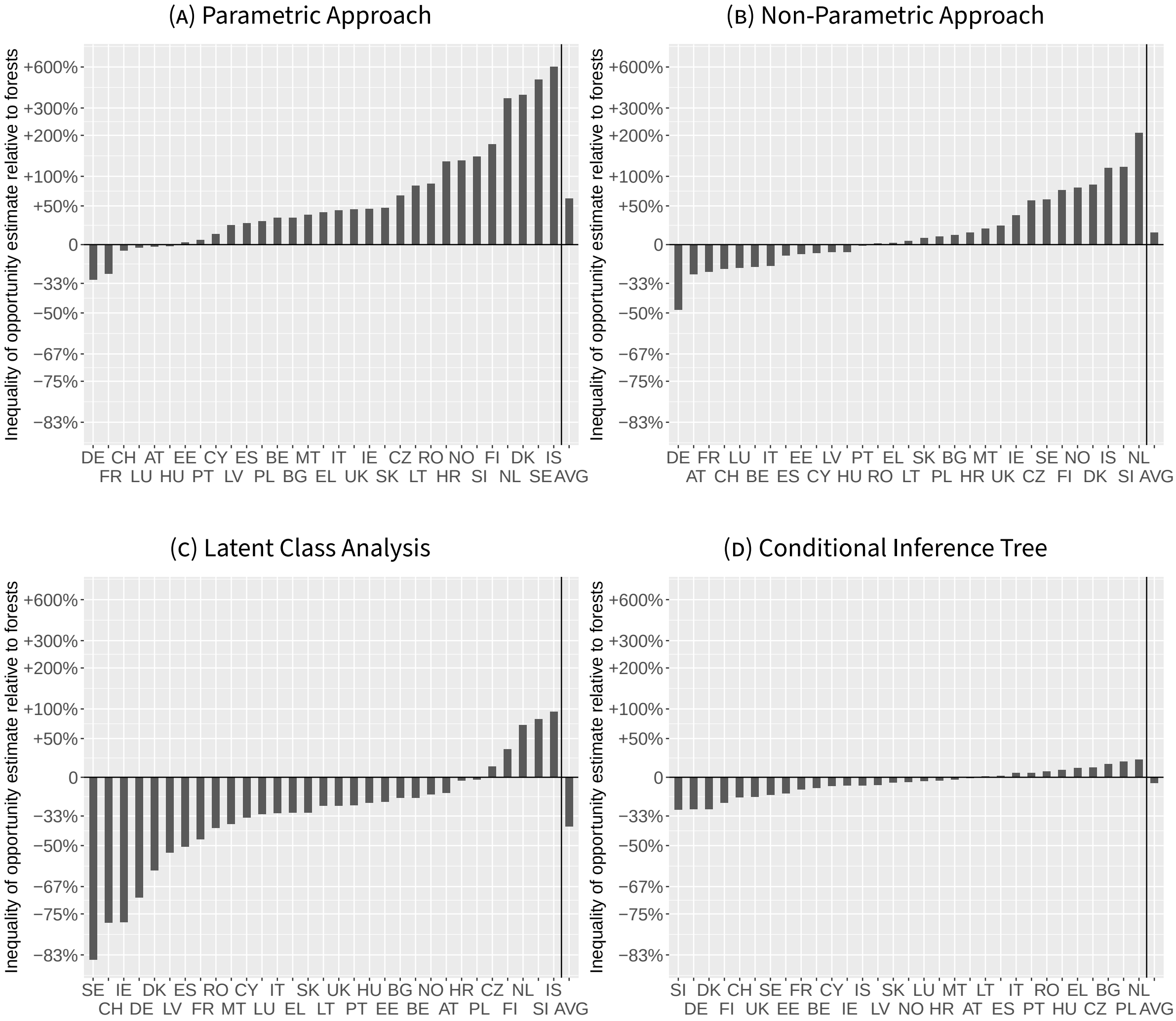
Appendix B.6 Alternative Inequality Indexes

Figure B.1 – Correlation of Estimates by Method (GE(0))



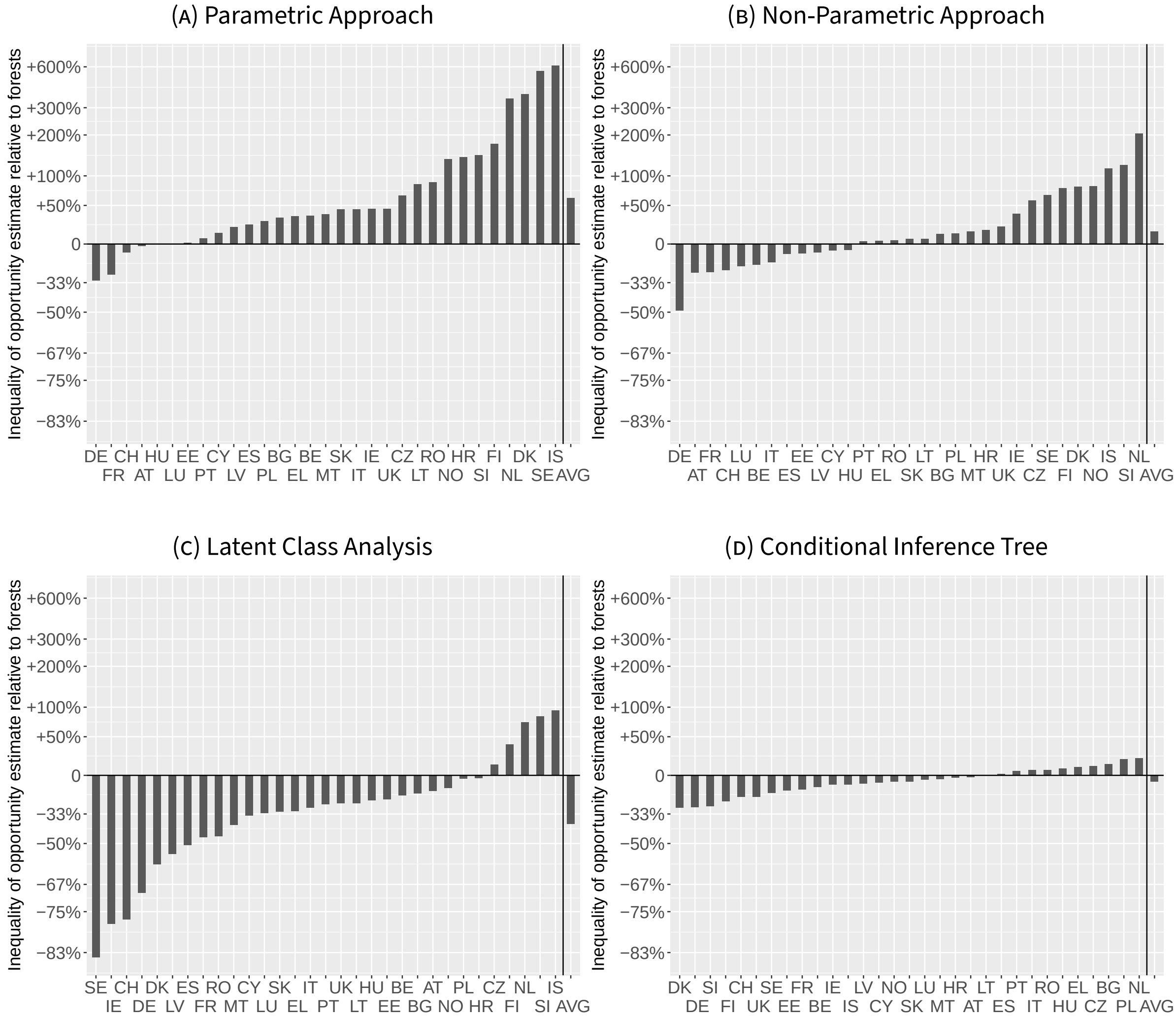
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: In each panel, the y-axis shows the inequality of opportunity estimate from the method in question divided by the inequality of opportunity estimate from forests, displayed on a logarithmic scale. Country-estimates above the black line indicate an overestimation of inequality of opportunity relative to the random forest benchmark. Reversely, country-estimates below the black line indicate an underestimation of inequality of opportunity relative to the random forest benchmark. For all methods inequality of opportunity is measured by the GE(0) index in the estimated counterfactual distribution $\hat{y}^C$. The figure is top (bottom) coded at $+600\%$ (-83%).

Figure B.2 – Correlation of Estimates by Method (GE(1))



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: In each panel, the y-axis shows the inequality of opportunity estimate from the method in question divided by the inequality of opportunity estimate from forests, displayed on a logarithmic scale. Country-estimates above the black line indicate an overestimation of inequality of opportunity relative to the random forest benchmark. Reversely, country-estimates below the black line indicate an underestimation of inequality of opportunity relative to the random forest benchmark. For all methods inequality of opportunity is measured by the GE(1) index in the estimated counterfactual distribution $\tilde{y}^C$. The figure is top (bottom) coded at $+600\%$ (-83%).

Figure B.3 – Correlation of Estimates by Method (GE(2))

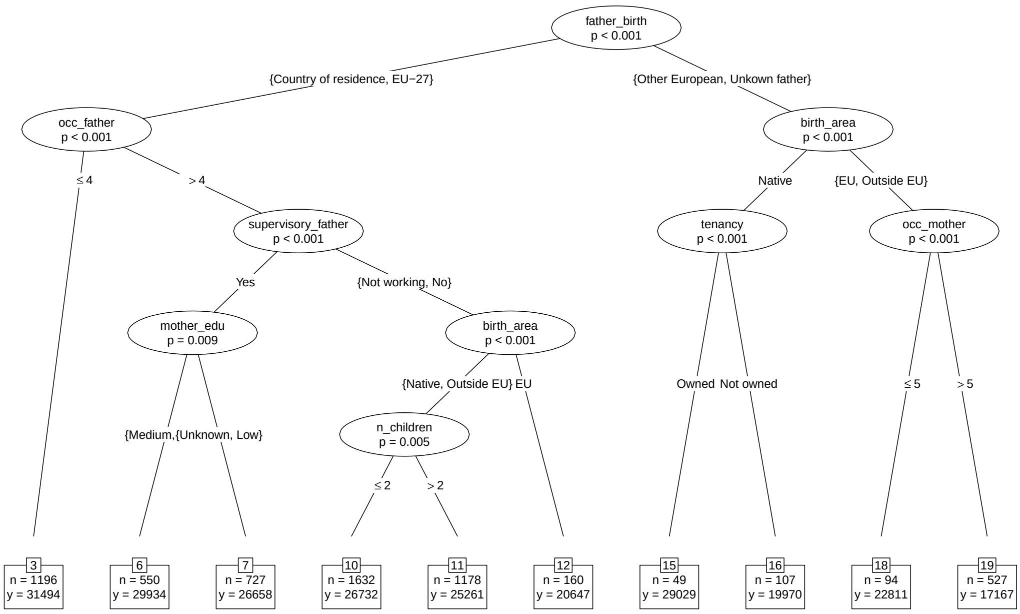


Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: In each panel, the y-axis shows the inequality of opportunity estimate from the method in question divided by the inequality of opportunity estimate from forests, displayed on a logarithmic scale. Country-estimates above the black line indicate an overestimation of inequality of opportunity relative to the random forest benchmark. Reversely, country-estimates below the black line indicate an underestimation of inequality of opportunity relative to the random forest benchmark. For all methods inequality of opportunity is measured by the GE(2) index in the estimated counterfactual distribution $\hat{y}^C$. The figure is top (bottom) coded at $+600\%$ (-83%).

Appendix B.7 Opportunity Structures

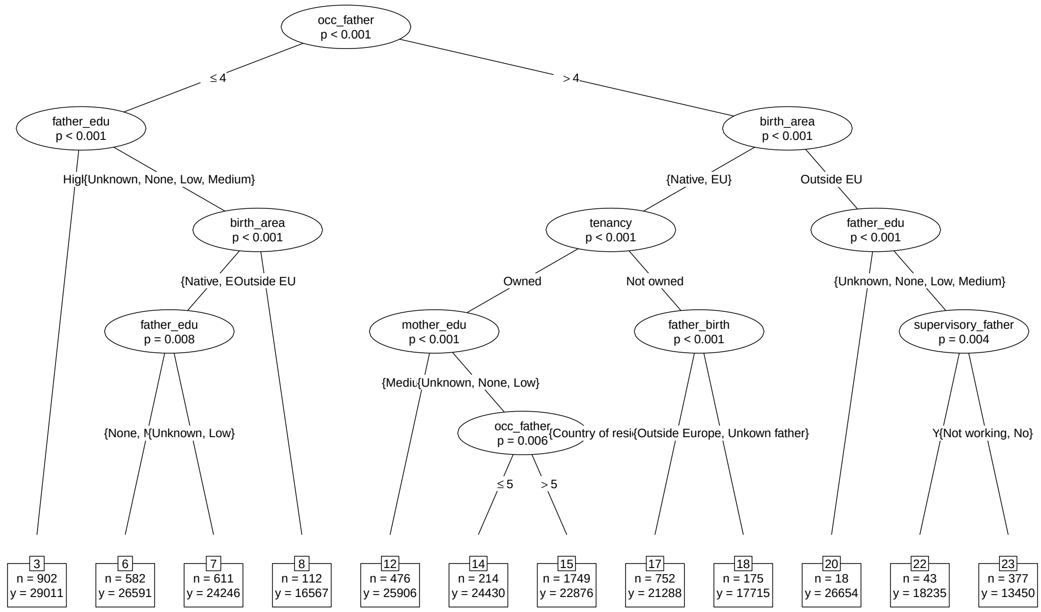
Trees

Figure B.4 – Opportunity Tree (Austria)



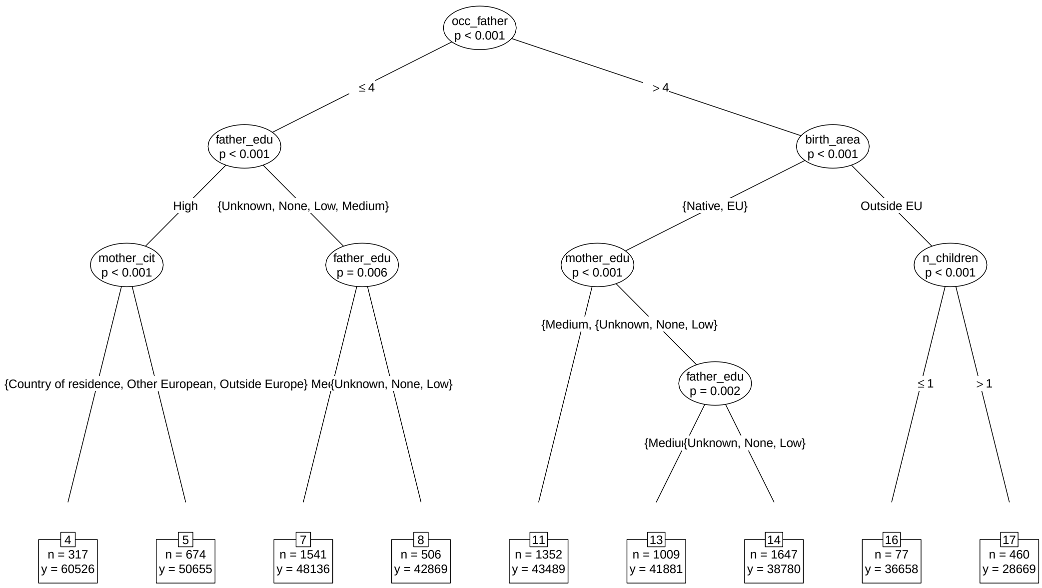
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Figure B.5 – Opportunity Tree (Belgium)



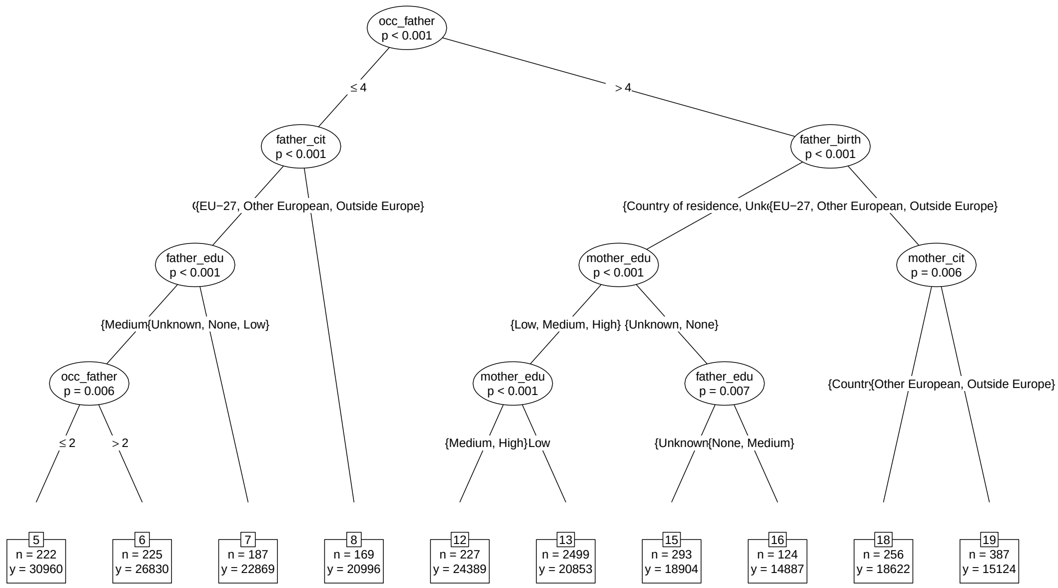
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.6 – Opportunity Tree (Switzerland)



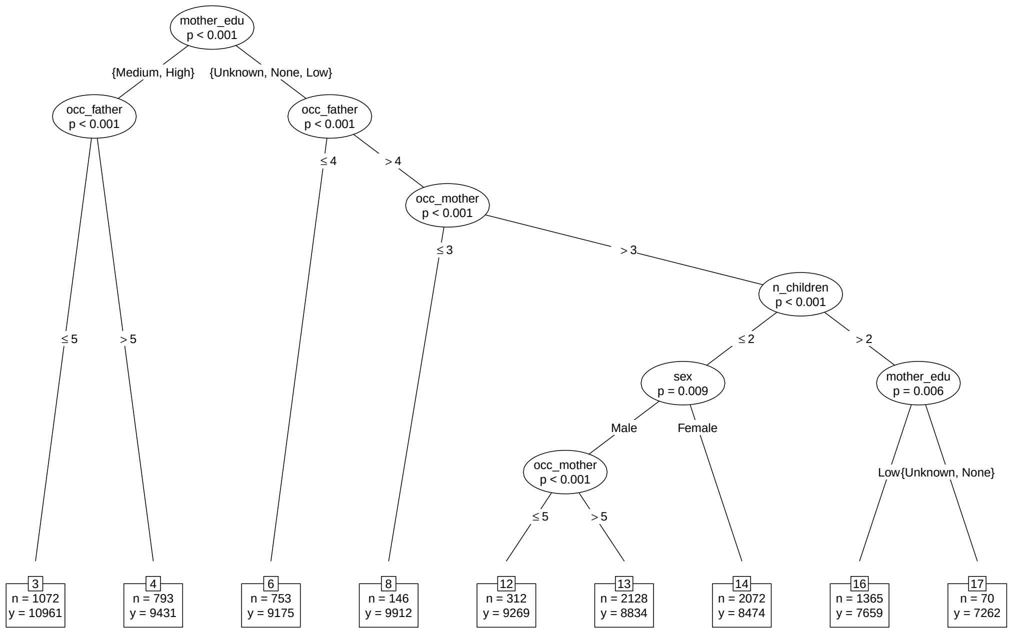
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\mathbf{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.7 – Opportunity Tree (Cyprus)



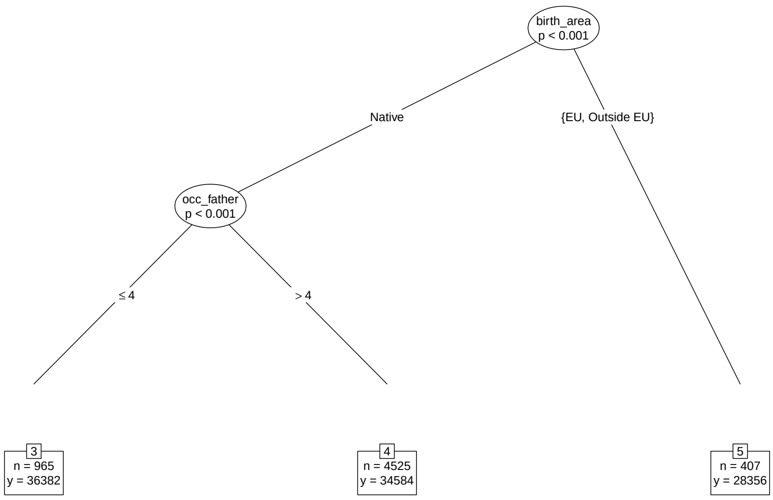
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.8 – Opportunity Tree (Czech Republic)



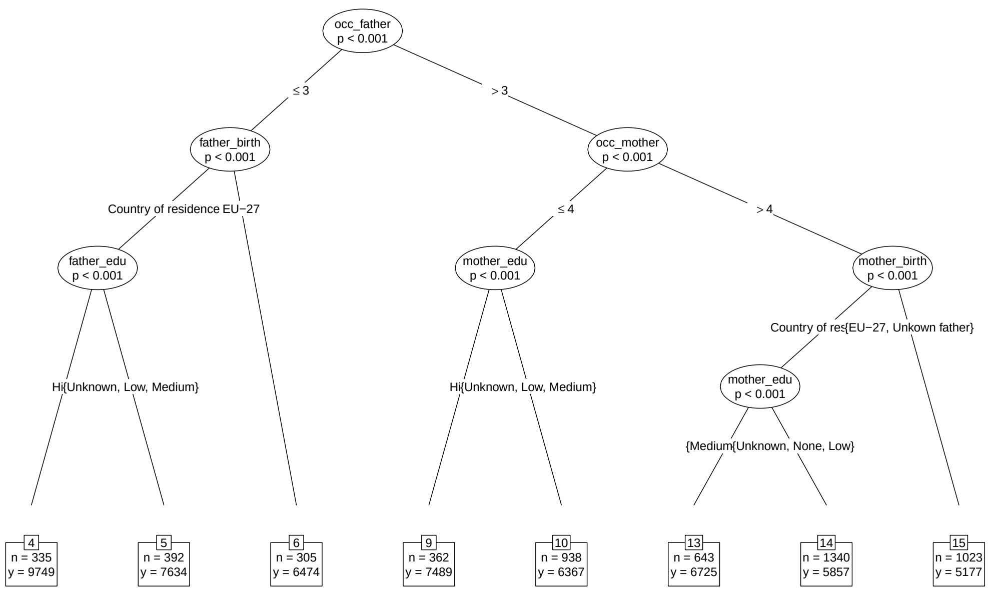
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Figure B.9 – Opportunity Tree (Denmark)



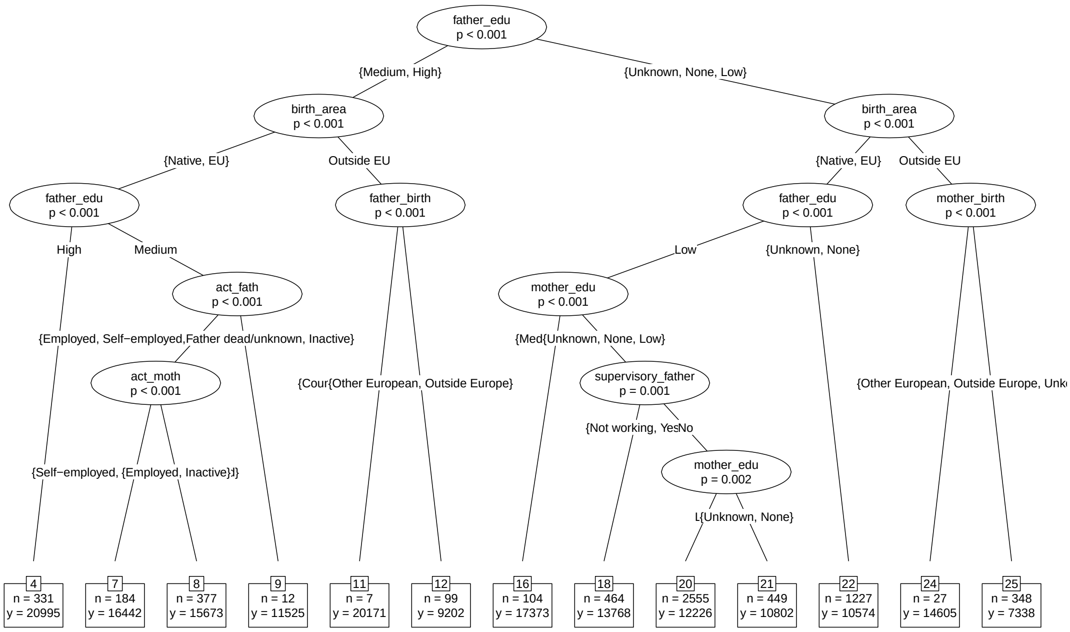
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Figure B.10 – Opportunity Tree (Estonia)



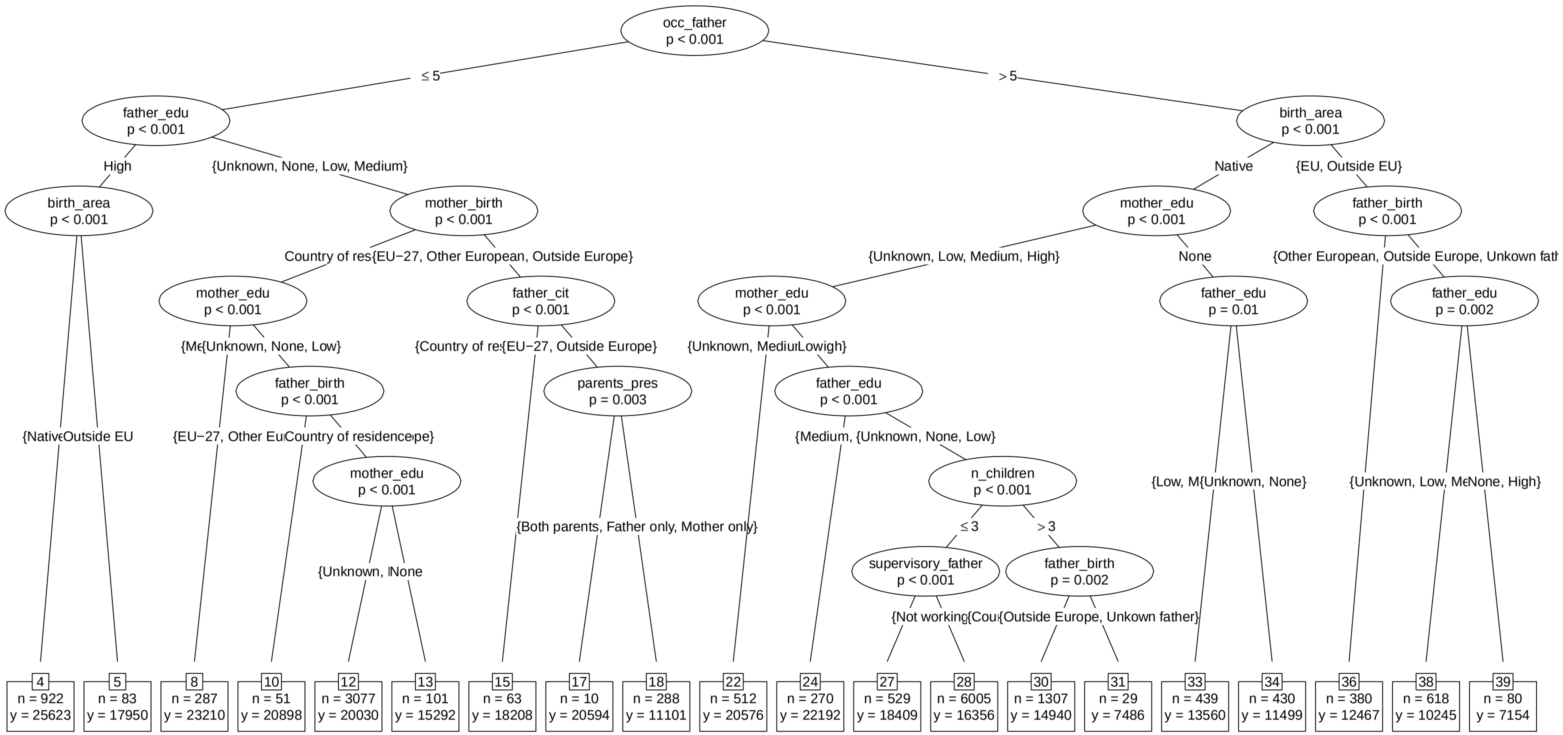
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Figure B.11 – Opportunity Tree (Greece)



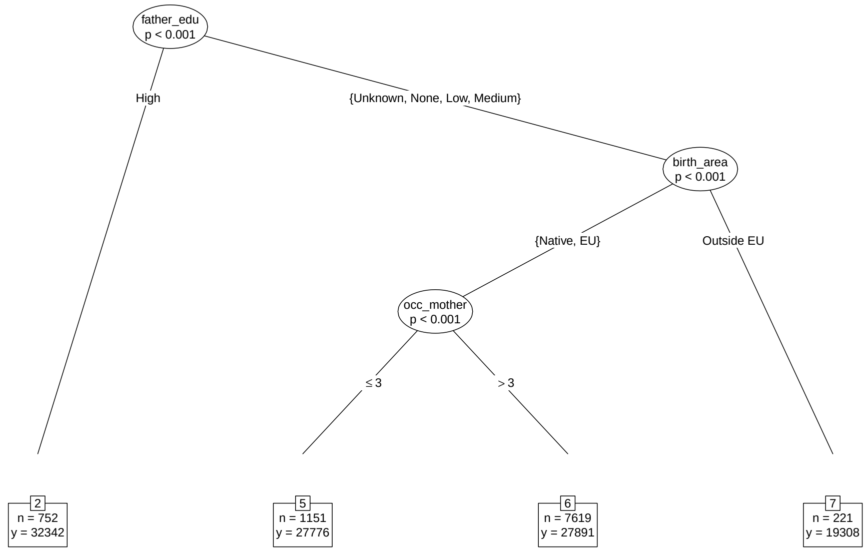
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.12 – Opportunity Tree (Spain)



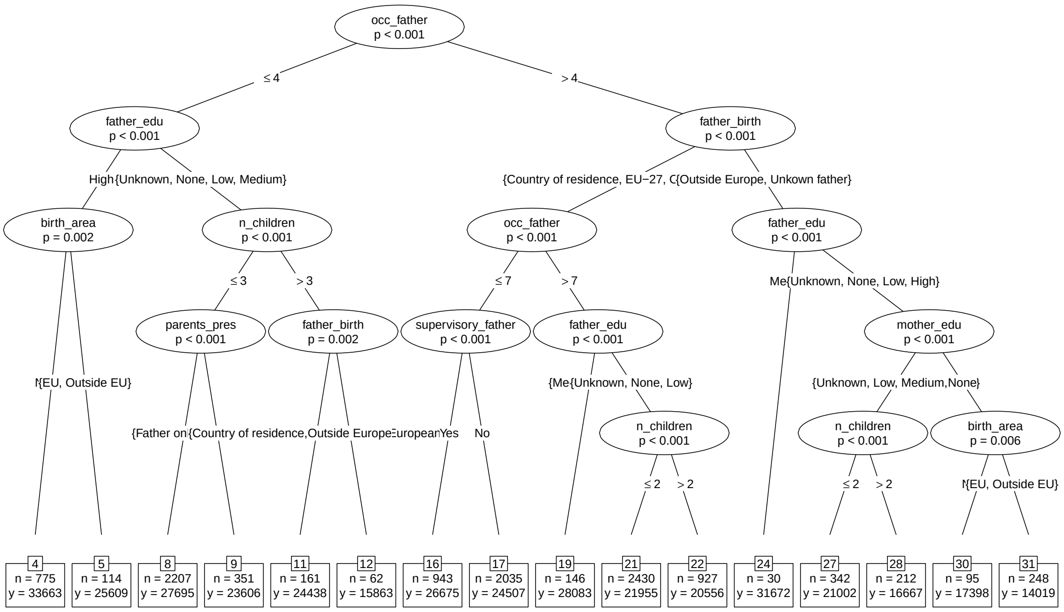
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.13 – Opportunity Tree (Finland)



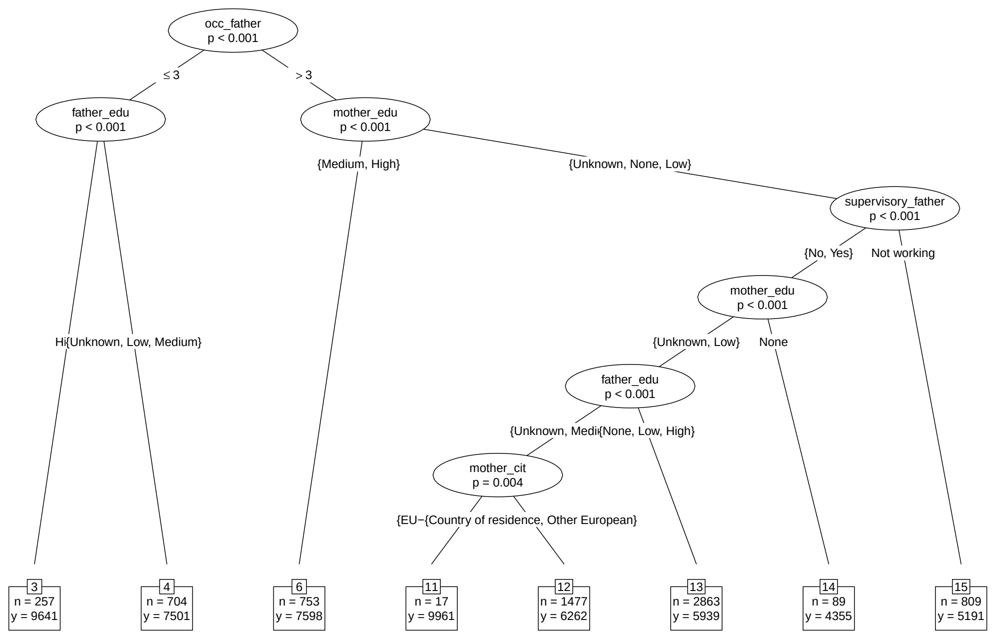
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.14 – Opportunity Tree (France)



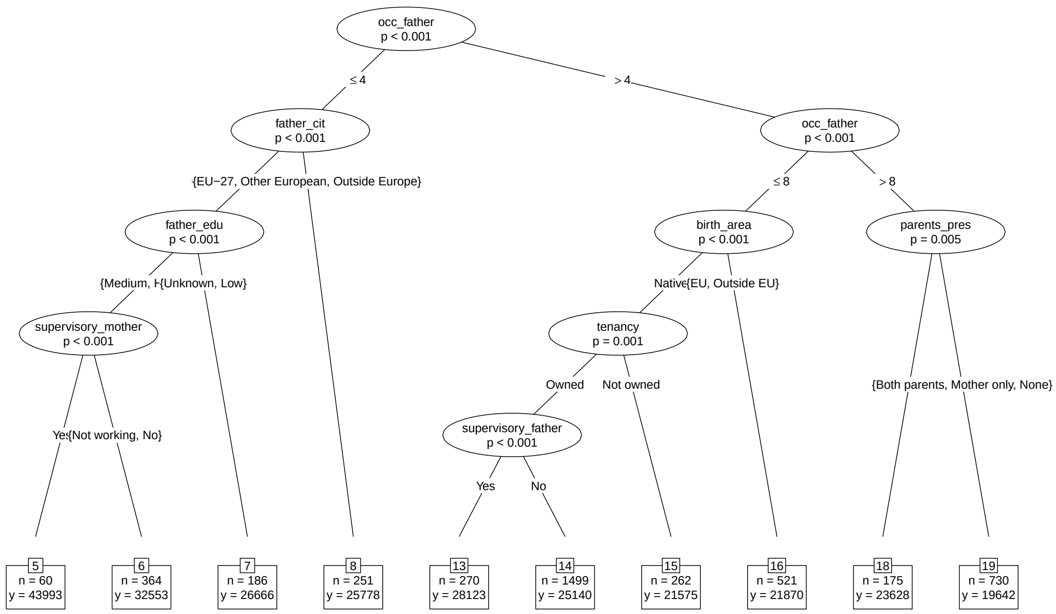
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.15 – Opportunity Tree (Croatia)



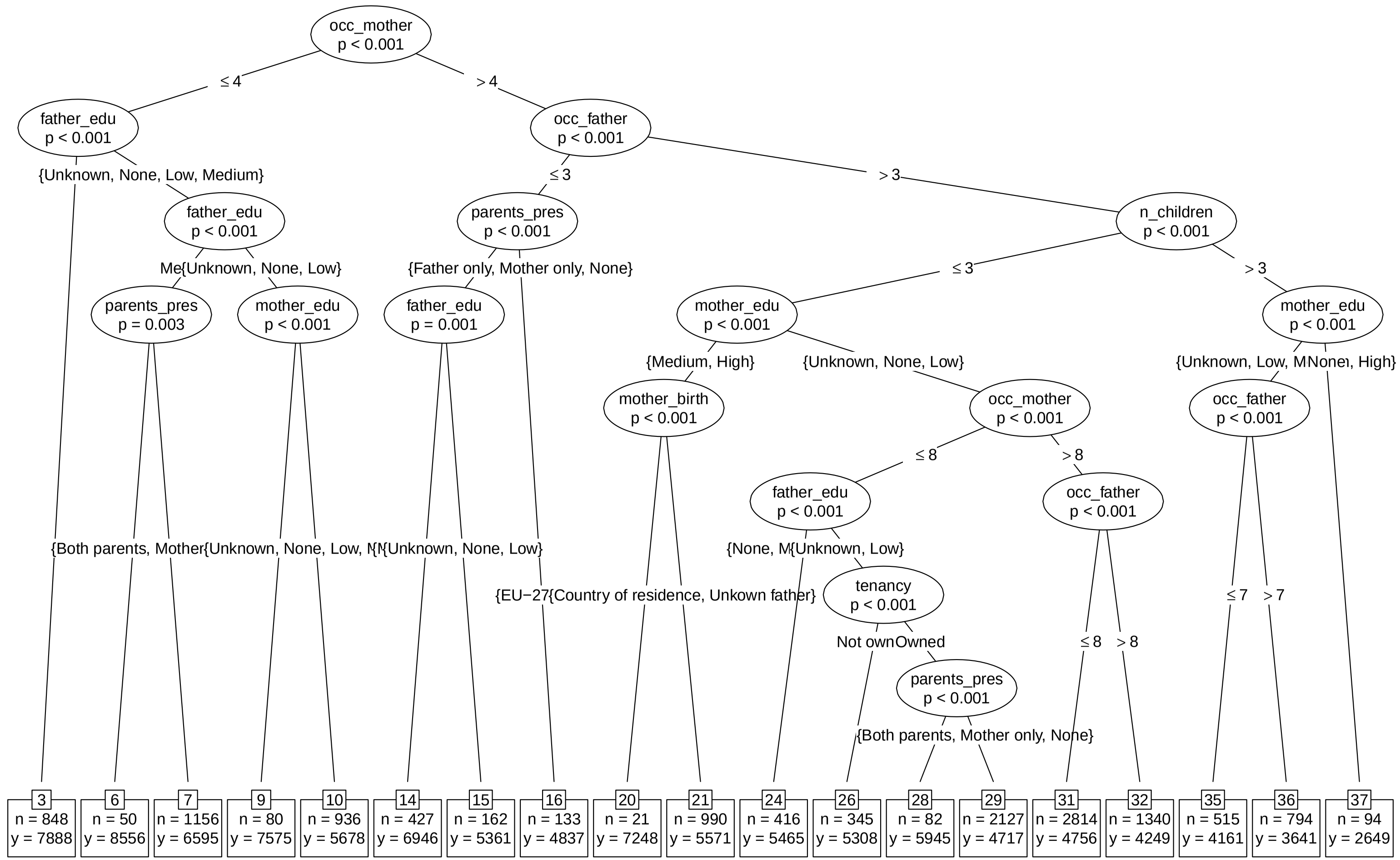
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\mathbf{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.16 – Opportunity Tree (Ireland)



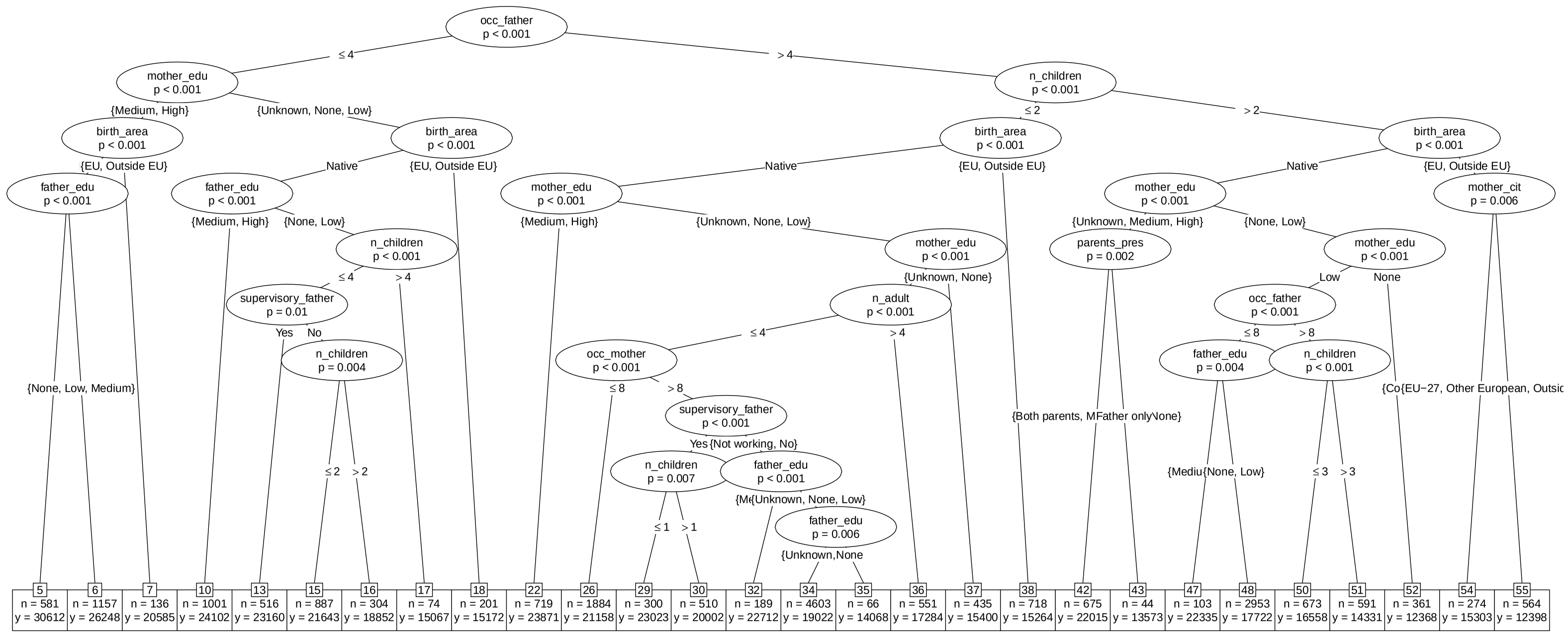
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Figure B.17 – Opportunity Tree (Hungary)



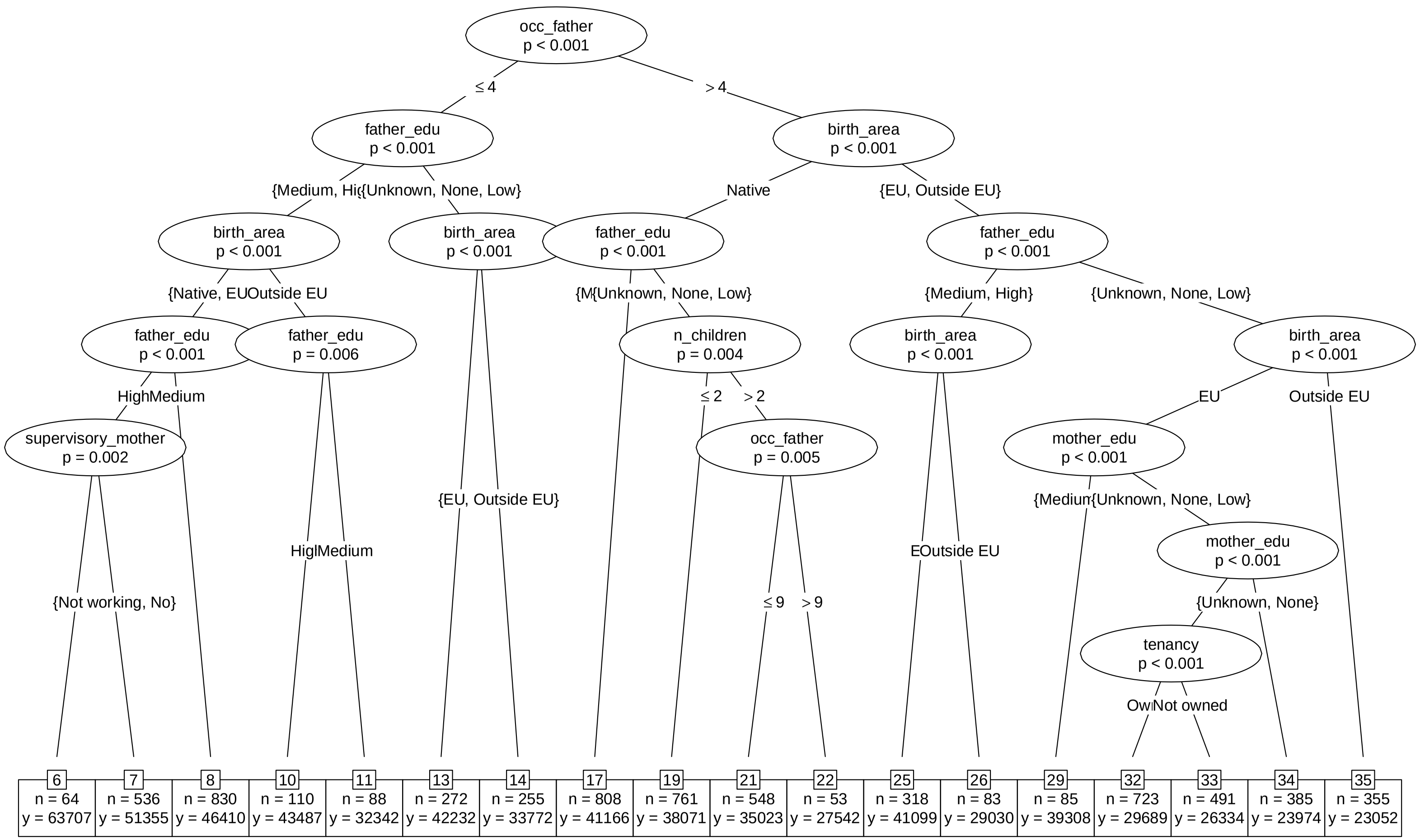
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.18 – Opportunity Tree (Italy)



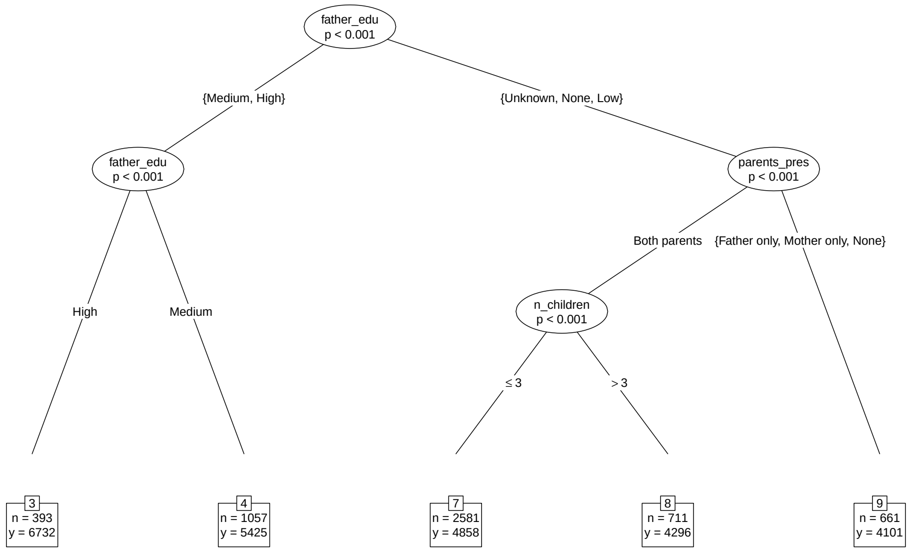
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.19 – Opportunity Tree (Luxembourg)



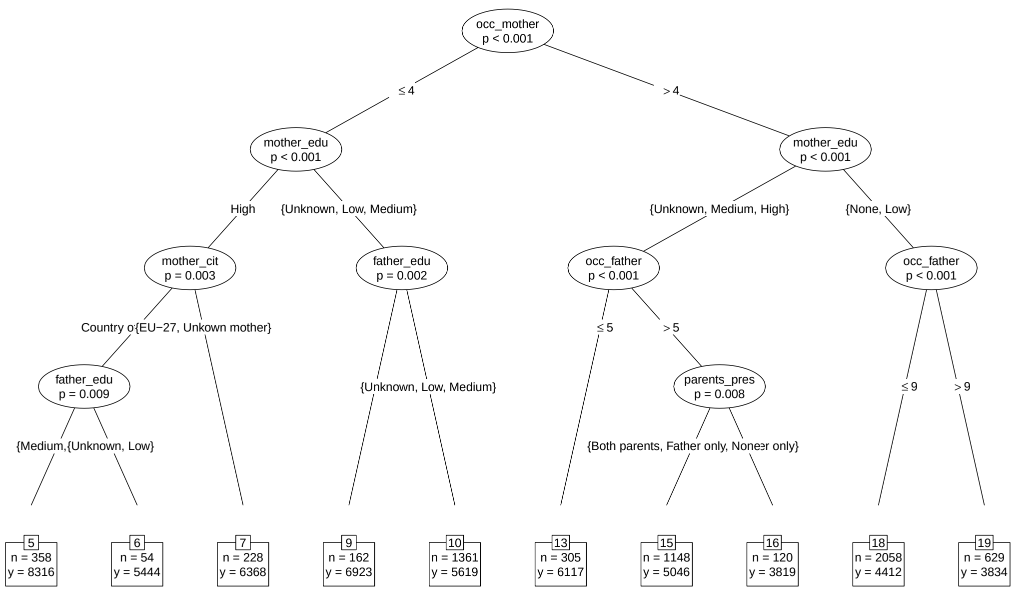
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Figure B.20 – Opportunity Tree (Lithuania)



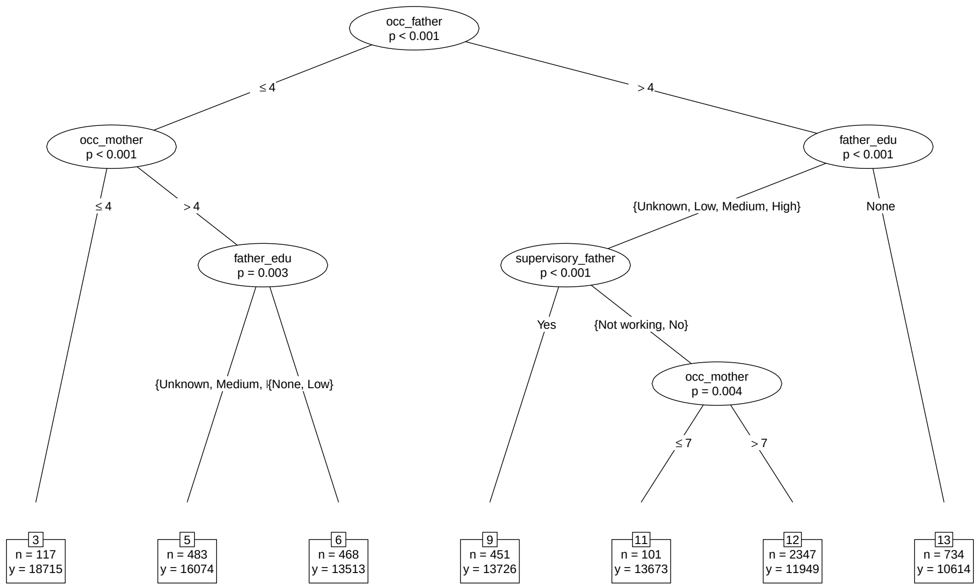
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\mathbf{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.21 – Opportunity Tree (Latvia)



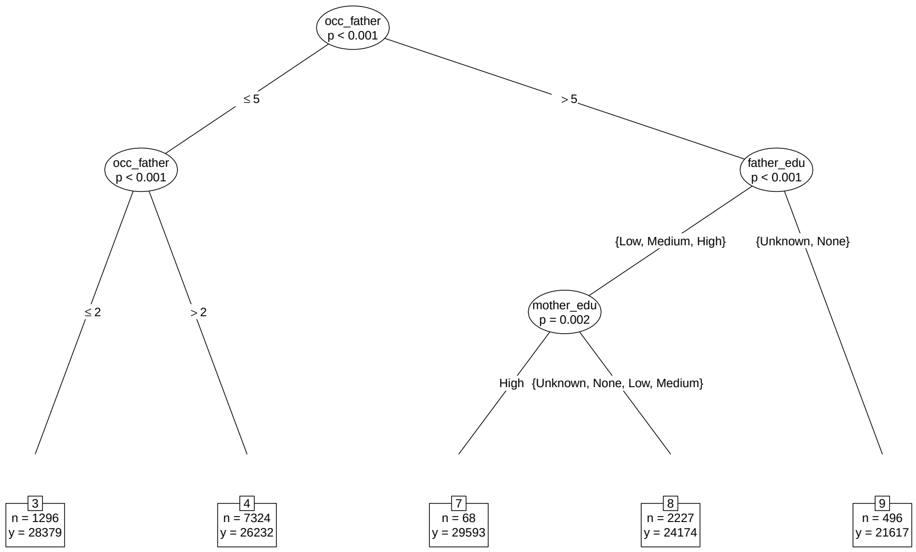
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.22 – Opportunity Tree (Malta)



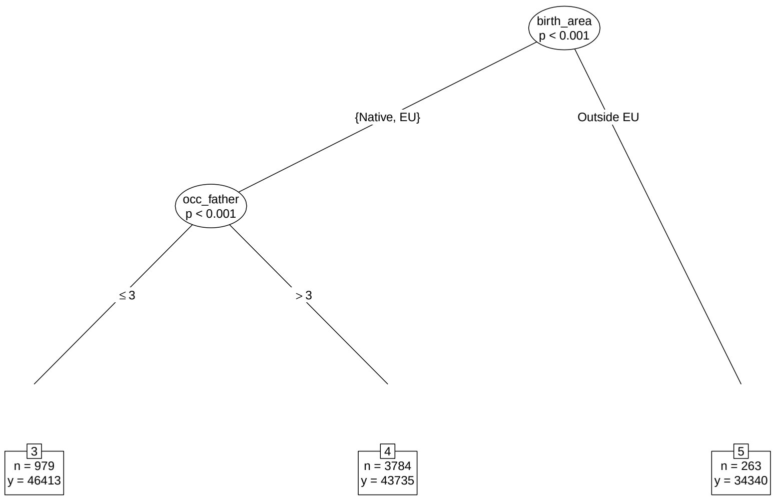
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.23 – Opportunity Tree (Netherlands)



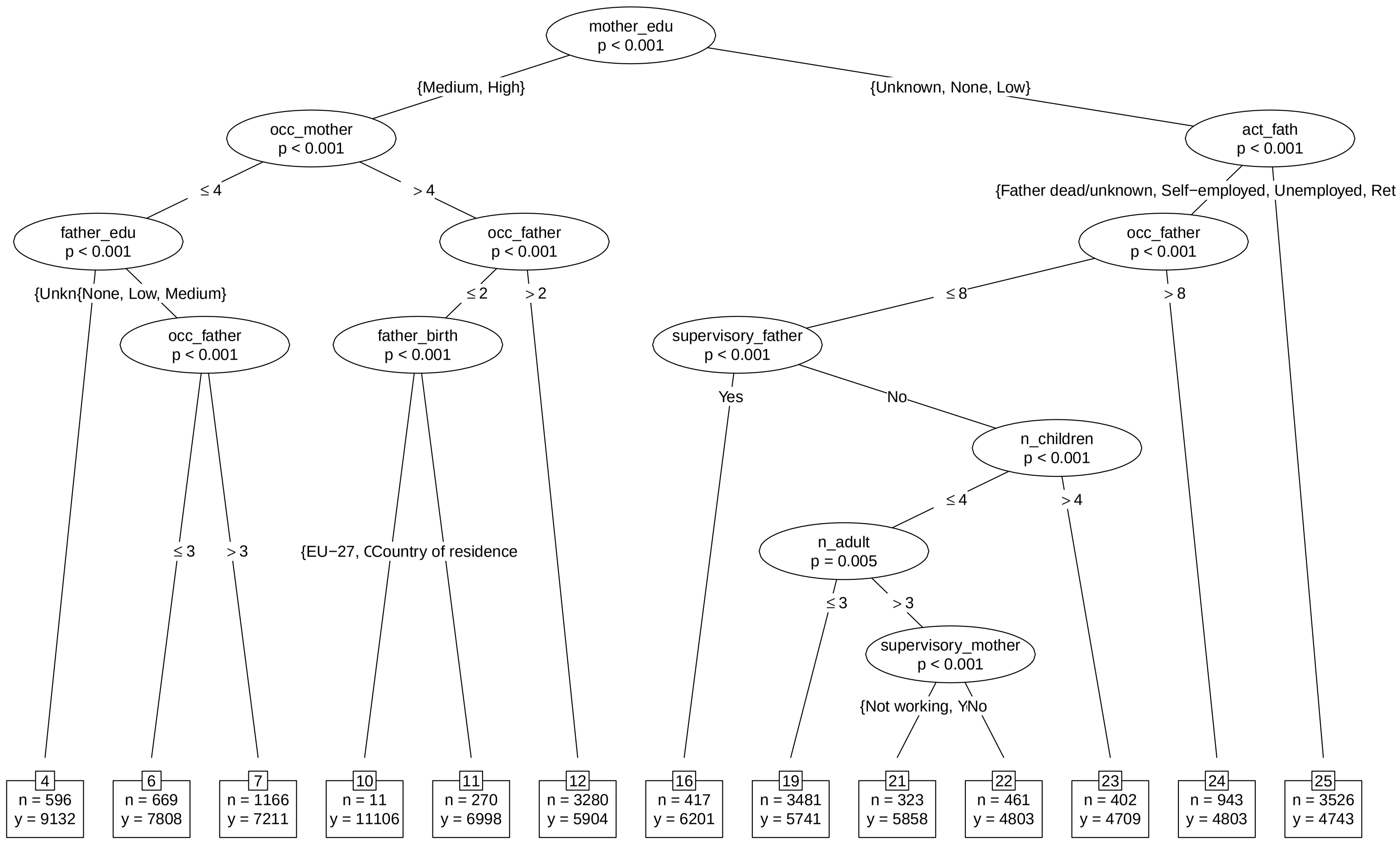
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Figure B.24 – Opportunity Tree (Norway)



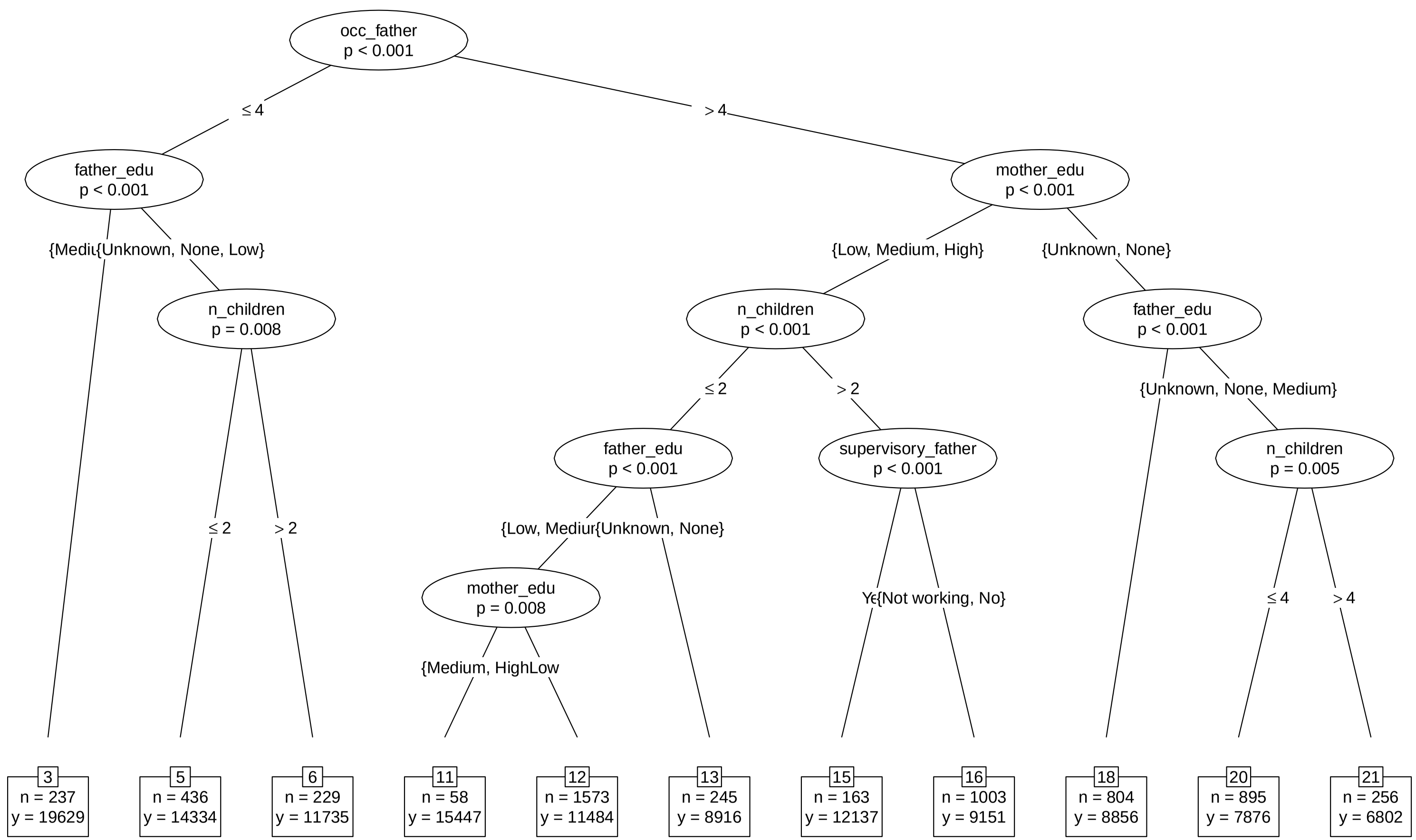
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.25 – Opportunity Tree (Poland)



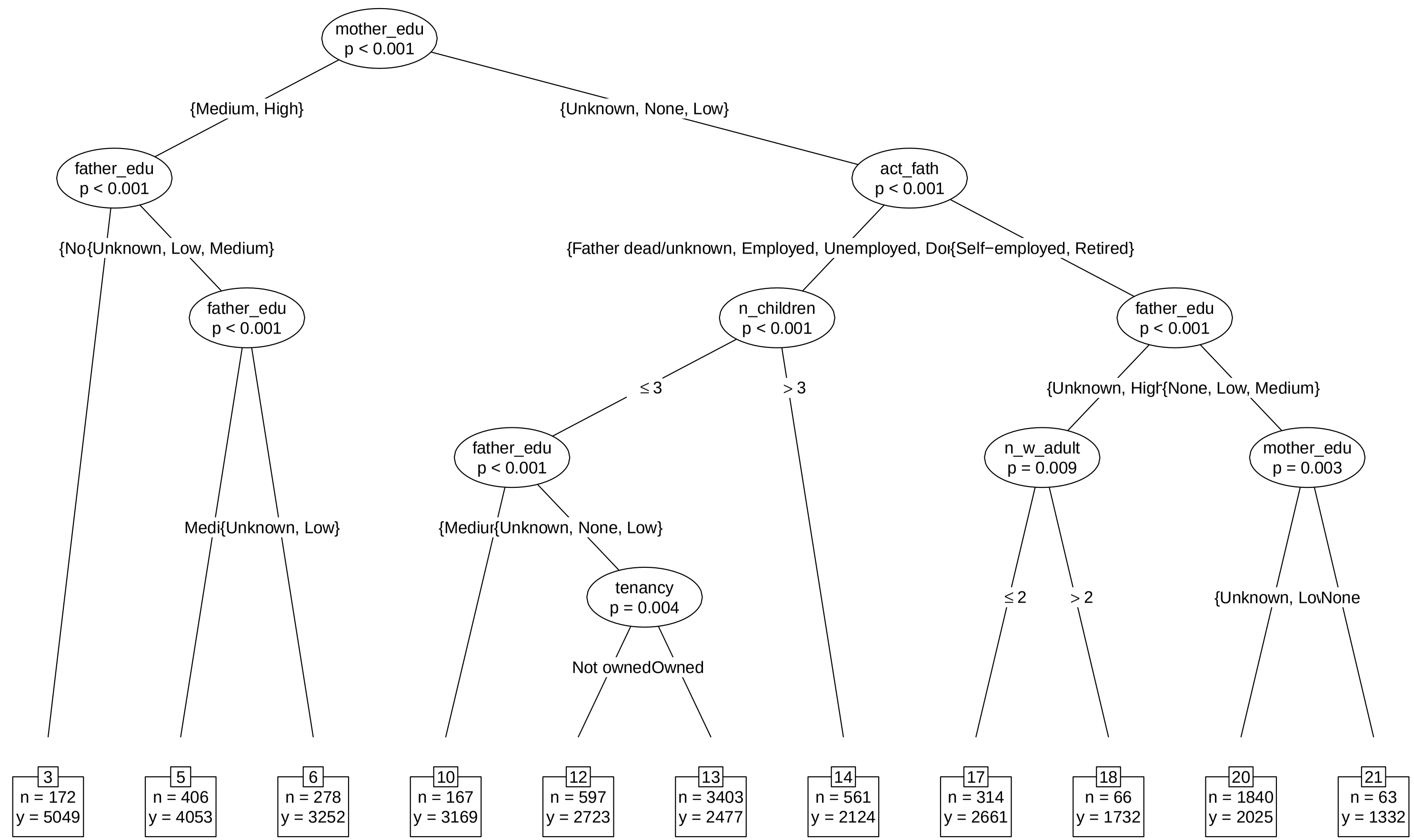
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Figure B.26 – Opportunity Tree (Portugal)



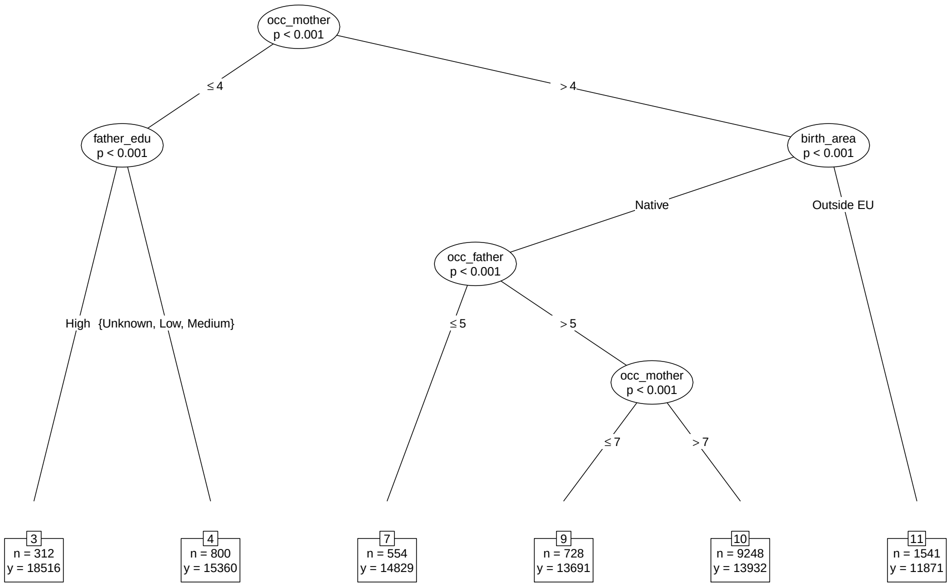
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.27 – Opportunity Tree (Romania)



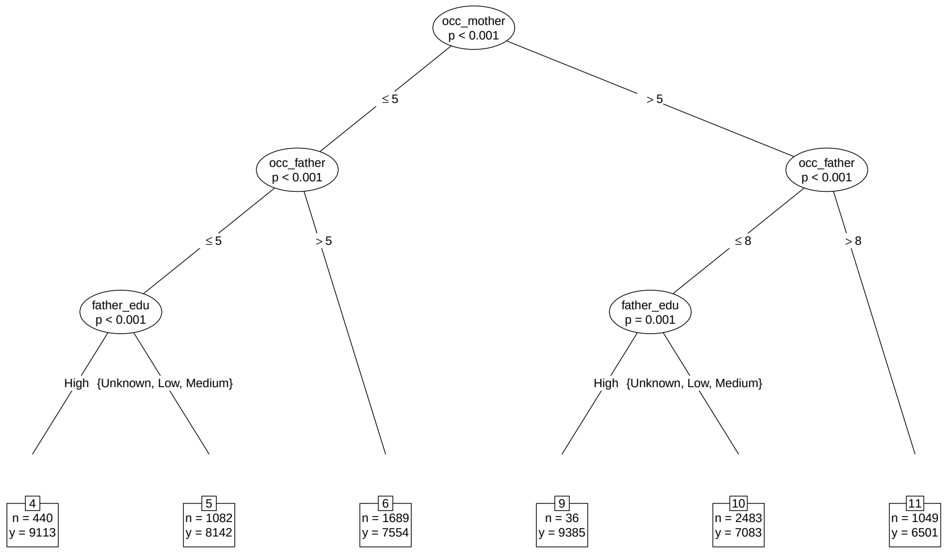
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.28 – Opportunity Tree (Slovenia)



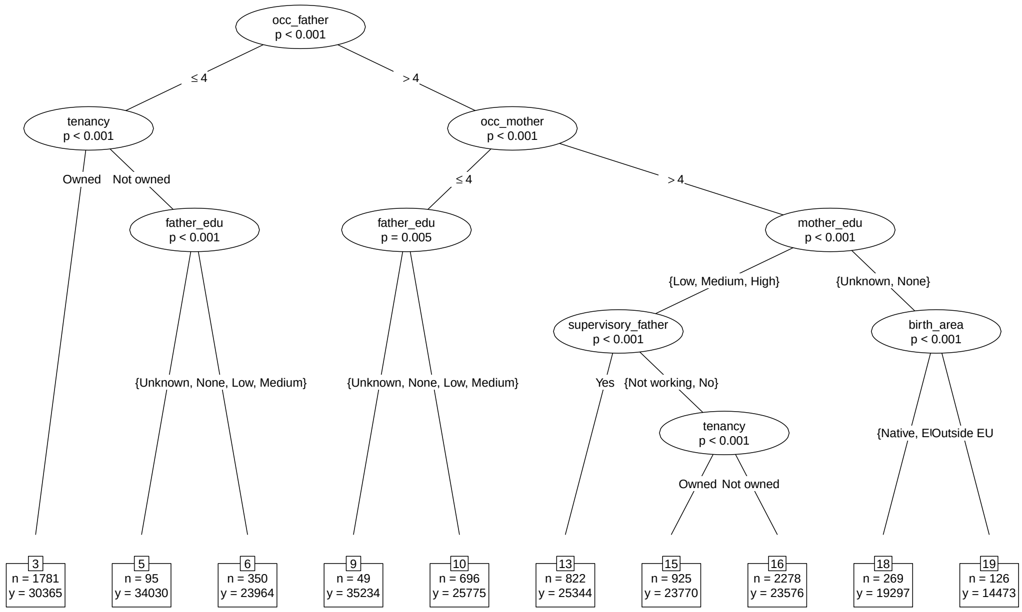
Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\mathbf{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

FIGURE B.29 – Opportunity Tree (Slovakia)



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances Ω used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

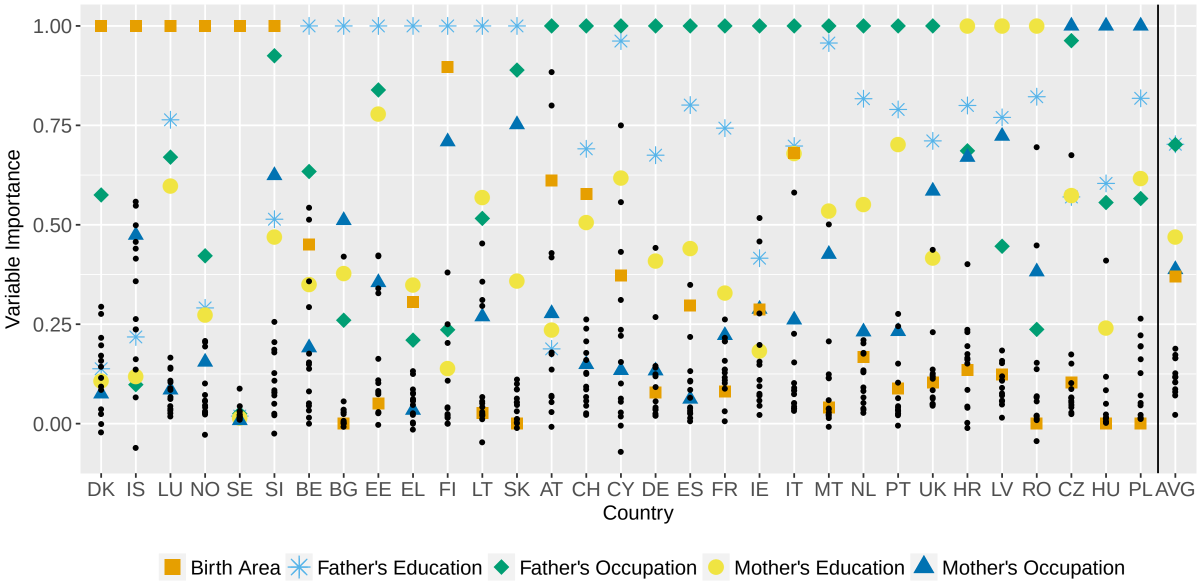
Figure B.30 – Opportunity Tree (United Kingdom)



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: The tree is constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\mathbf{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1. Ellipses indicate splitting points, while the rectangular boxes indicate terminal nodes. Within each ellipse we indicate the splitting variable as well as the p-value associated with the respective split. The first number inside the terminal nodes indicates the population share belonging to the circumstance type, while the second number shows the respective estimate of the conditional expectation y^C.

Forests

Figure B.31 – Variable Importance Plot



Data: EU-SILC 2011 cross-sectional (rev.5, June 2015).
Note: Each dot shows the importance of a particular circumstance variable ω^p. Variable importance is measured by the decrease in MSE^{OOB} after permuting ω^p such that it is orthogonal to y. The importance measure is standardized such that the circumstance with the greatest importance in each country equals 1. The forests are constructed by the conditional inference algorithm (Section 2.3). The set of observed circumstances $\boldsymbol{\Omega}$ used to construct the conditional inference tree is detailed in Table 2.1.

3 The Parental Wage Gap and the Development of Socio-emotional Skills in Children

3.1 Introduction

How does the expansion of labor market opportunities for women relative to men affect the socio-emotional development of their children? Throughout the post-World War II period, the convergence of wages and labor market participation rates of men and women has been a shared element of labor markets in many industrialized societies (Blau and Kahn, 2017; Olivetti and Petrongolo, 2016). In response to changing economic incentives, heterosexual couples with children have adjusted their time-use and spending patterns, henceforth leading to marked changes in the way they invest into the skill formation of their children (Aguiar and Hurst, 2007; Kornrich and Furstenberg, 2013). While these long-run trends are well-documented, there is currently no study that causally links the convergence of labor market opportunities across gender groups in the parental generation to the skill formation of children in the following generation.

In this paper, I study how changes in the parental wage gap influence the process of children's skill formation by focusing on socio-emotional skills as measured by the Big Five personality inventory.

My research design relies on two main features. First, I use the 2005–2017 waves of the German Socio-economic Panel (GSOEP) to construct a sample of $6,070$ siblings aged 2-17 for whom I observe measures of the Big Five inventory at the same age but in different calendar years. This sample allows me to implement a within-family sibling design (e.g. Løken et al., 2012) in which I rule out confounding effects through time constant factors that are specific to families when their children are of a particular age. For example, think of two families that have different preferences for the mother to stay home while their children are under school age. If the Big Five personality traits are affected by different care arrangements in this age period, a comparison across families would confound the effect of the parental wage gap on child development with family differences in childcare preferences. However, a focus on within-family variation rules out such confounding effects.

Second, comparisons across siblings at the same age may still reflect parental labor supply responses that are endogenous to the skill development of their children. For example, think of a mother of two that responds to the behavioral problems of one of her children by switching to a lower paying but less time consuming job. If such an adjustment has a spillover effect on the skill development of her second child, the effect of intra-family changes in the parental wage gap on child development would be confounded by reversed causality. To circumvent such concerns, I draw on a shift-share design to construct potential wages that reflect variation in the sex- and education-specific labor demand across commuting zones in Germany (Goldsmith-Pinkham et al., 2020). The general idea of shift-share designs is to predict group-specific wages based on sectoral shocks ("shift") and the historic employment shares of sectors in the respective group ("share"). As a consequence of replacing actual wages with potential wages, within-family changes in the parental wage gap reflect temporal variation in the labor market incentives for mothers and fathers that is plausibly exogenous to within-family decision-making.

This study makes two contributions in relation to the existing literature. First, the production of child skills can be conceived as a function of monetary investments (Akee et al., 2018; Dahl and Lochner, 2012; Løken et al., 2012; Milligan and Stabile, 2011) and time investments by the parents (Del Boca et al., 2017; Del Bono et al., 2016; Fiorini and Keane, 2014; Hsin and Felfe, 2014).[1] The existing literature studies the provision of these resources by focusing on mothers as the primary caretaker and by-and-large neglects the dynamics of family decision-making within the context of two-parent households.[2] However, the investigation of these dynamics is important. Even in an age of declining marriage and increasing divorce rates, 73% (65%) of all German (US-American) children live in a household with their two married parents (Federal Statistical Office, 2020; Livingston, 2018). Furthermore, the well-documented changes in relative labor market incentives for men and women suggest strong shifts in how these households allocate monetary and time resources across various activities that potentially affect the skill development of their children. In this paper I close this gap by studying how

[1] This stylized representation of the skill formation process focuses on the family context. It is incomplete as it omits other important input factors that are not directly linked to intra-family decision-making, including the quality of schools (Chetty et al., 2014a; Jackson, 2019), neighborhoods (Agostinelli et al., 2020; Chetty et al., 2016a) and individual natural endowments (Black et al., 2020; Papageorge and Thom, 2020). See Almond et al. (2018) and Heckman and Mosso (2014) for recent overviews.

[2] In particular the trade-off between the provision of monetary and time resources by mothers has garnered increased interest in the recent literature on child development (Agostinelli and Sorrenti, 2018; Nicoletti et al., 2020).

changes in the relative wages of parents influence family decisions with respect to labor market participation and childcare arrangements, and the extent to which these choices have an influence on the skill development of their children. Closest to this ambition are the papers of Del Boca et al. (2014) and Bruins (2017). Del Boca et al. (2014) provide a structural model of child development in which both mothers and fathers provide time investments, the benefits of which are balanced against the financial resources generated through increased labor market participation. In contrast to their paper, I focus on the development of socio-emotional skills of children in Germany instead of the development of cognitive skills in the US. Bruins (2017) uses a shift-share design to investigate the impact of gender convergence on parental time investments. In comparison to her paper, I tighten the identification approach by combining the shift-share design with a within-family sibling comparison. Furthermore, while having more detailed information on parental time-use, her data sources do not avail measures of child development. Hence, in comparison to Bruins (2017) I provide direct evidence on how changes in the relative wages of parents affect the process of skill formation in their children.

Second, next to cognitive skills and health, socio-emotional skills are a dimension of human capital that matters for a variety of important life outcomes.[3] In view of this importance, social scientists have dedicated increased attention to the causal factors that underlie the formation of these skills. In the context of families, these factors include the home environment (Carneiro et al., 2013), monetary resources (Akee et al., 2018), parental time investments (Agostinelli and Sorrenti, 2018) and parenting styles (Deckers et al., 2020). In this paper, I contribute to this literature by investigating how changes in the relative labor market incentives for mothers and fathers influence the socio-emotional development of children as measured by the Big Five inventory (Widiger, 2018).

Guided by a stylized model of collective household decision-making, my empirical analysis proceeds in three steps. First, I analyze the labor market adjustments of households in response to changes in the relative wages of mothers and fathers. In this step, I pay particular attention to changes in hours worked as well as the consequential labor market earnings of mothers, fathers and the overall availability of financial resources at the household level. Sec-

[3] The exact definition of socio-emotional skills is contested (Humphries and Kosse, 2017). They are oftentimes interpreted as a residual dimension of skills not captured by test scores and may include various economic preferences as well as personality traits. In this work I draw on the Big Five personality taxonomy to measure socio-emotional skills. Among others, recent work analyzes the impact of the Big Five personality traits on schooling decisions (Almås et al., 2016), job search behavior (Flinn et al., 2020), matching in marriage markets (Dupuy and Galichon, 2014), task productivity (Cubel et al., 2016), and longevity (Savelyev, 2020).

ond, I analyze how households reorganize the provision of childcare in response to changes in the relative wages of mothers and fathers. In this step, I pay particular attention to hours of care provision by mothers and fathers, and changes in total parental care provision as opposed to the use of extra-parental care providers. Third, I analyze the effect of changes in the relative wages of mothers and fathers on the development of the Big Five personality traits of their children. This last step establishes a reduced-form causal effect of changes in the parental wage gap on the formation of socio-emotional skills in children. The previous steps allow me to interpret these results in light of the mechanisms that are emphasized in the literature on collective household decision-making (R. Blundell et al., 2005; Browning et al., 2014; Cherchye et al., 2012; Knowles, 2012).

My findings can be summarized as follows. First, both fathers and mothers are characterized by a positive own-wage elasticity of labor supply: They both increase their labor hours in response to increasing potential wages. However, mothers and fathers tend to react differently to changes in the potential wages of their partners. While mothers tend to decrease their labor supply in response to positive wage shocks of their partners, the labor supply of fathers is insensitive to changes in the potential wages of their partners. As a consequence, the effect of closing parental wage gaps on the financial positions of households depends on whether the convergence is driven by wage gains of mothers or wage losses of fathers. If the former, gender convergence in wages leads to an expansion of total household resources since the labor supply of fathers does not adjust to the gains of mothers. If the latter, there is no effect on total household resources since women tend to substitute for the losses of fathers. In both cases, however, closing parental wage gaps lead to an increase in the relative share of financial resources controlled by mothers.

Second, the gendered asymmetry in cross-wage elasticities is also reflected in the way households adjust their childcare arrangements in response to changes in the relative wages of mothers and fathers. Wage gains of fathers lead to an increase in the hours of care provided by mothers and a decrease in the probability that the child is subject to extra-parental care provision. This response is consistent with the abovementioned finding that mothers decrease their labor supply in response to positive wage shocks of their partners. To the contrary, neither maternal nor paternal hours of care provision react to the relative wage gains of mothers. Mothers maintain the time they devote to their children in spite of their increasing engagement in the labor market. Descriptive analyses of German time-use diaries suggest that the constancy of maternal care provision results from shifting the timing of maternal time investments into the afternoon hours after they return from work. However, they substitute for their absence during the day by increasingly relying on informal childcare arrangements.

Third, in spite of the previously described changes in the financial positions and time-use of households, changes in the intra-household gender wage gap do not have an effect on the socio-emotional development of children. I can exclude at the 95% level of statistical significance that a €1 decrease in the intra-household hourly wage gap leads to shifts larger than 0.254 standard deviations in any of the Big Five dimensions. To put these numbers into perspective I compare them to existing evidence on the effects of various interventions on the Big Five inventory. For example, Akee et al. (2018) find that an unconditional cash transfer program worth $\$3,500$ per annum, decreased neuroticism in children of the Eastern Band of Cherokee Indians by 0.381 standard deviations. When it comes to a €1 decrease in the parental hourly wage gap in Germany I can rule out effects that are less than half of this size.[4]

These findings have important implications for economic policy-making. On the one hand, increasing gender equality has become a prominent goal for public policy in recent years.[5] On the other hand, one may oppose such policies as the increasing labor market participation of mothers could potentially exert adverse effects on the skill development of their children. The evidence presented in this work is not consistent with such claims.

To be sure, my identification strategy does not allow me to causally separate the impacts of the different channels of parental adjustments on child development. Instead I provide causal estimates for an omnibus treatment that shifts the time-use and financial positions of both mothers and fathers simultaneously. Furthermore, I analyze the average effects of these adjustments across children aged 2–17.[6] Therefore, my findings do not contradict existing work showing alternative care arrangements to be imperfect substitutes for the quality of care provided by mothers (e.g. Baker et al., 2019a). Nor do my findings contradict existing work that demonstrates the existence of sensitive age periods in which decreases in the time investments of mothers could have detrimental consequences for child development

[4] I furthermore show that a €1 decrease in the intra-family gap of hourly wages wage leads to a €$2,922$ increase of annual family earnings. The two interventions are thus broadly comparable in terms of their effects on family resources.

[5] In Germany, recent policy initiatives with the explicit goal to foster the economic convergence of men and women include the introduction of a 30% quota on supervisory boards of publicly traded companies in 2016 and the Pay Transparency Act from 2017. Similar policy initiatives exist in other industrialized countries as well, see for example Baker et al. (2019b), Bennedsen et al. (2020), Bertrand et al. (2018), and Gregory-Smith et al. (2014).

[6] I do provide heterogeneity analyses with respect to child age in section 3.5.5.

(Carneiro et al., 2015; Danzer and Lavy, 2018; Del Boca et al., 2017; Nicoletti et al., 2020). However, my work shows that across the life-cycle of German children, potentially existing quality gaps between the time investments provided by mothers and the time investments provided by other actors in the process of child development are small enough to be offset by the increase of total household resources and the relative increase of monetary resources controlled by mothers.

The remainder of this paper is structured as follows. In section 3.2, I present a stylized model of non-unitary household decision-making to guide the empirical analysis. Section 3.3 introduces the main data sources and details the construction of the relevant samples and variables. After outlining my identification strategy in section 3.4, I present the results of my analysis in section 3.5. Section 3.6 concludes.

3.2 Theoretical Background

Assume mothers and fathers indexed by $g \in \{m, p\}$ derive utility from consumption c_g and the development of their child C. Maternal utility is specified as follows:[7]

$$\begin{aligned} U_m(c_m, C) = \underbrace{w_m h_m - I_m + \delta_m(w_p h_p - I_p)}_{=c_m} \\ + \beta_m \underbrace{\ln[\alpha_m(1-h_m) + \alpha_p(1-h_p) + \gamma(I_m + I_p)]}_{=C} \end{aligned} \tag{82}$$

The consumption value c_m depends on private consumption – defined as the difference between individual earnings ($w_m h_m$) and the personal allocation of monetary resources to children (I_m) – and a spillover from her partner's consumption evaluated at a discount factor of δ_m.

Child development C depends on time investments of both mothers and fathers ($1 - h_m$, $1 - h_p$) and monetary investments $I_m + I_p$. Among others, the latter includes the purchase of extra-parental care services during the working time of parents. The productivities of these input factors are defined by the parameters $\alpha_m, \alpha_p, \gamma$.[8]

[7] Paternal utility is the exact mirror case.

[8] Note that C does not necessarily correspond to a production function for the development of specific cognitive or socio-emotional skills (Cunha et al., 2010). First, parenting decisions may involve mixed objectives including both the child's contemporary well-being as well as endowing it with the skills necessary to succeed in life (Doepke et al., 2019). Second, even if parents were to target a particular child skill, they may have mis-perceptions about

Parents dispose of one unit of time and make two individual decisions: First, how much time to spent in the labor market ($0 \leq h_g \leq 1$), where each increase in labor supply decreases the time available for childcare activities. Second, how much money to invest into the development of their child ($0 \leq I_g \leq w_g h_g - \bar{z}_g$), where each increase in child investments reduces the budget for private consumption and $\bar{z}_g$ specifies its desired minimum floor.

For the sake of the following illustration, I impose a number of restrictions on the set of exogenously given parameters $w_g, \delta_g, \alpha_g, \gamma, \beta_g, \bar{z}_g$. First, in line with evidence on the continued existence of gender wage gaps (Blau and Kahn, 2017, see also Figure 3.1), I assume $w_p > w_m$. Second, parents may place different discount factors on the value of their partner's consumption. Consistent with evidence on male breadwinner norms I impose $0 \leq \delta_p < \delta_m \leq 1$ (Bertrand et al., 2015, see also Figure 3.2). Third, the quality of maternal care is generally perceived as dominating alternative care arrangements including paternal and extra-parental care (Baker et al., 2019a; Del Boca et al., 2014, see also Figure 3.2). Therefore I impose $\alpha_m > \gamma > \alpha_p$. Fifth, mothers and fathers may differ in the utility value they place on child development β_g and the required minimum amount of private consumption $\bar{z}_g$. In line with the spending patterns documented in S. J. Lundberg et al. (1997), I impose $\beta_m > \beta_p$ and $w_p > \bar{z}_p > \bar{z}_m = 0$.[9]

Parents take the decisions of their partner as given and maximize their individual utilities while observing the budget constraints on working hours ($0 \leq h_g \leq 1$) and monetary investments into their children ($0 \leq I_g \leq w_g h_g - \bar{z}_g$). The first order condition for each parent yields:

$$\begin{gathered} w_g = \frac{\beta_g \alpha_g}{C}; \;\; 1 = \frac{\beta_g \gamma}{C}; \\ h_g \lambda_g = 0; \;\; (1 - h_g)\eta_g = 0; \;\; (w_g h_g - \bar{z}_g - I_g)\psi_g = 0; \;\; I_g \phi_g = 0. \end{gathered} \tag{83}$$

Observing the set of restrictions introduced above, we can distinguish six cases that vary in terms of i) the relative emphasis that parents put on the development of their children

the actual technology that produces the relevant trait (Attanasio et al., 2019; Cunha et al., 2013). For my purposes it is sufficient that the resources that are subject to the parental optimization calculus are relevant for the production of socio-emotional skills. This assumption is backed by the large body of literature showing the relevance of monetary resources and parental time investments for the development of socio-emotional skills (see among others Agostinelli and Sorrenti, 2018; Akee et al., 2018).

9 I also assume that $w_g \gamma \neq \alpha_g$, i.e. that time investments at home and time spent in the labor market are not equally productive in fostering the development of the child. This restriction limits the set of possible solutions by forcing at least one parent to be at a corner solution.

($w_g \lesseqgtr \beta_g$) and ii) their relative productivity in providing the necessary inputs via time or monetary investments ($w_g\gamma \lesseqgtr \alpha_g$). Table 3.1 shows the respective solutions to the household problem.

Table 3.1 – Overview of Model Solutions

Panel (a): $w_p > \beta_p;\ w_m < \beta_m$				
$w_m\gamma < \alpha_m, w_p\gamma \lesseqgtr \alpha_p$	$1-h_p=0;$	$I_p=0;$	$1-h_m=1;$	$I_m=0$
$w_m\gamma > \alpha_m, w_p\gamma > \alpha_p$	$1-h_p=0;$	$I_p=0;$	$1-h_m=0;$	$I_m=w_m$
Panel (b): $w_p > \beta_p;\ w_m > \beta_m$				
$w_m\gamma < \alpha_m, w_p\gamma \lesseqgtr \alpha_p$	$1-h_p=0;$	$I_p=0;$	$1-h_m=\frac{\beta_m}{w_m};$	$I_m=0$
$w_m\gamma > \alpha_m, w_p\gamma > \alpha_p$	$1-h_p=0;$	$I_p=0;$	$1-h_m=0;$	$I_m=\beta_m$
Panel (c): $w_p < \beta_p;\ w_m < \beta_m$				
$w_m\gamma < \alpha_m, w_p\gamma < \alpha_p$	$1-h_p=1-\frac{\bar{z}_p}{w_p};$	$I_p=0;$	$1-h_m=1;$	$I_m=0$
$w_m\gamma > \alpha_m, w_p\gamma > \alpha_p$	$1-h_p=0;$	$I_p=w_p-\bar{z}_p;$	$1-h_m=0;$	$I_m=w_m$

Panels (a) and (b) are similar in that fathers care strongly about their private consumption and put less emphasis on the development of their children ($w_p > \beta_p$). In these cases, the relevant inputs for the development of children are provided by mothers only. Panels (a) and (b) are different in the extent to which mothers care about their children as opposed to their private consumption ($w_m \lesseqgtr \beta_m$). Lastly, Panel (c) shows cases where both mothers and fathers care strongly for the development of their children ($w_g > \beta_g$). How do changes in the relative wages of mothers and fathers affect the provision of resources to the child in each of these scenarios?

In Panel (a) mothers care strongly for their child ($w_m < \beta_m$). If maternal wage rates are high enough and/or monetary investments are very conducive to child development ($w_m\gamma > \alpha_m$), she will work full time while purchasing the required inputs for the child in the market. In such a scenario, increases in maternal wage rates will lead to a one-to-one increase in monetary resources devoted to children. To the contrary, if maternal wages are low and/or monetary investments are relatively less productive than time investments ($w_m\gamma < \alpha_m$), she will care for the child at home with I_m and $1-h_m$ remaining unresponsive to changes in maternal wage rates.

In Panel (b) mothers care less strongly for their child ($w_m > \beta_m$) but still provide the entirety of the family's child investments. In such a scenario, the effect of changes in w_m on the resources devoted to children is ambiguous. If mothers perceive monetary investments as an inferior mode of child investment ($w_m\gamma < \alpha_m$), children will receive a decreasing share of maternal time as female wage rates and the opportunity cost of staying at home increase. To the contrary, if mothers prefer monetary investments, the child will receive an income bundle equal to β_m irrespective of the changes in w_m.

Panel (c) shows the cases where both mothers and fathers care strongly for the development of their children ($w_g > \beta_g$). Again we can distinguish two cases of how changes in the relative wages of mothers and fathers affect the provision of resources to the child. If wages are high enough and/or parental time investments are relatively more productive than monetary investments ($w_g\gamma > \alpha_g$), fathers will spent a minimum amount of time in the labor market to generate $\bar{z}_p$ while mothers specialize in home care for the children. Wage increases of fathers lead to a greater share of paternal time resources devoted to children since it takes less working time to satisfy their need for private consumption. In this case, resource allocations are unaffected by changes in w_m. To the contrary, if parents favor monetary investments, they will both work full time. Mothers invest their entire income into their children while fathers top up the maternal investments since the wage rate of mothers is too low to satisfy the paternal preferences for investments into the child. Hence, the resources devoted to children are again insensitive to changes in the wage rates of mothers. However, every increase of w_p that is in excess of $\bar{z}_p$ will lead to a one-to-one increase in the monetary resources devoted to children.

The solution of this stylized model illustrates that changes in the relative wages of mothers and fathers may impact both the amount and the mix of resources devoted to children. First, they alter the relative prices of private consumption and child investments for both mothers and fathers. Second, they alter the relative prices of important input factors for the development of children – time and money in particular. However, the illustration also highlights that gendered preferences for parental roles, i.e. β_g, $\bar{z}_g$, as well as beliefs about the productivity of different modes of child investments, i.e. α_g, γ, may insulate the resources devoted to children from changes in parental economic incentives

3.3 Context and Data

3.3.1 Gender Gaps in the Labor Market and at Home - The Case of Germany

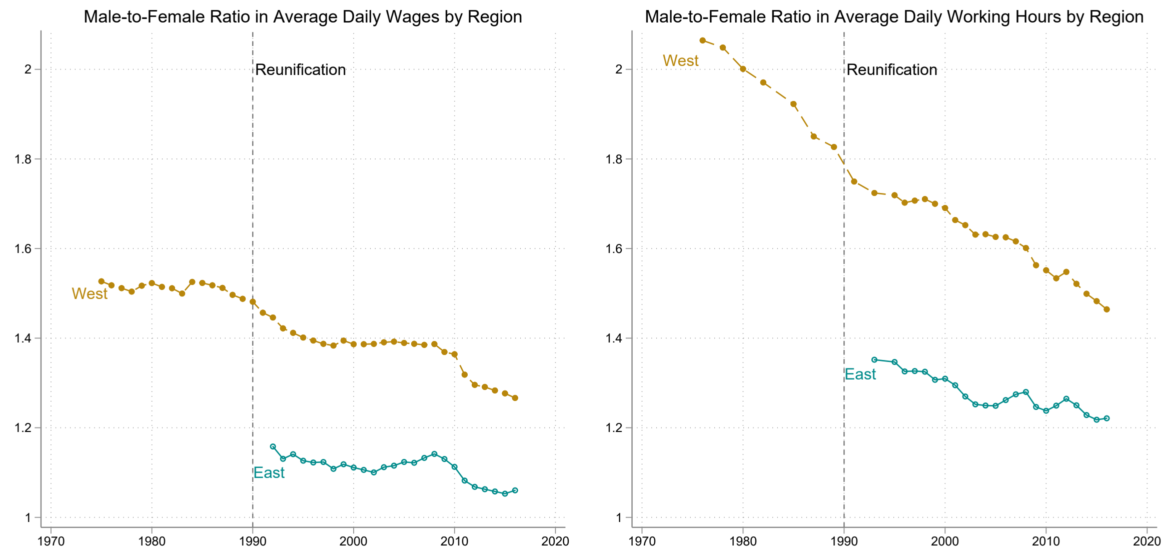
As in many industrialized societies, labor market outcomes for men and women in Germany have been converging in recent decades (Olivetti and Petrongolo, 2016). However, in spite of these strides towards gender equality, there still remain marked gender differences in labor market participation and home production, with the male breadwinner model being the norm among German households with children.

A particularity in the German context are the differences in gender roles between the former socialist East Germany and West Germany that continue to exist even three decades after reunification in 1990 (Boelmann et al., 2020; Lippmann et al., 2020). Figure 3.1 shows the development of the male-to-female ratios in average daily wages (daily working hours) over the time period 1975–2016 (1973–2016) separately for both regions. While there is a clear trend towards increased gender equality in both East and West, the remaining gender gap in daily wages (daily working hours) amounts to 27% (46%) in the West but only 6% (22%) in the East.

The legacy of the 41-year division is also reflected in gender role attitudes. In comparison to other industrialized countries, Germany as a whole is characterized by rather traditional gender norms (Kleven et al., 2019). However, this comparison masks important heterogeneity within the country. Figure 3.2 shows the evolution of preferences for the male breadwinner model and stated concerns about the adverse effects of working mothers on the development of children by region within Germany. While more conservative attitudes have been eroding over time, the two regions started to converge only recently when the trend towards more gender-equal attitudes plateaued in the East.

In recent years, Germany has implemented a number of policy reforms to foster gender equality and to support the reconciliation of family and work. In 2007, Germany introduced a new parental leave benefit with a 67% replacement rate of pre-birth earnings. The duration is 12 months with an additional 2 months – the so called “daddy months” – reserved for the partner of the primary caretaker (Raute, 2019). In addition, Germany has expanded the provision of center-based childcare significantly. Since 2013 the legal claim for publicly subsidized childcare has been extended from children aged 3–6 to all children aged one year and above (Felfe and Lalive, 2018). Current plans for the expansion of public childcare

FIGURE 3.1 – Development of the Unconditional Gender Wage/Hours Gap in Germany by Region, 1973-2016



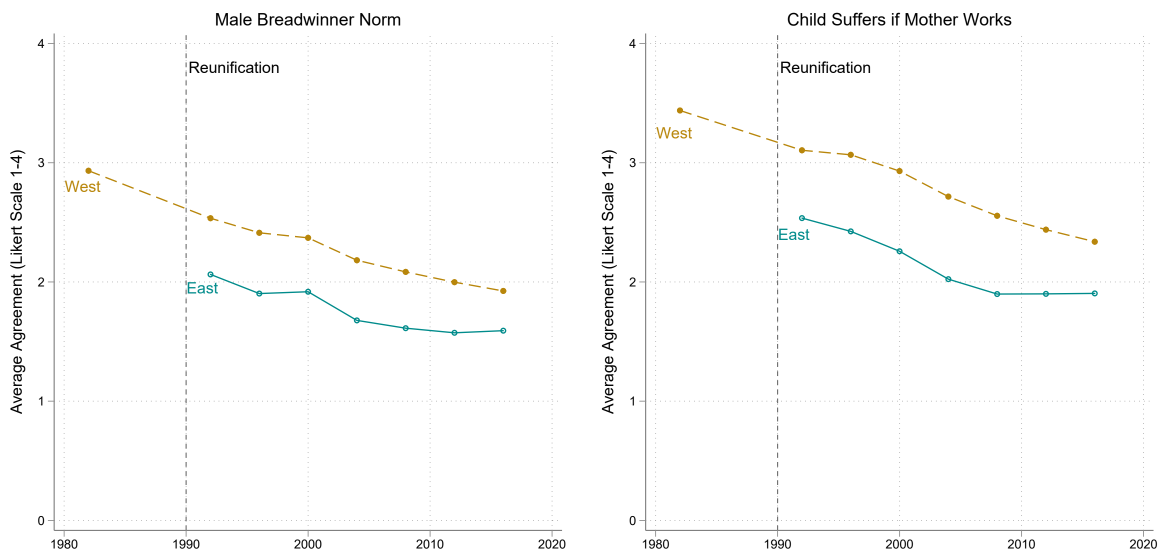
Data: Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This figure shows the development of the male-to-female ratio in mean daily wages (working hours) from 1975 to 2016 (1973–2016) by region in Germany. Daily wages are calculated for all SIAB observations aged 18–63 that are subject to social security contributions. Daily working hours are calculated for all MZ observations aged 18–63 by dividing their working hours in a typical work week by five. A detailed description of the underlying data sources is provided in section 3.3.2.

provision include a legal claim for afternoon care in elementary schools until 2025 (Federal Government of Germany, 2019). In contrast to these reform efforts, the German tax code is an inhibitor of increased gender equality since it combines the joint taxation of couples with a progressive schedule. It thus places high marginal tax rates on the secondary earner within a tax unit, i.e. females in the vast majority of cases (Bick and Fuchs-Schündeln, 2017).

3.3.2 Data

My research design combines a sibling comparison with a shift-share design to approximate within-family changes in the relative earnings potential between mothers and fathers. To operationalize this identification approach in the German context I rely on three principal data sources. The German Socio-economic Panel (GSOEP) provides the core data set in which I observe household responses to changes in the relative labor market incentives of mothers and fathers as well as measures of child development. The sample of the GSOEP, however, is too small to reliably calculate potential wages based on a shift-share design. Therefore, I use the Sample of Integrated Labour Market Biographies (SIAB) and the German Microcensus (MZ)

Figure 3.2 – Development of Gender Role Attitudes in Germany by Region, 1982-2016



Data: German General Social Survey (ALLBUS).
Note: Own calculations. This figure shows the development of gender role attitudes from 1982–2016 by region in Germany. Each data point reflects average agreement to the following statements among respondents aged 18–63 measured on a four point Likert scale : *People have different opinions about the role of women in the family and in bringing up children. For each of the statements on the card, please tell me whether you completely agree, tend to agree, tend to disagree, or completely disagree:* [Left-hand panel:] *It is much better for everyone concerned if the man goes out to work and the woman stays at home and looks after the house and children.* [Right-hand panel:] *A small child is bound to suffer if his or her mother goes out to work.*

to calculate hourly potential wages in gender times education times commuting zone cells ($2 \times 3 \times 96$) that are linked back to the GSOEP based on observable household characteristics.

The German Socio-economic Panel (GSOEP). Established in 1984, the GSOEP is an annual, nationally representative survey that covers approximately 15,000 private households and 25,000 individuals in its most recent waves (Goebel et al., 2019). Next to comprehensive information on socio-economic and demographic background characteristics, the GSOEP contains detailed information on financial positions, labor market participation, and the time-use of households and their members. Furthermore, there are dedicated questionnaires administered to primary caretakers and children themselves that allow me to construct established measurements for the socio-emotional development of children.

Guided by my empirical strategy, I restrict the GSOEP to intact families with two resident working age parents (18–63 years) who have at least two children for whom I observe the outcomes of interest at the same chronological age.[10] From 2005 onward, the GSOEP contains a mother-and-child questionnaire that includes a short scale for the personality development of children. From 2006 onward, the GSOEP contains a battery of self-reported personality questions that allow the derivation of analogous personality measures for older children. As a consequence, I restrict my analysis to the GSOEP waves covering the years 2005–2017. Following these restrictions, I obtain a sample of $6,070$ child-year observations and $2,833$ sibling groups for whom I provide descriptive statistics in Table 3.2.[11]

The resulting sample is gender-balanced. Only 1% of the sampled children have been born outside of Germany while 19% reside in the formerly socialist East.[12] On average, they are 8.6 years of age and the second-born child to their parents.

In my analysis I focus on the following set of variables. First, I analyze the labor market response of parents by reference to their working hours and earnings. Working hours are self-reported and I convert the provided variable on annual working hours into daily working hours by dividing with 260 days.[13] Earnings are self-reported, deflated to 2015 prices, and include all income from employment and self-employment in the year that precedes the survey wave. As shown in Table 3.2, there are marked gender gaps in the labor market outcomes of mothers and fathers in my sample. Fathers spend almost triple the time of mothers (8.4 vs. 3.0 hours/day) in the labor market and contribute four times the earnings of mothers to the financial resources of the household (51.2k vs. 12.5k €/year).

Second, I analyze the childcare response of parents by reference to the hours of care provided by both partners and the use of extra-parental care providers. Information on the hours of care are elicited from both partners separately and refer to a typical day in a work week. A comparison of the GSOEP with the German Time-Use Study (GTUS) suggests that the information on

[10] I define intact families as follows: Children below age 18 must i) live in the same household as their mother in all available waves, ii) refer to the same person as their mother figure in all available waves, iii) be either a biological child, adopted child or the child of the partner of the head of the household in which they reside. Following this definition, I allow for non-biological family relationships if they are characterized by a sufficient degree of stability over time. In section 3.5.4 I show that my results are robust to the exclusion of non-biological family ties.

[11] Note that the number of sibling groups is less than half the child-year observations since I allow for sibling groups that contain more than two siblings, i.e. triplets, quadruples etc., if they exist.

[12] In my baseline analysis I do not explicitly exclude children from the refugee over-samples that were added to the GSOEP in the waves of 2016 and 2017. However, as a consequence of my sample restrictions there are only 6 child-year observations from the refugee over-samples in my core sample. Excluding these observations does not change any of the results presented below.

[13] 260 days$\approx$12 months$\times$4.33 weeks/month$\times$5 days/week.

TABLE 3.2 – Summary Statistics

	N=6,070, Sibling Groups=2,833			
	Mean	SD	Min	Max
Children				
Female	0.49	0.50	0.00	1.00
Migration Background	0.01	0.10	0.00	1.00
East Germany	0.19	0.39	0.00	1.00
Age	8.64	5.24	2.00	17.00
Birth Rank	2.04	1.10	1.00	12.00
Formal Care	0.58	0.49	0.00	1.00
Informal Care	0.27	0.44	0.00	1.00
Openness	0.02	0.95	-4.05	2.12
Conscientousness	0.05	0.96	-3.39	1.92
Extraversion	-0.02	0.99	-3.89	1.79
Agreeableness	0.00	0.98	-3.76	2.02
Neuroticism	-0.03	0.97	-2.50	3.06
Mother				
Annual Earnings ('000 €)	12.47	18.73	0.00	576.00
Work Hours/Day	2.97	3.05	0.00	16.00
Childcare Hours/Day	6.50	4.62	0.00	16.00
Father				
Annual Earnings ('000 €)	51.23	45.39	0.00	672.00
Work Hours/Day	8.35	2.99	0.00	16.00
Childcare Hours/Day	1.99	2.31	0.00	16.00

Data: German Socio-economic Panel (GSOEP).
Note: Own calculations. This table shows summary statistics for the core analysis sample. The sample spans the years 2005–2017. It includes two-parent households aged 18–63 with at least two resident children aged 2–17 in year t who have non-missing information on the CZ of residence, parental education, parental working hours, parental child care hours and parental earnings in periods t and $t-1$. It only includes child-year observations with a valid measurement for at least one of the Big Five dimensions. Child-year observations without information on the child's sex, birth brank, migration background as well as the number of children in the household are subject to listwise deletion.

childcare is best understood as spending time with the child but not necessarily as a dedicated time investment into the child (see Table C.12 in the Appendix). I separate extra-parental care into formal and informal care. Formal care includes center-based childcare for children under six, after-school care for children aged six years and older, as well as the use of childminders outside of the parental household. Informal care includes care provision by the extended family, older siblings, friends, neighbors as well as paid in-home babysitters. As shown in Table 3.2, the gender gaps observed in the labor market reverse in the domain of childcare provision. Mothers invest more than triple the time of fathers into childcare activities. The use of external care providers is wide-spread with 58% (27%) of all children being exposed to some form of formal (informal) childcare.

Third, I analyze the impact of converging labor market opportunities on the socio-emotional development of children as measured by the Big Five dimensions of personality: openness, conscientousness, extraversion, agreeableness, neuroticism. The Big Five taxonomy evolved from the study of personality traits in Psychology and is derived by factor analysis on a battery of self-reported and/or observer-reported behaviors. While not without critics, it is the most widely accepted taxonomy of personality traits and has gained widespread traction in the economics literature.[14] In the GSOEP information on the Big Five dimensions are derived from assessments of the primary caretaker at child ages 2–3, 5–6, and 9–10 and child self-reports at ages 11–12, 13–15 and 17. These assessments are based on a battery of questions that rate the child in terms of various behaviors on a 10-point (7-point, in case of self-reports) Likert scale. Each question has a mapping into one of the Big Five dimensions.[15] I aggregate the questions additively such that higher values correspond to a higher expression of the underlying trait and standardize the resulting variables at each child age on the full sample to account for personality differences as children grow up. Table 3.2 shows that the sibling sample is slightly positively selected in terms of openness and conscientousness, and is characterized by lower levels of extraversion and neuroticism than the full sample.

Potential Wages. I approximate the differential changes in the labor market incentives for mothers and fathers by calculating potential wages for socio-demographic groups in Germany. While this section is dedicated to the construction of potential wages, I will elaborate on their econometric intuition in section 3.4. I use two data sets for the construction of potential wages.

The Sample of Integrated Labour Market Biographies (SIAB). The SIAB is an administrative data set compiled by the research institute of the Federal Employment Agency of Germany that contains a 2% random sample of Germans who are either employed, recipients of social

[14] See Almlund et al. (2011) and Borghans et al. (2008) for comprehensive overview articles. See also Table C.17 for short descriptions of each Big Five personalty dimension.
[15] See Table C.18 for an overview of the questions and their mapping into the Big Five dimensions.

benefits, or officially registered as job-seeking (Antoni et al., 2019).[16] The data is organized in spells and allows to trace the labor market biographies of the sampled individuals as long as they fall into one of the categories mentioned above. The latest version of the SIAB covers the time period 1975–2017 and contains information on socio-demographics, occupation, industry affiliation and daily wages. The data does not include self-employed workers and civil servants.

For the purpose of this study I restrict the SIAB to all spells in the time period 1995–2016 that refer to individuals of working age (18–63 years) and who are subject to social security contributions. Based on information about the individual's establishment, I aggregate the spells to job cells where each observation represents one job per individual in a particular year. As a result I obtain a data set with more than 12 million job observations ($N \approx 577,720$/year).[17] The SIAB contains information on daily wages that are right-censored at the cap for social security contributions. In my baseline analyses I impute the upper tail of the wage distribution by following the procedure proposed in Gartner (2005). However, in section 3.5.4 I show the robustness of my conclusions to a variety of different imputation assumptions.

The German Microcensus (MZ). The MZ is an annual household survey covering 1% of all German households. It contains information on family socio-demographics, living arrangements and labor force participation (GESIS, 2020). Importantly - and in contrast to the SIAB - the MZ contains information on working hours. For the purpose of this study I use the MZ waves 1995–2016. In order to match the sample composition of the SIAB, I restrict the MZ data to employed individuals of working age (18–63 years) while excluding individuals who are either self- or marginally employed.[18] As a result I obtain a data set with more than 3 million individual observations ($N \approx 166,849$/year). In my baseline analysis I use reports on individual working hours that refer to a typical work week of the respondent. However, in section 3.5.4 I show the robustness of my conclusions to alternative working hours definitions.

[16] In this study, I use the regional file SIAB-R 7517 which contains regional markers while cutting back on detail in other dimensions to preserve data confidentiality.
[17] I drop individuals who change their jobs more than three times per annum to exclude individuals with marginal labor force attachment.
[18] Tables C.15 and C.16 provide evidence that the resulting samples of the SIAB and the MZ are indeed comparable in terms of their socio-demographic, industry and occupation compositions.

Construction of Potential Wages. I combine the SIAB and the MZ to calculate potential wages for individuals according to a shift-share design. The general idea of shift-share designs is to predict group-specific wages based on sectoral shocks and the group's exposition to such shocks as approximated by the historic importance of the different sectors for the respective group.

I define *groups* by partitioning the German population into 576 cells that are pinned down by 2 expressions of gender, 3 education levels and 96 regional units. The low education group includes individuals with no more than a low-track secondary degree and without vocational training. The intermediate education group includes individuals with a low-track secondary degree and vocational training as well as individuals with a high-track secondary degree but no further tertiary education. The high education group consists of people with a tertiary education at the university level. The 96 regional units correspond to Germany's spatial planning regions. Spatial planning regions describe economic centers and their surroundings that are nested within the 16 federal states of Germany. Since commuting flows are an essential criterion for the definition of spatial planning regions, I will refer to them as commuting zones (CZ) in the following.

I define employment *sectors* by grouping employed individuals into 27×14 occupation-industry cells that are based on the German Classification of Occupations 2010 (KldB10) and the German Classification of Activities 2008 (WZ08).[19]

Based on these specifications, I calculate potential wages for individuals of gender g, with education level e, residing in region r, in year t as follows:

$$\hat{w}_{gert} = \sum_{j} \sum_{o} \underbrace{\frac{E^{oj}_{ger,1995}}{E_{ger,1995}}}_{(1)} \times \underbrace{w^{oj}_{t,-r}}_{(2)} . \tag{84}$$

Term (1) of equation (84) indicates the group-specific employment share of each industry-occupation cell in base year 1995. Term (2) of equation (84) indicates the leave-one-out average wage paid to individuals working in occupation o and industry j in year t at the national level. Hence, the group-specific potential wage $\hat{w}_{gert}$ is constructed as a weighted average across

[19] The cross-walks from the industry and occupation classification used in this paper to the German Classification of Occupations 2010 (KldB10) and the German Classification of Activities 2008 (WZ08) at the three digit level are accessible through the author's homepage.

the wages paid in the different sectors of the economy where the weights are given by the historic exposure of the groups to the respective sectors.

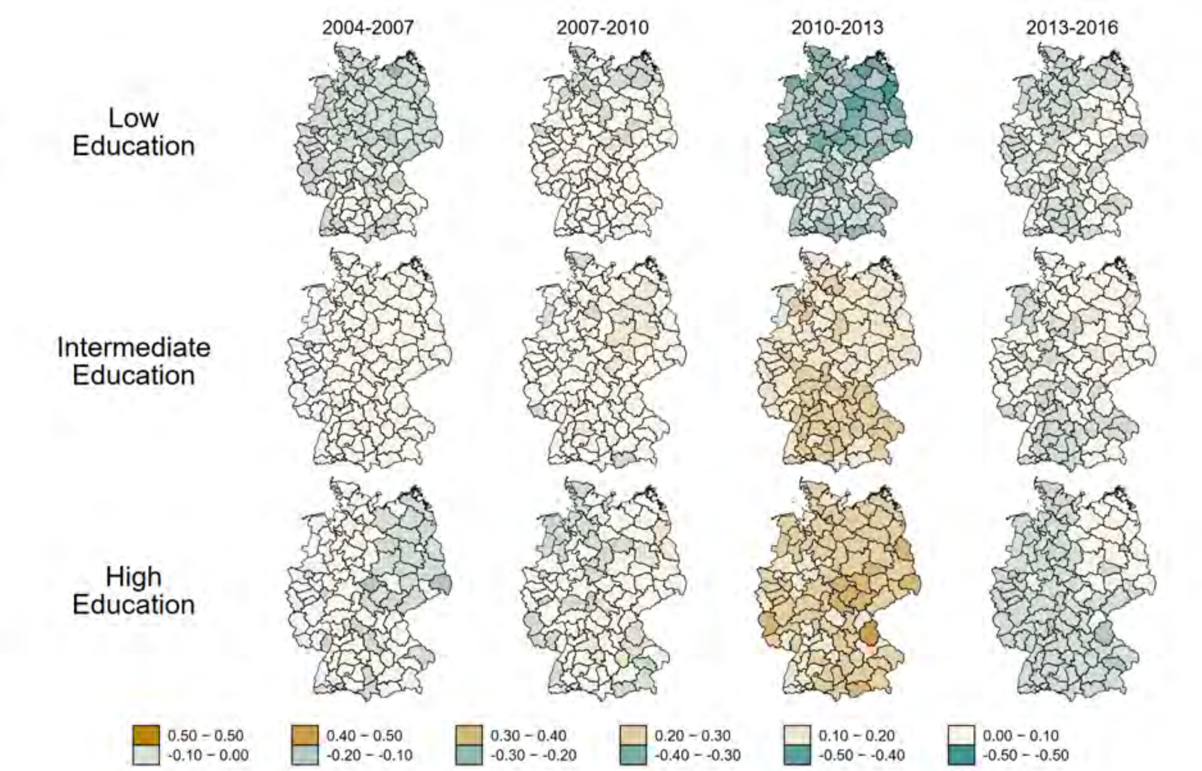
Specifically, I use the SIAB wave of 1995 to construct the group-specific employment share of each industry-occupation cell in base year 1995 (Term (1) of equation (84)). Tables C.13 and C.14 in the Supplementary Material document the differential sorting of gender and education groups into industries and occupations in 1995. For example, while almost a quarter of all low educated males worked in logistics occupations, an equally high share of low educated females worked in occupations related to facility management. The most important occupations for highly educated males are business administration and engineering, while their female analogues tend to work in nursing and teaching occupations. Furthermore, I use the SIAB waves 2004–2016 to construct the average wage paid to workers in each sector at the national level (Term (2) of equation (84)). However, the SIAB does not contain information on hourly wages. Therefore, I divide the average daily wage of individuals working in a particular sector in year t by the corresponding average daily working hours from the MZ.[20]

Figure 3.3 displays the change of the gender gap in potential wages by education group across the 96 CZs of Germany over the time period of my analysis (2005-2016). Blue areas indicate changes in favor of male wages, while red areas indicate changes in favor of female wages. There is strong heterogeneity in the evolution of gender gaps across regions and education groups, ranging from changes in hourly potential wages of € 0.40 to the advantage of females to changes of € 0.51 to the advantage of males.

Data Linkage. I match potential wages calculated from the SIAB and the MZ to the GSOEP sample based on an individual's expression in the group characteristics gender, education and CZ of residence. That is, for each year in the time period 2005–2017 GSOEP parents receive one out of 576 potential wages to approximate the respective parent's labor market incentives.

[20] Note that the MZ does not contain geographic information at the level of commuting zones. Hence, average daily wages at the national level that leave out a particular CZ are matched with average daily working hours at the national level that leave out the entire federal state in which the CZ is nested.

FIGURE 3.3 – Change in Gender Gap of Potential Hourly Wages by Education and Commuting Zone, 2004-2016



Data: Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This figure shows the change in the gender gap of potential wages from 2004 to 2016 in three-year windows by education level and commuting zone. Positive values (in red) indicate relative gains of females. Negative values (in blue) indicate relative losses of females. Potential wages are calculated according to equation (84). The 96 commuting zones are defined by the official territory definition of spatial planning regions of the Federal Office for Building and Regional Planning from 31.12.2017. Education is classified as follows – lower secondary degree without tertiary education (*Low*), lower secondary degree with vocational training or higher secondary degree without vocational training (*Intermediate*), university qualification (*High*).

3.4 Empirical Strategy

Identification Strategy. I am interested in the causal effect of the parental wage gap on the development of socio-emotional skills in children as well as the household decisions through which parents provide the input factors for the production of these skills. Let us denote the outcomes of interest by Y_{ifat} and the parental wage gap as the difference between maternal and paternal wages, $w^{\Delta}_{ifat}(= w^{m}_{ifat} - w^{p}_{ifat})$, respectively. Both variables of interest are measured when child i from family f is of age a in year t.

If w^{Δ}_{ifat} was randomly assigned across families and time we could estimate the sought-after average treatment effect with the following OLS regression:

$$Y_{ifat} = \alpha + \beta w^{\Delta}_{ifat} + \epsilon_{ifat}. \tag{85}$$

However, w^{Δ}_{ifat} is not randomly assigned and the identification assumption implicit in equation (85), namely that $Cov(\epsilon_{ifat}, w^{\Delta}_{ifat}) = 0$, may be violated through joint determinants of parental wages and child outcomes as well as reversed causality.

In response to the various threats to identification I estimate the following model instead:

$$Y_{ifat} = \alpha + \beta \hat{w}^{\Delta}_{ifat-1} + \gamma_{fa} + \tau_t + X'_{ifat}\delta + \epsilon_{ifat}. \tag{86}$$

First, I leverage the panel dimension of my data to construct a sibling sample in which I observe children from the same family f at the same child age a but in different calendar years t (see section 3.3 for details). This data structure allows me to include a vector of family times child age fixed effects, γ_{fa}, that absorbs all confounding factors nested in differences across families that are particular to a specific child age. Examples of confounding factors that are ruled out by the inclusion of γ_{fa} include family differences in gender norms (Boelmann et al., 2020; Lippmann et al., 2020), assortative matching (Eika et al., 2019), and genetic endowments (Demange et al., 2020).

Second, I include a vector of time fixed effects τ_t. As shown in Figure 3.1, the gender wage gap in Germany has a clear negative time trend. Hence, one may worry that within-family sibling comparisons confound the effect of changes in the parental wage gap with sibling birth order and parental age effects. The inclusion of τ_t takes care of both of these concerns. To see this, note that the inclusion of γ_{fa} fixes the age for the sibling comparison. Since a child's birth cohort is a linear combination of its age a and the year of observation t, the joint inclusion of γ_{fa} and τ_t excludes birth cohort effects as confounding factors (Black et al., 2018). Analogously, including γ_{fa} fixes the birth cohort of parents. Since parental age is a linear combination of their birth cohort and the year of observation t, the joint inclusion of γ_{fa} and τ_t excludes parental age effects as confounding factors (McGrath et al., 2014).

Third, I replace the observed wage difference in households, w^{Δ}_{ifat}, with the lagged difference in potential wages $\hat{w}^{\Delta}_{ifat-1}$. Observed wages are an endogenous proxy variable for the labor market incentives of mothers and fathers as parents may adjust their labor supply in response to the development of their children. Using potential wages along the lines of Bartik (1991) that reflect wage variation due to local labor demand instead of endogenous parental labor supply decisions addresses such concerns.[21]

Lastly, I include time-varying individual level controls X'_{ifat}. In my baseline specification X'_{ifat} consists only of $\hat{w}^{\Sigma}_{ifat-1} (= \hat{w}^{m}_{ifat-1} + \hat{w}^{p}_{ifat-1})$, i.e. the joint wage shock to mothers and fathers. Including $\hat{w}^{\Sigma}_{ifat-1}$ allows me to separate changes in the relative wages available to mothers and fathers from general shocks that affect the two partners simultaneously. In section 3.5.4 I show that my results are robust to richer specifications of X'_{ifat}.

Identifying Assumptions. Recently, the formal properties of shift-share designs have received increased attention in the methodological literature (Adão et al., 2019; Borusyak et al., 2019; Goldsmith-Pinkham et al., 2020; Jaeger et al., 2018). Exogenous variation in shift-share designs can originate from the exogenous assignments of the "shifters", i.e. term (2) of equation (84), or the "shares", i.e. term (1) of equation (84).[22] In this work I follow the interpretation suggested by Goldsmith-Pinkham et al. (2020) and discuss my identifying assumptions in terms of exogenously assigned sector shares in the base year 1995. In light of this interpretation, the construction of potential wages is reminiscent of a difference-in-differences design where term (2) of equation (84) defines the treatment and term (1) of equation (84) the treatment assignment. In analogy to the standard difference-in-differences design, my identifying assumption can be stated as follows:

$$
\begin{aligned}
& Cov\left(\epsilon_{ifat}, \frac{E^{oj}_{ger,1995}}{E_{ger,1995}} | \gamma_{fa}, \tau_t, X'_{ifat}\right) = 0, \\
& \forall\, (o, j) \in J \times O, \\
& \forall\, t \geq 1995 + 10.
\end{aligned}
\tag{87}
$$

[21] Shift-Share (or Bartik) designs have become widely adopted in the literature strands on household decision-making (Anderberg et al., 2015; Autor et al., 2019; Bertrand et al., 2015; Bruins, 2017; Schaller, 2016; Shenhav, 2020) and child development (Agostinelli and Sorrenti, 2018; Aizer, 2010; Lindo et al., 2018; Page et al., 2019).

[22] Find in the following a restatement of equation (84) for easy reference:

$$
\hat{w}_{gert} = \sum_j \sum_o \underbrace{\frac{E^{oj}_{ger,1995}}{E_{ger,1995}}}_{(1)} \times \underbrace{w^{oj}_{t,-r}}_{(2)} .
$$

In words: Conditional on the set of controls, the group-specific sector shares in 1995 need to be uncorrelated to the residuals of estimation equation (86). Note that i) the set of controls includes family times child age fixed effects, and that ii) the base year 1995 precedes the core time window of my investigation (2005–2017) by 10 years. Hence, the identifying assumption implies that group-specific industry shares in 1995 need to be uncorrelated to intra-family *changes* in the outcome of interest that lag the base year by at least a decade.

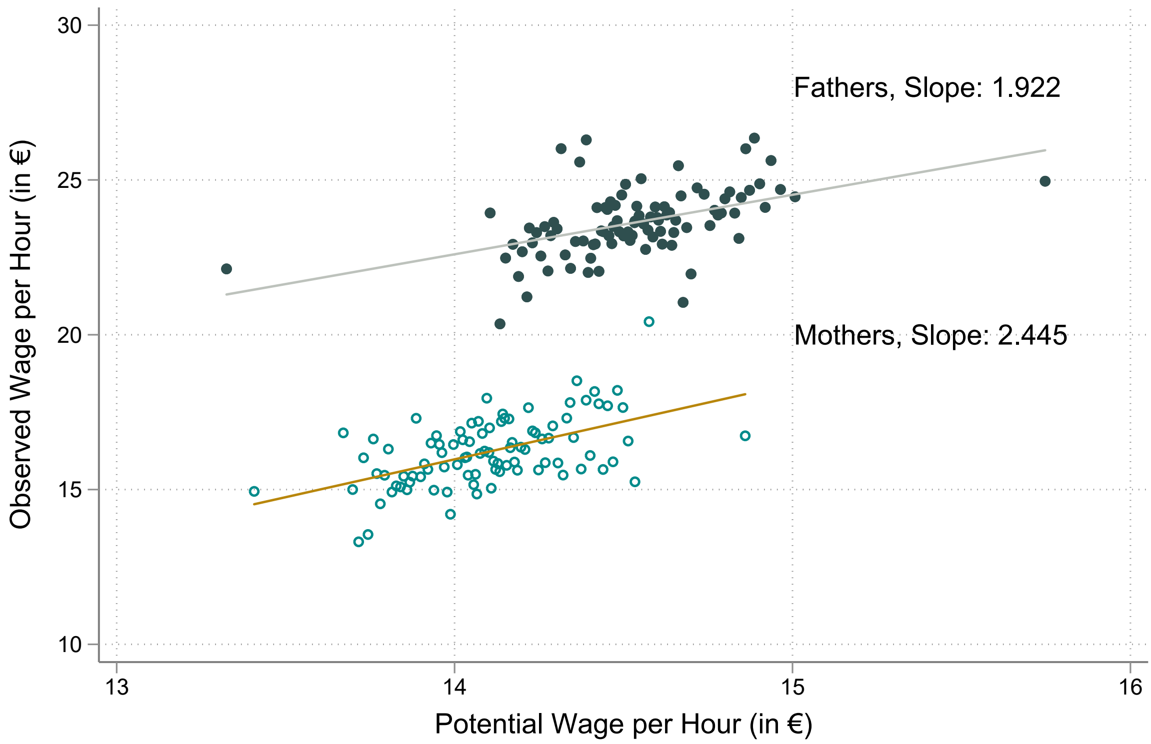
Evidence on Identifying Assumptions. I assess the plausibility of the discussed identifying assumptions in three steps. First, I illustrate the effects of the within-sibling FE design. For the sake of illustration, I draw a sample of sibling pairs from the core sample and partition them into a "high-shock" and a "low-shock" group depending on whether their value of $\hat{w}^{\Delta}_{ifat-1}$ exceeds the one of their sibling.[23] Panel (a) of Table 3.3 compares the resulting groups in terms of their individual characteristics. While both groups are comparable in many dimensions, there are statistically significant differences in terms of characteristics that are related to within-family cohort effects such as birth year, birth rank and parental age. However, as suggested in the discussion above, these differences vanish once I account for time fixed effects τ_t. Panel (b) of Table 3.3 compares the groups in terms of their exposure to differential labor market incentives for their parents. By construction the "high-shock" group is exposed to a significantly smaller gap in the potential wages of their parents. Importantly and in contrast to the sibling characteristics listed in Panel (a) these differences persist even when controlling for time fixed effects τ_t. The remaining intra-family differences in potential wages provide the identifying variation on which I base my estimates.

Second, I use the shift-share wages as a proxy for the labor market incentives of mothers and fathers. While the true potential wages for mothers and fathers are unobserved, I can validate this proxy by comparing it to the actual wages realized by mothers and fathers in the analysis sample. In Figure 3.4 I show the residual correspondence between potential wages and actual wages after accounting for family times child age fixed effects and collecting the data in centile bins of the respective potential wage variable. There is a strong correlation between intra-family changes in potential and observed wages which gives credence to the assumption that

[23] Note that this restriction to sibling pairs is implemented for illustrating the identification in terms of treatment and control groups. In Table C.19 I run the same test on the entire sample using regression analyses. Conditional on γ_{fa} and τ_t, $\hat{w}^{\Delta}_{ifat-1}$ does predict none of the 10 child characteristics at a significance level of 10%. Hence, the conclusions described in the main body of the text remain unaffected.

the shift-share wages are good proxies for the actual labor market opportunities available to mothers and fathers.

FIGURE 3.4 – Correlation of Within-Family Changes in Potential and Observed Wages



Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This figure shows the relationship between within-family changes in potential wages and within-family changes in observed wages by parental gender. It is constructed from the core sample described in Table 3.2 by partialling out the sibling times child age fixed effect γ_{fa} from actual wages and potential wages, respectively. The data is collapsed to gender-specific centile bins such that each data point reflects the average actual and potential wage within a centile bin of the gender-specific potential wage distribution.

Third, given the identification assumption stated in equation (87), the group-specific exposure to a particular sector in the base year can be interpreted as an instrument for the endogenous variable of interest. Hence, in my case the identification relies on $J \times O$ (14×27) instruments. To clarify the identification that underlies a particular shift-share design Goldsmith-Pinkham et al. (2020) propose a decomposition of the resulting estimates into just-identified instrumental variable coefficients and the corresponding "Rotemberg Weights". The latter indicate the importance of the individual sector shares for potential biases in the aggregate estimate. Tables C.1 and C.2 show the Rotemberg weights for the top ten industry times occupation cells by gender. For women most of the variation is accounted for by teachers and social workers employed in the educational sector ($\approx 31\%$) followed by sales occupations in retail ($\approx 6\%$) and cleaning and facility management occupations in the human health services industry ($\approx 5\%$). For men, the Rotemberg weights are much more dispersed across sectors with each of the top ten sectors accounting for less than ten but more than three percent. Most of the variation is accounted for by teachers and social workers employed in the educational sector ($\approx 10\%$), construction and civil engineering ($\approx 7\%$), as well as technical occupations in manufacturing ($\approx 7\%$). The importance of school teachers for the wage development of both women and men mirrors results for the US in the 1980–2010 period (Shenhav, 2020). In general, the distribution of the Rotemberg weights suggests a low sensitivity of my estimates to violations in the identification assumption for specific industry-occupation cells. The only notable exception is the importance of the school teacher category for the wage development of women. Hence, the causal interpretation of my results would be threatened if – conditional on controls – the region-and education-specific employment share of school teachers among women in base year 1995 would correlate with any features that predict intra-family variation in the outcomes of interest after the year 2005.

3.5 Results

I present the results of my analysis in three steps. First, I will present parental labor market responses towards the differential changes in labor market incentives across mothers and fathers. Second, I will present the childcare responses of these parents. In the third step, I present the reduced-form causal estimates of gender convergence in labor market incentives on the Big Five personality traits of the children in the affected families. Throughout the analysis, all coefficients represent responses to €1 increases in the respective wage variable. Columns indexed by Σ always indicate sums across mothers and father, while columns indexed by Δ always represent the difference between mothers and fathers.

TABLE 3.3 – Within-Family Variation of Characteristics by Treatment Status

		Sibling × Child Age FE Only			Sibling × Child Age FE + Year FE		
	N	Low Shock	High Shock	Δ	Low Shock	High Shock	Δ
Panel (a): Sibling Characteristics							
Female	4,960	0.451	0.478	0.027* (0.059)	0.458	0.480	0.022 (0.200)
Migration Background	4,960	0.016	0.014	-0.002 (0.273)	0.011	0.012	0.001 (0.783)
Birth Year	4,960	2003.266	2004.064	0.798*** (0.000)	2004.463	2004.463	-0.000 (0.999)
Birth Rank	4,960	1.571	1.809	0.238*** (0.000)	1.922	1.926	0.004 (0.761)
# of Siblings	4,960	1.847	1.846	-0.001 (0.532)	1.844	1.845	0.001 (0.699)
Birth Height (cm)	2,010	50.655	50.779	0.124 (0.220)	50.768	50.817	0.049 (0.680)
Birth Weight (kg)	2,022	3.238	3.271	0.033* (0.068)	3.262	3.279	0.017 (0.427)
Breastfed	1,810	0.912	0.904	-0.008 (0.317)	0.915	0.905	-0.010 (0.273)
Age Mother	4,960	37.826	38.624	0.798*** (0.000)	39.023	39.023	0.000 (1.000)
Age Father	4,960	41.000	41.798	0.798*** (0.000)	42.197	42.197	-0.000 (1.000)
Panel (b): Treatment Variables							
Parental Wage Gap	4,960	-0.630	-0.494	0.136*** (0.000)	-0.628	-0.493	0.135*** (0.000)
Wage Mother	4,960	14.038	14.089	0.051*** (0.000)	14.057	14.095	0.038*** (0.000)
Wage Father	4,960	14.668	14.582	-0.086*** (0.000)	14.685	14.588	-0.097*** (0.000)

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This table shows differences in sibling characteristics conditional on different control variables. Siblings are allocated to the *High Shock* (*Low Shock*) sample if they are subject to a higher (lower) value of $\hat{w}^{\Delta}_{ifat-1}(=\hat{w}^{m}_{ifat-1}-\hat{w}^{p}_{ifat-1})$ than their sibling counterpart. The left-hand panel controls for sibling times child age fixed effects γ_{fa}. The right-hand panel additionally controls for year fixed effects τ_t. For the sake of illustration the sample is restricted to sibling pairs. In Table C.19 I present analogous tests while allowing for larger sibling groups. Significance Levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Standard errors (in parenthesis) are clustered at the family level.

3.5.1 Labor Market Response

Table 3.4 displays the labor market response of households to changes in the relative wages of mothers and fathers as well as the ensuing effects on household earnings. Panel (a) separates

TABLE 3.4 – Parental Wage Gaps and Labor Market Responses

	Work Hours				Earnings			
	Mother	Father	Σ	Δ	Mother	Father	Σ	Δ
Panel (a): Wages by Parent								
Wage Mother	0.749*** (0.260)	0.246 (0.333)	0.995** (0.463)	0.504 (0.378)	5.209*** (1.523)	1.218 (1.663)	6.427** (2.607)	3.990** (1.837)
Wage Father	-0.157 (0.097)	0.450** (0.220)	0.292 (0.228)	-0.607** (0.252)	-0.974** (0.384)	1.557 (1.074)	0.583 (1.166)	-2.531** (1.116)
Panel (b): Parental Wage Gap								
Parental Wage Gap			0.351 (0.269)	0.555** (0.242)			2.922** (1.366)	3.261*** (0.953)
Sibling $\times$ Age FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
N	6,070	6,070	6,070	6,070	6,070	6,070	6,070	6,070
DV Mean	2.966	8.352	11.317	-5.386	12.473	51.227	63.701	-38.754
DV SD	3.045	2.985	4.309	4.219	18.729	45.386	50.596	47.554

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. All coefficients are estimated on the core sample described in Table 3.2. All regressions in Panel (b) control for $\hat{w}^{\Sigma}_{ifat-1}$ – the aggregate labor demand shock for family f in year $t-1$. The coefficient on the parental wage gap can thus be interpreted as a test of coefficient equality across maternal wages ($\hat{w}^{m}_{ifat-1}$) and paternal wages ($\hat{w}^{p}_{ifat-1}$), see Panel (a). Work hours are measured in hours per day. Earnings are measured in thousand € per year. Σ indicates the sum across parental outcomes. Δ indicates the difference between maternal and paternal outcomes. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The last two rows of the table list the mean and the standard deviation of the dependent variable that is displayed in the table header.

the effects by wage shocks to mothers and fathers. Note that the point estimates for the effects on total household labor supply (earnings) and the intra-household difference in parental labor supply (earnings) represent the horizontal sum and difference across the labor supply (earnings) effects on mothers and fathers, respectively.

Both mothers and fathers have a positive own-wage elasticity of labor supply. Conditional on the potential wage of their partner, mothers (fathers) respond to a €1 increase in their potential hourly wage by increasing their time in the labor market by 0.749 (0.450) hours per day. Thus, consistent with Bargain et al. (2014) the labor supply of partnered men in Germany is approximately two thirds as sensitive to variation in their own wages as the labor supply of women. To the contrary, men and women tend to respond asymmetrically to wage shocks of their partners. While mothers tend to reduce their labor supply in response to positive wage shocks of their partners, fathers respond positively – even though the latter effects are imprecisely estimated.[24] In combination, these responses have the effect that increases in maternal wages have a strong and statistically significant positive effect on household's total hours of work, while increases in paternal wages have a strong and statistically significant positive effect on the intra-household gender gap in hours worked: Conditional on paternal wages, a €1 increase in the potential wages of mothers leads parents to increase their combined labor supply by $0.995(= 0.749 + 0.246)$ hours per day. Conditional on maternal wages, a €1 increase in the potential wages of fathers increases the gap between maternal and paternal labor supply by $0.607(= 0.157 + 0.450)$ hours per day.

These labor supply responses are reflected in the availability of monetary resources and their distribution within households. Conditional on paternal wages, a €1 increase in the potential wages of mothers leads to an increase of joint labor market earnings by €$6,427(=$ €$5,209 +$ €$1,218)$ per year, while a €1 increase in the potential wages of fathers increases the intra-family earnings gap between mothers and fathers by €$2,531(= 974 +$ €$1,557)$ per annum.

Panel (b) summarizes the differential effect of wage shocks to mothers and fathers on household's working hours and earnings. I follow the specification of equation (86) and control for the combined wage shock $\hat{w}^{\Sigma}_{ifat-1}$ in order to separate the effect of changes in the relative wages available to mothers and fathers from general shocks that affect the two partners simultaneously. As a consequence, the point estimates on the parental wage gap $\hat{w}^{\Delta}_{ifat-1}$ amount to half the difference between the effects of maternal wages and paternal wages estimated in Panel (a). Furthermore, the coefficients can be interpreted as an F-test of whether wage shocks

[24] However, formally testing the equality of coefficients on $\hat{w}^{m}_{ijat-1}$ and $\hat{w}^{p}_{ijat-1}$ for both maternal and paternal labor supply, I can rule out a symmetric response of maternal work hours to her own and her partner's wage shocks at a statistical significance level of below 1%. To the contrary, I cannot rule out a symmetric response for fathers at any conventional level of statistical significance.

incurred by mothers and fathers have the same impact on the outcome of interest.[25] There is a statistically significant difference in the effect of maternal and paternal wage shocks on the parental gap in labor supply – however, there is no such differential effect on the total labor supply of households. In terms of earnings, a €1 decrease in the parental gender wage gap increases household resources from labor market earnings by €$2,922 (= 1/2($€$6,427 +$€$583))$ per year and decreases intra-household inequality by €$3,261 (= 1/2($€$3,990 +$€$2,531))$ per year. Relative wage gains of mothers thus translate into an increase of monetary resources at the household level and a corresponding increase in the total amount of monetary resources controlled by mothers. Both shifts may have a positive effect on the child development as monetary resources are an important input factor for the production of skills (e.g. Akee et al., 2018; Løken et al., 2012) and women tend to devote a higher share of their monetary resources to their children (e.g. S. J. Lundberg et al., 1997).

3.5.2 Childcare Response

Table 3.5 displays how households adjust their childcare arrangements in response to changes in the relative wages of mothers and fathers.

Panel (a) again shows a clear asymmetry between mothers and fathers. In line with their decrease of daily labor hours, mothers increase their childcare provision by 0.549 hours per day in response to a €1 increase in the hourly potential wages of their partner. This effect translates into an increase of 0.669 hours/day that the child is cared for at home, whereas there is 5.6 percentage point decrease in the probability that the family uses any non-parental care providers on a regular basis. The latter effect is especially driven by a 4.7 percentage point decrease in the use of formal care providers.

In contrast, changes in the potential wage of mothers do not lead to adjustments in the time that mothers care for their children. At first glance this finding seems to be at odds with the strong own-wage elasticity of maternal labor supply (0.749 hours/day, see Table 3.4). However, the analysis of Hsin and Felfe (2014) suggests that working mothers in the US are successful in protecting their time with children – especially in those activities that are conducive to child development. In Appendix Figure C.1 I provide descriptive evidence based on German time use diaries that support this explanation. The figure compares the share of mothers

[25] To see this, note that I estimate $y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$ in Panel (a) and $y = \gamma_1(x_1 - x_2) + \gamma_2(x_1 + x_2) + \eta$ in Panel (b). Hence, $1/2(\beta_1 - \beta_2) = \gamma_1$ and $\gamma_1 = 0 \iff \beta_1 = \beta_2$.

Table 3.5 – Parental Wage Gaps and Childcare Responses

	Parental Childcare				Non-Parental Childcare		
	Mother	Father	Σ	Δ	Any	Formal	Informal
Panel (a): Wages by Parent							
Wage Mother	0.087	0.079	0.166	0.008	-0.025	-0.067	0.113**
	(0.326)	(0.302)	(0.390)	(0.493)	(0.056)	(0.056)	(0.051)
Wage Father	0.549***	0.121	0.669**	0.428**	-0.056**	-0.047**	-0.049
	(0.204)	(0.127)	(0.265)	(0.212)	(0.026)	(0.019)	(0.035)
Panel (b): Parental Wage Gap							
Parental Wage Gap			-0.252	-0.210	0.016	-0.010	0.081***
			(0.238)	(0.283)	(0.032)	(0.031)	(0.031)
Sibling $\times$ Age FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
N	6,070	6,070	6,070	6,070	4,298	4,298	4,298
DV Mean	6.497	1.989	8.486	4.508	0.650	0.579	0.264
DV SD	4.621	2.308	5.689	4.582	0.477	0.494	0.441

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. All coefficients are estimated on the core sample described in Table 3.2. All regressions in Panel (b) control for $\hat{w}^{\Sigma}_{ifat-1}$ – the aggregate labor demand shock for family f in year $t-1$. The coefficient on the parental wage gap can thus be interpreted as a test of coefficient equality across maternal wages ($\hat{w}^{m}_{ifat-1}$) and paternal wages ($\hat{w}^{p}_{ifat-1}$), see Panel (a). Parental childcare hours are measured in hours per day. Non-parental childcare is measured as a binary variable indicating whether parents use the respective care arrangement. Σ indicates the sum across parental outcomes. Δ indicates the difference between maternal and paternal outcomes. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The last two rows of the table list the mean and the standard deviation of the dependent variable that is displayed in the table header.

and fathers involved in particular activities at each time of the day across the survey waves 2001/02 and 2012/13. Over time, there is an increasing share of mothers who report to be in employment during the business hours of the day and a corresponding decrease in the share of mothers who report to have their child present during these hours. However, from 2001/02 to 2012/13 there also is an increase in the share of mothers who report to spend time with their child in the early morning, afternoon and evening hours. This suggests that mothers compensate their absence during the work day by increasing interactions before and after work.[26]

[26] Furthermore, at no point of the day is there a decrease in the share of the mothers who report childcare to be their primary activity. If anything, there is a slight increase in the hours devoted to “intensive” childcare during the morning and afternoon hours. Appendix Figure C.2 shows that these upward shifts are driven by increases in personal care activities in the morning and increases of play and sports activities in the afternoon.

Panel (b) translates these (non-)responses into the aggregate effect of the parental wage gap. In view of the attenuated response of households to changes in maternal wages, there is no statistically significant effect of changes in the parental wage gap on the intra-household provision of childcare. However, a 1 € decrease in the parental wage gap leads to an 8.1 percentage point increase in the reliance on informal care providers. This shift may have a negative effect on the development of the affected children as informal childcare arrangements tend to be of lower quality than maternal care provision (e.g. Datta Gupta and Simonsen, 2010).

3.5.3 Socio-emotional Skills of Children

The previous sections have shown that increases in the relative wages of mothers lead to i) an increase of household financial resources, ii) an increase in the share of financial resources controlled by mothers and iii) an increase in the child's exposure to informal care arrangements. Table 3.6 shows how these changes at the household level affect the socio-emotional development of children. As previously, I separate by maternal and paternal wages in Panel (a) before translating these effects into the aggregate impact of changes in the parental wage gap in Panel (b).

First, with the exception of a marginally significant negative effect on children's openness, increases in maternal wages do not have a statistically significant effect on changes in any of the Big Five personality traits. This null finding may be explained by the different margins of household adjustments and their countervailing effects on child development. On the one hand, mothers respond to increases in their potential wages by spending more time outside the home and tend to replace their time with informal care providers. This substitution may have detrimental effects on children since informal childcare providers are oftentimes of lower quality than either maternal or center-based childcare (Datta Gupta and Simonsen, 2010). On the other hand, they do not adjust the total amount of time they spend with their children. Furthermore, the total amount of monetary resources in the household increases. Thus, as in Agostinelli and Sorrenti (2018) and Nicoletti et al. (2020), the effects of household's adjustment towards changes in maternal labor market incentives are not aligned and therefore attenuate the aggregate affect towards zero.

Second, increases in paternal wages do not have a statistically significant effect on changes in any of the Big Five personality traits. Wage increases of fathers lead to an increased involvement of mothers as the primary caretaker by substituting away from formal childcare

Table 3.6 – The Effect of Parental Wage Gaps on the Socio-emotional Skills of Children

	Openness	Conscient-ousness	Extra-version	Agree-ableness	Neuroticism
Panel (a): Wages by Parent					
Wage Mother	-0.176* (0.103)	0.075 (0.121)	-0.033 (0.104)	-0.085 (0.094)	0.170 (0.140)
Wage Father	-0.021 (0.060)	0.021 (0.046)	-0.074 (0.061)	-0.007 (0.056)	0.022 (0.107)
Panel (b): Parental Wage Gap					
Parental Wage Gap	-0.078 (0.061)	0.027 (0.067)	0.020 (0.061)	-0.039 (0.057)	0.074 (0.092)
Sibling × Age FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
N	5,999	6,049	6,039	6,032	4,346
DV Mean	0.026	0.055	-0.022	0.002	-0.028
DV SD	0.954	0.955	0.988	0.977	0.973

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. All coefficients are estimated on the core sample described in Table 3.2. All regressions in Panel (b) control for $\hat{w}^{\Sigma}_{ifat-1}$ – the aggregate labor demand shock for family f in year $t-1$. The coefficient on the parental wage gap can thus be interpreted as a test of coefficient equality across maternal wages ($\hat{w}^{m}_{ifat-1}$) and paternal wages ($\hat{w}^{p}_{ifat-1}$), see Panel (a). Short descriptions for each Big Five personality trait are provided in Table C.17. The Big Five personality traits are measured using the questionnaire batteries displayed in Table C.18. Dimension-specific responses are added and standardized to have $\mathcal{N}=(0,1)$ for each age group. Significance Levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Standard errors (in parenthesis) are clustered at the family level. The last two rows of the table list the mean and the standard deviation of the dependent variable that is displayed in the table header.

providers. This substitution may have positive effects on children if the quality of maternal care dominates its alternatives. However, formal childcare in Germany tends to be of high quality (e.g. Felfe and Lalive, 2018) which may cushion the associated gains of children. Furthermore, the relative wage gains of fathers do not have a discernible effect on total household resources. Thus, changes in paternal wage incentives lead to small adjustments in the quality and quantity of resources devoted to children attenuating the aggregate effect towards zero.

In sum, I find no evidence that changes in the parental wage gap have an impact on the socio-emotional development of children. To assess the precision of these null effects, I benchmark my estimates against the effect sizes found in other studies. In particular, I restrict this comparison to the preferred estimates from other (quasi-)experimental studies that take any dimension of the Big Five inventory as the outcome of interest and reject the null hypothesis of a zero effect at a statistical significance level of 5% or lower. Figure 3.5 shows the results of this comparison.

FIGURE 3.5 – Assessment of Effect Precision by Comparison to Other Interventions

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This figure shows the point estimates from Table 3.6 as well as the associated confidence intervals in comparsion to effects sizes from interventions studied in the extant literature.

For the majority of comparisons, I can comfortably exclude at the conventional levels of statistical significance that a €1 change in the relative wages of mothers and fathers affects child personality at a magnitude comparable to the effects found in the benchmark interventions. For example, Akee et al. (2018) find that an unconditional cash transfer program worth $\$3,500$ per annum, decreased neuroticism in children of the Eastern Band of Cherokee Indians by 0.381 SD. The lower bound of the 99% confidence interval on a €1 decrease of the parental wage gap, yields an effect of 0.162 SD, i.e. less than half of the aforementioned effect. Note that both interventions are broadly comparable in terms of their effects on total household resources since I have shown previously that a €1 decrease in the intra-family gap of hourly wages wage leads to a €$2,922$ increase of annual family earnings (Table 3.4). Other interventions are harder to compare in terms of the nature of the treatment. For example, Alan et al. (2019) show that 12-week à 2 hours/week curriculum intervention increased conscientousness in Turkish high-school students by 0.345 SD.[27] For a €1 decrease in the intra-family gap of hourly wages, I can exclude effects on conscientousness that are larger than 0.199 SD at a statistical confidence level of 99%.

In general, these comparisons suggest that the absence of evidence for a link between the wage convergence of mothers and fathers and children's socio-emotional skill development is not an artifact of lacking precision. To the contrary, my estimates are precise enough to comfortably exclude effects sizes that have been found with respect to other interventions in the extant literature. The only effects that consistently fall within the confidence bands of my estimates are the birth order effects estimated by Black et al. (2018). However, while these birth order effects are very precisely estimated, they are rather small in magnitude. Therefore, they do not threaten the conclusion that changes in the relative wages of mothers and fathers have a negligible effect on the socio-emotional skill development of their children.

3.5.4 Robustness

For each of the outcomes discussed above I conduct three sets of robustness checks, the results of which are displayed in Tables C.3–C.11 of the Appendix. First, I re-estimate all models under alternative constructions of the shift-share instruments (Tables C.3–C.5). Second, I re-estimate all models using different specifications for the set of control variables X'_{ifat} (Tables C.6–C.8). Lastly, I re-estimate all models under alternating sample restrictions (Tables C.9–C.11).

Alternative Shift-Share Instruments. In the baseline, I impute daily wages above the social security contribution limit by wage draws from a truncated log-normal distribution (Gartner, 2005). My results do not change if leaving the censored wages unchanged or uniformly replacing them with 150% of the social security contribution cap – an imputation technique commonly employed for top coded incomes in the Current Population Survey (CPS) (Autor et al., 2008; Shenhav, 2020). They are also unaffected when replacing the MZ variable for

[27] To be precise Alan et al. (2019) refer to the concept of grit, which, however, is highly related to conscientousness (Duckworth et al., 2007).

working hours in a typical work week with a variable that refers to working hours in the week that precedes the MZ data collection.

Shenhav (2020) proposes to extend the shift-share instrument by an updating term that accounts for intra-industry shifts in the occupation structure over time. Including this updating term has no discernible effect on my results. In contrast, using the most basic approach to the calculation of shift-share wages, where sectors are defined by industry instead of industry-occupation cells, leads to sizable divergences in point estimates and a simultaneous trebling of standard errors. This decrease in precision is driven by a reduction of the sector cells from $576 (= 27 \times 14)$ to 14. Such a reduced sectoral partition is too coarse to yield meaningful predictions for the group-specific wage development in Germany.

Lastly, the results are also robust to specifying the parental wage gap in terms of differences of log wages. While this transformation changes the interpretation of the coefficients, the relationships by-and-large hold at their previously estimated levels of statistical significance.

Additional Controls. In the baseline, I only control for economic shocks that affect the wage development of both partners, $\hat{w}^{\Sigma}_{ifat-1}$. However, my results remain unaffected when expanding X'_{ifat} by measures for the sibling's birth rank, migration background, the number of kids in the household, and the sibling's gender. This observation gives credence to the assumption that the assignment of wage shocks is orthogonal to intra-family variation in sibling characteristics.

The baseline estimates furthermore assume i) that families do not sort selectively into CZ across the time span of the sibling comparison, and ii) that parents do not selectively acquire additional education across the time span of the sibling comparison. As points of departure both assumptions are plausible. First, there is little residential movement across CZs among German families. Second, I focus on families with at least two children and who therefore most likely have finished their educational biographies. Indeed, only 3.1% of my sample are affected by intra-sibling variation in the CZ of residence or the educational status of their parents. However, to test both assumptions formally, I include vectors of CZ fixed effects as well as maternal and paternal education fixed effects in the set of control variables. My results remain unaffected.

Lastly, since 1996 every German family with children aged 3–6 has a legal entitlement for a place in publicly subsidized childcare. By 2013 this right had been expanded to children aged one year and older. Both legal provisions have led to massive expansions of public childcare that were characterized by strong regional heterogeneity in the speed of expansion. My identification would be threatened if the intra-family variation in potential wages would correlate with intra-family changes in the availability of public childcare slots. To address this concern I expand my baseline specification by adding separate controls for the CZ- and year-specific share of children aged 0–3 and 3–6 that attend publicly subsidized childcare.[28] The number of observations reduces slightly due to the non-availability of administrative data on childcare slots in the years 2005 and 2006. The results, however, remain unchanged.

Alternative Sample Restrictions. The baseline estimates are derived from a sample of stable families where I allow for changes in the partner of mothers as long as this partner is constant for the time period of the sibling comparison. Focusing on biological parents only reduces the sample by 238 observations but does not alter the results. Similarly, my results remain unaffected when restricting the sample to married parents only.

My sample shrinks significantly by list-wise deleting entries without information on the child's Big Five personality traits. While this restriction is necessary for the investigation of socio-emotional skills, I can estimate the parental labor market response and the household's childcare response on a validation sample that has more than four time the size of my core data sample ($N = 28,380$). However, even in this expanded sample the results remain comparable to my baseline estimates.

3.5.5 Heterogeneity

The average effects presented thus far may mask i) heterogeneity in the way households react to changes in relative wage incentives, and ii) differences in the effects of these allocation decisions across children with different characteristics. For example, the stylized model presented in section 3.2 suggests that parental beliefs and norms may insulate investments into children from economic incentives (see also Ichino et al., 2020). Furthermore, it is well-

[28] Demand for public childcare strongly exceeds its supply. Actual enrollment therefore is a suitable proxy for the availability of childcare slots (Felfe and Lalive, 2018). See Figure C.3 for an overview map that displays the regional heterogeneity in the speed of childcare expansion.

documented that children have a differential sensitivity towards parental investments, for example depending on their age (Del Boca et al., 2017) and gender (Bertrand and Pan, 2013).

In the following, I study the existence of heterogeneous effects across child and parental characteristics by estimating the following model:

$$\begin{aligned} Y_{ifat} = \alpha + \beta \hat{w}^{\Delta}_{ifat-1} + \psi \hat{w}^{\Sigma}_{ifat-1} \\ + \beta^H (\hat{w}^{\Delta}_{ifat-1} \times I^H) + \psi^H (\hat{w}^{\Sigma}_{ifat-1} \times I^H) \\ + \gamma_{fa} + \tau_t + X'_{ifat}\delta + \epsilon_{ifat}, \end{aligned} \tag{88}$$

where I^H indicates a binary indicator variable in heterogeneity dimension H.

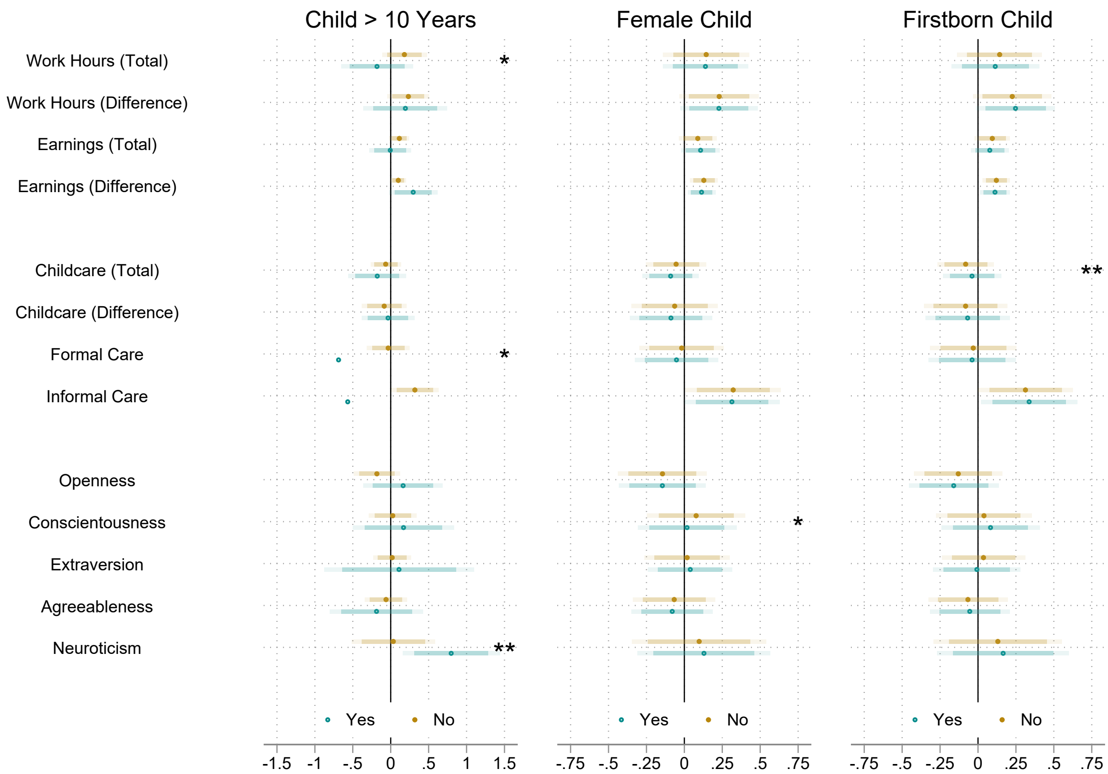
Figures 3.6 and 3.7 summarize the results of this heterogeneity analysis. In particular, for each outcome I plot the marginal effects of increases in $\hat{w}^{\Delta}_{ifat-1}$ as well as the corresponding confidence bands by group characteristic. These marginal effects indicate whether increases in $\hat{w}^{\Delta}_{ifat-1}$ yield a group-specific effect that is statistically different from zero. Furthermore, I add significance stars for the parameter β^H to indicate whether effects across groups are statistically significant from each other. To facilitate the graphical representation I standardize all outcome and wage shock variables to have mean zero and standard deviation one.

Child Characteristics. Figure 3.6 shows heterogeneous effects of maternal and paternal wage shocks by child age ($\lesseqgtr 10$ years), sex and birth order.

Kleven et al. (2019) show that women stabilize their labor force attachment at older child ages. Consistent with this observation, the effect of changes in the parental wage gap on total household labor supply is less pronounced if children are ten years of age and older. In the German school system age ten marks the transition from primary to secondary school. From this age on, there is no widely available formal childcare option and it is plausible that informal childcare arrangements decrease in importance as children grow into adolescents. Consistent with this fact, the previously detected increase in the use of informal child care arrangements is exclusively driven by children aged ten years and younger.[29]

[29] Since the use of childcare above age ten is infrequent, the respective coefficients are noisily estimated and I omit the corresponding confidence bands from the graphical representation to increase its visual clarity.

FIGURE 3.6 – Effect Heterogeneity across Child Characteristics



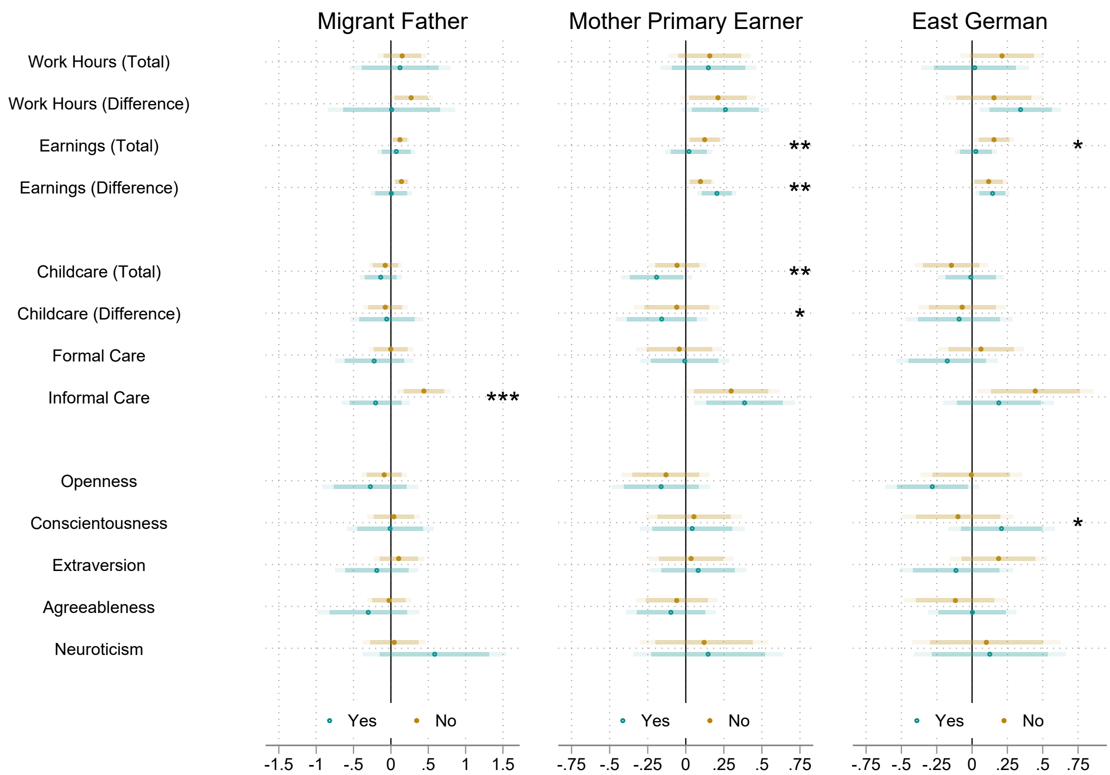
Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This figure shows heterogeneous effects of the parental gap in potential wages ($\hat{w}^{\Delta}_{ifat-1}$) across a selected set of child characteristics. Each data point shows the marginal effect of $\hat{w}^{\Delta}_{ifat-1}$ estimated from equation (88) for the binary characteristic indicated in the subfigure header. The dark and light shaded bars indicate the 95% and 99% confidence interval, respectively. In the leftmost panel, confidence intervals on *Formal Care* and *Informal Care* are omitted for better visualization. Stars indicate the statistical significance level of the interaction coefficient β^H from equation (88). All outcome variables as well as $\hat{w}^{\Delta}_{ifat-1}$ and $\hat{w}^{\Sigma}_{ifat-1}$ are standardized to have $\mathcal{N} = (0,1)$. All coefficients are estimated on the core sample described in Table 3.2. Standard errors are clustered at the family level. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Apart from these child age effects, however, household's adjustments to changes in their relative labor market incentives do not vary strongly with the characteristics of their child. In particular, parental responses are by-and-large consistent regardless of whether the child is male or female and whether the child is the firstborn or a higher-order sibling.

Similarly, there is little heterogeneity in the way parental wage shocks affect the socio-emotional skill development of their children. Decreases in the parental wage gap lead to a slightly stronger increase in conscientousness if the child is male. The marginal effect, however, remains indistinguishable from zero. Furthermore, there is a stronger increase in neuroticism if the child is ten years and older. However, this is the only subgroup for which I detect a non-zero marginal effect of decreasing parental wage gaps on children's socio-emotional skills. Otherwise the null effect of decreases in the parental wage gap on children's personality persists for all Big Five dimensions across all three child characteristics.

Parental Characteristics. A more diverse picture emerges for differences in parental characteristics. Figure 3.7 shows heterogeneous effects of maternal and paternal wage shocks by paternal migration background, by whether the mother was the household's primary earner in year $t-1$, or whether the family resides in the Eastern part of Germany. Each of these characteristics may be interpreted as a proxy variable for gender identity norms.

FIGURE 3.7 – Effect Heterogeneity across Parental Characteristics



Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This figure shows heterogeneous effects of the parental gap in potential wages ($\hat{w}^{\Delta}_{ifat-1}$) across a selected set of parental characteristics. Each data point shows the marginal effect of $\hat{w}^{\Delta}_{ifat-1}$ estimated from equation (88) for the binary characteristic indicated in the subfigure header. The dark and light shaded bars indicate the 95% and 99% confidence interval, respectively. Stars indicate the statistical significance level of the interaction coefficient β^H from equation (88). All outcome variables as well as $\hat{w}^{\Delta}_{ifat-1}$ and $\hat{w}^{\Sigma}_{ifat-1}$ are standardized to have $\mathcal{N}=(0,1)$. All coefficients are estimated on the core sample described in Table 3.2. Standard errors are clustered at the family level. Significance Levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Consistent with this interpretation, we observe that parental differences in labor hours and earnings react less to decreasing parental wage gaps if the father has a migration background. As a consequence, these families need to rely less on the use of informal care arrangements in response to such shocks.

In households in which the mother represents the primary earner, gender identity norms may be less binding. Consistent with this hypothesis, these households seem to react stronger in line with economic incentives: A decrease in the parental wage gap leads to a stronger decrease in the parental earnings difference, and a stronger decrease in both total care provision and the gender difference in parental care.

The regional patterns of gender gaps and gender norms displayed in Figures 3.1 and 3.2 suggest that Eastern and Western German families react differently to gendered changes in labor market incentives. Indeed, for Eastern German families decreases in the parental wage gap lead to a statistically significant decrease in the parental difference of hours worked. This is not the case for Western German families. In contrast, Western German families respond to decreases in the parental wage gap by a stronger increase in total hours of work. This suggests, that these households are characterized by a more positive paternal labor supply response to the wage increases of mothers. Furthermore, we observe that the increasing use of informal care arrangements is more strongly driven by Western German families which reflects the wider availability of formal childcare in the Eastern part of Germany.

In general, these results are consistent with Ichino et al. (2020) who show that Swedish couples react less strongly to changes in the net-of-tax wage rate if they belong to a group that adheres to more traditional gender norms. However, in spite of the differential responses of these households to their relative labor market incentives, there is little heterogeneity in the way parental wage shocks affect the socio-emotional development of children. The slight decrease in openness observed in Table 3.6 is driven by Eastern German children. Furthermore, decreases in the parental wage gap lead to a slightly stronger increase in conscientousness if the child is from an Eastern German family. The marginal effect, however, remains indistinguishable from zero.

3.6 Conclusion

In this paper I study the effect of converging parental wages on the socio-emotional development of their children. Thereby, I connect the literature branches on intra-household decision-making and child development. While the former has extensively studied household responses to changes in the gender wage gap (e.g. Eckstein et al., 2019; Knowles, 2012), the latter has focused on parental inputs and their effect on child development (e.g. Agostinelli and Sorrenti, 2018; Nicoletti et al., 2020).

I find that relative wage gains of mothers *increase* i) household's total financial resources, ii) the share of financial resources controlled by mothers, and iii) the use of informal care providers. To the contrary, I find *no effects* on i) the total hours of care provided by parents and ii) the share of parental care provided by mothers or fathers. Drawing on time-use data, I provide suggestive evidence that the latter effects are explained by mothers that compensate children for their increased absence during the business hours with increased attention in the morning and the afternoon after they return from work. In sum, I find no effects of converging parental wages on the socio-emotional skill development of their children as measured by the Big Five inventory. These null effects are estimated precisely enough to comfortably exclude the effect sizes of various interventions analyzed in the existing literature at the conventional levels of statistical significance.

Fostering gender equality and promoting the development of children are both prominent goals of family policy that are oftentimes thought to be in conflict with each other. The evidence presented in this paper suggests that increasing gender equality in the labor market does not have to come at the cost of child development. Yet, there are a number of qualifications that should be borne in mind. First, Germany provides childcare institutions that are of relatively high quality. Similar investigations in country contexts in which there is a larger quality gap between maternal care and its alternatives may lead to different conclusions. Second, mothers increase their labor market participation while maintaining their time investments into children. Such a "second shift" (Hochschild and Machung, 1990) of unpaid work may impose additional strain on mothers. Thus, resolving the trade-off between gender equality in the labor market and child development may actually come at the cost of adverse affects on maternal mental and physical health. An in-depth investigation such effects, however, is left for future research.

Appendix C.1 Rotemberg Weights

Table C.1 – Top 10 Rotemberg Weights, Mothers

	Rotemberg Weights		Coefficient	
Occupation/Industry	α_{io}	Share in %	β_{io}	95% CI
Teachers & Social Care Workers in Education	0.41	30.96%	2.52	[-1.00,6.00]
Sales Occ. in Wholesale and Retail	0.08	6.00%	7.83	[3.00,15.00]
Facility Management in Human Health Services	0.06	4.70%	4.99	[2.00,8.00]
Financial Services in Finance and Insurance	0.06	4.36%	-6.35	[-24.00,6.00]
Facility Management in Information, Communication, Business Services	0.05	4.08%	3.92	[0.00,7.00]
Facility Management in Public Administration	0.05	3.46%	5.97	[2.50,9.50]
Facility Management in Education	0.03	2.62%	7.07	[4.00,10.50]
Textile & Leather Processing in Manufacturing: Food/Textiles/Other	0.03	2.23%	7.16	[4.00,12.50]
Sales Occ. in Manufacturing: Food/Textiles/Other	0.03	1.98%	7.75	[3.00,17.00]
Logistics Occ. in Wholesale and Retail	0.02	1.88%	5.84	[3.00,9.00]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This table shows the 10 industry-occupation cells with the highest Rotemberg weights for mothers. The Rotemberg weights, α_{io}, are calculated on the core sample described in Table 3.2 using the programming routine provided by Goldsmith-Pinkham et al. (2020). The share of each Rotemberg weight is calculated by dividing α_{io} with $\sum_i \sum_o [\alpha_{io} | \alpha_{io} \geq 0]$. β_{io} reflects the coefficient on $\hat{w}^m_{ifat-1}$ from a just-identified 2SLS regression of maternal labor income on $\hat{w}^m_{ifat-1}$ while controlling for sibling times child age fixed effects γ_{fa} and year fixed effects τ_t. $\hat{w}^m_{ifat-1}$ is instrumented with the group-specficic sector share in base year 1995 ($E^{oj}_{ger,1995}/E_{ger,1995}$). The associated confidence interval is the weak instrument robust confidence interval based on the method of Chernozhukov and Hansen (2008) over the range $-30 - 30$.

Table C.2 – Top 10 Rotemberg Weights, Fathers

	Rotemberg Weights		Coefficient	
Occupation/Industry	α_{io}	Share in %	β_{io}	95% CI
Teachers & Social Care Workers in Education	0.12	9.53%	2.18	[-2.00,6.50]
Building Construction in Construction	0.09	6.72%	2.93	[-2.50,9.00]
Engineering Occ. in Manufacturing: Electronics/Vehicles/Machinery	0.08	6.64%	-2.93	[-14.00,7.00]
Logistics Occ. in Transportation and Storage	0.06	4.32%	3.59	[-1.00,9.00]
Business Adminsistration in Manufacturing: Electronics/Vehicles/Machinery	0.05	4.25%	-0.12	[-5.50,5.00]
Logistics Occ. in Wholesale and Retail	0.05	4.07%	3.05	[-1.50,8.00]
Building Services in Construction	0.05	4.02%	2.48	[-3.50,9.00]
Purchasing & Trading in Manufacturing: Electronics/Vehicles/Machinery	0.05	3.76%	2.71	[-6.50,12.50]
Financial Services in Finance and Insurance	0.05	3.69%	-3.57	[-30.00,16.50]
Interior Construction in Construction	0.04	3.01%	1.58	[-3.50,7.00]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This table shows the 10 industry-occupation cells with the highest Rotemberg weights for fathers. The Rotemberg weights, α_{io}, are calculated on the core sample described in Table 3.2 using the programming routine provided by Goldsmith-Pinkham et al. (2020). The share of each Rotemberg weight is calculated by dividing α_{io} with $\sum_i \sum_o [\alpha_{io} | \alpha_{io} \geq 0]$. β_{io} reflects the coefficient on $\hat{w}^p_{ifat-1}$ from a just-identified 2SLS regression of maternal labor income on $\hat{w}^p_{ifat-1}$ while controlling for sibling times child age fixed effects γ_{fa} and year fixed effects τ_t. $\hat{w}^p_{ifat-1}$ is instrumented with the group-specficic sector share in base year 1995 ($E^{oj}_{ger,1995}/E_{ger,1995}$). The associated confidence interval is the weak instrument robust confidence interval based on the method of Chernozhukov and Hansen (2008) over the range $-30 - 30$.

Appendix C.2 Robustness

C.2.1 Alternative Labor Demand Shocks

Table C.3 – Robustness Checks Labor Market Response: Alternative Labor Demand Shocks

	Labor Hours (Σ)	Labor Hours (Δ)	Earnings (Σ)	Earnings (Δ)
Panel (a): Baseline Effect				
Parental Wage Gap	0.351 (0.269) [6,070]	0.555** (0.242) [6,070]	2.922** (1.366) [6,070]	3.261*** (0.953) [6,070]
Panel (b): Robustness Checks				
Censored Wages (SIAB)	0.418 (0.318) [6,070]	0.759** (0.303) [6,070]	4.211** (1.690) [6,070]	4.045*** (1.226) [6,070]
CPS Imputation (SIAB)	0.369 (0.273) [6,070]	0.595** (0.248) [6,070]	3.082** (1.386) [6,070]	3.457*** (0.969) [6,070]
Hours Last Week (MZ)	0.323 (0.232) [6,070]	0.493** (0.211) [6,070]	2.423** (1.186) [6,070]	3.043*** (0.840) [6,070]
Updating Shenhav (2020)	0.340 (0.269) [6,070]	0.562** (0.242) [6,070]	2.950** (1.358) [6,070]	3.387*** (0.948) [6,070]
No Occupation	-0.516 (0.730) [6,070]	1.503** (0.619) [6,070]	2.688 (3.805) [6,070]	7.467** (3.254) [6,070]
Log Parental Wage Gap	4.816 (3.870) [6,070]	8.044** (3.482) [6,070]	31.167 (19.203) [6,070]	44.130*** (13.233) [6,070]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Work hours are measured in hours per day. Earnings are measured in thousand € per year. Σ indicates the sum across parental outcomes. Δ indicates the difference between maternal and paternal outcomes. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

Table C.4 – Robustness Checks Childcare Response: Alternative Labor Demand Shocks

	Childcare Hours (Σ)	Childcare Hours (Δ)	Non-parental Care	Formal Care	Informal Care
Panel (a): Baseline Effect					
Parental Wage Gap	-0.252 (0.238) [6,070]	-0.210 (0.283) [6,070]	0.016 (0.032) [4,298]	-0.010 (0.031) [4,298]	0.081*** (0.031) [4,298]
Panel (b): Robustness Checks					
Censored Wages (SIAB)	-0.327 (0.314) [6,070]	-0.259 (0.358) [6,070]	0.029 (0.041) [4,298]	-0.005 (0.038) [4,298]	0.098** (0.039) [4,298]
CPS Imputation (SIAB)	-0.273 (0.246) [6,070]	-0.220 (0.289) [6,070]	0.017 (0.033) [4,298]	-0.009 (0.031) [4,298]	0.080** (0.032) [4,298]
Hours Last Week (MZ)	-0.225 (0.202) [6,070]	-0.134 (0.238) [6,070]	0.017 (0.028) [4,298]	-0.005 (0.027) [4,298]	0.065** (0.027) [4,298]
Updating Shenhav (2020)	-0.258 (0.227) [6,070]	-0.193 (0.270) [6,070]	0.014 (0.032) [4,298]	-0.012 (0.030) [4,298]	0.081*** (0.031) [4,298]
No Occupation	-1.352* (0.810) [6,070]	-0.732 (0.787) [6,070]	-0.061 (0.107) [4,298]	-0.111 (0.093) [4,298]	0.134 (0.090) [4,298]
Log Parental Wage Gap	-3.815 (3.533) [6,070]	-2.713 (4.169) [6,070]	0.195 (0.471) [4,298]	-0.131 (0.455) [4,298]	1.128** (0.442) [4,298]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Parental childcare hours are measured in hours per day. Non-parental childcare is measured as a binary variable indicating whether parents use the respective care arrangement. Σ indicates the sum across parental outcomes. Δ indicates the difference between maternal and paternal outcomes. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

Table C.5 – Robustness Checks Socio-emotional Skill Development: Alternative Labor Demand Shocks

	Openness	Conscient-ousness	Extra-version	Agree-ableness	Neuroticism
Panel (a): Baseline Effect					
Parental Wage Gap	-0.078 (0.061) [5,999]	0.027 (0.067) [6,049]	0.020 (0.061) [6,039]	-0.039 (0.057) [6,032]	0.074 (0.092) [4,346]
Panel (b): Robustness Checks					
Censored Wages (SIAB)	-0.090 (0.078) [5,999]	0.056 (0.082) [6,049]	0.051 (0.080) [6,039]	-0.040 (0.074) [6,032]	0.086 (0.114) [4,346]
CPS Imputation (SIAB)	-0.076 (0.063) [5,999]	0.028 (0.068) [6,049]	0.032 (0.063) [6,039]	-0.035 (0.059) [6,032]	0.071 (0.094) [4,346]
Hours Last Week (MZ)	-0.046 (0.056) [5,999]	0.032 (0.057) [6,049]	0.029 (0.054) [6,039]	-0.025 (0.050) [6,032]	0.050 (0.080) [4,346]
Updating Shenhav (2020)	-0.077 (0.061) [5,999]	0.027 (0.067) [6,049]	0.015 (0.061) [6,039]	-0.043 (0.057) [6,032]	0.079 (0.091) [4,346]
No Occupation	-0.003 (0.192) [5,999]	0.223 (0.195) [6,049]	-0.104 (0.194) [6,039]	0.109 (0.162) [6,032]	0.162 (0.264) [4,346]
Log Parental Wage Gap	-1.259 (0.840) [5,999]	0.315 (0.944) [6,049]	-0.021 (0.840) [6,039]	-0.400 (0.780) [6,032]	1.295 (1.264) [4,346]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Short descriptions for each Big Five personality trait are provided in Table C.17. The Big Five personality traits are measured using the questionnaire batteries displayed in Table C.18. Dimension-specific responses are added and standardized to have $\mathcal{N} = (0, 1)$ for each age group. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

C.2.2 Alternative Control Variables

Table C.6 – Robustness Checks Labor Market Response: Additional Controls

	Labor Hours (Σ)	Labor Hours (Δ)	Earnings (Σ)	Earnings (Δ)
Panel (a): Baseline Effect				
Parental Wage Gap	0.351 (0.269) [6,070]	0.555** (0.242) [6,070]	2.922** (1.366) [6,070]	3.261*** (0.953) [6,070]
Panel (b): Robustness Checks				
Additional Child Controls	0.300 (0.276) [6,070]	0.559** (0.239) [6,070]	2.784** (1.390) [6,070]	3.017*** (0.916) [6,070]
CZ & Parental Education FE	0.080 (0.337) [6,070]	0.821** (0.323) [6,070]	3.675** (1.529) [6,070]	3.938*** (1.371) [6,070]
Childcare Availability	0.392 (0.268) [5,747]	0.602** (0.250) [5,747]	2.913** (1.387) [5,747]	3.538*** (0.967) [5,747]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Work hours are measured in hours per day. Earnings are measured in thousand € per year. Σ indicates the sum across parental outcomes. Δ indicates the difference between maternal and paternal outcomes. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

Table C.7 – Robustness Checks Childcare Response: Additional Controls

	Childcare Hours (Σ)	Childcare Hours (Δ)	Non-parental Care	Formal Care	Informal Care
Panel (a): Baseline Effect					
Parental Wage Gap	-0.252 (0.238) [6,070]	-0.210 (0.283) [6,070]	0.016 (0.032) [4,298]	-0.010 (0.031) [4,298]	0.081*** (0.031) [4,298]
Panel (b): Robustness Checks					
Additional Child Controls	-0.203 (0.208) [6,070]	-0.197 (0.273) [6,070]	0.014 (0.032) [4,297]	-0.015 (0.030) [4,297]	0.081*** (0.031) [4,297]
CZ & Parental Education FE	-0.023 (0.411) [6,070]	-0.240 (0.357) [6,070]	-0.007 (0.031) [4,298]	-0.043 (0.033) [4,298]	0.088*** (0.032) [4,298]
Childcare Availability	-0.240 (0.248) [5,747]	-0.127 (0.289) [5,747]	0.008 (0.033) [4,159]	-0.016 (0.031) [4,159]	0.080** (0.032) [4,159]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Parental childcare hours are measured in hours per day. Non-parental childcare is measured as a binary variable indicating whether parents use the respective care arrangement. Σ indicates the sum across parental outcomes. Δ indicates the difference between maternal and paternal outcomes. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

Table C.8 – Robustness Checks Socio-emotional Skill Development: Additional Controls

	Openness	Conscient-ousness	Extra-version	Agree-ableness	Neuroticism
Panel (a): Baseline Effect					
Parental Wage Gap	-0.078 (0.061) [5,999]	0.027 (0.067) [6,049]	0.020 (0.061) [6,039]	-0.039 (0.057) [6,032]	0.074 (0.092) [4,346]
Panel (b): Robustness Checks					
Additional Child Controls	-0.076 (0.060) [5,999]	0.034 (0.066) [6,049]	0.026 (0.061) [6,039]	-0.024 (0.057) [6,032]	0.070 (0.095) [4,344]
CZ & Parental Education FE	-0.077 (0.077) [5,999]	0.082 (0.089) [6,049]	-0.037 (0.088) [6,039]	-0.039 (0.079) [6,032]	0.115 (0.095) [4,346]
Childcare Availability	-0.093 (0.064) [5,680]	0.024 (0.070) [5,726]	0.040 (0.062) [5,716]	-0.032 (0.058) [5,709]	0.071 (0.095) [4,233]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Short descriptions for each Big Five personality trait are provided in Table C.17. The Big Five personality traits are measured using the questionnaire batteries displayed in Table C.18. Dimension-specific responses are added and standardized to have $\mathcal{N} = (0,1)$ for each age group. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

C.2.3 Alternative Sample Restrictions

Table C.9 – Robustness Checks Labor Market Response: Alternative Sample Restrictions

	Labor Hours (Σ)	Labor Hours (Δ)	Earnings (Σ)	Earnings (Δ)
Panel (a): Baseline Effect				
Parental Wage Gap	0.351 (0.269) [6,070]	0.555** (0.242) [6,070]	2.922** (1.366) [6,070]	3.261*** (0.953) [6,070]
Panel (b): Robustness Checks				
Biological Parents Only	0.365 (0.271) [5,832]	0.573** (0.242) [5,832]	2.962** (1.378) [5,832]	3.188*** (0.958) [5,832]
Married Parents Only	0.359 (0.299) [5,622]	0.487* (0.267) [5,622]	2.935* (1.507) [5,622]	3.300*** (1.019) [5,622]
Validation Sample	0.393* (0.214) [28,380]	0.170 (0.240) [28,380]	1.806** (0.752) [28,380]	2.199*** (0.800) [28,380]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Work hours are measured in hours per day. Earnings are measured in thousand € per year. Σ indicates the sum across parental outcomes. Δ indicates the difference between maternal and paternal outcomes. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

Table C.10 – Robustness Checks Childcare Response: Alternative Sample Restrictions

	Childcare Hours (Σ)	Childcare Hours (Δ)	Non-parental Care	Formal Care	Informal Care
Panel (a): Baseline Effect					
Parental Wage Gap	-0.252 (0.238) [6,070]	-0.210 (0.283) [6,070]	0.016 (0.032) [4,298]	-0.010 (0.031) [4,298]	0.081*** (0.031) [4,298]
Panel (b): Robustness Checks					
Biological Parents Only	-0.231 (0.237) [5,832]	-0.163 (0.284) [5,832]	0.013 (0.032) [4,190]	-0.012 (0.031) [4,190]	0.081*** (0.031) [4,190]
Married Parents Only	-0.213 (0.258) [5,622]	-0.157 (0.303) [5,622]	0.035 (0.036) [3,966]	0.008 (0.035) [3,966]	0.077** (0.033) [3,966]
Validation Sample	0.093 (0.170) [28,380]	0.196 (0.264) [28,380]	0.032 (0.021) [24,238]	0.023 (0.021) [24,238]	0.042** (0.017) [24,238]

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Parental childcare hours are measured in hours per day. Non-parental childcare is measured as a binary variable indicating whether parents use the respective care arrangement. Σ indicates the sum across parental outcomes. Δ indicates the difference between maternal and paternal outcomes. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

Table C.11 – Robustness Checks Socio-emotional Skill Development: Alternative Sample Restrictions

	Openness	Conscient-ousness	Extra-version	Agree-ableness	Neuroticism
Panel (a): Baseline Effect					
Parental Wage Gap	-0.078 (0.061) [5,999]	0.027 (0.067) [6,049]	0.020 (0.061) [6,039]	-0.039 (0.057) [6,032]	0.074 (0.092) [4,346]
Panel (b): Robustness Checks					
Biological Parents Only	-0.087 (0.061) [5,767]	0.022 (0.067) [5,814]	0.022 (0.061) [5,804]	-0.047 (0.058) [5,799]	0.078 (0.092) [4,121]
Married Parents Only	-0.045 (0.065) [5,555]	0.039 (0.070) [5,606]	0.047 (0.067) [5,593]	-0.017 (0.059) [5,589]	0.067 (0.097) [4,099]
Validation Sample	—	—	—	—	—

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. Short descriptions for each Big Five personality trait are provided in Table C.17. The Big Five personality traits are measured using the questionnaire batteries displayed in Table C.18. Dimension-specific responses are added and standardized to have $\mathcal{N} = (0, 1)$ for each age group. Significance Levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parenthesis) are clustered at the family level. The number of observations is indicated in brackets.

Appendix C.3 Additional Tables

Table C.12 – Comparison GSOEP and GTUS, Work and Childcare Hours per Day in 2001/02 and 2012/13

	GSOEP		GTUS	
	2001/02	2012/13	2001/02	2012/13
Mother				
Work Hours/Day	3.1	3.1	3.0	3.7
Childcare Hours/Day	5.8	5.6	5.0	5.6
Intensive Childcare Hours/Day	.	.	2.4	2.7
Father				
Work Hours/Day	7.9	6.7	7.3	7.3
Childcare Hours/Day	1.5	1.8	2.2	2.6
Intensive Childcare Hours/Day	.	.	0.9	1.1

Data: German Socio-economic Panel (GSOEP), German Time-Use Study (GTUS).
Note: Own calculations. This table compares working time and childcare time variables across the GSOEP and the GTUS. The samples include two-parent households aged 18–63 with at least one resident child aged 2–17. Work hours and childcare hours are measured in hours per day. The analysis is based on week days Monday through Friday only. *Childcare Hours/Day* in the GTUS capture any activity with the child present. *Intensive Childcare Hours/Day* capture any time when respondents refer to childcare as their primary activity.

Table C.13 – Industry Employment Shares by Education and Sex, 1995

	Male			Female		
	Low	Inter–mediate	High	Low	Inter–mediate	High
Agriculture/Mining/Utilities	6.1	4.6	3.2	1.5	1.7	1.5
Manufacturing: Food/Textiles/Other	11.0	8.4	4.5	12.9	7.3	3.3
Manufacturing: Raw Materials/Metals/Chemicals	19.1	11.5	7.8	9.0	3.6	3.5
Manufacturing: Electronics/Vehicles/Machinery	12.0	12.5	13.7	10.5	4.1	3.3
Construction	13.5	19.1	6.4	1.4	2.9	2.3
Wholesale and Retail	9.3	13.7	10.1	12.3	20.8	11.7
Transportation and Storage	6.6	7.3	3.2	2.0	3.7	2.1
Accommodation and Food Services	4.7	1.9	1.0	6.6	3.6	1.5
Information, Communication, Business Services	8.4	8.4	20.3	11.6	10.8	17.8
Finance and Insurance	0.6	2.4	6.1	2.7	4.4	7.0
Public Administration	4.3	4.8	6.1	8.2	9.9	10.3
Education	0.6	0.9	6.2	3.5	3.9	12.2
Human Health Services	1.7	2.5	7.1	13.1	17.8	17.7
Other	2.1	2.0	4.3	4.7	5.5	5.7

Data: Sample of Integrated Labour Market Biographies (SIAB).
Note: Own calculations. This table shows the employment share of each industry among employees aged 18–63 in 1995 by sex and education. Education is classified as follows – lower secondary degree without tertiary education (*Low*), lower secondary degree with vocational training or higher secondary degree without vocational training (*Intermediate*), university qualification (*High*).

Table C.14 – Occupation Employment Shares by Education and Sex, 1995

	Male			Female		
	Low	Inter–mediate	High	Low	Inter–mediate	High
Raw Material & Plastic Processing	7.7	2.6	0.4	3.4	0.6	0.1
Metal Processing	13.9	8.1	1.4	4.4	0.7	0.2
Machine-Building Occ.	3.7	7.4	6.2	2.5	0.6	0.4
Engineering Occ.	5.0	14.2	17.2	6.1	3.0	3.6
Food Processing	5.1	3.0	0.5	9.2	2.6	0.4
Construction Planning	0.1	0.6	5.9	0.0	0.1	1.6
Building Construction	13.0	10.0	1.1	0.4	0.4	0.1
Logistics Occ.	22.4	14.9	3.7	9.0	4.0	1.2
Facility Management	5.4	2.9	0.8	24.1	4.3	0.7
Sales Occ.	1.4	2.5	1.5	6.4	13.8	2.5
Business Adminsistration	2.2	8.5	22.4	10.3	33.4	34.8
Financial Services	0.2	1.9	5.0	1.2	3.2	5.9
Doctors Assistants	0.0	0.0	0.2	0.6	4.3	1.9
Nursing Occ	0.6	1.7	3.3	4.2	14.3	13.7
Medical Care Occ.	0.0	0.2	3.9	0.1	1.6	8.0
Teachers & Social Care Workers	0.2	0.5	7.1	4.6	2.2	12.3

Data: Sample of Integrated Labour Market Biographies (SIAB).
Note: Own calculations. This table shows the employment share of each occupation among employees aged 18–63 in 1995 by sex and education. Education is classified as follows – lower secondary degree without tertiary education (*Low*), lower secondary degree with vocational training or higher secondary degree without vocational training (*Intermediate*), university qualification (*High*).

TABLE C.15 – Comparison SIAB and MZ, Socio-demographics by Year

	1995		2005		2015	
	SIAB	MZ	SIAB	MZ	SIAB	MZ
Age, Average in Employed Population						
Age	38.4	38.4	40.3	39.9	41.9	41.9
Female, Employment Share in %						
Male	57.4	55.1	55.4	52.5	53.7	53.4
Female	42.6	44.9	44.6	47.5	46.3	46.6
Education, Employment Share in %						
Low	10.9	13.1	8.0	12.7	6.4	9.7
Intermediate	72.8	67.4	68.1	62.2	60.2	58.4
High	16.3	19.5	23.9	25.0	33.4	31.8
Federal State, Employment Share in %						
Schleswig-Holstein	2.9	3.3	3.0	3.5	3.0	2.9
Saarland	1.3	1.1	1.3	1.1	1.2	1.1
Berlin	4.7	4.3	4.0	3.8	4.3	3.8
Brandenburg	3.3	3.5	2.7	3.3	2.6	3.2
Mecklenburg-Vorpommern	2.3	2.5	1.9	2.0	1.8	1.8
Sachsen	6.1	6.2	5.2	5.7	5.1	5.1
Sachsen-Anhalt	3.6	3.7	2.9	3.3	2.6	2.9
Thüringen	3.2	3.6	2.7	3.1	2.7	2.9
Hamburg	2.7	2.0	2.9	2.1	3.0	1.8
Niedersachsen	8.0	8.4	8.3	7.8	8.6	10.1
Bremen	1.3	0.8	1.2	0.7	1.3	0.7
Nordrhein-Westfalen	20.5	19.9	21.2	19.6	20.6	19.2
Hessen	7.5	7.1	8.0	7.8	7.8	7.8
Rheinland-Pfalz	4.1	4.9	4.3	4.9	4.3	4.7
Baden-Württemberg	13.2	13.1	14.2	14.0	14.1	14.0
Bayern	15.1	15.6	16.3	17.2	17.0	17.7

Data: Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This table shows the socio-demographic composition of the SIAB and the MZ in the years 1995, 2005 and 2015. All statistics are calculated on the sample of employees aged 18–63. The MZ is restricted to match the sample characteristics of the SIAB by excluding the marginally employed (<10h/week), civil servants, and self-employed individuals. Education is classified as follows – lower secondary degree without tertiary education (*Low*), lower secondary degree with vocational training or higher secondary degree without vocational training (*Intermediate*), university qualification (*High*).

Table C.16 – Comparison SIAB and MZ, Employment Structure by Year

	1995		2005		2015	
	SIAB	MZ	SIAB	MZ	SIAB	MZ
Occupation: Employment Share in %						
Agriculture/Mining/Utilities	3.3	5.0	2.7	3.9	2.3	2.9
Finance and Insurance	3.6	4.1	3.7	4.2	3.1	3.6
Public Administration	7.0	6.9	6.0	5.8	5.2	5.5
Education	3.2	3.9	3.5	4.3	3.7	4.5
Human Health Services	9.1	9.4	11.6	12.3	13.2	11.4
Other	3.7	4.7	4.0	5.0	3.8	3.9
Manufacturing: Food/Textiles/Other	7.7	9.1	6.2	6.8	5.2	6.1
Manufacturing: Raw Materials/Metals/Chemicals	8.5	9.2	7.6	7.6	6.6	6.6
Manufacturing: Electronics/Vehicles/Machinery	9.3	8.9	9.4	10.2	8.5	11.1
Construction	10.6	10.2	6.4	6.6	5.5	6.9
Wholesale and Retail	15.1	14.6	14.8	14.7	14.1	15.5
Transportation and Storage	5.1	4.5	5.4	4.7	5.4	5.4
Accommodation and Food Services	2.7	2.3	3.0	3.0	3.4	3.4
Information, Communication, Business Services	11.1	7.2	15.7	10.7	20.0	13.3
Industry: Employment Share in %						
Agriculture/Forestry/Farming/Gardening	2.1	2.3	1.6	1.8	1.4	1.7
Construction Planning	1.0	0.7	0.8	0.7	0.8	0.9
Building Construction	5.2	3.4	2.8	2.0	2.3	2.3
Interior Construction	1.9	1.7	1.3	1.2	1.2	1.3
Building Services	1.5	2.1	1.1	1.5	1.7	1.7
Natural Science Occ.	1.8	1.6	1.7	1.5	1.4	1.6
IT Occ.	1.0	1.2	1.9	2.1	2.2	2.7
Logistics Occ.	9.7	9.9	9.5	9.0	10.3	8.9
Facility Management	4.3	2.8	4.3	3.2	4.0	3.9
Purchasing & Trading	2.4	1.4	2.6	1.8	2.8	2.4
Sales Occ.	6.0	7.0	6.1	7.4	6.9	7.7
Raw Material & Plastic Processing	1.9	1.0	1.6	0.8	1.5	0.9
Tourism Services	1.5	1.8	1.8	2.6	2.4	2.5
Business Administration	18.9	20.4	20.6	20.7	19.5	20.5
Financial Services	2.7	2.6	2.9	2.9	2.3	2.6
Doctors Assistants	1.5	1.5	1.8	1.9	1.9	1.7
Nursing Occ	6.5	7.5	8.4	9.3	10.0	7.9
Medical Care Occ.	1.5	1.6	2.0	2.2	2.6	2.2
Teachers & Social Care Workers	2.6	2.6	2.8	2.7	3.0	3.5
Artistic Occ.	0.9	0.9	0.9	1.1	1.0	1.1
Wood & Paper Processing	1.7	1.5	1.3	1.0	0.9	1.0
Media Design	0.9	0.8	0.8	0.7	0.7	0.7
Metal Processing	4.8	6.6	4.5	5.0	4.0	3.5
Machine-Building Occ.	4.3	4.5	4.3	4.7	4.3	5.4
Engineering Occ.	9.4	9.2	9.1	9.2	7.8	8.7
Textile & Leather Processing	1.0	1.0	0.6	0.6	0.4	0.4
Food Processing	2.9	2.2	2.8	2.3	2.6	2.4

Data: Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This table shows the employment structure of the SIAB and the MZ in the years 1995, 2005 and 2015. All statistics are calculated on the sample of employees aged 18–63. The MZ is restricted to match the sample characteristics of the SIAB by excluding the marginally employed (<10h/week), civil servants, and self-employed individuals.

Table C.17 – Definition of Big Five Dimensions

Dimension	Definition
Openness	... the tendency to be open to new aesthetic, cultural, or intellectual experiences.
Conscientousness	... the tendency to be organized, responsible, and hardworking.
Extraversion	... the tendency to be outgoing, gregarious, sociable, and openly expressive.
Agreeableness	... the tendency to act in a cooperative, unselfish manner.
Neuroticism	... a chronic level of emotional instability and proneness to psychological distress.

Note: Short Definitions from the APA Dictionary of Psychology.

Table C.18 – Big Five Scales in the GSOEP by Age Group

Age Group/ Likert Scale	Big Five Dimension	Questions
2-3 years 11 point Likert		*How would you rank your child in comparison to other children of the same age? My child is ...*
	O	quick at learning new things - needs more time
	C	focused - easily distracted
	E	shy - outgoing
	A	obstinate - obedient
	N	–
5-6 years 9-10 years 11 point Likert		*How would you rank your child in comparison to other children of the same age? My child is ...*
	O	not that interested - hungry for knowledge understands quickly - needs more time
	C	tidy - untidy focused - easy to distract
	E	talkative - quiet withdrawn - sociable
	A	good-natured - irritable obstinate - compliant
	N	self-confident - insecure fearful - fearless
11-12 years 13-15 years 17 years 7 point Likert		*People can have many different qualities – some are listed below. You will probably think that some of these are completely true of you whereas others are not at all. And with some of them, you might not be sure. I am someone who is ...*
	O	original, someone who comes up with new ideas someone who values artistic, aesthetic experiences imaginative eager for knowledge
	C	a thorough worker somewhat lazy effective and efficient in completing tasks
	E	communicative and talkative outgoing, sociable reserved
	A	sometimes a bit rude to others forgiving considerate and kind to others
	N	a worrier nervous relaxed, able to deal with stress

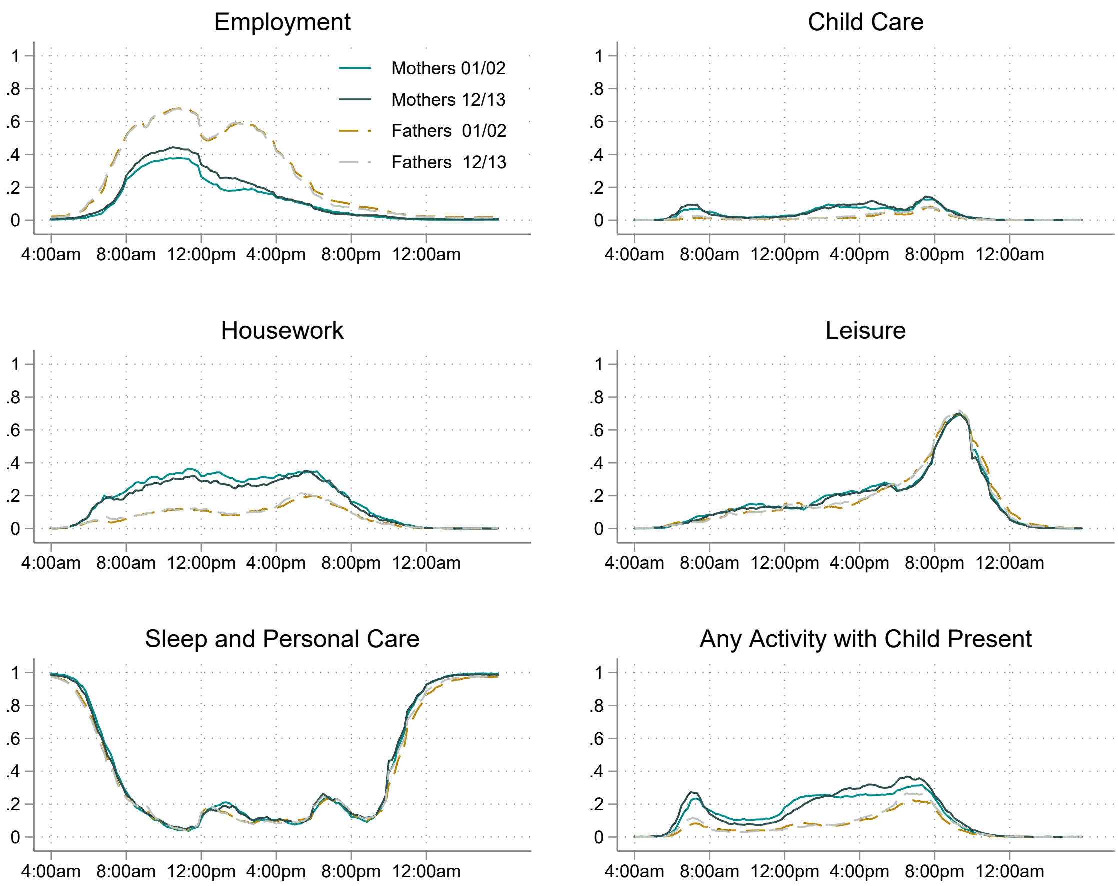
Table C.19 – Within-Family Correlation of Wage Shocks and Child Characteristics

Sibling Characteristic	N	Sibling $\times$ Child Age FE Only	Sibling $\times$ Child Age FE + Year FE
Female	6,070	0.026 (0.031)	0.022 (0.032)
Migration Background	6,070	0.003 (0.005)	0.007 (0.005)
Birth Year	6,070	0.847*** (0.133)	0.000 (0.000)
Birth Rank	6,070	0.275*** (0.052)	-0.003 (0.026)
# of Siblings	6,070	-0.002 (0.005)	-0.001 (0.005)
Birth Height (cm)	2,539	0.138 (0.213)	0.139 (0.215)
Birth Weight (kg)	2,553	0.020 (0.038)	0.012 (0.038)
Breastfed	2,341	-0.019 (0.017)	-0.017 (0.017)
Age Mother	6,070	0.847*** (0.133)	0.000 (0.000)
Age Father	6,070	0.847*** (0.133)	0.000 (0.000)

Data: German Socio-economic Panel (GSOEP), Sample of Integrated Labour Market Biographies (SIAB), Microcensus (MZ).
Note: Own calculations. This table shows correlations between $\hat{w}^{\Delta}_{ifat-1}(=\hat{w}^{m}_{ifat-1}-\hat{w}^{p}_{ifat-1})$ and sibling characteristics conditional on different control variables. The left-hand panel controls for sibling times child age fixed effects γ_{fa}, only. The right-hand panel additionally controls for year fixed effects, τ_t. Significance Levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Standard errors (in parenthesis) are clustered at the family level.

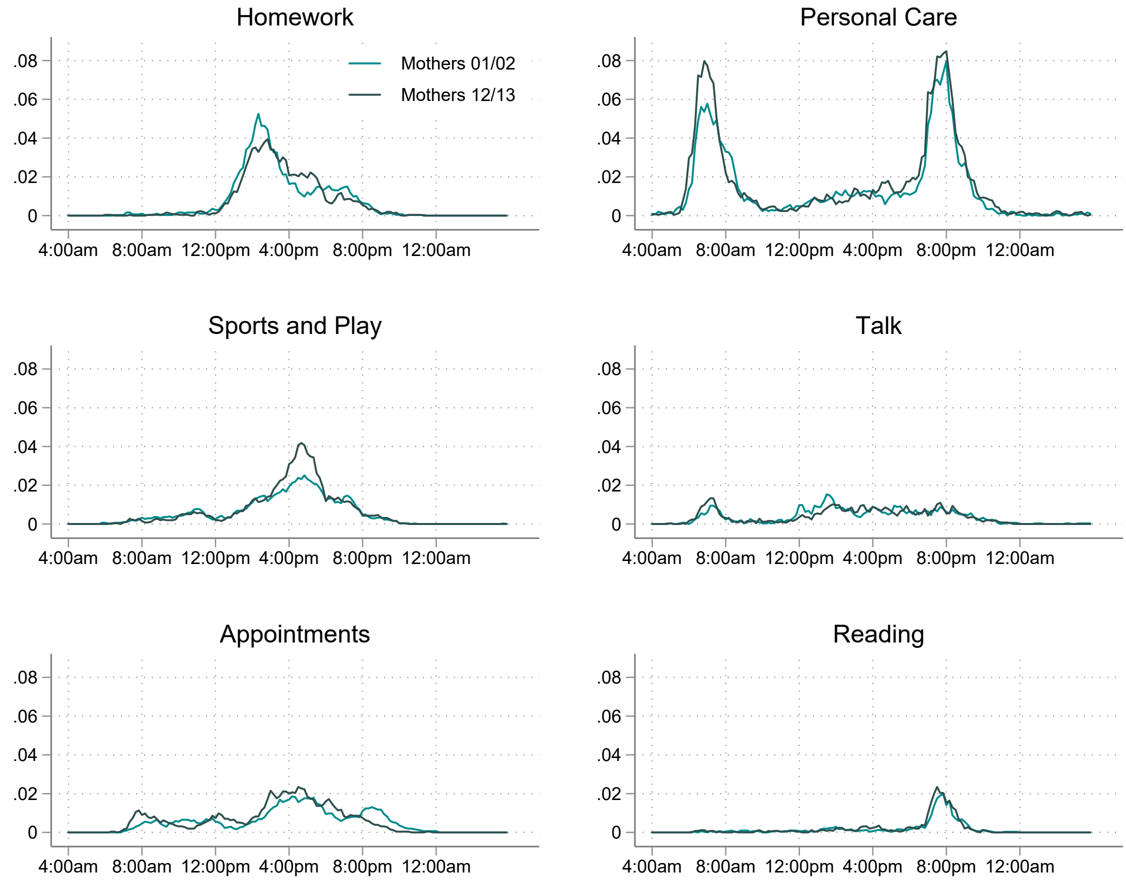
Appendix C.4 Additional Figures

Figure C.1 – Time-Use of Parents in Germany by Gender, 2001/02 and 2012/13



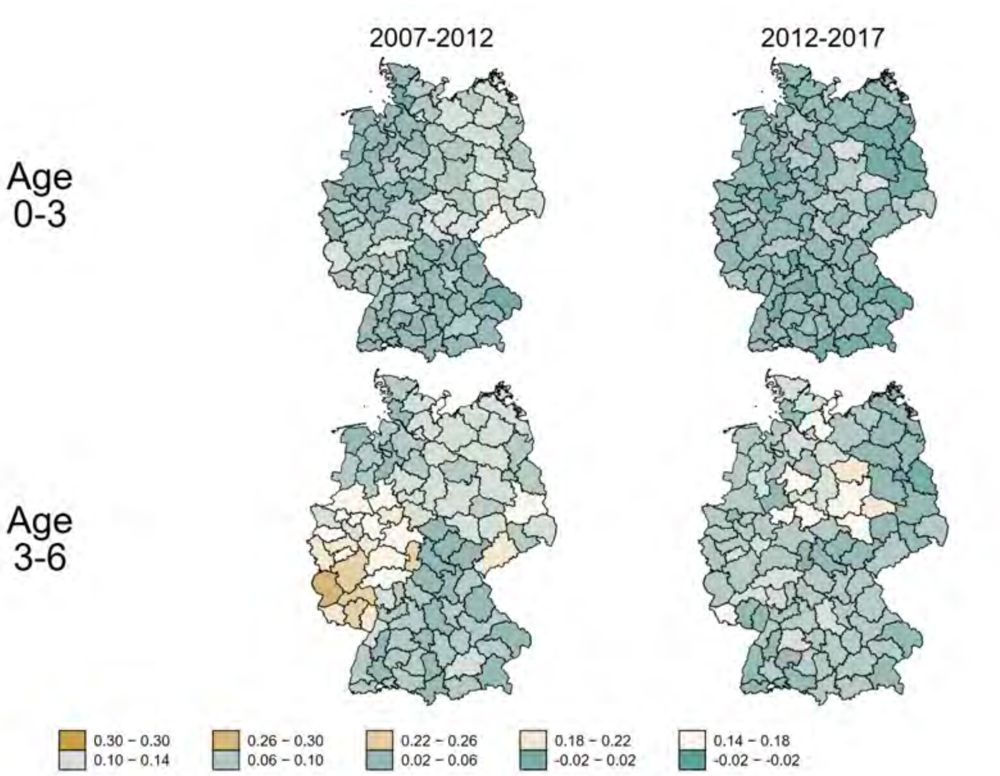
Data: German Time-Use Study (GTUS).
Note: Own calculations. This figure compares the share of mothers and fathers involved in a particular activity for each 10 minute time window of the day across the survey waves 2001/02 and 2012/13. The sample includes two-parent households aged 18–63 with at least one resident child aged 2–17 ($N = 3,065$ in 2001/02 and $N = 2,558$ in 2012/13). The analysis is based on week days Monday through Friday only. For each time of the day the shares across the first five panels sum to 100%. The panel titled *Any Activity with Child Present* represents the share of mothers and fathers who indicate the presence of one of their children in either of the activities represented on the first five panels.

Figure C.2 – Childcare Activities of Mothers in Germany, 2001/02 and 2012/13



Data: German Time-Use Study (GTUS).
Note: Own calculations. This figure compares the share of mothers involved in a particular childcare activity for each 10 minute time window of the day across the survey waves 2001/02 and 2012/13. The sample includes two-parent households aged 18–63 with at least one resident child aged 2–17 ($N = 3,065$ in 2001/02 and $N = 2,558$ in 2012/13). The analysis is based on week days Monday through Friday only.

Figure C.3 – Change in Full Day Childcare Availability by Child Age and Commuting Zone, 2007-2017



Data: Federal Statistical Office of Germany.
Note: Own calculations. This figure shows the change in the share of children attending full day childcare from 2007 to 2017 in five-year windows by child age and commuting zone. The 96 commuting zones are defined by the official territory definition of spatial planning regions of the Federal Office for Building and Regional Planning from 31.12.2017.

Bibliography

Aaberge, R. and A. Brandolini (2015). “Multidimensional Poverty and Inequality”. *Handbook of Income Distribution*. Ed. by A. B. Atkinson and F. Bourguignon. Vol. 2B. Amsterdam: Elsevier. Chap. 3, pp. 141–216.

Aaberge, R., M. Mogstad, and V. Peragine (2011). “Measuring long-term inequality of opportunity”. *Journal of Public Economics* 95 (3–4), pp. 193–204.

Aaronson, D. and B. Mazumder (2008). “Intergenerational Economic Mobility in the United States, 1940 to 2000”. *Journal of Human Resources* 43 (1), pp. 139–172.

Adão, R., M. Kolesár, and E. Morales (2019). “Shift-Share Designs: Theory and Inference”. *Quarterly Journal of Economics* 134 (4), pp. 1949–2010.

Agostinelli, F., M. Doepke, G. Sorrenti, and F. Zilibotti (2020). “It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects”. *National Bureau of Economic Research Working Paper Series* 27050.

Agostinelli, F. and G. Sorrenti (2018). “Money vs. Time: Family Income, Maternal Labor Supply, and Child Development”. *HCEO Working Paper Series* 2018-017.

Aguiar, M. and E. Hurst (2007). “Measuring Trends in Leisure: The Allocation of Time Over Five Decades”. *Quarterly Journal of Economics* 122 (3), pp. 969–1006.

Aizer, A. (2010). “The Gender Wage Gap and Domestic Violence”. *American Economic Review* 100 (4), pp. 1847–59.

Akee, R., W. Copeland, E. J. Costello, and E. Simeonova (2018). “How Does Household Income Affect Child Personality Traits and Behaviors?” *American Economic Review* 108 (3), pp. 775–827.

Alan, S., T. Boneva, and S. Ertac (2019). “Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit”. *Quarterly Journal of Economics* 134 (3), pp. 1121–1162.

Alesina, A. and P. Giuliano (2011). “Preferences for Redistribution”. *Handbook of Social Economics*. Ed. by A. Bisin and J. Benhabib. Vol. 1. San Diego: North Holland. Chap. 4, pp. 93–132.

Alesina, A., S. Stantcheva, and E. Teso (2018). “Intergenerational Mobility and Preferences for Redistribution”. *American Economic Review* 108 (2), pp. 521–554.

ALLEN, R. C. (2017). “Absolute Poverty: When Necessity Displaces Desire”. *American Economic Review* 107 (12), pp. 3690–3721.

ALMÅS, I., A. W. CAPPELEN, J. T. LIND, E. Ø. SØRENSEN, and B. TUNGODDEN (2011). “Measuring unfair (in)equality”. *Journal of Public Economics* 95 (7–8), pp. 488–499.

ALMÅS, I., A. W. CAPPELEN, K. G. SALVANES, E. Ø. SØRENSEN, and B. TUNGODDEN (2016). “What Explains the Gender Gap in College Track Dropout? Experimental and Administrative Evidence”. *American Economic Review* 106 (5), pp. 296–302.

ALMLUND, M., A. L. DUCKWORTH, J. HECKMAN, and T. KAUTZ (2011). “Personality Psychology and Economics”. *Handbook of The Economics of Education*. Ed. by E. A. HANUSHEK, S. MACHIN, and L. WOESSMANN. Vol. 4. Elsevier, pp. 1–181.

ALMOND, D., J. CURRIE, and V. DUQUE (2018). “Childhood Circumstances and Adult Outcomes: Act II”. *Journal of Economic Literature* 56 (4), pp. 1360–1446.

ANDERBERG, D., H. RAINER, J. WADSWORTH, and T. WILSON (2015). “Unemployment and Domestic Violence: Theory and Evidence”. *Economic Journal* 126 (597), pp. 1947–1979.

ANDERSON, E. S. (1999). “What Is the Point of Equality?” *Ethics* 109 (2), pp. 287–337.

ANDREOLI, F., T. HAVNES, and A. LEFRANC (2018). “Robust inequality of opportunity comparisons: Theory and application to early-childhood policy evaluation”. *Review of Economics and Statistics* Forthcoming.

ANDREONI, J., D. AYDIN, B. A. BARTON, B. D. BERNHEIM, and J. NAECKER (2019). “When Fair Isn’t Fair: Understanding Choice Reversals Involving Social Preferences”. *Journal of Political Economy* Forthcoming.

ANTONI, M., A. GANZER, and P. vom BERGE (2019). “Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975 - 2017”. *FDZ Data Report* 04-2019.

ARNESON, R. J. (1989). “Equality and equal opportunity for welfare”. *Philosophical Studies* 56 (1), pp. 77–93.

– (2018). “Four Conceptions of equal opportunity”. *Economic Journal* 128 (612), F152–F173.

ATHEY, S. (2018). “The Impact of Machine Learning on Economics”. *The Economics of Artificial Intelligence: An Agenda*. Ed. by A. K. AGRAWAL, J. GANS, and A. GOLDFARB. Chicago: University of Chicago Press. Chap. 21.

ATHEY, S. and G. W. IMBENS (2019). “Machine Learning Methods That Economists Should Know About”. *Annual Review of Economics* 11, pp. 685–725.

ATKINSON, A. B., A.-C. GUIO, and E. MARLIER, eds. (2017). *Monitoring social exclusion in Europe. 2017 Edition*. Luxembourg: Publications Office of the European Union.

Atkinson, A. B. and T. Piketty, eds. (2007). *Top Incomes over the Twentieth Century: A Contrast Between European and English-Speaking Countries*. Oxford: Oxford University Press.

Attanasio, O., F. Cunha, and P. Jervis (2019). “Subjective Parental Beliefs. Their Measurement and Role”. *National Bureau of Economic Research Working Paper Series* 26516.

Autor, D. H. and D. Dorn (2013). “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market”. *American Economic Review* 103 (5), pp. 1553–1597.

Autor, D. H., D. Dorn, and G. Hanson (2019). “When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men”. *American Economic Review: Insights* 1 (2), pp. 161–78.

Autor, D. H., D. Dorn, L. F. Katz, C. Patterson, and J. V. Reenen (2020). “The Fall of the Labor Share and the Rise of Superstar Firms”. *Quarterly Journal of Economics*, Forthcoming.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008). “Trends in U.S. Wage Inequality: Revising the Revisionists”. *Review of Economics and Statistics* 90 (2), pp. 300–323.

Bach, M., J. Koebe, and F. Peter (2019). “Long Run Effects of Universal Childcare on Personality Traits”. *DIW Discussion Papers* 1815.

Baker, M., J. Gruber, and K. Milligan (2019a). “The Long-Run Impacts of a Universal Child Care Program”. *American Economic Journal: Economic Policy* 11 (3), pp. 1–26.

Baker, M., Y. Halberstam, K. Kroft, A. Mas, and D. Messacar (2019b). “Pay Transparency and the Gender Gap”. *National Bureau of Economic Research Working Paper Series* 25834.

Bargain, O., K. Orsini, and A. Peichl (2014). “Comparing Labor Supply Elasticities in Europe and the United States: New Results”. *Journal of Human Resources* 49 (3), pp. 723–838.

Barry, B. (2005). *Why Social Justice Matters*. Cambridge: Polity Press.

Bartik, T. J. (1991). *Who Benefts from State and Local Economic Development Policies?* Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.

Bayer, P. and K. K. Charles (2018). “Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940”. *Quarterly Journal of Economics* 133 (3), pp. 1459–1501.

Bell, R. and Y. Koren (2007). “Lessons from the Netflix prize challenge”. *ACM SIGKDD Explorations* 9 (2), pp. 75–79.

Bennedsen, M., E. Simintzi, M. Tsoutsoura, and D. Wolfenzon (2020). “Do Firms Respond to Gender Pay Gap Transparency?” *Journal of Finance*, Forthcoming.

Bertrand, M., S. E. Black, S. Jensen, and A. Lleras-Muney (2018). “Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway”. *Review of Economic Studies* 86 (1), pp. 191–239.

Bertrand, M., E. Kamenica, and J. Pan (2015). “Gender Identity and Relative Income within Households”. *Quarterly Journal of Economics* 130 (2), pp. 571–614.

Bertrand, M. and J. Pan (2013). “The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior”. *American Economic Journal: Applied Economics* 5 (1), pp. 32–64.

Biau, G. and E. Scornet (2016). “A random forest guided tour”. *Test* 25 (2), pp. 197–227.

Bick, A. and N. Fuchs-Schündeln (2017). “Quantifying the Disincentive Effects of Joint Taxation on Married Women’s Labor Supply”. *American Economic Review* 107 (5), pp. 100–104.

Björklund, A. and M. Jäntti (1997). “Intergenerational Income Mobility in Sweden Compared to the United States”. *American Economic Review* 87 (5), pp. 1009–1018.

Björklund, A., M. Jäntti, and J. E. Roemer (2012a). “Equality of opportunity and the distribution of long-run income in Sweden”. *Social Choice and Welfare* 39 (2-3), pp. 675–696.

Björklund, A., J. Roine, and D. Waldenström (2012b). “Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity?” *Journal of Public Economics* 96 (5-6), pp. 474–484.

Black, S. E., P. Devereux, P. Lundborg, and K. Majlesi (2020). “Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviors”. *Review of Economic Studies* 87 (4), pp. 1683–1725.

Black, S. E. and P. J. Devereux (2011). “Recent Developments in Intergenerational Mobility”. *Handbook of Labour Economics*. Ed. by D. Card and O. Ashenfelter. Vol. 4B. Amsterdam: Elsevier. Chap. 16, pp. 1487–1541.

Black, S. E., E. Grönqvist, and B. Öckert (2018). “Born to Lead? The Effect of Birth Order on Noncognitive Abilities”. *Review of Economics and Statistics* 100 (2), pp. 274–286.

Blau, F. D. and L. M. Kahn (2017). “The Gender Wage Gap: Extent, Trends, and Explanations”. *Journal of Economic Literature* 55 (3), pp. 789–865.

Blundell, J. and E. Risa (2019). “Income and family background: Are we using the right models?” *mimeo*.

Blundell, R., P.-A. Chiappori, and C. Meghir (2005). “Collective Labor Supply with Children”. *Journal of Political Economy* 113 (6), pp. 1277–1306.

Boelmann, B., A. Raute, and U. Schönberg (2020). “Wind of Change? Cultural Determinants of Maternal Labor Supply”. *IAB Discussion Paper* 3020.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. t. Weel (2008). “The Economics and Psychology of Personality Traits”. *Journal of Human Resources* 43 (4), pp. 972–1059.

Borusyak, K., P. Hull, and X. Jaravel (2019). “Quasi-Experimental Shift-Share Research Designs”. *National Bureau of Economic Research Working Paper Series* 24997.

Bossert, W. and M. Fleurbaey (1995). “Redistribution and compensation”. *Social Choice and Welfare* 13 (1), pp. 343–355.

Bourguignon, F., F. H. G. Ferreira, and M. Walton (2006). “Equity, efficiency and inequality traps: A research agenda”. *Journal of Economic Inequality* 5 (2), pp. 235–256.

Bourguignon, F., F. H. G. Ferreira, and M. Menéndez (2007). “Inequality of Opportunity in Brazil”. *Review of Income and Wealth* 53 (4), pp. 585–618.

Bratberg, E., J. Davis, B. Mazumder, M. Nybom, D. D. Schnitzlein, and K. Vaage (2017). “A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US”. *Scandinavian Journal of Economics* 119 (1), pp. 72–101.

Braun, S. T. and J. Stuhler (2018). “The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany”. *Economic Journal* 128 (609), pp. 576–611.

Breiman, L., J. Friedman, C. Stone, and R. Olshen (1984). *Classification and Regression Trees*. Belmont: Taylor & Francis.

Breiman, L. (2001). “Random Forests”. *Machine Learning* 45 (1), pp. 5–32.

Brewer, M., B. Etheridge, and C. O’Dea (2017). “Why are Households that Report the Lowest Incomes So Well-off?” *Economic Journal* 127 (605), F24–F49.

Browning, M., P.-A. Chiappori, and Y. Weiss (2014). *Economics of the Family*. Cambridge: Cambridge University Press.

Bruins, M. (2017). “Women’s economic opportunities and the intra-household production of child human capital”. *Labour Economics* 44, pp. 122–132.

Brunori, P., F. H. G. Ferreira, M. A. Lugo, and V. Peragine (2013). “Opportunity-Sensitive Poverty Measurement”. *IZA Discussion Papers* 7818.

Brunori, P., P. Hufe, and D. Mahler (2020). “The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees and Forests”. *R&R Scandinavian Journal of Economics*.

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

Burkhauser, R. V., S. Feng, S. P. Jenkins, and J. Larrimore (2012). “Recent Trends in Top Income Shares in the United States: Reconciling Estimates from March CPS and IRS Tax Return Data”. *Review of Economics and Statistics* 94 (2), pp. 371–388.

Butrica, B. A. and R. V. Burkhauser (1997). “Estimating Federal Income Tax Burdens for Panel Study of Income Dynamics (PSID) Families Using the National Bureau of Economic Research

TAXSIM Model". *Maxwell Center for Demography and Economics of Aging Studies Program Paper* 12.

Cappelen, A. W., C. Cappelen, and B. Tungodden (2018). "Second-best fairness under limited information: The trade-off between false positives and false negatives". *NHH Department of Economics Discussion Paper* 18.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden (2007). "The Pluralism of Fairness Ideals: An Experimental Approach". *American Economic Review* 97 (3), pp. 818–827.

Cappelen, A. W., J. Konow, E. Ø. Sørensen, and B. Tungodden (2013a). "Just Luck: An Experimental Study of Risk-Taking and Fairness". *American Economic Review* 103 (4), pp. 1398–1413.

Cappelen, A. W., K. O. Moene, E. Ø. Sørensen, and B. Tungodden (2013b). "Needs Versus Entitlements—An International Fairness Experiment". *Journal of the European Economic Association* 11 (3), pp. 574–598.

Cappelen, A. W. and B. Tungodden (2017). "Fairness and the proportionality principle". *Social Choice and Welfare* 49 (3–4), pp. 709–719.

Carneiro, P., K. V. Løken, and K. G. Salvanes (2015). "A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children". *Journal of Political Economy* 123 (2), pp. 365–412.

Carneiro, P., C. Meghir, and M. Parey (2013). "Maternal Education, Home Environments, and the Development of Children and Adolescents". *Journal of the European Economic Association* 11 (1), pp. 123–160.

Checchi, D. and V. Peragine (2010). "Inequality of opportunity in Italy". *Journal of Economic Inequality* 8 (4), pp. 429–450.

Checchi, D., V. Peragine, and L. Serlenga (2016). "Inequality of Opportunity in Europe: Is There a Role for Institutions?" *Inequality: Causes and Consequences*. Ed. by L. Cappellari, S. W. Polachek, and K. Tatsiramos. Vol. 43. Emerald Insight. Chap. 1, pp. 1–44.

Cherchye, L., B. De Rock, and F. Vermeulen (2012). "Married with Children: A Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure Information". *American Economic Review* 102 (7), pp. 3377–3405.

Chernozhukov, V. and C. Hansen (2008). "The reduced form: A simple approach to inference with weak instruments". *Economics Letters* 100 (1), pp. 68–71.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014a). "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood". *American Economic Review* 104 (9), pp. 2633–79.

Chetty, R., N. Hendren, and L. F. Katz (2016a). “The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment”. *American Economic Review* 106 (4), pp. 855–902.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014b). “Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States”. *Quarterly Journal of Economics* 129 (4), pp. 1553–1623.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014c). “Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility”. *American Economic Review* 104 (5), pp. 141–47.

Chetty, R., N. Hendren, F. Lin, J. Majerovitz, and B. Scuderi (2016b). “Childhood Environment and Gender Gaps in Adulthood”. *American Economic Review* 106 (5), pp. 282–88.

Cohen, G. A. (1989). “On the Currency of Egalitarian Justice”. *Ethics* 99 (4), pp. 906–944.

Corak, M. (2013). “Income Inequality, Equality of Opportunity, and Intergenerational Mobility”. *Journal of Economic Perspectives* 27 (3), pp. 79–102.

Corak, M. and P. Piraino (2011). “The Intergenerational Transmission of Employers”. *Journal of Labor Economics* 29 (1), pp. 37–68.

Cowell, F. A. (1985). “Measures of Distributional Change: An Axiomatic Approach”. *Review of Economic Studies* 52 (1), pp. 135–151.

Cowell, F. A. and M.-P. Victoria-Feser (1996). “Robustness Properties of Inequality Measures”. *Econometrica* 64 (1), pp. 77–101.

Cubel, M., A. Nuevo-Chiquero, S. Sanchez-Pages, and M. Vidal-Fernandez (2016). “Do Personality Traits Affect Productivity? Evidence from the Laboratory”. *Economic Journal* 126 (592), pp. 654–681.

Cunha, F., I. Elo, and J. Culhane (2013). “Eliciting maternal beliefs about the technology of skill formation”. *National Bureau of Economic Research Working Paper Series* 19144.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). “Estimating the Technology of Cognitive and Noncognitive Skill Formation”. *Econometrica* 78 (3), pp. 883–931.

Dahl, G. B. and L. Lochner (2012). “The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit”. *American Economic Review* 102 (5), pp. 1927–56.

Dahmann, S. C. and S. Anger (2018). “Cross-Fertilizing Gains or Crowding Out? Schooling Intensity and Noncognitive Skills”. *HCEO Working Paper* 2018-065.

Danzer, N. and V. Lavy (2018). “Paid parental leave and children’s schooling outcomes”. *Economic Journal* 128 (608), pp. 81–117.

Datta Gupta, N. and M. Simonsen (2010). “Non-cognitive child outcomes and universal high quality child care”. *Journal of Public Economics* 94 (1), pp. 30–43.

Davis, J. and B. Mazumder (2019). “The Decline in Intergenerational Mobility After 1980”. *Federal Reserve Bank of Chicago Working Paper Series* 2017-05.

Decerf, B. (2017). “Why not consider that being absolutely poor is worse than being only relatively poor?” *Journal of Public Economics* 152 (8), pp. 79–92.

Deckers, T., A. Falk, F. Kosse, P. Pinger, and H. Schildberg-Hörisch (2020). “Socio-Economic Status and Inequalities in Children’s IQ and Economic Preferences”. *Journal of Political Economy*, Forthcoming.

Del Boca, D., C. Flinn, and M. Wiswall (2014). “Household Choices and Child Development”. *Review of Economic Studies* 81 (1), pp. 137–185.

Del Boca, D., C. Monfardini, and C. Nicoletti (2017). “Parental and Child Time Investments and the Cognitive Development of Adolescents”. *Journal of Labor Economics* 35 (2), pp. 565–608.

Del Bono, E., M. Francesconi, Y. Kelly, and A. Sacker (2016). “Early Maternal Time Investment and Early Child Outcomes”. *Economic Journal* 126 (596), F96–F135.

Demange, P. A. et al. (2020). “Investigating the Genetic Architecture of Non-Cognitive Skills Using GWAS-by-Subtraction”. *bioRxiv*.

Derenoncourt, E. and C. Montialoux (2019). “Minimum Wages and Racial Inequality”. *mimeo*.

DiCiccio, T. J. and B. Efron (1996). “Bootstrap Confidence Intervals”. *Statistical Science* 11 (3), pp. 189–212.

Doepke, M., G. Sorrenti, and F. Zilibotti (2019). “The Economics of Parenting”. *Annual Review of Economics* 11 (1), pp. 55–84.

Duckworth, A. L., C. Peterson, M. D. Matthews, and D. R. Kelly (2007). “Grit: perseverance and passion for long-term goals.” *Journal of Personality and Social Psychology* 92 (6), pp. 1087–1101.

Dupuy, A. and A. Galichon (2014). “Personality Traits and the Marriage Market”. *Journal of Political Economy* 122 (6), pp. 1271–1319.

Dworkin, R. (1981a). “What is Equality? Part 1: Equality of Welfare”. *Philosophy & Public Affairs* 10 (3), pp. 185–246.

– (1981b). “What is Equality? Part 2: Equality of Resources”. *Philosophy & Public Affairs* 10 (4), pp. 283–345.

Eckstein, Z., M. Keane, and O. Lifshitz (2019). “Career and Family Decisions: Cohorts Born 1935–1975”. *Econometrica* 87 (1), pp. 217–253.

EIKA, L., M. MOGSTAD, and B. ZAFAR (2019). "Educational Assortative Mating and Household Income Inequality". *Journal of Political Economy* 127 (6), pp. 2795–2835.

EUROSTAT (2013). *European household income by groups of households*. Luxembourg: Publications Office of the European Union.

FARAVELLI, M. (2007). "How context matters: A survey based experiment on distributive justice". *Journal of Public Economics* 91 (7–8), pp. 1399–1422.

FCOWELL (2016). "Inequality and Poverty Measures". *Oxford Handbook of Well-Being and Public Policy*. Ed. by M. D. ADLER and M. FLEURBAEY. Oxford: Oxford University Press. Chap. 4, pp. 82–125.

FEDERAL GOVERNMENT OF GERMANY (2019). *We want greater equality of opportunity for all children*. URL: `https://www.bundesregierung.de/breg-en/issues/ganztagsausbau-1691264` (visited on 10/12/2020).

FEDERAL STATISTICAL OFFICE (2020). *Haushalte und Familien – Familien mit minderjährigen Kindern in der Familie nach Lebensform und Kinderzahl*. URL: `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Tabellen/2-5-familien.html` (visited on 07/29/2020).

FEENSTRA, R. C., R. INKLAAR, and M. P. TIMMER (2015). "The Next Generation of the Penn World Table". *American Economic Review* 105 (10), pp. 3150–3182.

FELFE, C. and R. LALIVE (2018). "Does early child care affect children's development?" *Journal of Public Economics* 159, pp. 33–53.

FERREIRA, F. H. G. and J. GIGNOUX (2011). "The Measurement of Inequality of Opportunity: Theory and an Application to Latin America". *Review of Income and Wealth* 57 (4), pp. 622–657.

FERREIRA, F. H. G. and V. PERAGINE (2016). "Individual Responsibility and Equality of Opportunity". *Oxford Handbook of Well-Being and Public Policy*. Ed. by M. D. ADLER and M. FLEURBAEY. Oxford: Oxford University Press. Chap. 25, pp. 746–784.

FIORINI, M. and M. P. KEANE (2014). "How the Allocation of Children's Time Affects Cognitive and Noncognitive Development". *Journal of Labor Economics* 32 (4), pp. 787–836.

FISMAN, R., I. KUZIEMKO, and S. VANNUTELLI (2018). "Distributional preferences in larger groups: Keeping up with the Joneses and keeping track of the tails". *mimeo*.

FLETCHER, J. M. and S. SCHURER (2017). "Origins of Adulthood Personality: The Role of Adverse Childhood Experiences". *The B.E. Journal of Economic Analysis & Policy* 17 (2), p. 20150212.

FLEURBAEY, M. (2008). *Fairness, Responsibility, and Welfare*. Oxford: Oxford University Press.

Fleurbaey, M. (1995). “The requisites of equal opportunity”. *Social Choice, Welfare and Ethics*. Ed. by W. A. Barnett, H. Moulin, M. Salles, and N. J. Schofield. Cambridge: Cambridge University Press. Chap. 2, pp. 37–55.

– (2007). “Poverty as a Form of Oppression”. *Freedom from Poverty as a Human Right - Who Owes What to the Very Poor?* Ed. by T. Pogge. Oxford: Oxford University Press.

Fleurbaey, M. and F. Maniquet (2011). *A Theory of Fairness and Social Welfare*. Cambridge: Cambridge University Press.

Fleurbaey, M. and F. Maniquet (2006). “Fair Income Tax”. *Review of Economic Studies* 73 (1), pp. 55–83.

Fleurbaey, M. and F. Maniquet (2018). “Optimal Income Taxation Theory and Principles of Fairness”. *Journal of Economic Literature* 56 (3), pp. 1029–79.

Fleurbaey, M. and V. Peragine (2013). “Ex Ante Versus Ex Post Equality of Opportunity”. *Economica* 80 (317), pp. 118–130.

Flinn, C., P. Todd, and W. Zhang (2020). “Personality Traits, Job Search and the Gender Wage Gap”. *HCEO Working Paper Series* 2020-010.

Fong, C. (2001). “Social preferences, self-interest, and the demand for redistribution”. *Journal of Public Economics* 82 (2), pp. 225–246.

Fortin, N., T. Lemieux, and S. Firpo (2011). “Decomposition Methods in Economics”. *Handbook of Labor Economics*. Ed. by D. Card and O. Ashenfelter. Vol. Volume 4, Part A. Amsterdam: Elsevier. Chap. 1, pp. 1–102.

Foster, J. E. (1998). “Absolute versus Relative Poverty”. *American Economic Review* 88 (2), pp. 335–341.

Frankfurt, H. (1987). “Equality as a Moral Ideal”. *Ethics* 98 (1), pp. 21–43.

Friedberg, R., J. Tibshirani, S. Athey, and S. Wager (2018). “Local Linear Forests”. *arXiv* 1807.11408.

Friedman, J., T. Hastie, and R. Tibshirani (2009). *The elements of statistical learning*. New York: Springer.

Fuentes, J. R. and E. E. Leamer (2019). “Effort: The Unrecognized Contributor to US Income Inequality”. *National Bureau of Economic Research Working Paper Series* 26421.

Gaertner, W. and L. Schwettmann (2007). “Equity, Responsibility and the Cultural Dimension”. *Economica* 74 (296), pp. 627–649.

García, J. L., J. J. Heckman, and A. L. Ziff (2018). “Gender Differences in the Benefits of an Influential Early Childhood Program”. *European Economic Review* 109, pp. 9–22.

GARTNER, H. (2005). “The imputation of wages above the contribution limit with the German IAB employment sample”. *FDZ Methodenreport* 2, pp. 1–8.

GELBACH, J. B. (2016). “When Do Covariates Matter? And Which Ones, and How Much?” *Journal of Labor Economics* 34 (2), pp. 509–543.

GESIS (2020). *The Microcensus*. URL: `https://www.gesis.org/en/missy/metadata/MZ/` (visited on 07/29/2020).

GOEBEL, J., M. M. GRABKA, S. LIEBIG, M. KROH, D. RICHTER, C. SCHRÖDER, and J. SCHUPP (2019). “The German Socio-Economic Panel (SOEP)”. *Jahrbücher für Nationalökonomie und Statistik* 239 (2), pp. 345–360.

GOLDSMITH-PINKHAM, P., I. SORKIN, and H. SWIFT (2020). “Bartik Instruments: What, When, Why, and How”. *American Economic Review* 110 (8), pp. 2586–2624.

GREGORY-SMITH, I., B. G. MAIN, and C. A. O’REILLY III (2014). “Appointments, Pay and Performance in UK Boardrooms by Gender”. *Economic Journal* 124 (574), F109–F128.

GUVENEN, F. and G. KAPLAN (2017). “Top Income Inequality in the 21st Century: Some Cautionary Notes”. *Quarterly Review* 38 (1), pp. 2–15.

HEATHCOTE, J., F. PERRI, and G. L. VIOLANTE (2010a). “Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006”. *Review of Economic Dynamics* 13 (1), pp. 15–51.

HEATHCOTE, J., K. STORESLETTEN, and G. L. VIOLANTE (2010b). “The Macroeconomic Implications of Rising Wage Inequality in the United States”. *Journal of Political Economy* 118 (4), pp. 681–722.

HECKMAN, J. J. and S. MOSSO (2014). “The Economics of Human Development and Social Mobility”. *Annual Review of Economics* 6 (1), pp. 689–733.

HILGER, N. (2019). “The Great Escape: Intergenerational Mobility in the United States, 1930-2010”. *mimeo*.

HOCHSCHILD, A. and A. MACHUNG (1990). *The Second Shift*. New York: Avon Books.

HOTHORN, T., K. HORNIK, and A. ZEILEIS (2006). “Unbiased Recursive Partitioning: A Conditional Inference Framework”. *Journal of Computational and Graphical Statistics* 15 (3), pp. 651–674.

HSIN, A. and C. FELFE (2014). “When Does Time Matter? Maternal Employment, Children’s Time With Parents, and Child Development”. *Demography* 51 (5), pp. 1867–1894.

HUFE, P., A. PEICHL, J. E. ROEMER, and M. UNGERER (2017). “Inequality of Income Acquisition: The Role of Childhood Circumstances”. *Social Choice and Welfare* 143 (3-4), pp. 499–544.

Hülle, S., S. Liebig, and M. J. May (2018). "Measuring Attitudes Toward Distributive Justice: The Basic Social Justice Orientations Scale". *Social Indicators Research* 136 (2), pp. 663–692.

Humphries, J. E. and F. Kosse (2017). "On the interpretation of non-cognitive skills – What is being measured and why it matters". *Journal of Economic Behavior & Organization* 136, pp. 174–185.

Ichino, A., M. Olsson, B. Petrongolo, and P. S. Thoursie (2020). "Economic incentives, home production and gender identity norms". *IZA Discussion Paper* 12391.

Jackson, C. K. (2019). "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes". *Journal of Political Economy* 126 (5), pp. 2072–2107.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018). "Shift-Share Instruments and the Impact of Immigration". *National Bureau of Economic Research Working Paper Series* 24285.

Johnson, D. S., K. A. McGonagle, V. A. Freedman, and N. Sastry (2018). "Fifty Years of the Panel Study of Income Dynamics: Past, Present, and Future". *The ANNALS of the American Academy of Political and Social Science* 680 (1), pp. 9–28.

Jolliffe, D. and E. B. Prydz (2016). "Estimating international poverty lines from comparable national thresholds". *Journal of Economic Inequality* 14 (2), pp. 185–198.

Jusot, F., S. Tubeuf, and A. Trannoy (2013). "Circumstances and Efforts: How Important is their Correlation for the Measurement of Inequality of Opportunity in Health?" *Health Economics* 22 (12), pp. 1470–1495.

Kanbur, R. and A. Snell (2018). "Inequality Measures as Tests of Fairness". *Economic Journal* Forthcoming.

Kanbur, R. and A. Wagstaff (2016). "How Useful Is Inequality of Opportunity as a Policy Construct?" *Inequality and Growth: Patterns and Policy*. Ed. by K. Basu and J. E. Stiglitz. Vol. 1. London: Palgrave Macmillan UK. Chap. 4, pp. 131–150.

Kleven, H. J., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019). "Child Penalties across Countries: Evidence and Explanations". *AEA Papers and Proceedings* 109, pp. 122–126.

Kleven, H. J., C. Landais, J. E. Søgaard, and J. Egholt (2018). "Children and Gender Inequality: Evidence from Denmark". *American Economic Journal: Applied Economics* 11 (4), pp. 181–209.

Knowles, J. A. (2012). "Why are Married Men Working So Much? An Aggregate Analysis of Intra-Household Bargaining and Labour Supply". *Review of Economic Studies* 80 (3), pp. 1055–1085.

König, F. (2019). "Superstar Effects and Market Size: Evidence from the Roll-Out of TV". *mimeo*.

Konow, J. (2003). “Which Is the Fairest One of All? A Positive Analysis of Justice Theories”. *Journal of Economic Literature* 41 (4), pp. 1188–1239.

Konow, J. and L. Schwettmann (2016). “The Economics of Justice”. *Handbook of Social Justice Theory and Research*. Ed. by C. Sabbagh and M. Schmitt. New York: Springer. Chap. 5, pp. 83–106.

Kornrich, S. and F. Furstenberg (2013). “Investing in Children: Changes in Parental Spending on Children, 1972–2007”. *Demography* 50, pp. 1–23.

Krawczyk, M. (2010). “A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution”. *Journal of Public Economics* 94 (1–2), pp. 131–141.

Kreisman, D. and M. A. Rangel (2015). “On the Blurring of the Color Line: Wages and Employment for Black Males of Different Skin Tones”. *Review of Economics and Statistics* 97 (1), pp. 1–13.

Lang, K. and J.-Y. K. Lehmann (2012). “Racial Discrimination in the Labor Market: Theory and Empirics”. *Journal of Economic Literature* 50 (4), pp. 959–1006.

Lee, C.-I. and G. Solon (2009). “Trends in Intergenerational Income Mobility”. *Review of Economics and Statistics* 91 (4), pp. 766–772.

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

Leigh, A. (2007). “How Closely Do Top Income Shares Track Other Measures of Inequality?” *Economic Journal* 117 (524), F619–F633.

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

Lindo, J. M., J. Schaller, and B. Hansen (2018). “Caution! Men not at work: Gender-specific labor market conditions and child maltreatment”. *Journal of Public Economics* 163, pp. 77–98.

Lippert-Rasmussen, K. (2001). “Egalitarianism, Option Luck, and Responsibility”. *Ethics* 111 (3), pp. 548–579.

– (2011). “Luck-Egalitarianism: Faults and collective choice”. *Economics and Philosophy* 27 (02), pp. 151–173.

Lippmann, Q., A. Georgieff, and C. Senik (2020). “Undoing Gender with Institutions. Lessons from the German Division and Reunifcation”. *Economic Journal*, Forthcoming.

Livingston, G. (2018). *The Changing Profile of Unmarried Parents*. Pew Research Center.

Løken, K. V., M. Mogstad, and M. Wiswall (2012). “What Linear Estimators Miss: The Effects of Family Income on Child Outcomes”. *American Economic Journal: Applied Economics* 4 (2), pp. 1–35.

Lundberg, S., R. A. Pollak, and J. Stearns (2016). “Family Inequality: Diverging Patterns in Marriage, Cohabitation, and Childbearing”. *Journal of Economic Perspectives* 30 (2), pp. 79–102.

Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997). “Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit”. *Journal of Human Resources* 32 (3), pp. 463–480.

Magdalou, B. and R. Nock (2011). “Income distributions and decomposable divergence measures”. *Journal of Economic Theory* 146 (6), pp. 2440–2454.

Mare, R. D. (2011). “A Multigenerational View of Inequality”. *Demography* 48 (1), pp. 1–23.

McGrath, J. J., L. Petersen, E. Agerbo, O. Mors, P. B. Mortensen, and C. B. Pedersen (2014). “A Comprehensive Assessment of Parental Age and Psychiatric Disorders”. *JAMA Psychiatry* 71 (3), pp. 301–309.

Meyer, B. D., D. Wu, V. R. Mooers, and C. Medalia (2019). “The Use and Misuse of Income Data and Extreme Poverty in the United States”. *National Bureau of Economic Research Working Paper* 25907.

Meyer, B. D. and W. K. C. Mok (2019). “Disability, earnings, income and consumption”. *Journal of Public Economics* 171, pp. 51–69.

Meyer, B. D., W. K. Mok, and J. X. Sullivan (2015). “The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences”. *mimeo*.

Meyer, B. D. and J. X. Sullivan (2012). “Identifying the Disadvantaged: Official Poverty, Consumption Poverty, and the New Supplemental Poverty Measure”. *Journal of Economic Perspectives* 26 (3), pp. 111–135.

Miller, D. (1996). “Two Cheers for Meritocracy”. *Journal of Political Philosophy* 4 (4), pp. 277–301.

Milligan, K. and M. Stabile (2011). “Do Child Tax Benefits Affect the Well-being of Children? Evidence from Canadian Child Benefit Expansions”. *American Economic Journal: Economic Policy* 3 (3), pp. 175–205.

Mittag, N. (2019). “Correcting for Misreporting of Government Benefits”. *American Economic Journal: Economic Policy* 11 (2), pp. 142–64.

Morgan, J. N. and J. A. Sonquist (1963). “Problems in the Analysis of Survey Data, and a Proposal”. *Journal of the American Statistical Association* 58 (302), pp. 415–434.

Mullainathan, S. and J. Spiess (2017). “Machine Learning: An Applied Econometric Approach”. *Journal of Economic Perspectives* 31 (2), pp. 87–106.

Nicoletti, C., K. G. Salvanes, and E. Tominey (2020). “Mothers Working during Preschool Years and Child Skills. Does Income Compensate?” *IZA Discussion Paper* 13079.

Nybom, M. and J. Stuhler (2016). “Biases in Standard Measures of Intergenerational Income Dependence”. *Journal of Human Resources* 52 (3), pp. 800–825.

Olivetti, C. and B. Petrongolo (2016). “The Evolution of the Gender Gap in Industrialized Countries”. *Annual Review of Economics* 8, pp. 405–434.

Ooghe, E. and A. Peichl (2015). “Fair and Efficient Taxation under Partial Control”. *Economic Journal* 125 (589), pp. 2024–2051.

Page, M., J. Schaller, and D. Simon (2019). “The Effects of Aggregate and Gender-Specific Labor Demand Shocks on Child Health”. *Journal of Human Resources* 54 (1), pp. 37–78.

Palomino, J. C., G. A. Marrero, and J. G. Rodríguez (2019). “Channels of inequality of opportunity: The role of education and occupation in Europe”. *Social Indicators Research* 143, pp. 1045–1074.

Papageorge, N. W. and K. Thom (2020). “Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study”. *Journal of the European Economic Association* 18 (3), pp. 1351–1399.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge: The Belknap Press of Harvard University Press.

Piketty, T. and E. Saez (2003). “Income Inequality in the United States, 1913–1998”. *Quarterly Journal of Economics* 118 (1), pp. 1–41.

Piketty, T., E. Saez, and G. Zucman (2018). “Distributional National Accounts: Methods and Estimates for the United States”. *Quarterly Journal of Economics* 133 (2), pp. 553–609.

Pistolesi, N. (2009). “Inequality of opportunity in the land of opportunities, 1968–2001”. *Journal of Economic Inequality* 7 (4), pp. 411–433.

Ramos, X. and D. Van de gaer (2016). “Empirical Approaches to Inequality of Opportunity: Principles, Measures, and Evidence”. *Journal of Economic Surveys* 30 (5), pp. 855–883.

Raute, A. (2019). “Can Financial Incentives Reduce the Baby Gap? Evidence from a Reform in Maternity Leave Benefits”. *Journal of Public Economics* 169, pp. 203–222.

Ravallion, M. and S. Chen (2003). “Measuring pro-poor growth”. *Economics Letters* 78 (1), pp. 93–99.

Rawls, J. (1971). *A Theory of Justice*. Cambridge: The Belknap Press of Harvard University Press.

Roemer, J. E. (1993). “A Pragmatic Theory of Responsibility for the Egalitarian Planner”. *Philosophy & Public Affairs* 22 (2), pp. 146–166.

– (1998). *Equality of Opportunity*. Cambridge: Harvard University Press.

Roemer, J. E. and A. Trannoy (2015). “Equality of Opportunity”. *Handbook of Income Distribution*. Ed. by A. B. Atkinson and F. Bourguignon. Vol. 2A. Amsterdam: Elsevier. Chap. 4, pp. 217–300.

– (2016). “Equality of Opportunity: Theory and Measurement”. *Journal of Economic Literature* 54 (4), pp. 1288–1332.

Roine, J. and D. Waldenström (2015). “Long-Run Trends in the Distribution of Income and Wealth”. *Handbook of Income Distribution*. Ed. by A. B. Atkinson and F. Bourguignon. Vol. 2A. Handbook of Income Distribution. Amsterdam: Elsevier. Chap. 7, pp. 469–592.

Saez, E. and S. Stantcheva (2016). “Generalized Social Marginal Welfare Weights for Optimal Tax Theory”. *American Economic Review* 106 (1), pp. 24–45.

Savelyev, P. (2020). “Conscientiousness, Extraversion, College Education, and Longevity of High-Ability Individuals.” *Journal of Human Resources*, Forthcoming.

Schaller, J. (2016). “Booms, Busts, and Fertility: Testing the Becker Model Using Gender-Specific Labor Demand”. *Journal of Human Resources* 51 (1), pp. 1–29.

Sen, A. (1980). “Equality of What?” *The Tanner Lectures on Human Values*. Ed. by S. McMurrin. Vol. 1. Cambridge: Cambridge University Press. Chap. 6.

Shenhav, N. (2020). “Lowering Standards to Wed? Spouse Quality, Marriage, and Labor Market Responses to the Gender Wage Gap”. *Review of Economics and Statistics*, Forthcoming.

Shorrocks, A. F. (2012). “Decomposition procedures for distributional analysis: a unified framework based on the Shapley value”. *Journal of Economic Inequality* 11 (1), pp. 99–126.

Smeeding, T. M. (2018). “The PSID in Research and Policy”. *The ANNALS of the American Academy of Political and Social Science* 680 (1), pp. 29–47.

Solon, G. (1992). “Intergenerational Income Mobility in the United States”. *American Economic Review* 82 (3), pp. 393–408.

Song, X., C. G. Massey, K. A. Rolf, J. P. Ferrie, J. L. Rothbaum, and Y. Xie (2019). “Long-term decline in intergenerational mobility in the United States since the 1850s”. *Proceedings of the National Academy of Sciences*.

Strobl, C., A.-L. Boulesteix, A. Zeileis, and T. Hothorn (2007). “Bias in random forest variable importance measures: Illustrations, sources and a solution”. *BMC bioinformatics* 8 (1), p. 25.

U.S. Bureau of the Census (Dec. 18, 2019). *Poverty of People, by Region: 1959 to 2018*. url: https://www2.census.gov/programs-surveys/cps/tables/time-series/historical-poverty-people/hstpov9.xls (visited on 12/18/2019).

Van de gaer, D. (1993). "Equality of Opportunity and Investment in Human Capital". PhD thesis. University of Leuven.

van den Boom, B., A. Halsema, and V. Molini (2015). "Are We Confusing Poverty with Preferences?" *World Bank Policy Research Working Paper* 7247.

Varian, H. R. (2014). "Big Data: New Tricks for Econometrics". *Journal of Economic Perspectives* 28 (2), pp. 3–27.

Vita, A. da (2007). "Inequality and Poverty in a Global Perspective". *Freedom from Poverty as a Human Right - Who Owes What to the Very Poor?* Ed. by T. Pogge. Oxford: Oxford University Press.

Vosters, K. (2018). "Is the simple Law of Mobility really a law? Testing Clark's hypothesis". *Economic Journal* 612, F404–F421.

Vosters, K. and M. Nybom (2017). "Intergenerational Persistence in Latent Socioeconomic Status: Evidence from Sweden and the United States". *Journal of Labor Economics* 35 (3), pp. 869–901.

Weinzierl, M. (2014). "The promise of positive optimal taxation: normative diversity and a role for equal sacrifice". *Journal of Public Economics* 118, pp. 128–142.

Widiger, T. A., ed. (2018). *The Oxford Handbook of the Five Factor Model*. Oxford: Oxford University Press.

Wimer, C., L. Fox, I. Garfinkel, N. Kaushal, and J. Waldfogel (2016). "Progress on Poverty? New Estimates of Historical Trends Using an Anchored Supplemental Poverty Measure". *Demography* 53 (4), pp. 1207–1218.

Zheng, B. (1993). "An axiomatic characterization of the Watts poverty index". *Economics Letters* 42 (1), pp. 81–86.

Ziliak, J. P. (2006). "Understanding Poverty Rates and Gaps: Concepts, Trends, and Challenges". *Foundations and Trends in Microeconomics*. Ed. by W. K. Viscusi. Vol. 1. Hanover: now Publishers Inc., pp. 127–197.

Curriculum Vitae

Paul Hufe

Munich Graduate School of Economics

www.paulhufe.net   paulhufe

Research Fields: Labor Economics, Public Economics, Normative Economics

Research Topics: Inequality Measurement, Intergenerational Mobility, Human Capital Formation

EDUCATION

2017 – 2021	Ph.D. Economics	LMU (University of Munich)
2014 – 2017	Ph.D. Economics	University of Mannheim
2011 – 2013	M.A. Philosophy & Economics	University of Bayreuth
2008 – 2011	B.Sc. Business/Economics	WHU Vallendar

RESEARCH VISITS

March/April 2020	NHH Bergen	Host: Bertil Tungodden
Feb – Oct 2019	Princeton University	Host: Marc Fleurbaey
March 2017/Nov 2018	Cornell University	Host: Ravi Kanbur