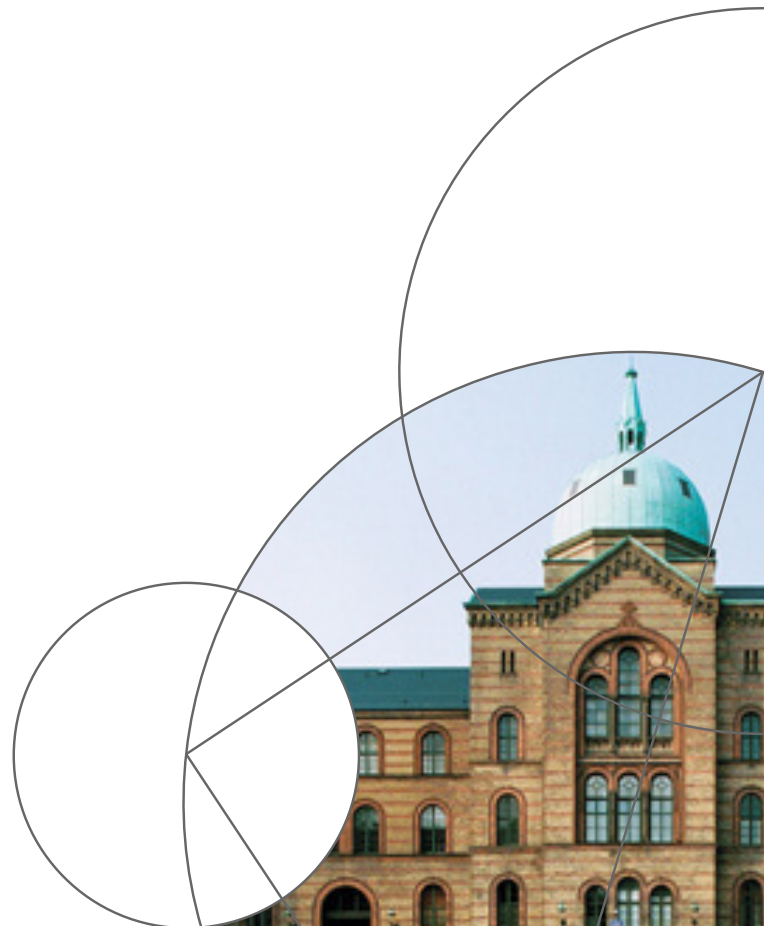




# Financial Problems, Fairness Views and Crime:

Evidence from new combinations of administrative records and online surveys and experiments

Kristoffer Balle Hvidberg



# **PdD Dissertation**

Submitted: July 31, 2021

Advisors: Claus Thustrup Kreiner and Søren Leth-Petersen

University of Copenhagen

Faculty of Social Science

Department of Economics

Center for Economic Behavior and Inequality

*To Ellen and Johannes*

# Contents

Acknowledgements	ii
English summary	iii
Danish summary	v
1 Field of Study and Financial Problems: How economics reduces the risk of default	1
2 Social Position and Fairness View	60
3 Preferences predict who commits crime among young men	158



# Acknowledgements

First, I want to thank my advisors, Claus Thustrup Kreiner and Søren Leth-Petersen. They have both helped and supported me since I started working for them as a research assistant and throughout my PhD studies. It has been a privilege working with them both.

Second, I want to thank Stefanie Stantcheva for inviting me to visit Harvard University as a Fulbright scholar. The visit was an amazing experience and collaborating with her has been a great source of inspiration.

Third, I also want to thank all my colleagues at CEBI. It has been a supportive and inspiring place to work and I have missed the lunches, coffee breaks and Friday CEBEER the past one and a half years.

Finally, I want to thank my friends and family, in particular Stine. All the best moments the past four and a half years have not been related to the work presented here.

# English summary

This PhD dissertation consists of three self-contained chapters. All three chapters leverage new combinations of administrative records and data from online surveys and experiments.

Chapter 1 builds on and repeats text from my master's thesis. The chapter documents how extensive economic education can reduce the risk of getting into financial problems by comparing people who enter business and economics programs with people who enter other higher education programs. To identify the causal effect, I exploit GPA admission thresholds that quasi-randomize applicants near the thresholds into different programs. I find that admission to an economics education significantly reduces the probability of loan default by one-half. This large reduction in the default probability is associated with changes in financial behavior, but it is not associated with differences in the level or stability of people's income.

Chapter 2 is co-authored with Claus Thustrup Kreiner and Stefanie Stantcheva. In this chapter, we link survey data containing Danish people's perceptions of where they rank in various reference groups and fairness views with administrative records on their income history, life events, and reference groups. People know their income positions well, but believe others are closer to themselves than they really are. The perceived fairness of inequalities is strongly related to current social position, moves with shocks to social position (e.g., unemployment or promotions), and changes when people are experimentally shown their actual positions. People view inequalities within education group and co-workers as most unfair, but underestimate inequality the most exactly within these reference groups.

Chapter 3 is co-authored with Thomas Epper, Ernst Fehr, Claus Thustrup Kreiner, Søren Leth-Petersen and Gregers Nytoft Rasmussen. In this chapter, we revisit a key topic in Social Science, namely who commits crime and why which is important for the design of crime

prevention policy. In theory, people who commit crime face different social and economic incentives for criminal activity than other people or they evaluate the costs and benefits of crime differently because they have different preferences. Empirical evidence on the role of preferences is scarce. Theoretically, risk tolerant, impatient and self-interested people are more prone to commit crime than risk averse, patient and altruistic people. We test these predictions with a unique combination of data where we use state-of-the art methods in experimental economics to elicit the preferences of young men and link this experimental data to their criminal records. In addition, our data allow us to control extensively for other characteristics such as cognitive skills, socio-economic background and self-control problems. We find that preferences are strongly associated with actual criminal behavior. Impatience and, in particular, risk tolerance are still strong predictors when we include the full battery of controls. Crime propensities are 8-10 percentage points higher for the most risk tolerant individuals compared to the most risk averse. This effect is half the size of the effect of cognitive skills, which is one of the best predictors of criminal behavior. Looking into different types of crime, we find that preferences predict property offences, while self-control problems predict violent, drug and sexual offences.

# Danish summary

Denne ph.d.-afhandling består af tre selvstændige kapitler. Alle tre kapitler benytter sig af nye kombinationer af administrative data og data fra online spørgeskemaundersøgelser og eksperimenter.

Kapitel 1 bygger på og gentager tekst fra mit speciale. I kapitlet viser jeg, at omfattende økonomisk uddannelse kan reducere risikoen for at komme i finansielle problemer. Jeg sammenligner personer, der er optaget på erhvervsøkonomiske og økonomiske uddannelser med personer, der er optaget på andre videregående uddannelser. For at identificere den kausale effekt udnytter jeg adgangskvotienter baseret på karaktergennemsnit, der kvasi-randomiserer ansøgere ind på forskellige uddannelser, hvis deres karaktergennemsnit er tæt på de pågældende adgangskvotienter. Jeg finder, at optagelse på en økonomisk uddannelse halverer sandsynligheden for at komme i restance på et lån. Denne store reduktion i sandsynligheden er associeret med ændringer i ansøgernes finansielle adfærd, men den er ikke associeret med forskelle i niveauet eller stabiliteten i deres indkomst.

Kapitel 2 er skrevet sammen med Claus Thustrup Kreiner og Stefanie Stantcheva. I kapitlet kombinerer vi data fra en spørgeskemaundersøgelse af danskeres opfattelse af, hvor de rangerer inden for forskellige referencegrupper og deres syn på rimeligheden af ulighed med administrative data vedrørende deres indkomsthistorie, livsbegivenheder og referencegrupper. Vi finder, at personerne er forholdsvis velinformede om deres indkomstpositioner men tror, at andre er tættere på dem selv, end de i virkeligheden er. Opfattelsen af rimeligheden af ulighed viser sig at være stærkt relateret til deres nuværende sociale position og ændrer sig, hvis der sker en pludselig markant ændring i deres sociale position (eksempelvis arbejdsløshed eller forfremmelse), eller når de får oplyst deres faktiske positioner. Ulighed inden for egen

uddannelsesgruppe og blandt kolleger opfattes som mest urimelig, men netop inden for disse referencegrupper undervurderer folk uligheden mest.

Kapitel 3 er skrevet sammen med Thomas Epper, Ernst Fehr, Claus Thustrup Kreiner, Søren Leth-Petersen og Gregers Nytoft Rasmussen. I kapitlet undersøger vi et kernespørgsmål inden for samfundsvidenskab, nemlig hvem begår kriminalitet og hvorfor, hvilket har stor betydning for udformningen af præventiv kriminalitetspolitik. I økonomisk teori har personer, der begår kriminalitet, andre sociale og økonomiske incitamenter til kriminel aktivitet end personer, der ikke begår kriminalitet eller også vurderer de fordelene og ulemperne ved kriminel aktivitet anderledes, fordi de har andre præferencer. Der findes kun meget lidt evidens for præferencernes betydning. Teoretisk set er personer, der er mere risikotolerante, utålmodige og egennyttige mere tilbøjelige til at begå kriminalitet end personer, der er risikouvillige, tålmodige og altruistiske. Vi tester disse hypoteser med en unik kombination af data, hvor vi bruger de mest avancerede metoder inden for eksperimentel økonomi til at klarlægge unge mænds præferencer og linker denne eksperimentelle data til deres straffeforhold. Derudover giver vores data mulighed for at kontrollere for andre karakteristika såsom kognitive evner, socioøkonomisk baggrund samt problemer med selvkontrol. Vores resultater viser en klar sammenhæng mellem præferencer og kriminel adfærd. Utålmodighed og især risikotolerance prædikerer stadig klart kriminalitet, når vi inkluderer samtlige af vores kontrolvariabler. Sandsynligheden for kriminalitet er 8-10 procentpoint højere for de mest risikotolerante personer sammenlignet med de mest risikouvillige personer. Denne effekt svarer til halvdelen af effekten af kognitive evner, hvilket er en af de bedst prædikerende variable for kriminel adfærd. Når vi ser nærmere på forskellige typer af kriminalitet, finder vi, at præferencer prædikerer ejendomsforbrydelser, mens problemer med selvkontrol prædikerer overtrædelse af loven om euforiserende stoffer samt volds- og seksualforbrydelser.

# Chapter 1

**Field of Study and Financial Problems:**

**How economics reduces the risk of default<sup>1</sup>**

---

<sup>1</sup>The chapter has previously, in a preliminary version, constituted my master's thesis, which after two years in the PhD program was handed in for assessment. Thus the chapter is an extension of the master's thesis. Since I handed in the mater's thesis, I have added additional years of data, carried out a joint estimation and included a more comprehensive analysis of heterogeneous effects and potential channels.

## 1.1 Introduction

Financial decision-making in households has received growing interest in recent years. For instance, OECD argues that ill-informed financial decision-making has “tremendous adverse effects on both personal, and ultimately, global finance” (OECD, 2016). The interest often concerns the debt behavior of households and how well people manage and service their debt. Loan defaults and delinquencies have great consequences for the debtors, who can lose credit access, and for the creditors, who can lose their money (Kreiner et al., 2020). More generally, defaults have important adverse effects on the credit market and can lead to credit rationing (Stiglitz and Weiss, 1981). Economic education is often suggested as a way of improving financial decision-making, but evidence on the causal effect of economic education on financial outcomes is sparse (Beshears et al., 2018).

I investigate how field of study causally affects the probability of getting into financial problems and, in particular, whether studying economics can reduce the risk of default. The main challenge in identifying the causal effect is that the observed correlations can be driven entirely by students self-selecting into fields of study based on preferences or skills that also drive their financial behavior.

To address this self-selection problem, I leverage a unique combination of administrative third-party reported data on applications and admissions to higher education and on the universe of personal loans in Denmark. This combination enables me to identify the causal effect of field of study by exploiting that the Danish system of admission to higher education generates locally unpredictable grade point average (GPA) admission thresholds. Applicants with a GPA close to the admission thresholds are effectively randomized into different fields of study.

Intuitively, I compare applicants who are very similar but are admitted to different fields of study due to slightly different GPAs from upper secondary school. For instance, imagine two applicants who would both prefer to study political science (which belongs to the *Social science* field of study) if their GPA is above the admission threshold of this specific program. Both applicants have economics (which belongs to the *Business and economics* field of study) as their alternative choice. The admission thresholds of the two programs are based on the

GPA distribution of all applicants to these programs, and in the two applicants' year of application, economics gets a lower GPA admission threshold than political science. The applicants' GPAs are both very close to the admission threshold of political science, but one is just above the threshold and one is just below. Due to this small difference between their GPAs, one applicant is admitted to *Social science* (the preferred field of study) and the other is admitted to *Business and economics* (the alternative field of study). The example illustrates that the admission process quasi-randomizes applicants near the thresholds into different fields of study.

I find that admission to *Business and economics* significantly reduces the risk of default more than 10 years after the year of application. In particular, I find that applicants who have *Business and economics* as their alternative field of study are 6-8%-points less likely to default if they are just below the GPA admission threshold of their preferred field of study. This effect is more than half of the average probability of default for all first time applicants, which is 11%.

To investigate potential channels that can explain this finding, I estimate the causal effect of admission to *Business and economics* on outcomes related to labor market performance and financial behavior. For the applicants who have *Business and economics* as their alternative field, I find no changes in income, unemployment or self-employment around the admission threshold. Instead, I find changes in financial behavior and I show that the applicants admitted to *Business and economics* are less likely to hold debt and are less likely to be liquidity constrained. This pattern suggests that economic education reduces the probability of getting into financial problems by changing financial behavior and not by increasing the level or stability of income.

This paper is the first to show that extensive economic education can substantially reduce the risk of getting into financial problems. A large body of literature studies the association between economic education and a range of financial outcomes (Fernandes et al., 2014; Miller et al., 2015; Kaiser and Menkhoff, 2017, 2020; Kaiser et al., 2020). Many studies in this literature evaluate interventions in the context of lower levels of education and outside the formal education system, and many rely on financial literacy assessments, self-efficacy or hypothetical or self-reported behaviors (Bruhn et al., 2016; Skimmyhorn, 2016; Lusardi



et al., 2017; Bover et al., 2018). My paper is closest related to three papers that all exploit variation in high school requirements across U.S. states to study the causal effects of economic education on financial outcomes. First, Brown et al. (2016) use an event study design and find that additional mathematics training and financial education improves financial outcomes whereas additional economics training impairs the same outcomes. Second, Cole et al. (2016) exploit state-level variation in exposure to math and personal finance courses and find that additional mathematics training improves financial outcomes while personal finance courses have no effect. Third, Urban et al. (2020) use a difference-in-difference approach in which they compare students from three U.S. states that mandated financial education in high school with students from synthetic control states. They find that financial education improves financial outcomes but show that the effect size varies substantially across the states.

Summing up the evidence on the effect of education and information interventions on financial outcomes, including the three studies above, Beshears et al. (2018) conclude that the effects “are often small or null and appear to depreciate rapidly with time.” The current paper differs from the existing literature by studying a very extensive educational intervention and showing that economic education can indeed have substantial and long-lasting effects on financial outcomes. Furthermore, I evaluate the long-term effects of education and because I can also link the data on financial problems with data on several labor market outcomes, liabilities and peers, I can provide suggestive evidence on the channels that are important for my findings. The implication of my findings is not that everyone should study economics at university level, but rather that an extensive intervention can improve financial decision-making even for individuals with an affluent background and a relatively low baseline probability of getting into financial problems.

The paper is also related to a study by Christiansen et al. (2008). They show that graduating from an economics education affects the likelihood of participating in the stock market. My paper differs from their study by investigating how economic education affects the probability of getting into financial problems using a different research design. I replicate their finding on stock market participation using my identification strategy, which further supports the hypothesis that studying economics causally affects financial behavior.

Finally, this paper is also related to a growing body of literature that use GPA admission thresholds for the identification of educational effects, but I am the first to show that field of study in higher education have important impacts on financial behavior. Kirkeboen et al. (2016) use Norwegian data to estimate the early labor market payoff to field of study and institution and outline the importance of controlling for the alternative field of study in the estimation. The Danish admission data have also previously been used to study how admission to the first choice or preferred field affects educational outcomes, earnings, the timing of family formation and the gender gap in earnings (Heinesen, 2018; Humlum et al., 2017; Daly et al., 2021; Andersen et al., 2020). Other studies use Swedish (Öckert, 2010; Dahl et al., 2020a,b), Finnish (Silliman and Virtanen, 2021), Norwegian (Kirkebøen et al., 2021) and non-Scandinavian data (Altmejd et al., 2021; Hastings et al., 2014) to study labor market returns, family spillovers and assortative mating.

The paper proceeds as follows. Section 1.2 presents the institutional background and methodology and section 1.3 presents the data. Section 1.4 presents the results and discusses the potential channels that can explain the findings. Finally, section 1.5 concludes.

## **1.2 Institutional Background and Methodology**

### **1.2.1 Admission to post-secondary education**

In Denmark post-secondary education is free of charge and most students are eligible to public support from the State Educational Grant. Most students that start on a bachelor program at one of the eight Danish universities continue on a master program directly after obtaining the bachelor degree. Generally, it therefore requires 5 years of study to obtain a university degree. At the university colleges it normally takes 3.5 years to obtain a professional bachelors degree while the academy profession programs from the business academies requires 2 years of study.

The admission to higher education programs normally requires an Upper Secondary School Leaving Certificate and the admission process is administered by the Coordinated Admission under the Ministry of Higher Education and Science. The applicants can apply for and rank up to eight programs and each program is a combination of specific program and institution.

Admission to the programs is allocated through either the Quota 1 system or the Quota 2 system. The majority of slots are allocated through the Quota 1 system where the applicants are ranked based on their GPA from upper secondary school. The best ranked applicant gets his or her preferred choice, the second best ranked applicant gets his or her highest available choice and so on. The number slots is limited in most programs and if the number of applicants exceeds the number of slots, admission is restricted. This implies that applicants with a GPA above a certain threshold will be admitted to a particular program and applicants with a GPA below the threshold will be offered another program if any. It is important to notice that the applicants cannot know the specific thresholds at the time of application. Thereby, the Quota 1 admissions process generates locally unpredictable GPA thresholds that effectively randomize applicants near the thresholds into different programs and fields of study.

The Quota 2 admissions are allocated by the educational institutions based on criteria they select. These can be work experience, grades in particularly relevant subjects etc. If students apply for a program through the Quota 2 system, but fulfil the Quota 1 requirements, they will be admitted to the program through Quota 1. For a more detailed description of the admission process see Heinesen (2018).

### 1.2.2 Fuzzy Regression Discontinuity Design

The institutional setting described in the previous section enables me to estimate the causal effect of admission to a particular field of study on the probability of loan default using a fuzzy regression discontinuity design.

Imagine we have a group of individuals,  $i = 1, \dots, N$ , who all have the preferred field  $f_i^p = j$  and the alternative field  $f_i^a = k$ . The effect on an outcome,  $y$ , of being admitted to field  $j$  instead of field  $k$  can then be estimated by 2SLS for this sample:

$$D_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_i + \tilde{\beta}_2 x_i T_i + \tilde{\beta}_3 T_i + u_i \quad (1.1)$$

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i T_i + \beta_3 D_i + \varepsilon_i \quad (1.2)$$

where (1.1) is the first stage and (1.2) is the second stage. In the equations above  $x_i$  is the running variable, the distance to the GPA threshold.  $D_i$  is a dummy that equals 1 if individual  $i$  is admitted to field  $j$  and  $T_i$  is a dummy indicating whether  $i$ 's GPA is above the threshold of  $i$ 's preferred field,  $j$ .

## 1.3 Data

This section first provides an overview of the different sources of data I combine and how I select the sample used in the estimations. Second, it defines field of study and financial problems and gives a graphical illustration of the research design.

### 1.3.1 Data Sources

I combine third party reported Danish administrative data from the Coordinated Admission, Statistics Denmark and the Danish Tax Authorities in order to estimate the effects of field of study, and I link the data to individuals using a personal identifier.

From the Coordinated Admission, I have all applications for higher education programs in Denmark from 1993 to 2006. For each year and each applicant, I have information on their applications to different programs and institutions and how they rank their choices. From the Coordinated Admission, I also have information on the GPA threshold of each study program in the same period. The data does not include the individuals' GPAs from upper secondary school, but for the majority of the applicants, I obtain this information from Statistics Denmark. From Statistics Denmark, I also have information on income, assets, education, employment and demographic variables which I use in sensitivity analyses and to investigate potential channels that explain my main finding.

Finally, I also use data from the Danish Tax Authorities on the universe of personal loans. This data contains information on loan defaults from 2003 to 2016 and I use this data to determine whether individuals are in financial problems.

### 1.3.2 Sample selection and summary statistics

I study individuals who applied for a higher education program between 1993 and 2006. For the oldest cohort, I have loan information from 10 years after the year of application and for the youngest cohort I have loan information until 10 years after the year of application. In this period, I observe 1,364,480 applications from 531,033 individuals. I focus on first time applicants from age 18 to 30 and this leaves me with 427,885 applicants.

In order to implement the research design described in section 1.2.2, I can only use applicants with a binding GPA admission threshold in their local program ranking. The local program ranking is the subset of the applicant's total program ranking where the GPA effectively determines which program he or she is admitted to. This means I drop an applications if (i) a higher ranked program has a lower threshold, (ii) an applicant's GPA is below the threshold of a lower ranked program, (iii) the applicant's GPA is above the thresholds of at least two higher ranked programs, (iv) or if there is a binding threshold for a higher ranked program (see examples in Appendix Table A-1.1). I refer to the highest ranked program as the *Preferred* program (and later field of study) and the lowest ranked program as the *Alternative*. As in Kirkeboen et al. (2016), the preferred program is not necessarily the first priority program but rather the highest ranked program in the local program ranking where the GPA threshold is binding. This leaves me a sample of 53,882 applicants (see an overview of the selection process in table A-1.2).

Finally, the sample used in the analysis consists of the applicants whose preferred program and alternative program are within different fields of study (18,236) and who completes at least one higher education program within 10 years from the year of application and not two programs from different fields of study. I make this restriction in order to increase the probability that admission actually leads to studying, but I show that the results are robust to the inclusion of non-completing applicants and applicants that complete several fields. This gives me a sample of 14,181 applicants.

Table 1.1 shows summary statistics for all first time applicants, applicants with a binding GPA threshold and the analysis sample. In the table, we see that the sample used in the analysis is slightly younger than the pool of all first time applicants and the share of male applicants is 36.3% compared to 40.5% for all first time applicants. We also see that the

Table 1.1: Summary statistics

	1st time applicants		Binding threshold		Sample	
	Mean	SD	Mean	SD	Mean	SD
Age	21.8	2.4	21.4	1.8	21.1	1.8
Male (%)	40.5	49.1	32.2	46.7	36.3	48.1
GPA	8.3	1.0	8.6	0.9	8.8	0.8
1st priority threshold	5.8	4.1	8.8	0.7	9.1	0.6
Offered rank	1.2	0.6	1.6	1.0	1.8	1.1
Number of applications	1.9	1.2	3.1	1.3	3.2	1.4
Income rank	58.2	28.0	58.6	28.0	63.2	28.5
Father's income rank	67.0	28.4	69.2	28.2	70.7	28.2
Mother's income rank	47.3	26.0	51.0	26.5	52.7	27.1
Father has master (%)	13.3	33.9	18.9	39.2	22.8	41.9
Mother has master (%)	6.4	24.5	9.6	29.4	11.9	32.3
Default (%)	11.1	31.4	9.1	28.8	8.0	27.1
Observations	427885		53882		14181	

*Notes:* The *1st time applicants* are all applicants observed in the data. This includes applicants to business academies that use another admission system than universities and university colleges. The *Binding threshold* group are applicants whose preferred program has a binding GPA threshold. The *Sample* are the applicants whose preferred program has a binding GPA threshold and whose preferred program and alternative program are within different fields of study. *Income rank* is within cohort rank based on total income measured 10 years after application. Father's and mother's income rank are measured when they are 45 years old. *Default* is measured 10 to 23 years after application. Appendix Table A-1.3 shows the exact number of observations for each variable.

analysis sample have a higher average GPA from upper secondary school. Their GPA is almost 0.6 standard deviations higher. This is because the majority of the sample applicants apply for university programs whereas the group of all first time applicants also include applicants that apply for the shorter academy profession programs. This selection is also reflected in the applicants' backgrounds where we see that the sample applicants come from more advanced backgrounds when we look at the parents' incomes and educational levels. We also see that 10 years after the year of application, the analysis sample have a within cohort income rank that is 5 ranks higher than the average rank for all first time applicants.

When we look at the application pattern, the sample applicants apply for more programs, 3.2 on average compared to 1.9 for all first time applicants. Despite that the sample applicants have a higher GPA from upper secondary school, they are admitted to programs that they rank lower. This is because they apply for more competitive programs. Their first priority program has an average threshold of 9.1 where it is only 5.8 for all applicants.

In the final row, the table shows the default rate for the three groups, which is the indicator for being in financial problems. Here we see that the applicants used in the analysis have a lower probability of experiencing default. One explanation could be the fact that they end up having higher incomes themselves, or because their parents on average have higher incomes. The important take away from the table is that the analysis sample is an advantaged group of applicants. Therefore we might a priori not even expect field of study to affect the probability of default since it is already low for this group and because they come from relatively affluent backgrounds.

### 1.3.3 Fields of study

I define 8 fields of study. I use the broad fields (level 4) of the Danish International Standard Classification of Education (ISCED) classification provided by Statistics Denmark (2020) as a starting point, but make some adjustments for two reasons: The first is that there is no broad ISCED field with a focus on business and economics. The second reason is that I want the fields to resemble the fields used in the previous field of study literature to ease comparison.

Here it is worth emphasizing the difference between *fields* and *programs*. If we take economics as an example, then a specific program would be Economics at the University of Copenhagen. If this program has more applicants than slots, this generates a GPA admission threshold for the program. In the ISCED classification, this program is a part of the detailed field (level 2) *Economics*, the narrow field (level 3) *Social and behavioural sciences* and the broad field (level 4) *Social sciences, journalism and information*. This means that a field does not have single GPA threshold, but the applicants have different thresholds for the same preferred field of study depending on what specific program they prefer.

Appendix Table A-1.4 illustrates how I construct the 8 fields of study based on the ISCED classifications. The main field in this paper is the *Business and economics* field. To construct this, I use pool the narrow field *Business and administration* with the detailed fields *Economics*, *Agricultural economics* and *Mathematical economics*. The programs in the *Business and economics* field are very similar to the programs within economics, accounting or finance used by Chetty et al. (2014).

Table 1.2: Combinations of preferred and alternative field of study

	Alternative field of study							Tot.
	B&E	Law	HAA	Educ.	Welf.	SocSci	STEM	
Preferred field								
Bus. & Econ.		47	344	< 25	< 25	214	162	767
Law	806		490	28	71	150	127	1672
Hum., Art & Arch.	196	79		227	238	591	444	1775
Education	41	< 25	338		416	92	89	976
Welfare	163	55	410	425		162	489	1704
Social Science	730	477	2593	260	429		327	4816
Medicine	127	121	161	< 25	518	73	1061	2061
Total	2063	779	4336	940	1672	1282	2699	13771

*Notes:* The table is based on all applicants whose preferred field of study has a binding threshold and whose preferred and alternative field of study differ. The rows show the number of applicants in the sample that prefer each field and the columns show the number of applicants that have each field as their alternative. The “< 25” cells are excluded from the totals.

No direct link exists between the programs in the application data from the Coordinated Admission and the DISCED programs in the educational data from Statistics Denmark. To establish this link, I use all applicants admitted to a each program in the Coordinated Admission data and investigate which program they are enrolled in six months later according to the educational data from Statistics Denmark. In each year, I then establish the link by determining the most common program according to Statistics Denmark among the admitted applicants.

It is important to notice that the program composition of the fields can be different based on whether the field is the preferred or the alternative. Appendix Table A-1.5 shows the most common programs for each field of study depending on whether the field of study is the preferred or alternative. For instance, if *Business and economics* is the preferred field of study then the most common programs are *Business economics and language* (29%), *Business economics* (15%) and *Language and international marketing* (9%). On the other hand, if *Business and economics* is the alternative field then the most common programs are *Economics* (38%), *Business economics and law* (22%) and *Business economics* (20%).

Table 1.2 shows the sample applicants’ preferred and alternative fields of study. I drop applicants who have *STEM* as their preferred field or *Medicine* as their alternative, since there are only 265 and 98 applicants in these groups. If we again focus on the *Business and economics* field, then the first column shows that 2063 applicants have this field as their



alternative field. Almost 75% of these applicants have *Law* or *Social science* as their preferred field of study. This is important to keep in mind when we interpret the effects of admission to *Business and economics* when this is the alternative field in section 1.4.1. In section 1.4.1, I will also analyse the effect by preferred field of study for these applicants. If *Business and economics* is the preferred field, then 73% have *Humanities, art and architecture* or *Social science* as their alternative field.

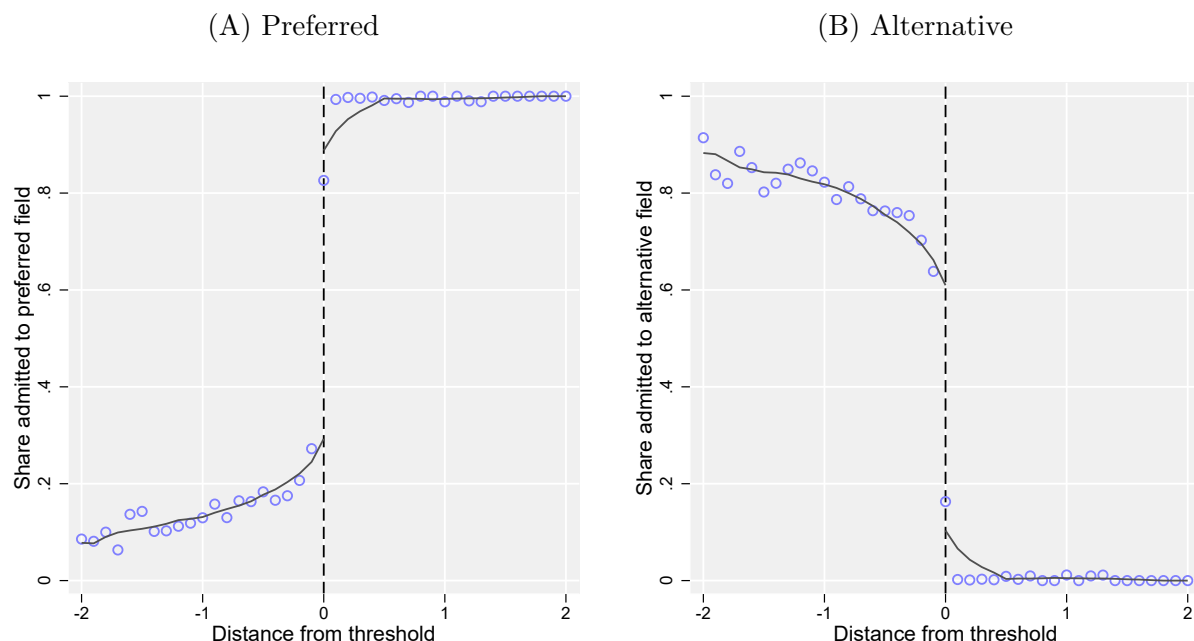
### 1.3.4 Graphical illustration of research design

Figure 1.1 shows the change in the applicants' probability of being admitted to their preferred or alternative field of study based on the distance to the GPA admission threshold of their preferred field. Appendix Figure A-1.1 also shows sharp discontinuities in enrolment after one year and completion of a program from the preferred or alternative field of study within 10 years of application.

Panel A of Figure 1.1 shows that the probability of admission to the preferred field is not zero if an applicant's GPA falls below the threshold. As discussed in section 1.2.1 this is due to the Quota 2 system, where admission is not only determined by GPA. The probability of admission increases with the GPA below the threshold because the GPA is also taken into consideration in the Quota 2 system.

The panel also shows that almost all applicants with a GPA strictly above the threshold are admitted to the preferred field while it is around 80% for the applicants who are exactly at the threshold. There are two explanations for this. First, the GPA is measured with one decimal's precision in the data, but the educational institutions may have more precise information than this. Second, if all applicants at the threshold cannot be admitted, it is decided either by lottery or age who will be admitted. Only some programs use the age criterion, where they admit the oldest applicants. In the graphical illustrations, I will not take this into account, but in all regressions I will use the age criterion where it is possible to characterize whether an applicant is above or below the threshold. Furthermore, I show that the main result is robust to using a donut regression discontinuity design where I drop all applicants exactly at the threshold from the estimation.

Figure 1.1: Share of applicants admitted to preferred or alternative field of study



*Notes:* The figures are based on all applicants whose preferred field of study has a binding threshold and whose preferred and alternative field of study differ. The y-axes show the probability of admission to either the preferred (Panel A) or alternative (Panel B) field of study. The x-axes show the distance to the GPA admission threshold (0) of the preferred field of study. The bin width in the two panels is the smallest possible, 0.1. The local linear polynomials have a bandwidth of 0.5.

Panel B of Figure 1.1 shows the probability of being admitted to the alternative field. It clearly mirrors Panel A and also shows a sharp discontinuity in the probability of being admitted to the alternative field at the GPA threshold.

Both of these clear discontinuities enable me to estimate the causal effect of admission to a particular field on different outcomes. I exploit whether an individual is above or below the GPA threshold as an instrument for admission to the preferred or alternative field in a fuzzy regression discontinuity design as described in section 1.2.2.

Appendix Figure A-1.2 shows the distribution of the applicants' distances to the threshold of the preferred field of study. The figure shows no evidence of manipulation of the running variable, i.e. the applicants cannot sort themselves above the threshold in order to be admitted to their preferred field. Similarly, Appendix Table A-1.6 shows that formal manipulation tests do not indicate manipulation either.

### 1.3.5 Financial problems

As described in section 1.3.1, I use data from the Danish Tax Authorities to measure whether people are in financial problems. For each personal loan, the data indicates if the debtor is at least 60 days late with payments on the loan at the end of the year. This is reported by banks and other financial intermediaries to the tax authorities in order for them to verify that tax deductions for interest payments are correct.

I generate an indicator that equals one if an applicant has a loan in default at some point in time 10 years or later after the year of application. This means that I observe all application cohorts from 1993 to 2006 in the default data at least one year, but I observe the oldest cohort up to 23 years after the year of application.

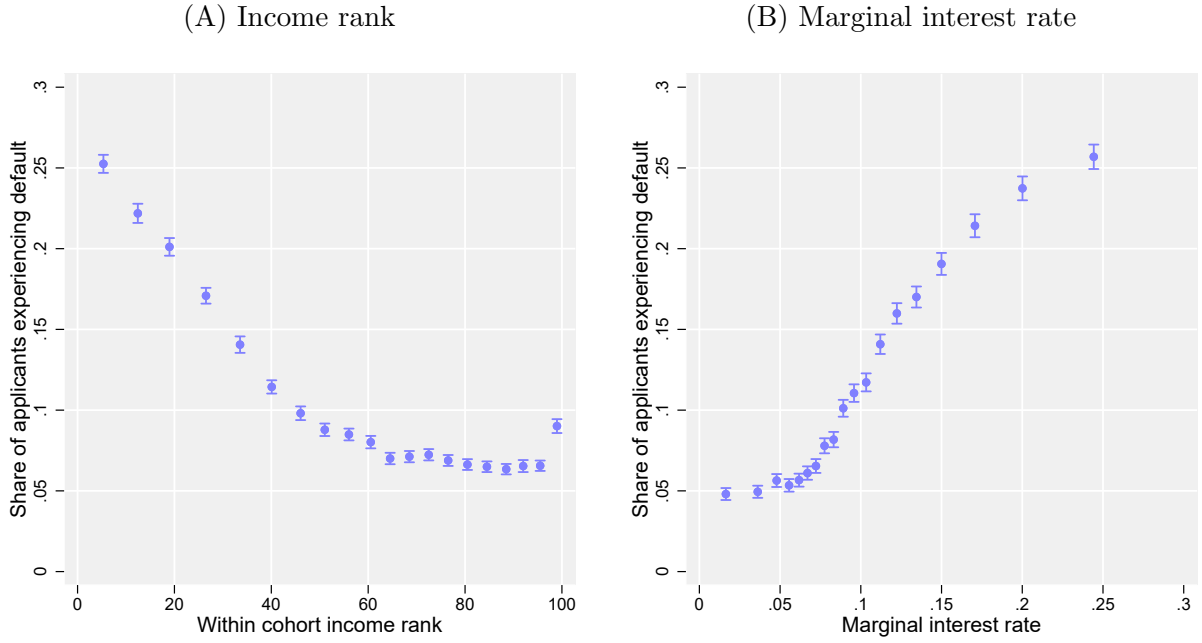
This quantification of financial problems is similar to the method used by Kreiner et al. (2020) based on the same source of data from 2004 to 2011. They also use information on financial problems from two credit bureau companies and conclude that their findings using the data from The Danish Tax Authorities are not confined to this specific measure of financial problems.

Panel A of Appendix Figure A-1.3 shows how the share of applicants in default varies across completed fields of study from 7.0% for *Medicine* to 10.3% for *Humanities, Art and Architecture*. For *Business and Economics* the share of applicants who experience default is 9.6%. It is important to note that there are large differences between detailed and narrow fields within the broad fields of study. For instance, for *Economics* the average probability of default is 6.0% while it is 9.8% for *Business and administration*.

Panel B of Appendix Figure A-1.3 shows that the outstanding amounts on accounts in default are non-trivial. 40,927 first time applicants had debt in default in 2016. The median outstanding amount in default for these applicants is 13,506 DKK (approximately 1800 Euro) and the distribution is highly right skewed with a mean outstanding amount in default of 147,452 DKK (almost 20,000 Euro).

Panel C and D of Appendix Figure A-1.3 show the evolution of default after application. Both panels show that the default probability is low right after application, increases during studies and early work life and somewhat stabilizes after 10 to 12 years.

Figure 1.2: Correlation between default and income rank and marginal interest rate



*Notes:* The figures are based on all first time applicants. Default status is observed from 10 years after application until 23 years after application for the oldest application cohorts. Within cohort income rank is measured 10 years after application while the marginal interest is the average marginal interest rate from 7 to 9 years after application. There are the same number of applicants in the 20 bins in each panel. In the right panel, I have left out the outlier top bin.

Figure 1.2 shows how the probability of experiencing default co-varies with income and the marginal interest rate. These are two potential channels that could be affected by field of study and explain why field of study affects the probability of default. A third channel is unemployment which Kreiner et al. (2020, p. 250) show affects the default probability.

Panel A shows that applicants who are below the median income within their cohort 10 years after application are more likely to experience default and the probability increases the lower they are ranked in the income distribution.

In Panel B, I follow Kreiner et al. (2019) and use the marginal interest rate as a continuous measure of how liquidity constrained the applicants are. The marginal interest rate is defined as the highest interest rate an individual pays on a single loan in a year and the interest rate is calculated as the interest paid during the year divided by the mean of the outstanding amount in the beginning of the year and at the end of the year. The marginal interest rate is also generated using the data from the Danish Tax Authorities.

The panel shows that applicants who are more liquidity constrained 7 to 9 years after application according to this measure, are also more likely to experience default 10 years or later after the year of application.

The correlations presented in this section show that the default indicator is a valid measure of financial problems, that these financial problems are non-trivial and that they vary across fields of study. In section 1.4.1, I will present results on the causal effect of field of study on the probability of getting into financial problems, and in section 1.4.2, I explore the three potential channels income, financial behavior and labor market outcomes and discuss the role of peers.

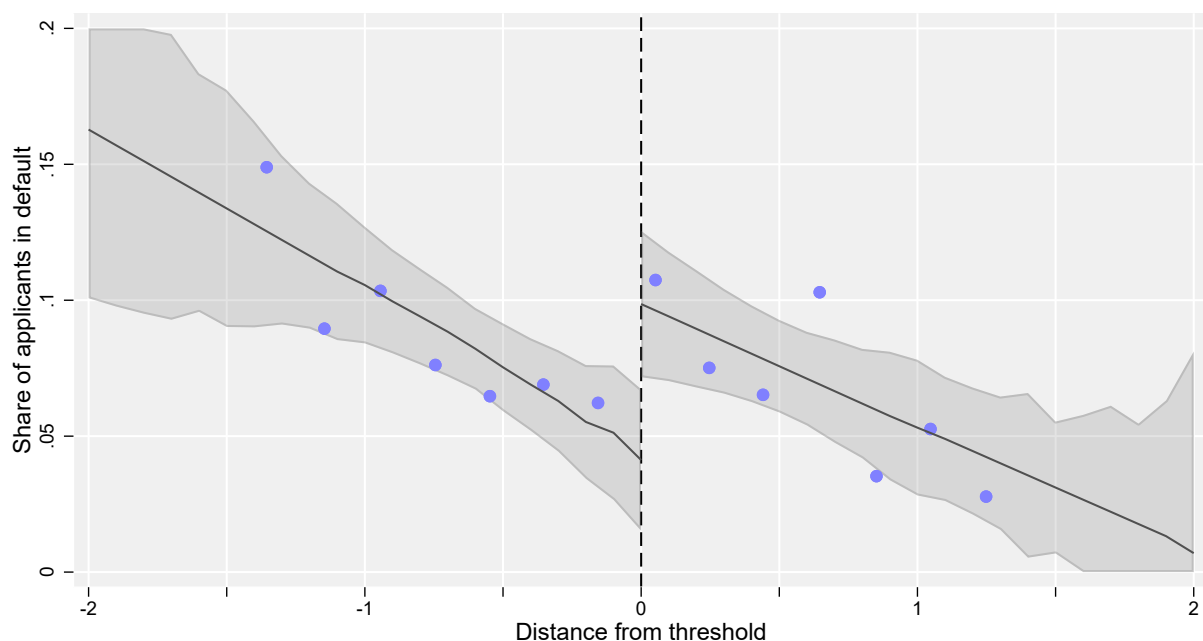
## 1.4 Results

### 1.4.1 Economics and financial problems

In this section, I present evidence on the effect of admission to *Business and economics* for applicants who have this as their alternative field. The models estimated are very similar to equation 1.1 and 1.2, but instead of only using applicants who have preferred field  $f_i^p = j$  and alternative field  $f_i^a = k$ , I first pool all preferred fields and compare them to *Business and economics*. This group of applicants is particularly interesting, since the applicants would actually prefer to study something else than *Business and economics*, but they are “pushed” into this field if their GPAs are below the admission thresholds to their preferred fields. In section 1.4.1, I explore the effect by preferred field as in equation 1.1 and 1.2 for applicants who prefer *Social science* or *Law*.

Figure 1.3 shows the graphical reduced form evidence on the effect of admission to *Business and economics* on the probability of default. The figure shows that a higher GPA is associated with a lower probability of default and the effect is similar above and below the threshold. Exactly at the threshold, there is a clear jump in the probability of default, such that the applicants who are just above the threshold of their preferred field of study are approximately 6%-points more likely to experience default than the applicants who are

Figure 1.3: Admission to *Business and economics* and the probability of default



*Notes:* The figure is based on applicants with *Business and economics* as their alternative field of study and another field as their preferred. On the x-axis, 0 is the admission threshold of the preferred field of study. Therefore, applicants with GPAs below the threshold are more likely to be admitted to *Business and economics*. The y-axis shows the probability of default 10 years or more after application. The local linear polynomials have a bandwidth of 2 and I use a rectangular kernel. The grey area indicates the 95% confidence intervals for the local linear polynomials. The confidence bands are winsorized at 0 and 0.2. The bin width is 0.2, but I only plot bins with at least 30 observations.

just below the threshold. The 95% confidence bands also indicate that this jump in the probability is statistically significant.<sup>2</sup>

Panel (A) of Appendix Figure A-1.5 shows the first stage, namely the probability of admission to the *Business and economics* field based on the distance to the GPA threshold of the preferred field. The figure shows a pattern very similar to Panel B of Figure 1.1: The applicants above the threshold are very unlikely to be admitted to the *Business and economics* field, whereas the applicants just below the threshold are much more likely to be admitted to *Business and economics*.

In line with the graphical evidence in Appendix Figure A-1.5, column (1) of Table 1.3 shows that being above the threshold reduces the probability of admission to *Business and economics* with 71.2%-points. Column (2) shows the estimated reduced form effect of being

<sup>2</sup>Appendix Figure A-1.4 and shows different versions of Figure 1.3. Panel B shows the figure with four bins of roughly the same size on each side of the threshold, Panel D shows the figure with a smaller bin width of 0.1, which is the smallest bin width possible, and Panel F shows the figure with a local linear polynomial with a bandwidth of 1.

Table 1.3: Admission to *Business and economics* and the probability of default

	Admission (%)	Probability of default (%)				
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	-71.2*** (2.3)	5.5*** (2.0)				
$\mathbb{1}(\text{Admission B\&E})$			-7.7*** (2.8)	-8.1*** (2.9)	-7.8*** (2.8)	-8.2*** (2.8)
$N$	1983	1983	1983	1983	1983	1983
Preferred field FE				✓	✓	✓
Sex, YOA and age					✓	✓
Income						✓

*Notes:* In the estimations, I use applicants with *Business and economics* as their alternative field of study and another field as their preferred for whom I observe all the control variables. Column (1) and (2) show the first stage and reduced form effects in %-points estimated with OLS. In column (3)-(6), I instrument admission to *Business and economics* with the indicator for being above the threshold of the preferred field in 2SLS estimations. *Preferred field FE* are fixed effects for the preferred field of study. *Sex* is a male indicator, *YO*A is year of application fixed effects, and *age* are indicators for age in the year of application. *Income* is included as dummies for five levels of the average total income 7 to 9 years after the year of application. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

above the threshold on the default probability as Figure 1.3. The effect is a 5.5%-point increase in the probability of default if an applicant is admitted to his or her preferred field of study instead of *Business and economics*.

Column (3) shows the 2SLS estimate of the effect of admission to *Business and economics* on the probability of default. I find a 7.7%-point decrease in the probability of default which is the local average treatment effect of admission. Given an average default probability of 8.0% for this pool of applicants, it is a substantial decrease and the graphical evidence in Figure 1.3 suggest a local reduction of one half in the default probability near the threshold.

In column (4)-(6), I include preferred field fixed effects, flexible dummy controls for the predetermined characteristics sex, year of application and age at application, and finally flexible dummies for average income levels 7 to 9 years after the year of application. The 2SLS estimates show that adding these additional controls does not change the estimated effect in column (3), which indicates that applicants just above and just below the threshold are very similar. This is confirmed by Panel B of Appendix Figure A-1.5 which shows no evidence on covariate imbalance around the threshold using the predicted probability of default based on predetermined characteristics. As discussed in section 1.3.5, income itself is

potentially affected by field of study and is therefore a “bad control”, but it is reassuring to see that the findings are robust to the inclusion of the income controls. To avoid controlling for income measured at the same time as default, I control for the income level 7 to 9 years after application which is before I measure default. In section 1.4.2, I investigate the income channel further and show that there are no income changes for these applicants around the threshold.

Appendix Table A-1.7 shows that the estimated effects in Table 1.3 are robust to using all observations, i.e. dropping the bandwidth, using a smaller bandwidth of 1, using a triangular kernel instead of a rectangular kernel, using quadratic polynomials for the running variable, and excluding observations exactly at the threshold in a donut regression discontinuity design.

I explore the sensitivity of the estimates to the choice of bandwidth further in Appendix Figure A-1.6. The figure shows the reduced form effect of being above the threshold as in column (2) of Table 1.3 with varying bandwidth and demonstrates that the estimated effects are relatively stable across bandwidth choices. The confidence interval widens as the bandwidth gets smaller but the estimated effect is significant on the 10%-level until a bandwidth of 0.9, which is close to the bandwidth suggested by the McCrary (2008) procedure (0.84). The bandwidth suggested by the Calonico et al. (2017) procedure is 0.47. At this bandwidth, the estimated effect is not significant but the point estimate is similar to the estimate using the McCrary bandwidth.

Appendix Table A-1.8 shows that the results are also robust to using the definition of economics from Chetty et al. (2014) and to including non-completers, who never complete a higher education. The table also shows that the estimated effect is significant if I only look at default in 2016 (the only year where I observe the default indicator for all applicants) instead of pooling all observed years. The effect is still significant on the 10%-level if I restrict to accounts in default with an outstanding amount of more than 4000 DKK (500 Euro).

As a final validity check, I investigate whether I detect significant jumps in the default probability at placebo thresholds and for placebo fields. In Appendix Figure A-1.7, I use the same specification as in column (2) of Table 1.3 and vary the threshold from -1 to 1 in steps of 0.1. In Panel A, I use a bandwidth of 2 and the panel shows that there are no significant jumps in the probability of default for placebo thresholds far from the true threshold. Right



below the true threshold, the jumps at the placebo thresholds are significant at the 10%-level. Therefore, I shrink the bandwidth window to 0.9 in Panel B (the smallest bandwidth for which I detect a significant jump at the true threshold with a P-value of 0.060). The point estimate at the true threshold is almost unchanged but the point estimates at the placebo thresholds decrease and all become insignificant at the 10%-level.

Next, I investigate if I find similar effects of being admitted to other fields of study than *Business and economics*, the “placebo fields”. I do this by estimating the effect of admission to each field of study where I pool applicants who prefer the field and have it as the alternative. I generate a variable,  $Z_i$ , that equals 1 if an applicant prefers field  $j$  and is above the threshold,  $T_i = 1$ , or if an applicant has field  $j$  as the alternative and is below the threshold,  $T_i = 0$ . I then estimate the following equation:

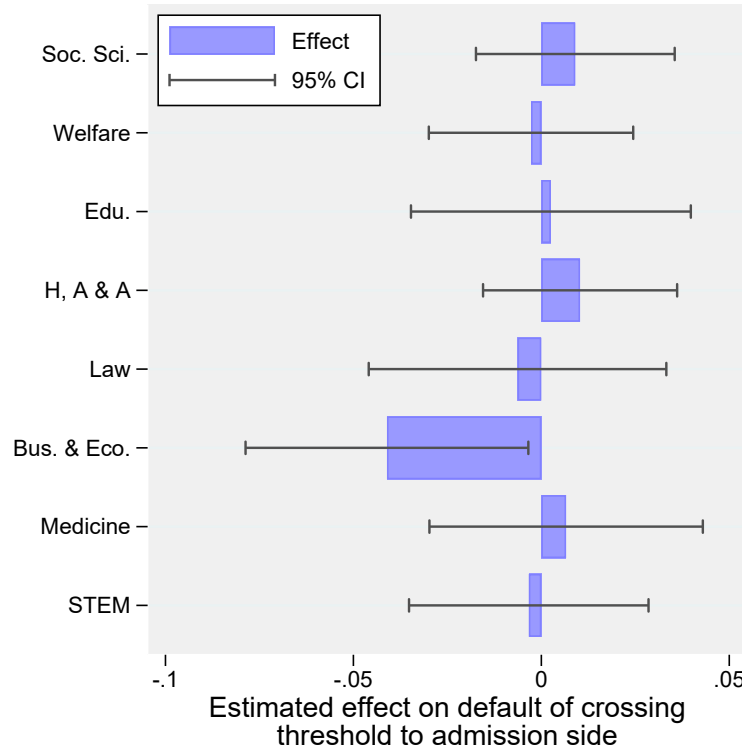
$$y_i = \beta_1 x_i + \beta_2(x_i \times p_i) + \beta_3(x_i \times T_i) + \beta_4(x_i \times p_i \times T_i) + \beta_5 T_i + \beta_6 Z_i + \phi f_i^p + \psi f_i^a + \varepsilon_i \quad (1.3)$$

where I allow for different effects of the running variable above and below the threshold and depending whether field  $j$  is preferred ( $p_i = 1$ ) or the alternative ( $p_i = 0$ ). I also include fixed effects for preferred and alternative field,  $f_i^p$  and  $f_i^a$ .

Figure 1.4 shows the estimated  $\beta_6$  from separate estimations of equation 1.3 field by field. These estimates can be interpreted as the reduced form effect of crossing the threshold to the admission side of the given field of study, no matter if it is the preferred or alternative. The figure shows that the only field where crossing the the threshold to the admission side has a significant effect on the default probability is *Business and economics*. Crossing the threshold to the admission side of the other fields of study has no significant effect and the point estimates are all close to 0.

To conclude, the estimated effects of admission to *Business and economics* are very consistent across the different specifications with a reduced form effect of being above the threshold of the preferred field of 4-6%-points and a local average treatment effect of 6-8%-points. At the threshold, the probability of default is reduced by almost one half. This effect size is in line with the effects found by Urban et al. (2020) who estimate the effect of financial education in three US states and find that the probability of being 30 days and 90+ days

Figure 1.4: Admission to different fields of study and the probability of default



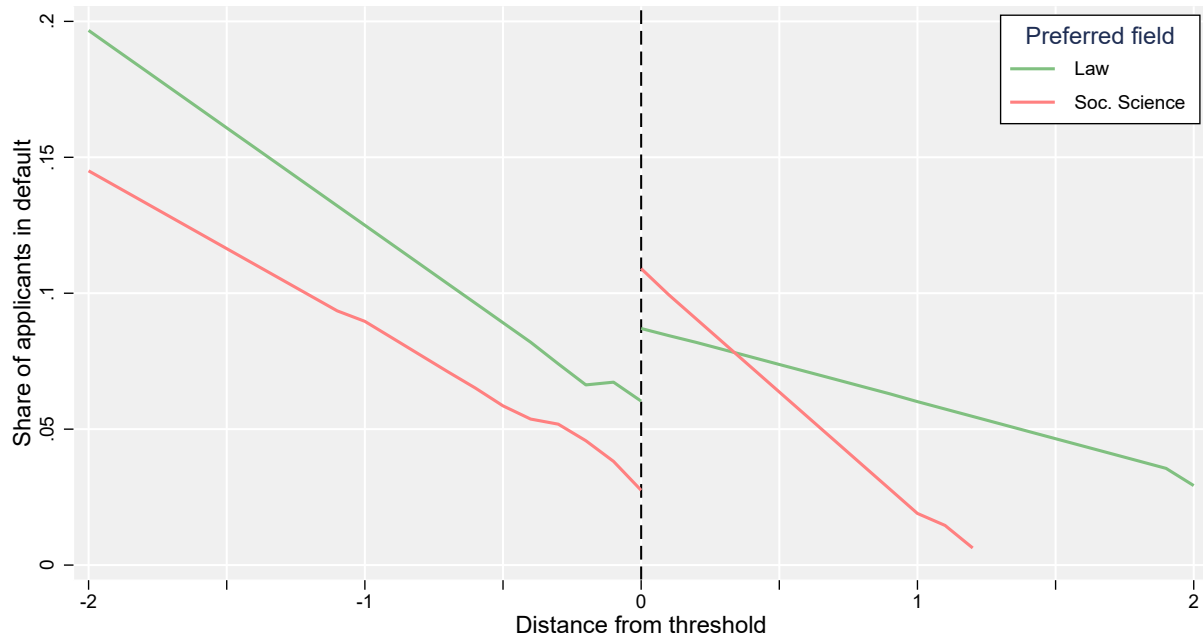
*Notes:* The bars plot the estimated reduced form effect of crossing the threshold to the admission side of each field of study. Each field's effect is estimated separately as in equation (1.3) where I pool applicants who have the field  $j$  as the preferred field of study, and another as their alternative, with applicants who have the field  $j$  as the alternative field of study, and another field as their preferred.

behind on an account decreases. The effects differ by state and increase the longer the policy has been implemented, but the declines are between 30% and 50% of the baseline rates and up to 95% for the 90+ days delinquency outcome.

The other studies using high school curriculum reforms have found smaller effects of education on the probability of getting into financial problems. Brown et al. (2016) find that requiring a one-year increase in math education for graduation in high school reduces the probability of having had a balance in collections with 0.6%-points from a baseline of 40%. Cole et al. (2016) estimate the effect of increasing high school math requirements and find a decrease in the probability of foreclosure of 0.4%-points from a baseline of 9%.

Even among the three studies that all rely on high school curriculum changes for identification, the magnitude of the treatment effects differ, which emphasizes that the specific contexts are very important. Since the treatment I study in this paper is very extensive compared to the studies cited above and involves education within both mathematics, economics and

Figure 1.5: Admission to *Business and economics* vs. *Law* and *Social science* and default



*Notes:* The figure is based applicants who have *Business and economics* as their alternative field of study and either *Law* or *Social Science* as their preferred field. The y-axis shows the probability of default 10 years or more after application. The x-axis shows the distance to the GPA admission threshold (0) of the preferred field of study. The local linear polynomials use rectangular kernels with a bandwidth of 2.

personal finance at the same time, it is not surprising that the effects I find are larger. One interpretation of the differences is that numerical training and financial knowledge are complements and providing education within both areas at the same time improves financial behavior. Another interpretation is that students in higher education institutions are better at transferring the financial education into financial decision-making than high school students. In any case, the results in this section show that extensive economic education can have large and lasting effects on financial behavior.

### Effect by preferred field of study

In table 1.2, we saw that a large share of the applicants, who have *Business and economics* as their alternative field, have *Law* or *Social science* as their preferred field. Figure 1.5 shows the effect of being above the threshold on the probability of default as in Figure 1.3 but separately for applicants who have *Law* or *Social science* as their preferred field of study.

For both groups of applicants, the figure shows an increase in the default probability right around the threshold but the jump is larger for the applicants who have *Social science* as their

preferred field. Appendix Table A-1.9 shows the formal estimation of the size of the jump at the threshold and the table shows that the jump in the probability is only statistically significant for the applicants with *Social science* as their preferred field.<sup>3</sup>

Figure 1.5 also shows negative slopes for all the local linear polynomials as in Figure 1.3. Below the threshold, the slopes are similar since the majority of applicants are admitted to the same field but there is still a level difference. This suggests that applicants with the same alternative field are very likely different “types” if they have different preferred fields. The two groups also have a different composition of programs within the *Business and economics* field. For instance, the most frequent program for applicants with *Law* as the preferred field is *Business economics and commercial law* (42%) while the most frequent program for applicants who prefer *Social science* is *Economics* (52%).

Next, I extend the analysis by looking at all combinations of preferred and alternative fields.<sup>4</sup> I jointly estimate the reduced form effect for each combination by estimating the following equation:

$$y_i = \sum_a \beta_{1a}(x_i \times f_i^a) + \sum_p \beta_{2p}(x_i \times T_i \times f_i^p) + \sum_a \sum_{p \neq a} (\beta_{3ap}(T_i \times f_i^a \times f_i^p) + \pi_{ap}(f_i^a \times f_i^p)) + \varepsilon_i \quad (1.4)$$

I do not allow for separate effects of the running variable above and below the threshold for all combinations of fields due to the limited number of observations. Instead, I assume that the effect of the running variable below the threshold is the same for each alternative field no matter what the preferred field is. Similarly, I assume that the effect of the running variable above the threshold is the same for each preferred field no matter what the alternative field is. Still, I do allow for level differences for each combination of fields.

The estimates of the different  $\beta_{3ap}$  are presented in Table 1.4.<sup>5</sup> The first column shows the estimated effects of being above the thresholds to different preferred fields for applicants with *Business and economics* as their alternative. All point estimates are positive for this group

<sup>3</sup>Column (1) and (4) of Appendix Table A-1.9 show that the first stage reduction in the probability of being admitted to *Business and economics* is almost 25%-points larger for the *Social science* group. This is because more applicants are admitted to *Law* through Quota 2 (10.5%) than to *Social science* (5.9%).

<sup>4</sup>Appendix Figure A-1.8 shows the distribution of the F-statistics from “first stage” regressions for each combination of preferred and alternative field. For three combinations (*Education* preferred and either *STEM*, *Business and economics* or *Social science* as alternative) the F-statistic is below 10. Therefore, I leave out these three additional field combinations from the estimation.

<sup>5</sup>In the appendix, table A-1.10 shows the corresponding local average treatment effects from a joint two-stage least square estimation.

Table 1.4: Joint estimation of the effect on default of being above the GPA threshold

	Alternative field						
	B&E	Law	HAA	Educ.	Welf.	Soc.	STEM
Preferred field							
Bus. & Econ.		-9.2 (8.5)	-2.4 (3.6)			-4.4 (4.8)	1.9 (5.3)
Law	4.2* (2.5)		-7.5** (3.0)	-6.0 (10.7)	2.0 (6.8)	6.3 (5.0)	-10.2** (5.1)
Hum., A. and A.	4.4 (4.4)	4.6 (7.1)		-1.9 (4.4)	-0.2 (4.4)	-1.4 (3.0)	2.1 (3.2)
Education			1.5 (3.7)		0.2 (3.5)		
Welfare	8.7* (5.3)	6.7 (7.7)	-5.1 (3.2)	-2.0 (3.8)		-2.4 (5.0)	1.3 (3.2)
Social Science	7.9*** (2.7)	-0.5 (3.7)	-2.7* (1.5)	2.8 (4.4)	2.9 (3.5)		5.7 (3.5)
Medicine	11.9** (6.0)	8.0 (5.7)	3.5 (4.6)		-1.8 (3.0)	6.3 (6.7)	0.2 (2.4)
<i>N</i>	1948	764	4185	858	1530	1125	2490
Default rate	8.1	7.9	9.0	6.9	6.4	9.0	6.7

*Notes:* The table shows the reduced form effects in %-points on default of being above the GPA threshold in a joint estimation of all combinations of preferred and alternative field of study. In the estimation, I use all applicants whose preferred field of study has a binding threshold and whose preferred and alternative field of study differ. The number of observations in the estimation is 12,900. *N* indicates the number of applicants in the estimation with the alternative field of study in the column. The *Default rate* indicates the probability of experiencing default for the applicants who have the field of study in the column as their alternative field of study. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

of applicants meaning that admission to *Business and economics* reduces the probability of default no matter what the preferred field is, and most estimates are significant on the 10%-level. For the other fields we do not see any systematic patterns related to the probability of default.

### Heterogeneous effects of admission

Appendix Table A-1.11 investigates whether the treatment effects differ for men and women, by parental background, by time of application and by program.

Panel A shows the effect of being above the threshold for men and women who have *Business and economics* as their alternative field, separately, and I find that the effect is largest for men. The effect is statistically significant on the 5%-level for men, but not for women. In a pooled regression where the dummy for being above the threshold of the preferred field is interacted with the male dummy, the table again shows that the effect is largest for men, but the difference between men and women is not statistically significant.

Men have a higher baseline probability of experiencing default, which might explain the difference, but men and women also have different application patterns. If we look at the three most common specific programs that the applicants apply for within the *Business and economics* field, the share of male applicants is around 60% for *Economics* and *Business economics* while it is only 43% for *Business economics and law*.

Panel B shows the effect split by whether the applicant's father has a higher education or not. The table shows that effects are very similar for the two groups. The point estimate is slightly higher for applicants whose father has a higher education but the estimates are not significantly different from each other and both are significantly different from 0 at the 10% level.

Next, Panel C splits the sample into applicants who applied before and after year 2000. The panel shows that the effect seems to be driven by the application cohorts who applied before 2000. Again, this may be because they have a higher baseline probability of default by construction since I observe them for a longer period in the data from the Danish Tax Authority. If we again look at the specific programs, the same pattern appears as when I split by gender: The cohorts who applied before 2000 are overrepresented in *Economics* and *Business economics* (55%), while they are underrepresented in *Business economics and law* (38%).<sup>6</sup>

Looking at the composition of specific programs across the groups studied so far, the effect seems to be strongest when *Economics* is the alternative. Panel D examines this directly and shows that indeed the effect is strongest for those who have *Economics* as their alternative. Again, it is important to keep in mind that there are differences in the preferred field of study between those who have *Economics* as the alternative and those who have for instance

---

<sup>6</sup>As described in the notes for Appendix Table A-1.2, from 2000 and onwards I observe applicants who went to commercial upper secondary schools.

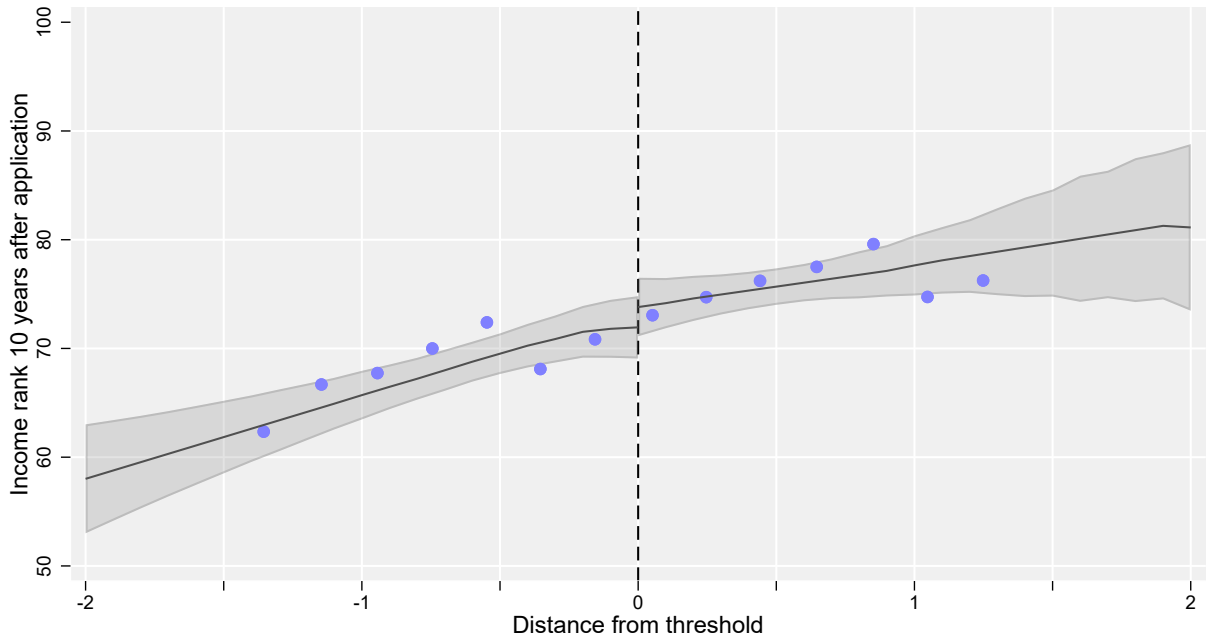
*Business Economics and law* as the alternative. Furthermore, a concern is whether income can explain this finding. In the next section, I show this is not the case.

### 1.4.2 Channels

#### Income

A key concern for the interpretation of the findings on the default probability is whether studying *Business and economics* leads to a higher income and that this drives the reduced the risk of default. For instance, Bleemer and Mehta (2021) use a regression discontinuity design to estimate the return to majoring in economics when this is the preferred major at the University of California, Santa Cruz, and find a 58% increase in the students' early career wages. This is in a different educational context and economics is the preferred field while the alternative is often sociology or psychology. I focus on applicants who have *Business and economics* as their alternative field and would often prefer to study political science or law, which are both high paying master degrees in the Danish context.

Figure 1.6: Admission to *Business and economics* and income rank



*Notes:* The figure is based on applicants with *Business and economics* as their alternative field of study and another field as their preferred. The y-axis shows the income rank within birth cohort 10 years after application. The x-axis shows the distance to the GPA admission threshold (0) of the preferred field of study. The local linear polynomials use rectangular kernels with a bandwidth of 2. The grey area indicates the 95% confidence interval. The bin width is 0.2, but I only plot bins with at least 30 observations.

In Table 1.3, I controlled directly for income, but Figure 1.6 shows the reduced form effect of being above the threshold on income rank 10 years after application when *Business and economics* is the alternative field. The figure shows no indication of a discontinuity around the threshold. Appendix Table A-1.12 shows the estimated effects on labor income, total income, household income, income rank and on the probability of being in the top or bottom of the income distribution. I find no statistically significant effects for any of these measures of income.<sup>7</sup> I also look at different sub periods (1 to 9 years after application, which includes income during studies, and 10 to 12 years after application), but I find no evidence that being admitted to the *Business and economics* field leads to a higher income when this field is the alternative.

## Financial behavior

A potential channel that could explain why admission to *Business and economics* reduces the risk of getting into financial problems is that it affects the financial behavior of the applicants. For instance, in order to default, you need to have debt, and if you do not hold sufficient liquid assets, even small shocks to your personal economy could make you unable to service your debt.

Table 1.5 shows the effect of being above the threshold of the preferred field on the probability of being liquidity constrained and the debt behavior of the applicants with *Business and economics* as the alternative. Column (1) shows the reduced form effect on the probability of having bank deposits smaller than one month of total income. The table shows that being above the threshold increases the probability of having a low level of bank deposits with 7.8%-points where 46.2% of the applicants on average have a low level of bank deposits. Column (2) shows the effect on the probability of having a high marginal interest rate (greater than 30%). I find an increase of 6.0%-points at the threshold, which is a large effect given that only 4.1% have such a high marginal interest rate. Both of these estimates indicate that admission to *Business and economics* makes the applicants less liquidity constrained 10 years after application.

---

<sup>7</sup>Appendix Table A-1.13 shows that the effect is not different for those applicants who have the *Economics* as their alternative



Table 1.5: Admission to *Business and economics* and financial behavior

	Low bank deposit (1)	High marg. int. rate (2)	Bank debt (3)	Debt-to- inc. ratio (4)	Has pub. debt (5)	Stock part. (6)
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	7.8** (3.6)	6.0*** (2.0)	5.8** (2.7)	4.2 (3.5)	4.4* (2.5)	-6.2* (3.3)
$N$	1959	1069	1959	1629	1952	1959
Average outcome	46.2	4.1	83.2	22.6	13.1	27.6

*Notes:* The table shows the estimated reduced form effects in %-points on financial behavior outcomes of being above the threshold of the preferred field. In the estimations, I use applicants with *Business and economics* as the alternative field of study and another field as their preferred. All outcomes are dummy indicators. The outcomes in column (1)-(4) and (6) are measured 10 years after application. In column (5), *Has pub. debt* is measured from 10 years after application. *Low bank deposit* equals one if an applicant's bank deposit is smaller than one month of total income. *High marg. int. rate* equals one if an applicant's marginal interest rate is greater than 30%. *Bank debt* equals one if an applicant has bank debt. *Debt-to-inc. ratio* equals one if the applicant's debt-to-income ratio is larger than one. *Has pub. debt* equals one if an applicant has public debt. *Stock part.* equals one if an applicant holds stocks with a value in excess of 1500 DKK (200 Euro) in 2015 prices. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

In column (3), the outcome is an indicator for whether the applicants have non-mortgage bank debt and I find an increase of 5.8%-points at the threshold. Column (4) shows the effect on the probability of having a high debt-to-income ratio (larger than 1) conditional on having debt. The estimate is positive but insignificant. This means that being above the threshold increases the probability of having bank debt, but conditional on having bank debt there is not a significant effect on how much debt the applicants have relative to their incomes.

In column (5), the outcome is an indicator that equals one if an applicant has public debt more than 10 years after application. This is because public debt is only recorded from 2013 and onwards. Public debt is interest-bearing outstanding amounts that a person owes to the government, e.g. unpaid taxes, and could also be interpreted as a sign of financial problems in itself. Here, the table also shows an increase in the probability of having this type of debt at the threshold of 4.4%-points, which is significant at the 10%-level.

In the final column (6), I follow Christiansen et al. (2008) and estimate the effect of crossing the threshold on a dummy indicating stock market participation. Similar to their results, I find that admission to the *Business and economics* field increases stock market participation, since being above the threshold reduces the probability of participation with 6.2%-points. This suggests that studying economics affects financial behavior broadly, and not only affects debt behavior. It could also indicate that the applicants become more willing

to take risks which could have an adverse effect on the probability of getting into financial problems. The fact that they still have a lower probability of default suggests that they indeed become better at managing and servicing their debt.

In total, Table 1.5 shows that being admitted to *Business and economics* affects the financial behavior of the applicants. The applicants are less likely to be liquidity constrained and also less likely to take on debt. Both of these findings can potentially contribute to explain why the applicants admitted to *Business and economics* are less likely to experience financial problems.

### **Labor market outcomes**

The level of income, as studied in section 1.4.2, is not the only way income could be related to financial problems. Another possibility is that fluctuating income or an unstable source of income can increase the risk of default.<sup>8</sup>

Table 1.6 explores how being above the threshold affects different labor market outcomes for applicants with *Business and economics* as their alternative field. Column (1), (2) and (3) show the reduced form effects on different measures of unemployment. In column (1), the outcome is an indicator that equals one if the applicant experienced more than 3 months of unemployment in year 10 after application. The outcome in column (2) is a dummy that equals one if the applicant was unemployed ultimo November 10 years after application, and in column (3), the outcome is the total number of months with unemployment from 7 to 9 years after the year of application. None of these measures of unemployment seem to change around the threshold. Taken together with the estimated insignificant effects on labor income in Appendix Table A-1.12, these results suggest that for the applicants with *Business and economics* as their alternative there are no changes in the unemployment risk around the threshold.

Column (4) shows the effect of being above the threshold on the probability of being self-employed 10 years after application. If applicants admitted to *Business and economics* are less likely to be self-employed this could indicate that they have a more stable source of

---

<sup>8</sup>For instance, Kreiner et al. (2020) show that unemployment shocks increase the probability of default.

Table 1.6: Admission to *Business and economics* and labor market outcomes

	Unemployment			Self-employed	Private sector	Finance & insurance
	> 3 m.	Nov.	Cum. 7-9 y.			
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	-0.7 (1.7)	0.8 (1.1)	25.2 (28.3)	-0.9 (0.7)	-10.3*** (3.9)	-4.1** (2.0)
$N$	1893	1893	1988	1893	1689	1706
Average outcome	5.2	2.2	1.6	1.6	61.7	7.0

*Notes:* The table shows the estimated reduced form effects in %-points on labor market outcomes of being above the threshold of the preferred field. In the estimations, I use applicants with *Business and economics* as the alternative field of study and another field as their preferred. All outcomes are indicators measured 10 years after application except *Cum. 7-9 y.*, which is the cumulated number of months of unemployment 7 to 9 years after application. *> 3 m.* equals one if an applicant was unemployed for more than 3 months in the year. *Nov.* equals one if an applicant was unemployed in November. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

income, but the estimated effect shows this is not the case. The table also shows that these applicants are very unlikely to become self-employed in the first place (1.6%).

In column (5) and (6), I investigate which sectors the applicants are employed in 10 years after application. Column (5) shows the effect of being above the threshold on the probability of being employed in the private sector. Job security is often considered to be higher in the public sector,<sup>9</sup> so if applicants admitted to *Business and economics* are more likely to work in the public sector, this could mean that they have a more stable source of income and therefore are less likely to experience default. The estimated effect of being above the threshold is a 10.3%-point *reduction* in the probability of being employed in the private sector. This means that admission to *Business and economics* actually makes it significantly less likely that the applicant will be employed in the public sector, and therefore the job security explanation does not seem to explain the reduced risk of getting into financial problems.

Column (6) shows the effect on the probability of being employed in the *Finance and insurance* sector.<sup>10</sup> The table shows that being above the threshold reduces the probability of working in the *Finance and insurance* sector with 4.1%-points. This result and the result on the probability of working in the private sector in column (5) shows that field of study does affect the applicants' future work lives. The effect of studies and subsequent work

<sup>9</sup>See for instance Luechinger et al. (2010).

<sup>10</sup>The sector is defined by the DB07 classification provided by Statistics Denmark.

experience cannot be disentangled in this setup, since they are co-determined.<sup>11</sup> If we believe that it requires more financial knowledge and understanding to get a job in the *Finance and insurance* sector, this result suggests that admission to the *Business and economics* fields of study actually increases financial knowledge and understanding.

The results from Table 1.6 show that employment or income stability does not seem to change for applicants around the threshold. Therefore, it is unlikely that these channels explain the change in the probability of default around the threshold. The table also shows that admission to *Business and economics* lead to more employment in the private sector and in the *Finance and insurance* sector.

## Peers

An applicants' field of study also affects who his or her peers are and this could potentially influence the probability of getting into financial problems. In this section, I explore whether admission to *Business and economics* affects which partner an applicant lives with 10 years after application and if it affects the background of the applicant's peers in the study program he or she is admitted to.

College can be considered as a marriage market and which field of study an applicant is admitted to could affect what his or her later partner's field of study is (Eika et al., 2019; Kirkebøen et al., 2021). If we look at all applicants who have different preferred and alternative fields of study, column (3) of Appendix Table A-1.14 shows that being above the threshold actually increases the probability that the partner 10 years later has an education within that same field of study. The same is not true if we look at the alternative field of study in column (4). Being above the threshold of the preferred field does not significantly decrease the probability that the partner has an education within the alternative field of study. This is also the case if we only look at the applicants who have *Business and economics* and their alternative field of study in column (5). This suggest that the reduced probability of default is not driven by an increase in the probability that you get an "economist" as partner later in life. Furthermore, the results indicate that potential peer effects are strongest

---

<sup>11</sup>As a robustness check, I drop all applicants who end up working in the *Finance and insurance* sector 10 years after application in column (3) of Appendix Table A-1.8. This reduces the estimated effect of studying *Business and economics*, but the effect remains significant on the 10%-level with a P-value of 0.078.

for those who are admitted to their preferred field of study and perhaps less important for those who are admitted to the alternative field.

In column (6) of Appendix Table A-1.14, I use the educational background of an applicant's peers in his or her study program as the outcome. In particular, I look at the share of peers who come from the mathematical track (instead of the linguistic track) in upper secondary school. I find that being above the threshold reduces the share of peers from the mathematical track with 9.0%-points. For these applicants the average share is 44.5% so this is a substantial reduction that is significant on the 1%-level. I cannot rule out that this change in the peer composition is a part of the explanation for the finding on financial problems but another way to interpret this result is that admission to *Business and economics* exposes an applicant to a more math intensive program that potentially increases the applicant's numeracy which is also important for financial understanding.

## 1.5 Conclusion

This paper investigates how field of study in higher education affects the probability of getting into financial problems, and, in particular, if studying economics can reduce the risk of default. Ill-informed financial decision-making has great consequences for both creditors and debtors but causal evidence on the effect of economic education on financial outcomes is sparse.

I link data on admission to higher education in Denmark with data on the universe of personal loans to estimate the causal effect of admission to the *Business and economics* field of study on the probability of experiencing default. I identify the causal effect by exploiting that the Danish system of admission generates locally unpredictable GPA admission thresholds, which enables me to compare similar applicants who are quasi-randomized into different fields of study due to small differences in their GPA from upper secondary school.

I find that applicants who have *Business and economics* as their alternative field of study but would prefer another field are 6-8%-points less likely to experience default if they are just below the GPA admission threshold of their preferred field. This is a sizeable effect since 11% of all first time applicants experience default. The evidence suggests that admission to

*Business and economics* reduces the probability of default no matter what the preferred field is.

I show that admission to *Business and economics* does not increase income which could explain why these applicants are less likely to default. This is because the most common preferred fields, *Social science* (namely the program *Political science*) and *Law*, lead to high incomes in the Danish labor market. I also present evidence that admission to *Business and economics* does not reduce the probability of becoming unemployed or lead to more stable employment.

Instead, I find changes in the financial behavior around the GPA admission threshold for the applicants who have *Business and economics* as their alternative field. These changes can potentially explain why the applicants have a lower default probability. I find that they are less likely to be liquidity constrained, i.e. they are less likely to have low levels of bank deposits and a high marginal interest rate. They are also less likely to have bank debt and public debt.

Therefore, I conclude that admission to the *Business and economics* field of study reduces the risk of getting into financial problems. This result shows that extensive economic education can improve financial decision-making and have long-lasting effects on financial outcomes even for individuals from a relatively affluent background.

## References

- Altmejd, A., A. Barrios-Fernández, M. Drlje, J. Goodman, M. Hurwitz, D. Kovac, C. Mulhern, C. Neilson, and J. Smith (2021). O Brother, Where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries. *The Quarterly Journal of Economics*.
- Andersen, S., P. D'Astous, J. Martínez-Correa, and S. H. Shore (2020). Cross-program differences in returns to education and the gender earnings gap. Working paper / center for economic analysis of risk (CEAR), CEAR, Georgia State University, Atlanta, GA.
- Beshears, J., J. J. Choi, D. Laibson, and B. C. Madrian (2018). Behavioral Household Finance. In B. D. Bernheim, S. DellaVigna, and D. Laibson (Eds.), *Handbook of Behavioral Economics: Foundations and Applications 1* (First ed.), Chapter Behavioral, pp. 177–276. Amsterdam: Elsevier.
- Bleemer, Z. and A. S. Mehta (2021). Will Studying Economics Make You Rich? A Regression Discontinuity Analysis of the Returns to College Major. *American Economic Journal: Applied Economics*, Forthcoming.
- Bover, O., L. Hospido, and E. Villanueva (2018). The Impact of High School Financial Education on Financial Knowledge and Choices: Evidence from a Randomized Trial in Spain. *IZA Discussion Paper No. 11265*.
- Brown, M., J. Grigsby, W. Van Der Klaauw, J. Wen, and B. Zafar (2016). Financial education and the debt behavior of the young. *Review of Financial Studies* 29(9), 2490–2522.
- Bruhn, M., L. d. S. Leão, A. Legovini, R. Marchetti, and B. Zia (2016). The impact of high school financial education: Evidence from a large- scale evaluation in brazil. *American Economic Journal: Applied Economics* 8(4), 256–295.
- Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). Rdrobust: Software for Regression-discontinuity Designs. *The Stata Journal* 17(2), 372–404.
- Chetty, R., J. N. Friedman, S. Leth-Petersen, T. H. Nielsen, and T. Olsen (2014). Active vs. Passive decisions and crowd-out in retirement savings accounts. *Quarterly Journal of Economics* 129(3), 1141–1219.
- Christiansen, C., J. S. Joensen, and J. Rangvid (2008). Are economists more likely to hold stocks? *Review of Finance* 12(3), 465–496.
- Cole, S., A. Paulson, and G. K. Shastry (2016). High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses. *Journal of Human Resources* 51(3), 656–698.
- Dahl, G. B., D.-O. Rooth, and A. Stenberg (2020a). Family spillovers in field of study. NBER working paper 27618, National Bureau of Economic Research, Cambridge, MA.
- Dahl, G. B., D.-O. Rooth, and A. Stenberg (2020b). Long-run returns to field of study in secondary school. NBER working paper 27524, National Bureau of Economic Research, Cambridge, MA.
- Daly, M., M. F. Jensen, and D. le Maire (2021). University admission and the similarity of fields of study. Unpublished manuscript.
- Eika, L., M. Mogstad, and B. Zafar (2019). Educational Assortative Mating and Household Income Inequality. *Journal of Political Economy* 127(6), 2795–2835.
- Fernandes, D., J. G. Lynch, and R. G. Netemeyer (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science* 60(8), 1861–1883.

- Hastings, J. S., C. A. Neilson, and S. D. Zimmerman (2014). Are some degrees worth more than others? Evidence from college admission cutoffs in Chile. NBER working paper 19241, National Bureau of Economic Research, Cambridge, MA.
- Heinesen, E. (2018). Admission to higher education programmes and student educational outcomes and earnings—Evidence from Denmark. *Economics of Education Review* 63(January), 1–19.
- Humlum, M. K., J. H. Kristoffersen, and R. Vejlin (2017). College admissions decisions, educational outcomes, and family formation. *Labour Economics* 48, 215–230.
- Kaiser, T., A. Lusardi, L. Menkhoff, and C. Urban (2020). Financial Education Affects Financial Knowledge and Downstream Behaviors. NBER working paper 27057, National Bureau of Economic Research, Cambridge, MA.
- Kaiser, T. and L. Menkhoff (2017). Does financial education impact financial literacy and financial behavior, and if so, when? *World Bank Economic Review* 31(3), 611–630.
- Kaiser, T. and L. Menkhoff (2020, October). Financial education in schools: A meta-analysis of experimental studies. *Economics of Education Review* 78, 101930.
- Kirkeboen, L., E. Leuven, and M. Mogstad (2021). College as a Marriage Market. NBER working paper 28688, National Bureau of Economic Research, Cambridge, MA.
- Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection. *Quarterly Journal of Economics* 131(3), 1057–1111.
- Kreiner, C. T., D. D. Lassen, and S. Leth-Petersen (2019). Liquidity constraint tightness and consumer responses to fiscal stimulus policy. *American Economic Journal: Economic Policy* 11(1), 351–379.
- Kreiner, C. T., S. Leth-Petersen, and L. C. Willerslev-Olsen (2020). Financial trouble across generations: Evidence from the universe of personal loans in Denmark. *Economic Journal* 130(625), 233–262.
- Luechinger, S., S. Meier, and A. Stutzer (2010). Why does unemployment hurt the employed? Evidence from the life satisfaction gap between the public and the private sector. *Journal of Human Resources* 45(4), 998–1045.
- Lusardi, A., A. Samek, A. Kapteyn, L. Glinert, A. Hung, and A. Heinberg (2017). Visual tools and narratives: New ways to improve financial literacy. *Journal of Pension Economics and Finance* 16(3), 297–323.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142(2), 698–714.
- Miller, M., J. Reichelstein, C. Salas, and B. Zia (2015). Can you help someone become financially capable? A meta-analysis of the literature. *World Bank Research Observer* 30(2), 220–246.
- Öckert, B. (2010). What’s the value of an acceptance letter? Using admissions data to estimate the return to college. *Economics of Education Review* 29(4), 504–516.
- OECD (2016). *PISA 2015 Assessment and Analytical Framework*. Paris: OECD Publishing / PISA.
- Silliman, M. and H. Virtanen (2021). Labor market returns to vocational secondary education. *American Economic Journal: Applied Economics*, Forthcoming.
- Skimmyhorn, W. (2016). Assessing financial education: Evidence from boot camp. *American Economic Journal: Economic Policy* 8(2), 322–343.



- Statistics Denmark (2020). DISCED-15, Field of education, current educations, v1:2020. <https://www.dst.dk/en/Statistik/dokumentation/nomenklaturer/disc15-fagomraade-igangvaerende-uddannelser/#> (accessed August 13, 2020).
- Stiglitz, J. E. and A. Weiss (1981). Credit Rationing in Markets with Imperfect Information. *American Economic Review* 71(3), 393–410.
- Urban, C., M. Schmeiser, J. Michael Collins, and A. Brown (2020). The effects of high school personal financial education policies on financial behavior. *Economics of Education Review* 78, 1–12.

## 1.A Appendix

Table A-1.1: Examples of program rankings at University of Copenhagen in 2006

	GPA	Priority	Program	Threshold	Field
<b>Ex. 1</b>	8.6				
		1	Law	8.5	Preferred
		2	Economics	0	Alternative
		3	Pol. science	9.7	Drop: Law has higher rank and lower threshold
<b>Ex. 2</b>	8.3				
		1	Pol. science	9.7	Drop: GPA below threshold of Law
		2	Law	8.5	Preferred
		3	Economics	0	Alternative
<b>Ex. 3</b>	10.0				
		1	Pol. science	9.7	Preferred
		2	Law	8.5	Alternative
		3	Economics	0	Drop: GPA above threshold for all programs
<b>Ex. 4</b>	9.5				
		1	Pol. science	9.7	Preferred
		2	Law	8.5	Alternative
		3	Economics	0	Drop: Lowest of two thresholds

*Notes:* The table shows made up examples of four applicants to programs at the University of Copenhagen in 2006. The first applicant has a grade point average of 8.6 from upper secondary school and has *Law* as her first priority, *Economics* as her second priority, and *Political science* as her third priority. The column labelled *Threshold* shows the (real) GPA thresholds for the three programs. The final column shows how the applications are used in the research design. For this applicant *Law* is the preferred program, *Economics* is the alternative program and the *Political science* application is discarded because the other programs have a higher rank and lower thresholds and therefore the applicant will never be admitted to *Political science*. The three remaining examples illustrate other cases where applications are discarded for different reasons.

Table A-1.2: Selection of sample

Restriction	Observations
First time applicants	427,885
Only one program has first priority	426,666
At least two applications in ordinary system	170,098
GPA threshold for first priority program	133,009
Not admitted to program with another admission system	129,320
Admitted to one program	98,688
GPA is not missing	81,350
GPA is above the GPA threshold of at least one program	67,522
Applicants with binding threshold	53,882

*Notes:* The large reduction from the restriction *At least two applications in ordinary system* is caused by the fact that many applicants only apply for one program and many only apply for programs that use another admission system. The *GPA is not missing* restriction reduces the sample size because GPA was not recorded for Upper secondary higher commercial examination before 2000. The sample size is similar to the sample sizes in other studies using the same data. For instance, Daly et al. (2021) have 46,213 observations that meet their selection criteria. They use data from 1996 to 2006, applicants that are 17 to 25 years old and only applications to university level programs whereas I include university college programs as well.

Table A-1.3: Summary statistics

	1st time applicants			Sample		
	Obs.	Mean	SD	Obs.	Mean	SD
Age	427885	21.8	2.4	14181	21.1	1.8
Male (%)	427885	40.5	49.1	14181	36.3	48.1
GPA	323751	8.3	1.0	14181	8.8	0.8
1st priority threshold	320740	5.8	4.1	13666	9.1	0.6
Offered rank	358047	1.2	0.6	14181	1.8	1.1
Number of applications	427885	1.9	1.2	14181	3.2	1.4
Income rank	412865	58.2	28.0	14022	63.2	28.5
Father's income rank	325715	67.0	28.4	11638	70.7	28.2
Mother's income rank	367522	47.3	26.0	13014	52.7	27.1
Father has master (%)	389010	13.3	33.9	13530	22.8	41.9
Mother has master (%)	400293	6.4	24.5	13829	11.9	32.3
Default (%)	422365	11.1	31.4	14170	8.0	27.1

*Notes:* The *1st time applicants* are all applicants observed in the data. This includes applicants to business academies that use another admission system than universities and university colleges. The *Sample* are the applicants whose preferred program has a binding GPA threshold and whose preferred program and alternative program are within different fields of study. *Income rank* is within cohort rank based on total income measured 10 years after application. Father's and mother's income rank are measured when they are 45 years old. *Default* is measured 10 to 23 years after application.

Table A-1.4: Construction of fields of study from DISCED classifications

Fields of Study	DISCED
Humanities, arts and architecture	Arts and humanities (L4) Architecture and town planning (L2)
STEM	Engeneering, manufacturing and construction (L4) Natural sciences, mathematics and statistics (L4) Information and communication technologies (L4)
Law	Law (L3)
Business and economics	Business and administration (L3) Economics (L2) Mathematical economics (L2) Agricultural economics (L2)
Social science	Social sciences, journalism and information (L4)
Education	Education (L4)
Welfare	Health and welfare (L4)
Medicine	Medicine (L2) Veterinary (L3)

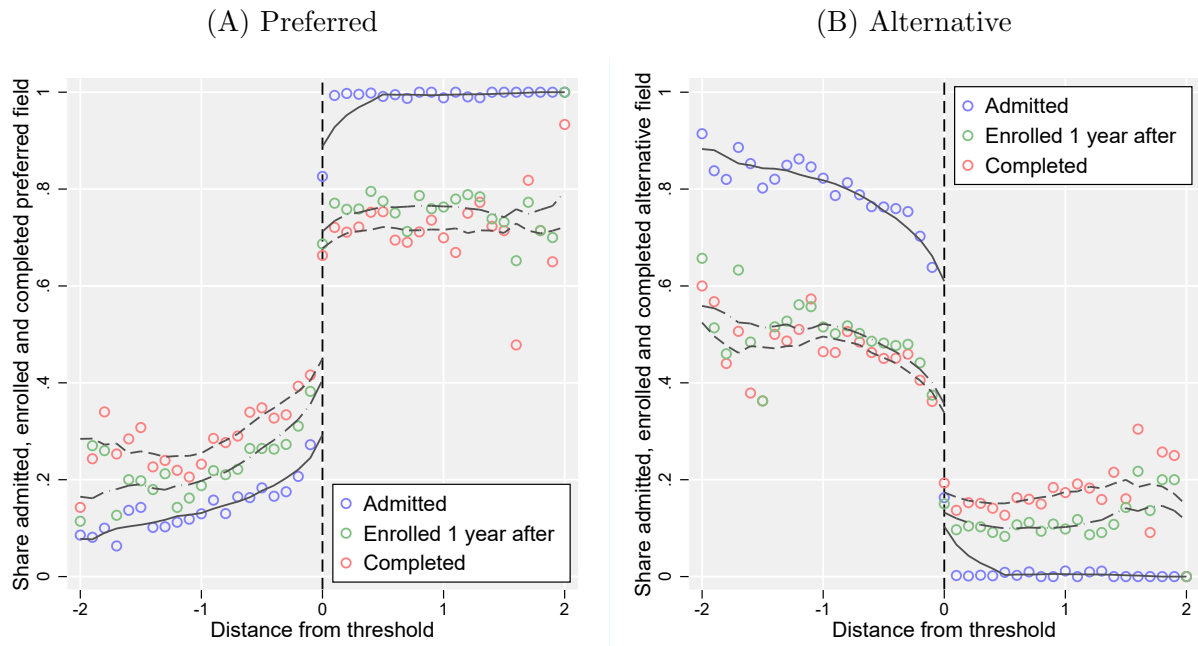
*Notes:* The parentheses indicate the DISCED level. Level 4 is ISCED Broad fields, level 3 is ISCED Narrow fields and level 2 is ISCED Detailed fields. Two level 4 (L4) DISCED groups are left out. These are *Agriculture, forestry, fisheries and veterinary* and *Services*. For the first group there are very few observations that do not belong to the narrow field *Veterinary* (included in *Medicine*) and for the latter group there are also very few observations, since the *Service*-educations are not at university or university college level. The DISCED name for *Mathematical economics* is *Inter-disciplinary programs involving mathematics and statistics*. The DISCED name for *Agricultural economics* is *Inter-disciplinary programs involving agriculture*.

Table A-1.5: Most common programs within the fields of study

Field of study	Preferred	Alternative
Hum., art & arch.	Humanistic basic education (22%) Architect (16%) History (8%)	History (12%) Business language (10%) Danish (7%)
STEM		Engineer (25%) Biology (15%) Natural science (15%)
Law	Law (100%)	Law (100%)
Bus. & econ.	Business econ. and language (29%) Business economics (16%) Language and int. marketing (9%)	Economics (38%) Bus. econ. and law (22%) Business economics (20%)
Social science	Political science (25%) Psychology (22%) Anthropology* (12%)	Soc. sci. basic edu. (39%) Librarian (13%) Admin./Soc. studies (10%)
Education	Teacher (84%) Audiologopedics (7%) Pedagogy (4%)	Teacher (91%) Pedagogy (6%) Audiologopedics (2%)
Welfare	Sports (22%) Dentist (17%) Social worker (14%)	Pedagogue (31%) Nurse (24%) Social worker (12%)
Medicine	Medical science (83%) Veterinary (16%)	

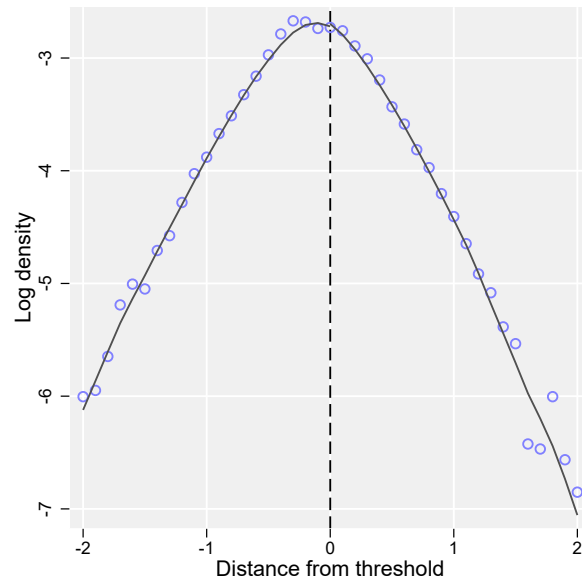
*Notes:* The *Preferred* column shows the most common programs when the field is the preferred field of study. The *Alternative* column shows the most common programs when the field is the alternative field of study. \**Anthropology* consists of two programs: *Anthropology* and *Humanities 2*. The latter was a program at Aarhus University from 1993 to 1998 where the enrolled applicants were divided into different programs. The mode of applicants accepted to the program studied *Anthropology* later in the year of application, but others studied philosophy or literature for instance.

Figure A-1.1: Admission, enrolment and completion for preferred or alternative field of study



*Notes:* The figures are based on all applicants whose preferred field of study has a binding threshold and whose preferred and alternative field of study differ. The bin width in the two panels is the smallest possible, 0.1. The local linear polynomials have a bandwidth of 0.5.

Figure A-1.2: Distribution of the distance to the GPA threshold



*Notes:* The figure is based on all applicants whose preferred field of study has a binding threshold and whose preferred and alternative field of study differ. The bin width in the two panels is the smallest possible, 0.1. The local linear polynomials have a bandwidth of 0.5.



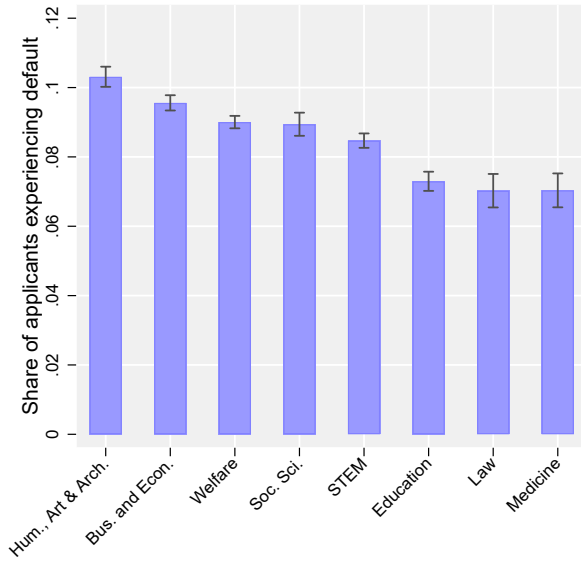
Table A-1.6: Test for manipulation of the running variable

	Stata	Test stat.	Std. Err. / P-val	Bandwidth
McCrary (2008)	DCdensity	-0.0167	0.037	0.68
Cattaneo et al. (2018)	rddensity	0.0571	0.9544	0.22/0.25

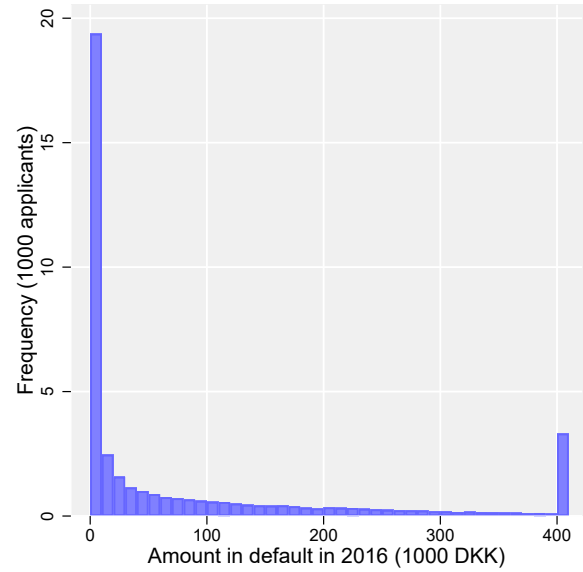
*Notes:* The tests are based on all applicants whose preferred field of study has a binding threshold and whose preferred and alternative field of study differ. The column *Stata* indicates which Stata program was used to implement the test. The column *Std. Err. / P-val* shows the standard error for the McCrary test and the P-value for the test proposed by Cattaneo, Jansson and Ma. For the latter test, the two calculated bandwidths are to the left and right of the threshold.

Figure A-1.3: Completed field of study, time since application and default

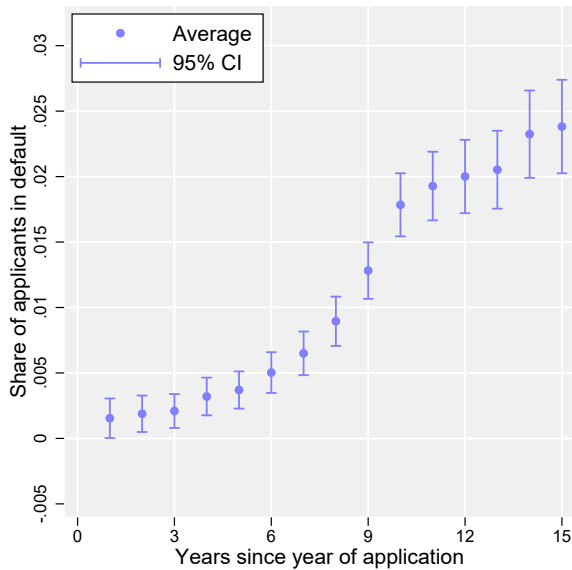
(A) Default by field of study



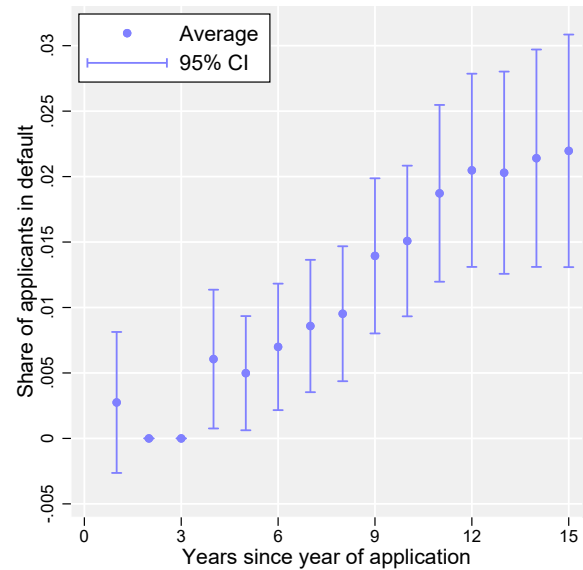
(B) Distribution of amount in default



(C) Default and time since application (applicants with binding GPA threshold)



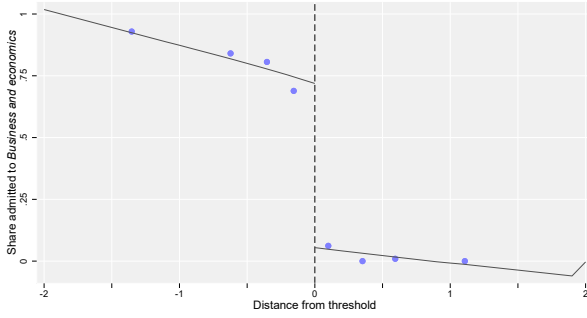
(D) Default and time since application (applicants with B&E as alternative)



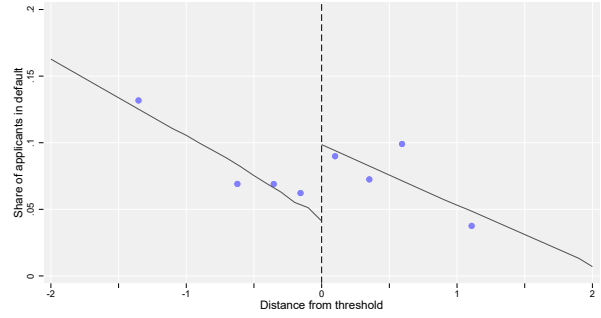
*Notes:* In the top left panel, default status is observed from 10 years after application until 23 years after application for the oldest application cohorts. The figure is based on first time applicants who complete a higher education degree. For these applicants the average probability of default is 8.9%. In the top right panel, the bin width is 10,000 DKK (1300 Euro). The amount in default is censored above 400,000 DKK. The bottom panels show the share of applicants with binding GPA thresholds that experience default in each year since their first year of application. The left panel includes all applicants with a binding threshold and different preferred and alternative fields of study and the right panel only includes applicants with *Business and Economics* as their alternative field of study.

Figure A-1.4: Admission to *Business and economics* and default

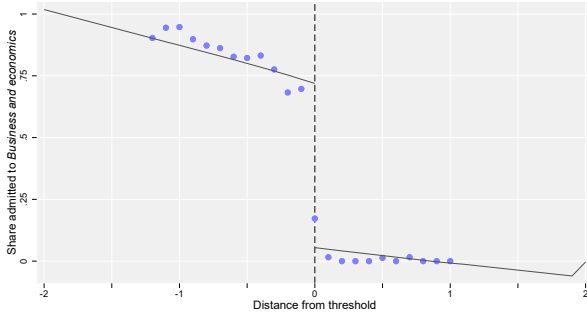
(A) First stage with 4 bins



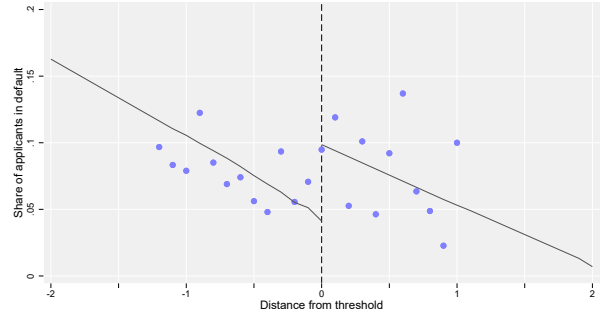
(B) Reduced form with 4 bins



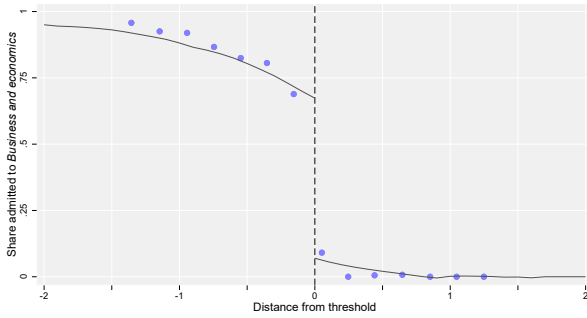
(C) First stage with smallest bin width



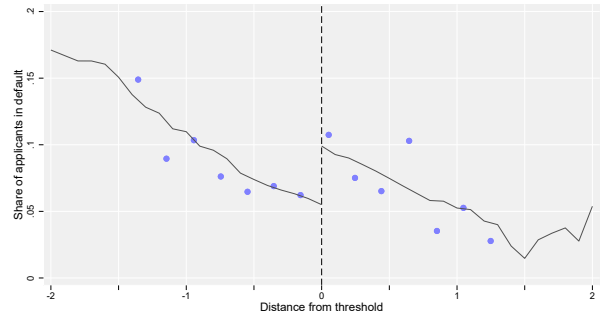
(D) Reduced form with smallest bin width



(E) First stage with a bandwidth of 1



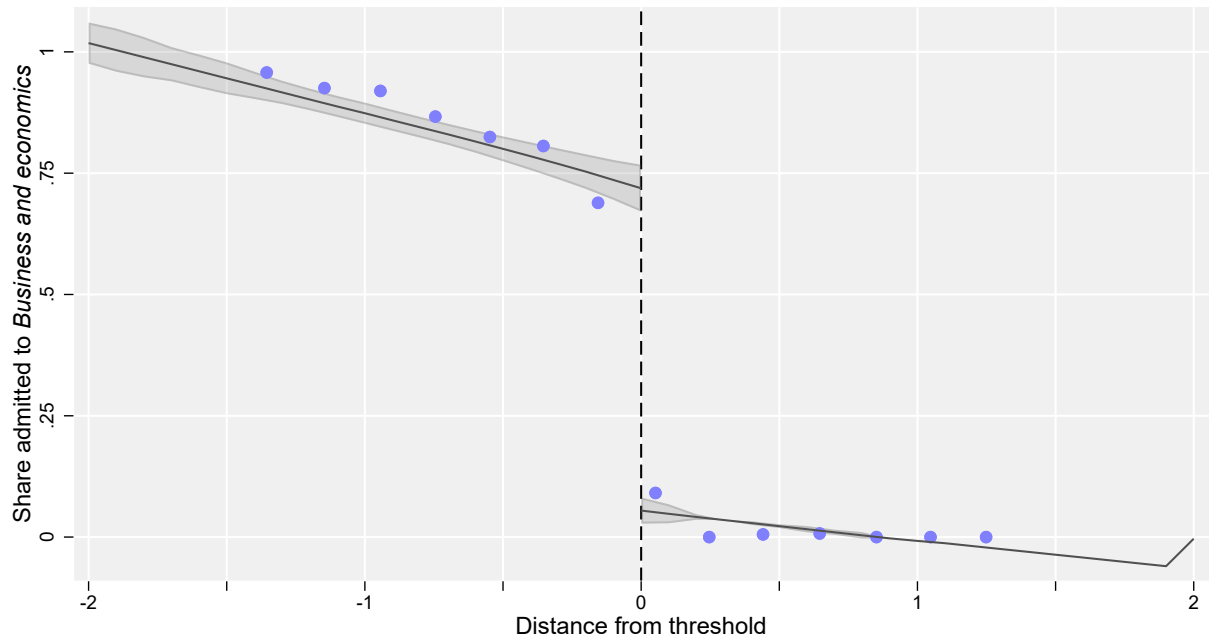
(F) Reduced form with a bandwidth of 1



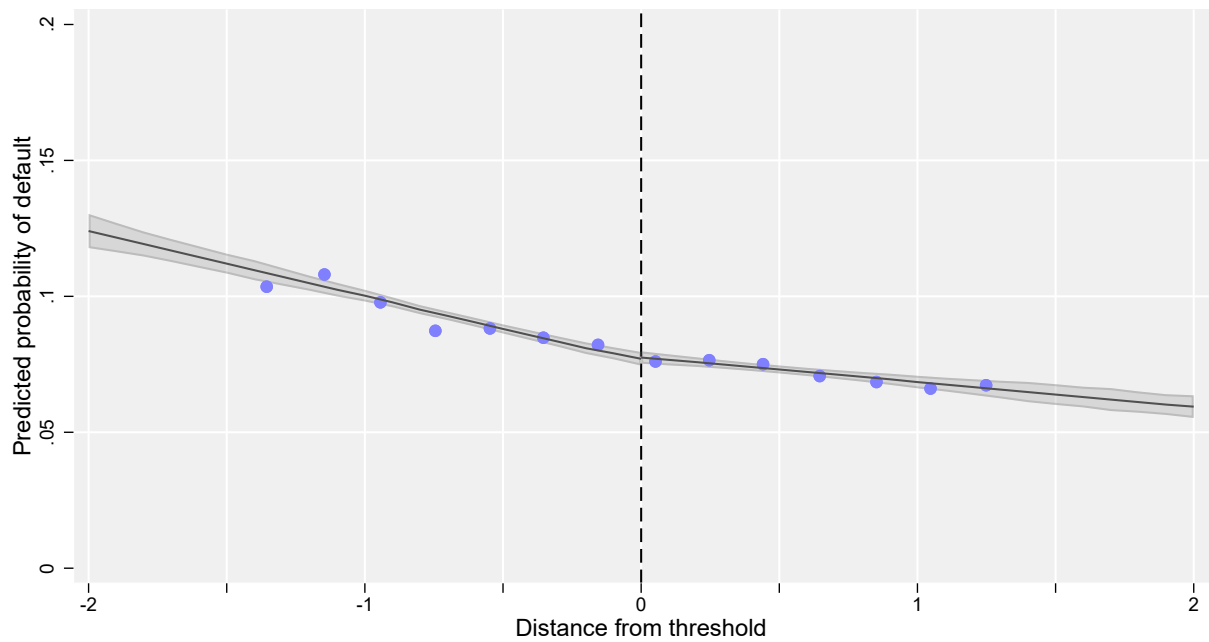
*Notes:* The figures are based on applicants with *Business and economics* as their alternative field of study and another field as their preferred. In the top panels, I use a bandwidth of 2 and the four bins have approximately the same number of observations. In the middle panels, the bin width is the smallest possible, 0.1, but I only plot bins with at least 30 observations. In the bottom panels, I use a bandwidth of 1 and the bin width is 0.2, but I only plot bins with at least 30 observations. All panels use rectangular kernels and I include individuals who are admitted to either their preferred or alternative field of study.

Figure A-1.5: Admission to *Business and Economics* and predicted probability of default

(A) First stage

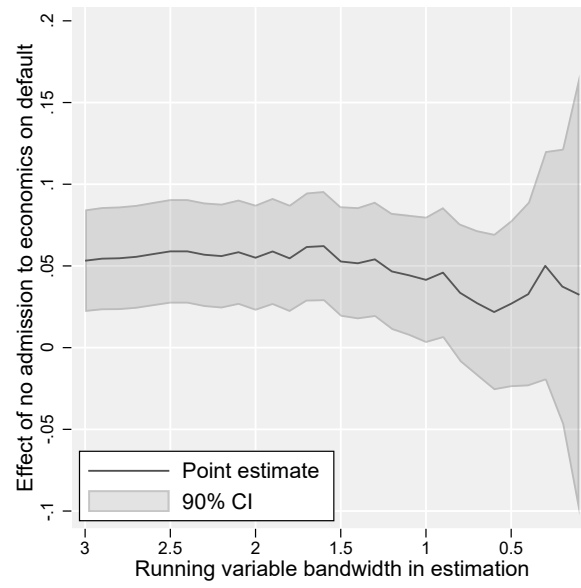


(B) Reduced form using predicted default



*Notes:* The figures are based on applicants who have *Business and economics* as their alternative field of study and another field as their preferred. In both panels, the local linear polynomials have a bandwidth of 2 and I use a rectangular kernel. The grey area indicates the 95% confidence interval. The bin width is 0.2, but I only plot bins with at least 30 observations. The top panel shows the first stage, namely how the probability of admission to *Business and economics* changes around the threshold of the preferred field of study. In the bottom panel, the probability of default is predicted using a probit model with sex, year of application, age and grade point average as the explanatory variables.

Figure A-1.6: *Business and economics* and default for different choices of bandwidth



*Notes:* The figure shows the estimated reduced form effect as in column (2) of Table 1.3 with different choices of bandwidth. The Stata program `DCdensity` suggests a bandwidth of 0.84 (McCrary, 2008) and `rdrobust` suggests a bandwidth of 0.47 (Calonico et al., 2017).

Table A-1.7: Admission to *Business and economics* and the probability of default

	Probability of default more than 10 years after application					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}(\text{GPA} \geq \text{threshold})$ :	-71.2***	-71.9***	-67.3***	-69.0***	-67.0***	-71.1***
Effect on admission	(2.3)	(2.1)	(2.9)	(2.6)	(3.0)	(2.3)
$\mathbb{1}(\text{GPA} \geq \text{threshold})$ :	5.5***	5.3***	4.1*	5.1**	5.6**	5.6***
Effect on default	(2.0)	(1.9)	(2.4)	(2.1)	(2.5)	(2.2)
Local average treatment effect	-7.7***	-7.4***	-6.2*	-7.4**	-8.4**	-7.9***
	(2.8)	(2.7)	(3.6)	(3.1)	(3.8)	(3.1)
$N$	1983	2013	1707	1983	2013	1867
Bandwidth	2		1	2		2
Functional form	Linear	Linear	Linear	Linear	Quadratic	Linear
Kernel	Rect.	Rect.	Rect.	Tri..	Rect.	Rect.
Donut RD						✓

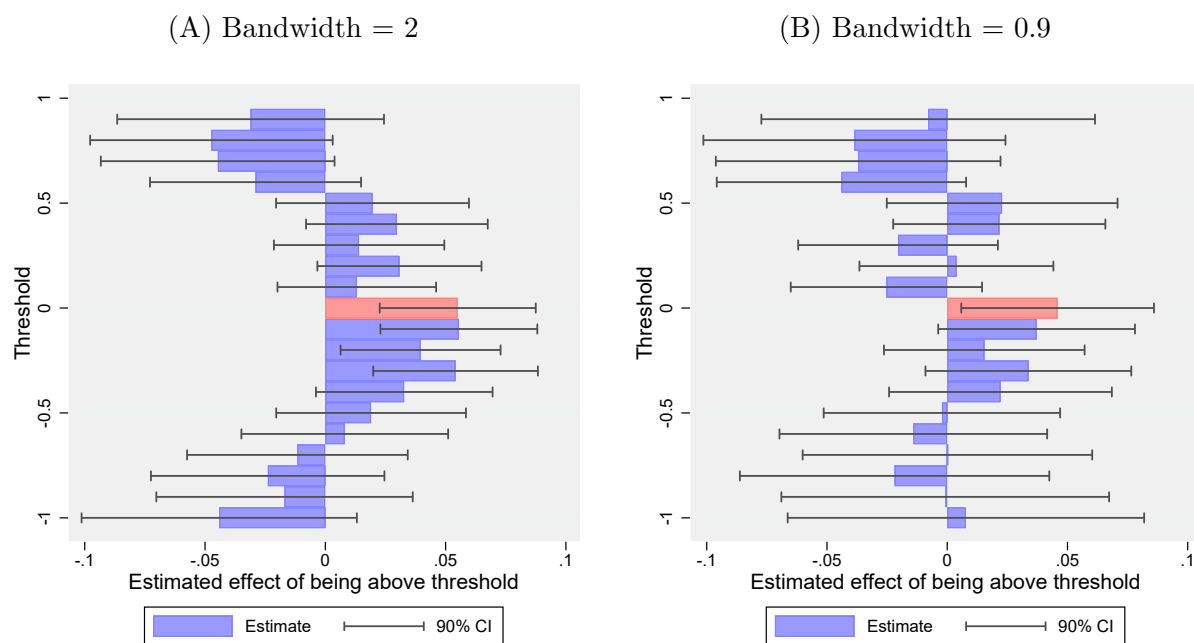
*Notes:* In the estimations, I use applicants with *Business and economics* as their alternative field of study and another field as their preferred for whom I observe all the control variables. The estimated effects are in %-points. The top row shows the first stage effects of being above the threshold on the probability of admission to *Business and economics*. The middle row shows the reduced form effects of being above the threshold on the probability of default. The bottom row shows the effects of admission to *Business and economics* on the probability of default using two stage least squares estimation where admission is instrumented using whether an applicant is above the threshold of the preferred field. *Bandwidth*, *Functional form* and *Kernel* indicate how the local linear polynomials are specified. *Donut RD* indicates that applicants exactly at the threshold are discarded in column (6). Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A-1.8: Admission to *Business and economics* and the probability of default

	Default +10 years after			Default in 2016	
	(1)	(2)	(3)	(4) $\geq 0$ DKK	(5) $\geq 4000$ DKK
$\mathbb{1}(\text{GPA} \geq \text{threshold})$ : Effect on admission	-72.4*** (1.9)	-70.3*** (2.3)	-69.4*** (2.7)	-71.2*** (2.3)	-71.2*** (2.3)
$\mathbb{1}(\text{GPA} \geq \text{threshold})$ : Effect on default	5.6*** (1.9)	6.0*** (2.0)	3.7* (2.1)	4.3** (1.9)	2.1* (1.3)
Local average treatment effect	-7.7*** (2.6)	-8.5*** (2.9)	-5.4* (3.0)	-6.1** (2.7)	-3.0* (1.8)
$N$	2665	1955	1589	1983	1983
Bandwidth	2	2	2	2	2
Functional form	Linear	Linear	Linear	Linear	Linear
Kernel	Rect.	Rect.	Rect.	Rect.	Rect.
Incl. non-completers	✓				
Alternative econ. def.		✓			
Excl. bankers			✓		

*Notes:* In the estimations, I use applicants with *Business and economics* as their alternative field of study and another field as their preferred for whom I observe all the control variables. The estimated effects are in %-points. In column (1)-(3), I use the same outcome as in Table 1.3. In column (4) and (5), I only use an indicator for default in 2016, and in column (5), I only use accounts in default with an outstanding amount of more than 4000 DKK (500 Euro). The top row shows the first stage effect of being above the threshold on the probability of admission to *Business and economics*. The middle row shows the reduced form effect of being above the threshold on the probability of default. The bottom row shows the effect of admission to *Business and economics* on the probability of default using two stage least squares estimation where admission is instrumented using whether an applicant is above the threshold of the preferred field. *Bandwidth*, *Functional form* and *Kernel* indicate how the local linear polynomials are specified. In column (1), I include non-completers, who never obtain a higher education degree. In column (2), I use the definition of *Business and economics* from Chetty et al. (2014). In column (3), I exclude applicants who end up working in the *Finance and insurance* sector. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Figure A-1.7: Admission to *Business and economics* and default at placebo thresholds



*Notes:* The figures show the estimated reduced form effect similar to column (2) of Table 1.3, but where I vary the threshold from -1 to 1 in steps of 0.1 which means that 0 on the y-axis is the “true” threshold.

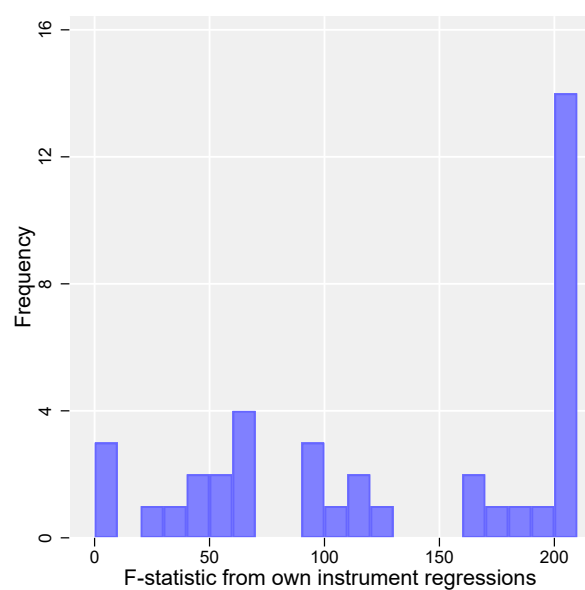


Table A-1.9: Admission to *Business and economics* vs. *Law* and *Social science* and default

	Law			Social Science		
	Admission	Default		Admission	Default	
	(1) OLS	(2) OLS	(3) 2SLS	(4) OLS	(5) OLS	(6) 2SLS
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	-57.1*** (4.0)	1.6 (3.3)		-82.2*** (3.4)	8.2** (3.3)	
$\mathbb{1}(\text{Admission B\&E})$			-2.8 (5.8)			-9.9** (4.0)
$N$	782	782	782	711	711	711

*Notes:* The table is based on applicants who have *Business and economics* as their alternative field of study and who prefer either *Law* or *Social science*. The estimated effects are in %-points. Column (1) and (3) show the first stage effect of being above the threshold on the probability of admission to *Business and economics*. Column (2) and (5) show the reduced form effects of being above the thresholds on the probability of default. Column (3) and (6) show the effect of admission to *Business and economics* on the probability of default from two stage least squares estimation where admission is instrumented using whether an applicant is above the thresholds. The estimations use a bandwidth of 2, linear polynomials on each side of the threshold and a rectangular kernel. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Figure A-1.8: Histogram of F-statistics from own instrument regressions



*Notes:* The F-statistics are from estimations of  $P(admission_i) = \beta_0 + \beta_1 x_i + \beta_2(x_i \times T_i) + \beta_3 T_i + \varepsilon_i$  for applicants to each combination of preferred and alternative field of study. The figure is censored above 200.

Table A-1.10: Joint estimation of the LATE on default of being above the GPA threshold

Preferred field	Alternative field						
	B&E	Law	HAA	Educ.	Welf.	Soc.	STEM
Bus. & Econ.		-10.7 (10.1)	-3.7 (5.2)			-7.4 (8.3)	2.8 (8.0)
Law	6.5* (3.7)		-9.9** (3.9)	-7.0 (11.8)	2.3 (8.7)	7.9 (6.5)	-11.6** (5.9)
Hum., A. and A.	6.5 (5.7)	5.1 (7.8)		-3.0 (6.5)	-0.3 (6.0)	-1.6 (3.6)	2.7 (4.1)
Education			2.5 (7.3)		0.7 (10.6)		
Welfare	11.3* (6.5)	7.0 (8.1)	-7.1 (4.3)	-3.7 (7.1)		-3.3 (7.1)	1.6 (3.8)
Social Science	10.1*** (3.3)	-0.5 (4.3)	-3.3* (1.8)	2.9 (5.3)	3.2 (3.9)		6.9 (4.3)
Medicine	13.4** (6.4)	9.2 (6.7)	3.7 (5.1)		-2.2 (3.7)	8.3 (8.9)	0.2 (2.6)
<i>N</i>	1948	764	4185	858	1530	1125	2490
Default rate	8.1	7.9	9.0	6.9	6.4	9.0	6.7

*Notes:* The table shows the local average treatment effects in %-points on default of being admitted to the preferred field in a joint estimation similar to the estimation in Table 1.4. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A-1.11: Heterogeneous effects of not being admitted to *Business and economics*

Panel A: Gender	Probability of default				
				Women	Men
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	5.4*** (2.0)	5.3*** (2.0)	4.0 (2.4)	4.0 (2.4)	6.5** (3.1)
Male	3.6*** (1.2)	2.3 (2.0)	0.9 (2.7)		
$\mathbb{1}(\text{GPA} \geq \text{threshold}) \times \text{Male}$			2.5 (3.9)		
<i>N</i>	1983	1983	1983	914	1069
Panel B: Father education					
				Low	High
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	5.4*** (2.0)	5.4*** (2.0)	4.9* (2.9)	4.9* (2.9)	5.9** (2.8)
High Edu. Father	0.3 (1.2)	-0.3 (2.0)	-0.8 (2.8)		
$\mathbb{1}(\text{GPA} \geq \text{thres.}) \times \text{High}$			1.0 (4.0)		
<i>N</i>	1916	1916	1916	935	981
Panel C: Period					
				1993-99	2000-06
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	5.5*** (2.0)	5.6*** (2.0)	2.4 (2.7)	9.4*** (3.1)	2.4 (2.7)
Before 2000	0.8 (1.2)	1.9 (2.1)	-1.8 (2.9)		
$\mathbb{1}(\text{GPA} \geq \text{thres.}) \times \text{Before 2000}$			7.0* (4.1)		
<i>N</i>	1983	1983	1983	918	1065
Panel D: Economics					
				Economics	Not econ.
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	5.5*** (2.0)	5.6*** (2.0)	2.2 (2.4)	11.4*** (3.5)	2.2 (2.4)
Economics	1.2 (1.3)	0.8 (2.2)	-4.1 (2.9)		
$\mathbb{1}(\text{GPA} \geq \text{thres.}) \times \text{Economics}$			9.2** (4.3)		
<i>N</i>	1983	1983	1983	762	1221
Slopes by group		✓	✓		

*Notes:* The table shows the reduced form effects in %-points of being above the threshold of the preferred field for applicants who have *Business and economics* as their alternative field and another field as their preferred. *Father education* is defined as high if it is at the level *Short cycle higher education* or higher. *Economics* is defined by the DISCED Detailed field. *Slopes by group* indicates whether I allow for the two groups to have different running variable slopes, both above and below the threshold. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A-1.12: Reduced form effect on income of being above the threshold when *Business and economics* is the alternative

	Labor income		Total income		Household inc.		Inc. rank		Top 10%		Bottom 50%	
	1-9	10-12	1-9	10-12	10	10	10	10	10	10	10	10
Years after application												
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	-0.1 (5.6)	12.2 (13.3)	1.4 (5.2)	15.9 (12.7)	3.4 (24.1)	1.8 (2.1)	0.4 (3.4)					
<i>N</i>	1986	1968	1986	1968	1959	1959	1959	1959	1959	1959	1959	1959
Average outcome	144.1	349.0	195.8	385.5	633.3	63.1	19.2					
Unit	1000 DKK	1000 DKK	1000 DKK	1000 DKK	1000 DKK	Rank	%-point					

*Notes:* *Labor income* and *Total income* and *Household inc.* are winsorized at *p*1 and *p*99 and are in 2015-prices averaged over the years indicated in the *Years after application* row. *Inc. rank* is within birth cohort rank based on total income. *Top 10%* and *Bottom 50%* are dummies for being in the top 10 or bottom 50% within birth cohort with respect to total income. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A-1.13: The effects on income of not being admitted to *Business and economics*

	Labor inc. 10-12 yr	Total inc. 10-12 yr	Inc. rank 10 yr	Top 10% 10 yr	Bottom 50% 10 yr
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	13.2 (17.0)	15.3 (16.2)	1.7 (2.5)	-0.6 (4.4)	-3.5 (3.7)
Economics	-27.1 (19.1)	-35.8** (18.3)	-4.4 (3.2)	-6.7 (4.9)	3.2 (4.8)
$\mathbb{1}(\text{GPA} \geq \text{thres.}) \times \text{Econ.}$	-2.6 (27.1)	2.0 (25.7)	0.3 (4.3)	3.3 (7.0)	1.1 (6.2)
$N$	1965	1965	1957	1957	1957

*Notes:* The table shows the effect of being above the threshold of the preferred field for applicants who have *Business and economics* as their alternative field and another field as their preferred. *Economics* is an indicator that equals 1 if an applicant's alternative program belongs to the DISCED detailed field *Economics*. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A-1.14: The effects of being above the threshold on partner and peers

	Has partner		Partners completed field of study			Peers from
	All (1)	B&E (2)	Same as pref. (3)	Same as alt. (4)	B&E (5)	math. track (6)
$\mathbb{1}(\text{GPA} \geq \text{threshold})$	-3.0** (1.2)	-3.0 (3.4)	4.2*** (1.2)	-1.7 (1.2)	-2.0 (3.6)	-9.0*** (1.0)
$N$	13826	1983	9415	9415	1321	1983
Average outcome	68.1	66.6	16.4	16.8	21.7	44.5

*Notes:* The table shows the reduced form effects in %-points of being above the threshold of the preferred field. In column (1), (3) and (4), I use all applicants with a binding threshold and different preferred and alternative fields. In column (2), (5) and (6), I only use applicant who have *Business and economics* as their alternative field. Robust standard errors in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## References

- Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). Rdrobust: Software for Regression-discontinuity Designs. *The Stata Journal* 17(2), 372–404.
- Cattaneo, M. D., M. Jansson, and X. Ma (2018). Manipulation testing based on density discontinuity. *Stata Journal* 18(1), 234–261.
- Chetty, R., J. N. Friedman, S. Leth-Petersen, T. H. Nielsen, and T. Olsen (2014). Active vs. Passive decisions and crowd-out in retirement savings accounts. *Quarterly Journal of Economics* 129(3), 1141–1219.
- Daly, M., M. F. Jensen, and D. le Maire (2021). University admission and the similarity of fields of study. Unpublished manuscript.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142(2), 698–714.



# Chapter 2

## **Social Position and Fairness View**

with Claus Thustrup Kreiner and Stefanie Stantcheva

## 2.1 Introduction

People’s social positions can affect their views on a range of issues. A long-standing literature on social status, economic decision-making, and subjective well-being shows that people care about their social positions relative to others (Duesenberry, 1949; Easterlin, 1974, 1995; Blanchflower and Oswald, 2004) and theoretical work in Political Economy and Public Economics highlights that social positions are important for fairness considerations and redistribution policy (Boskin and Sheshinski, 1978; Meltzer and Richard, 1981; Bénabou and Ok, 2001; Alesina and Angeletos, 2005). But how well do people actually know their own position relative to others and how does their social position affect their views on the fairness of inequality? How do they view (un)fairness of inequality within different reference groups and are they better or less well-informed about inequality and social position where it matters the most to them?

To answer these questions, we leverage a unique dataset constructed by linking responses from a custom survey of a large sample of people in Denmark born between 1969 and 1973 to detailed administrative data on their full income histories, life events, and true positions in the income distributions of different “reference groups.” The reference groups include large groups such as people from the same cohort and of the same gender, living in the same municipality, having the same education level, or working in the same sector, as well as smaller groups such as neighbors, co-workers in the same firm, family members, and former schoolmates. In the survey, we ask people about their knowledge of the income distributions in these reference groups, how fair they think income inequalities within these groups are, and about where they rank themselves within the various groups (i.e., their income or “social” position within each group).

The link between survey and administrative data enables us to explore how well people know their positions in various reference groups and the relationship between social positions and fairness views. Income in the survey is defined in the same way as on tax returns, implying that we can compute the true income positions from the tax returns of all people in the reference groups. The link also enables us to study how changes in social positions over the course of life, including changes due to unemployment, health shocks, and promotions,

affect fairness views. Finally, we show how a randomized treatment informing individuals of their true social positions shapes these views.

Our results can be grouped into three main sets of findings. First, respondents are overall well aware of the income distributions and their own social positions, even though they underestimate the degree of inequality by systematically believing that the income levels of others are closer to their own than they actually are. Starting from the perception of the overall income distribution, 45% of the respondents perceive the median income level of their cohort correctly with at most 10% error. For comparison, 70% recall their own income correctly within the same error band. Lower-income respondents tend to underestimate the median income level and also the 95th percentile level of their cohort. In contrast, higher-income respondents tend to overestimate these levels. Yet, at most positions in the income distribution, the average perception is within 5% of the actual median and within 10% of the actual 95th percentile. The most striking misperception is that people at the very top of the distribution (above the 95th percentile) overestimate the 95th percentile level by 50%.

We then consider the accuracy of respondents' perceptions of their own position in the income distribution. This is what is normally done in the literature but, as we show, this may exaggerate how inaccurate people are. Consistent with the systematic misperceptions of the P50 and P95 moments, lower-income respondents believe they are ranked higher in the distribution than they really are, while higher-income respondents believe they are ranked lower. These systematic misperceptions of own social position seem large at first glance, but they are in fact mainly due to a mechanical "center bias" that arises because ranks are bounded between 1 and 100. Indeed, people in the top of the distribution cannot overestimate their rank, while people in the bottom cannot underestimate it. This implies that perceived rank positions are not uniformly distributed like actual rank positions, but instead concentrated towards the center. By ranking the respondents by their perceived social positions, thereby creating a uniform rank distribution of perceptions and removing the mechanical center bias, we show that perceived and actual social positions line up almost perfectly. Put differently, people who have rank  $x$  in the actual distribution tend to also have rank  $x$  in the distribution of perceptions.

The second set of findings shows that fairness views on inequality across all the reference groups studied strongly depend on the current social positions of individuals. We show this link in three ways: First, we highlight that views on the fairness of inequalities are more strongly correlated with current social position than with historical (past) social positions. In contrast, political views are more weakly correlated with current social position and more strongly correlated with respondents' past social positions and even significantly correlated with the social position of their father when they were growing up. Second, we show that changes in social positions following shocks affect fairness views. Conditional on a detailed array of individual-level controls and starting social position, we find that the perceived fairness of inequality significantly declines with negative shocks (unemployment spells, hospitalization episodes, or disability) and increases with positive income shocks (promotions at work). Third, we exploit our randomized information treatment that informs individuals of their true positions in their reference groups. This information affects views on the fairness of inequality within all reference groups, but in an asymmetric manner. Those who overestimated their social position in any of the reference groups to start with believe inequality is more unfair when they are informed of their actual (lower) social position. In contrast, those who are told that they are ranked higher than they thought do not adjust their fairness views. In line with the overall correlation patterns, people's political views respond much less than fairness views to the information treatment and the real-life shocks.

Third, people view inequality within peers working in the same sector or with same education level as more unfair than inequality among peers of the same age, or same gender, or living in the same municipality. Yet, these are exactly the reference groups within which respondents tend to underestimate the degree of inequality the most and within which lower-income people strongly overestimate their own positions. More precisely, for all reference groups, respondents perceive quite accurately the median income level, but systematically underestimate the 95th percentile income level among their co-workers and people with same education. Lower-ranked individuals overestimate their social position mostly within their education group or work sector. For example, people at the 20th percentile among their co-workers on average think they are well above the 40th percentile, while people at the 20th percentile in their municipality believe they are around the 30th percentile. This pattern

also holds if we zoom in on smaller reference groups, namely co-workers within a firm instead of within sector and if we look at neighbors instead of people living in the same municipality. We also show that respondents perceive their social positions relative to former schoolmates surprisingly well; perceive their positions relative to siblings very well, but far from perfect; and have very little idea about the social positions of their parents when they were around the same age.

To sum up, on average, individuals seem well-aware of their social positions, and their positions are important for their fairness views on inequality. Moving up the social ladder makes people more tolerant of inequality, while moving down makes them less tolerant of it. By contrast, political views seem more stable. Respondents perceive inequality to be most unfair among people with the same education level and working in the same sector. Yet, these are also the reference groups in which lower-income people overestimate their position the most and within which respondents in general underestimate inequality the most.

**Related Literature.** Crucial for our results is the link between survey data on people’s perceptions and attitudes and information from administrative records on their real-life outcomes. Recent research has started to combine subjective information from surveys with objective information from administrative records (Almås et al., 2017; Kreiner et al., 2019; Andersen and Leth-Petersen, 2020; Epper et al., 2020). Related to our agenda, one previous study (Karadja et al., 2017) has merged survey data and administrative data to check the reported income of respondents against actual income. We go much further by using the administrative records to obtain information on many of the reference groups of the respondents, their income histories over the life cycle, and their experience of major life events.

One of our key contributions is to measure and compare perceptions and misperceptions of social position in different reference groups that vary by domain, size, and proximity to the respondent and to show their relationship to views on the fairness of inequality within these groups. Connected to our result on the perceived position within co-workers in the same firm, recent papers have analyzed the impacts on satisfaction and effort of within-firm or within-employer wage differences (Card et al., 2012; Cullen and Perez-Truglia, 2018a,b; Baker et al., 2019). Complementary to these studies, our new findings show that people care

more about income differences within co-workers, as compared to other reference groups, and that they particularly strongly misperceive inequality and their own income position within this reference group.

A second contribution is our analysis of the link between changes in social position and fairness views using the unique combination of information on individual income histories back in time, income shocks that shift social positions, and randomized information treatments. Previous literature has looked at the relationship between tastes for redistribution and living or growing up in different environments (Roth and Wohlfart, 2018; Giuliano and Spilimbergo, 2014; Malmendier and Nagel, 2011), but not on changes in social positions of the respondents.

Related to our information experiment, Kuziemko et al. (2015) show respondents information on the actual distribution of income in the U.S. and where they rank based on self-reported income in the survey, but are unable to study how it relates to misperceptions or to different reference groups. Cruces et al. (2013) find that those who overestimate their position in the overall national distribution tend to demand higher levels of redistribution when informed about their true position. Conversely, Karadja et al. (2017) show that those who underestimated their position demand less redistribution. Fehr et al. (2019) provide information about position in both the national and international distribution and find that only demand for national redistribution decreases with national relative income. Perez-Truglia (2020) studies a natural experiment in Norway that made tax records and incomes easily visible online. He finds that this transparency substantially increased the gap in happiness and life satisfaction between higher and lower income individuals, which is consistent with our findings that those who realize they are ranked lower than they thought perceive inequality as more unfair.

Our third contribution is to show that people overall are well-informed about social positions although they systematically underestimate inequality by believing others are closer to themselves than they really are. Our finding of a high accuracy of perceptions is in contrast to the few existing studies on people’s perceived ranking. Cruces et al. (2013) find an inverted S-shape between perceived own position and actual own position, similar to what we find before correcting for the center bias. Karadja et al. (2017) find that a majority of surveyed individuals misperceive their position in the income distribution and believe they are ranked

lower than they actually are. Fehr et al. (2019) find that respondents are misinformed about their positions in both the global and national income distributions.

We believe that the high accuracy of perceptions we detect is due to a number of methodological advances. We demonstrate how large misperceptions arise when asking people about their own social position because of a center bias in the elicitation and use a simple method to filter out this type of systematic error. In addition, eliciting people’s perceptions of the income distribution allows us to disentangle possible misperceptions along those dimensions from misperceptions of own position. We focus on people’s position relative to peers of their cohort, which neutralizes large differences due to life cycle effects. Indeed, as we show, small changes in the definition of the relevant age group (e.g., cohort, vs. all adults, or the full working population, or the total population including those below 18 and retirees) have large effects on the percentiles of the distributions, which makes it important to ask people about their position in very clear and well-defined groups. In addition, position within one’s cohort is arguably a more relevant measure of social position for fairness concerns than thinking about position relative to people of all ages. We also use a well-defined concept of income, by asking respondents in the survey about income as it appears on their last tax return and their corresponding perceptions about income positions. This allows us to distinguish misperceptions of social position from misreporting or misperception of own income. To make the elicitation procedure as precise as possible, we use video instructions with illustrations on income ladders to explain the concept of income positions and what respondents are asked to do, and a corresponding graphical interface where respondents report percentile levels (median, “P50”, and 95th percentile, “P95”), and their own positions in the distribution. Finally, our sample is an order of magnitude larger than existing studies, which implies that we can provide more precise evidence.

Our paper is also broadly related to empirical work documenting that people care about relative income and that their social positions shape their well-being (Easterlin, 2001; Easterlin et al., 2010; Clark and Oswald, 1996). Luttmer (2005) shows that holding own income constant, self-reported happiness declines as neighbors’ incomes increase. Using German panel data, Ferrer-i Carbonell (2005) finds that people’s income rank in their reference group is a determinant of happiness and well-being whose importance is comparable to that of their

own income. Kuziemko et al. (2014) highlight the role of “last-place aversion,” a particular form of relative position concerns whereby individuals particularly fear being ranked last. Charité et al. (2015) point out the importance of reference points, while we highlight the need to consider specific reference groups. Fisman et al. (2020) show that people care about inequality in a non-linear way relative to their own position, putting weight both on their nearest neighbors and on the top of the distribution. Using online surveys, Weinzierl (2014) demonstrates that people do not hold utilitarian preferences, but rather have other, mixed fairness views.

**Organization.** Section 2.2 describes our survey, the administrative data, and our sample. Section 2.3 analyzes respondents’ perceptions and misperceptions about the distributions of income and their own position in various reference groups. Section 2.4 studies the relationship between perceived social position and fairness views. Section 2.5 offers some concluding remarks.

## 2.2 Data Collection, Survey, and Administrative Data Linkage

### 2.2.1 Survey Sample and Link to Administrative Data

**Target Sample.** Assisted by Statistics Denmark, we conducted a large-scale survey in February and March 2019. We sent out survey invitations to a representative sample of 50,100 respondents, born in Denmark in the years from 1969 to 1973, randomly selected by Statistics Denmark. The respondents were 45 to 49 at the time of the survey and, hence, no longer enrolled in formal education, well into their careers with a large share of their lifetime income realized, but still quite far from retirement. We excluded immigrants because we ask people about histories, schoolmates, and parental positions, which are only available for Danish-born respondents.

**Survey Method.** Our survey method is original and leverages an official channel of communication of the Danish public authorities with citizens. The invitations were sent out through the secure website “Digital Post,” used to receive and read mail from public



authorities. By law, all citizens older than 15 have to have an electronic mailbox where they receive information from public institutions, for example tax and health authorities. Communications may also come from private companies, for instance salary statements from employers or account statements from banks. The use of this official channel of communication, together with the University of Copenhagen’s stamp, likely increased the credibility of our survey and experiment, and of the information provided to respondents, which sets the setting apart from lower-stakes survey environments. To incentivize respondents further, they were told that those who completed the full survey would be enrolled in a lottery for 100 gift cards with a value of 1,000 DKK (150 USD) each to be used in more than 150 chains of stores in Denmark.

The average time for completion of the survey was 33 minutes with a median time of 25 minutes (the full distribution of time spent on the survey can be seen in Appendix Figure A-2.5). Responses were linked by Statistics Denmark to the register data using the social security number (assigned to all Danes at birth), which ensures a precise and unique match.

**Testing for Selection into the Survey and Attrition.** Thanks to the register data, we can analyse selection into the survey. Indeed, we know the characteristics of the respondents who entered the survey, of those who completed it, of those who were sent an invitation but chose not to participate, and of those who were not sent an invitation at all. Table 2.1 shows summary statistics for our sample of people who received an invitation and completed the survey (column 1), and compares it to the characteristics of those who received an invitation to participate and started the survey, regardless of whether they completed it or not (column 2), the characteristics of the full Danish-born population in these cohorts, excluding non-Danish born people (column 3) and the full population in these cohorts, including immigrants (column 4). The invitee group of people who received an invitation to participate, regardless of whether they did start the survey or not, is not shown here as it is almost perfectly identical to the full Danish-born population excluding immigrants (column 3) in these cohorts, as should be the case given that they were randomly drawn from this group by Statistics Denmark. The final analysis sample of respondents who completed the survey has on average somewhat higher income and education levels than the full target population in column 3, but is representative in terms of region of residence, age, and gender. Compared

Table 2.1: Summary statistics: Sample compared to population

	Analysis sample (1)	Started survey (2)	Full population (excl. immigrants) (3)	Full population (4)
<b>Demographics</b>				
Male	0.51	0.47	0.51	0.50
Age	47.0	47.0	47.0	47.0
Married	0.63	0.61	0.57	0.58
Immigrant	0.00	0.00	0.00	0.13
Descendant	0.00	0.00	0.01	0.00
<b>Income Position</b>				
Income position	64.2	59.6	53.3	50.5
Bottom 50%	0.29	0.36	0.46	0.50
Middle 40%	0.54	0.50	0.43	0.40
Top 10%	0.17	0.14	0.11	0.10
<b>Education</b>				
Primary education	0.08	0.10	0.16	0.17
Upper secondary edu.	0.06	0.06	0.05	0.06
Vocational education	0.31	0.34	0.39	0.38
Short cycle higher edu.	0.09	0.08	0.07	0.07
Bachelor programs	0.27	0.26	0.20	0.20
Masters programs	0.19	0.17	0.13	0.13
<b>Socio Economic Status</b>				
Self-employed	0.04	0.04	0.06	0.06
Employee	0.90	0.87	0.80	0.77
Unemployed	0.01	0.02	0.02	0.02
Not in work force	0.05	0.07	0.12	0.14
Private Sector	0.66	0.65	0.69	0.70
<b>Regions</b>				
Copenhagen	0.31	0.30	0.30	0.32
Sealand	0.16	0.16	0.16	0.15
Southern Denmark	0.21	0.21	0.21	0.21
Middle Jutland	0.23	0.23	0.23	0.22
North Jutland	0.09	0.09	0.10	0.10
<b>Parents' Income</b>				
Mother's income position	53.1	52.1	50.5	50.2
Father's income position	53.3	52.4	50.8	50.5
Observations	9415	13686	339231	389863

*Notes:* *Full Population* is the full Danish population born between 1969 and 1973. *Full population (excl. immigrants)* is the population our contact sample was drawn from. This sample was provided by Statistics Denmark and is the full population excluding immigrants. *Started survey* are the respondents who started the survey. *Analysis sample* are respondents who completed the survey and are used in the analysis. All variables are indicators, except for the income positions, which are based on the percentile rank position within the cohort of the respondent.

to other surveys, the top of the income distribution is very well-represented. People from the top five percent of the income distribution make up almost 8% of our analysis sample.

The use of the official Digital Post channel are perhaps the reason we are able to sample extensively from the top of the income distribution, a group that is typically very hard to reach with standard survey methods.

Out of the 50,100 people invited from the population 13,686 clicked on the personal link in the invitation (column 2) and 10,089 completed the survey. After dropping respondents for whom the reported birth year or gender do not match the register data (19 respondents), who spent less than ten minutes answering the survey (50), who did not report their income as instructed in the survey for example by reporting monthly instead of annual income (343), had zero or negative income according to the register data or missing background register data (61) or who skipped one of our key questions (201), we have 9,415 respondents in total (column 1).<sup>1</sup> The response rate of 20% ( $=10,089/50,100$ ) is reasonably high when contacting a representative sample of new potential respondents that have never expressed a particular interest in taking surveys.<sup>2</sup>

Appendix Table A-2.1 highlights which characteristics predict the drop out rate and at which point respondents drop out. Out of those who start the survey, 6% dropped out at the consent page or are screened out for the reasons listed above; 10% drop out when having to report their income. Only 1% drop out after the treatment. This means that attrition is not selectively driven by the treatment, as confirmed by the insignificant coefficient on treatment status. Men, non-married, higher-income and more educated respondents are less likely to drop out.

---

<sup>1</sup>The completion rate of 74% ( $=10,089/13,686$ ) may seem low, but our invitees are not people who have signed up in advance to participate in survey panels as is the case in other settings. Instead, our potential respondents receive an invitation through the official Digital Post, which probably leads many to click on the survey link in order to learn more about this somewhat unusual for them invitation. Once people realize it is a research survey they are not obliged to answer and they have to report personal information some of them drop out. In regular survey settings where respondents have signed up to receive survey links, those not interested do not even click on the link to start with as there is no element of surprise for them. In our case this will appear as attrition, while in other settings, we will never get to see who did not click on the survey link to start with.

<sup>2</sup>For comparison, a recent study in Denmark invited similar cohorts by ordinary mail and reports a response rate of 13% (Epper et al., 2020).

### 2.2.2 Survey Outline

The survey consists of five blocks of questions and is available in full in Appendix 2.A.1. In addition, a consent page informs respondents about the use of their responses in accordance with the General Data Protection Regulation of the European Union and a conclusion section asks respondents whether they thought that the survey was left- or right-wing biased. 81% think the survey is neutral, 14% that it is left-wing biased and 5% that it is right-wing biased.

**Background and Political Views block.** This block contains questions on birth year, gender, educational attainment, and sector of employment. These answers are later used to inform respondents about their positions relative to other people in the same large reference groups (see Table 2.2 for a definition of each reference group). We also ask about voting behavior and attitudes towards economic policy:

*“Which party did you vote for in the last general election (in 2015)? [10 parties; Other; Did not vote; Do not wish to answer]”*

*“How would you describe your attitude on economic policy? [Very left-wing; Left-wing; Moderate; Right-wing; Very right-wing]”*

**Income block.** This block asks about the income of the respondent one year ago (earned in 2017) and includes wage income, self-employment income, and taxable income benefits and transfers (composed mainly of unemployment insurance benefits, disability benefits and social assistance). We ask separately about these three income components and with the sum of the components appearing on the screen (see an image of the exact formulation in Appendix Figure A-2.1). The breakdown of total income into smaller parts is done to help people report the correct income and to highlight that self-employment income and taxable benefits are included in total income. We include taxable benefits and transfers to reflect the fact that they contribute to income and leaving them out may lead us to wrongly rank individuals, e.g., individuals receiving UI benefits are in general better off economically both in the short run and in the long run than individuals receiving social assistance. Respondents are informed that it is important to report the income correctly and that they can see the

Table 2.2: Definition of reference groups

Reference group	Definition
Large reference groups	
Cohort	People born the same year.
Gender	People born the same year with the same gender.
Municipality	People born the same year currently living in the same municipality.
Educational level	People born the same year with the same level of education: basic school, upper secondary education, vocational education and training, short cycle higher education, bachelor degree and master or PhD degree. Uses the Danish DISCED education classification, which follows the international education classification ISCED.
Sector of work	People born the same year and working in the same sector: Construction, real estate, business services, finance and insurance, trade and transport, manufacturing, information and communication, culture, agriculture, public work. Uses the Danish Sector Codes DB07, which is a sub-classification of the NACE classifications of the EU.
Small reference groups	
Schoolmates	People born the same year who went to the same school the year they turned 15.
Co-workers	People working in the same workplace. Workplace is defined as a single address entity, e.g., for a firm with multiple locations, each location is a separate workplace.
Neighbors	For people living in an apartment, the neighbors are people from age 25 to 65 who live in the same stairwell. For people living in a house, the neighbors are people from age 25 to 65 who live on the same road.

amounts on their annual tax statement (available online). Our rationale for asking about income as it appears on the tax statement is to be able to base the analysis on a well-defined income concept that is both clear to the respondent and for which the true value can be verified in the register data. With the exception of self-employment income, the income components are third-party reported to the tax agency and pre-populated on the tax return. Tax evasion is in general low in Denmark and close to nil on third-party reported income components (Kleven et al., 2011).

To avoid making the survey too complicated and time consuming, we exclude capital income, deductions and tax payments. This is not an important issue for our analysis for two reasons. First, our narrower income definition makes up almost all of total income

as calculated by Statistics Denmark for most respondents, which includes capital income. Thus, the average across individuals of our narrower income concept relative to average total income according Statistics Denmark is 96.0%; the median income according to our definition represents 98.5% of the median total income according to Statistics Denmark.

Second, and crucially, Appendix Figure A-2.6 shows that the income rank positions based on total income line up almost perfectly with the the positions based on our income definition. In fact, this remains the case if we instead used a third definition of income, namely Statistics Denmark’s measure of “disposable income” that includes the imputed value of housing, interest deductions, and tax payments.

**Perceptions block.** This block elicits people’s perceptions about the median (hereafter, P50), the 95th percentile (hereafter, P95) and their own position in the distribution of each of the five large reference groups. The block starts with a video that uses a ladder and 100 stick people to explain the different positions in the income distribution. It states and illustrates, for instance, that the P50 is the income level for which 50% have a lower income and 50% have a higher income. The full script for and link to the video are in Appendix 2.A.2. After this video, we elicit respondents’ perception of the P50 and P95 incomes for their cohort (see Appendix Figure A-2.2). We then ask the respondents to use a horizontal slider to indicate their perceived P50 and P95 income levels for their municipality, education group, gender group, and sector (see Appendix Figure A-2.3). Respondents are subsequently prompted to place themselves within each of the five large reference groups using a vertical slider next to the illustrative ladder that was also used in the explanatory video (see Panel A of Figure 2.1). We also ask respondents about their parents’ positions in the income distribution of all the parents of the other people in their cohort and to compare their own income to that of their sibling(s). For neighbors, co-workers, and former schoolmates, we first asked the respondent about the perceived number of individuals in these reference group (denoted by  $N$ ) and then asked them to report their perceived income position on a horizontal slider going from 1 to  $N$  (see e.g., Figure A-2.4 for the co-worker question). For these small reference groups, it does not make sense to ask about moments of the distribution such as P50 and P95 as we do for the large reference groups.

**Treatment block.** This part is presented for the treatment group at this point in the survey and for the control group at the very end of the survey (so that it does not affect any of their answers). The treatment informs respondents about their true social positions. Based on the respondents' earlier answers to the questions in the background and income blocks, we interactively calculate their true positions in each of the five large reference groups. *For each of these five reference groups*, the treatment reminds people which position they had earlier responded to be in, shows them their actual position, and highlights how much higher or lower they are in the distribution compared to where they thought they were. Panel B of Figure 2.1 shows a screenshot from one of the treatment screens (in English, and only for the cohort reference group) for a fictitious respondent with a positive misperception of their own position within her cohort. In this example, the respondent indicated that they thought they were ranked at position 70; the treatment informs them that they are, in fact, ranked at position 57 and points out the misperception gap in red on the ladder and numerically in the text above. The misperceptions of the positions within the other four reference groups are shown in the same way.

An alternative treatment could construct many sub-samples and inform respondents about their true position in only a single group. However, informing respondents of their position in one reference group only may still make them update their beliefs about their position in the other groups, but in a way we do not control. By telling them their actual positions in each group, we are certain about the information provided. Due to Danish rules of conduct, we cannot show respondents their true position in the small reference groups (co-workers, neighbors, former schoolmates).

Appendix Table A-2.2 shows that the treatment and control groups are balanced in terms of observable respondent characteristics.

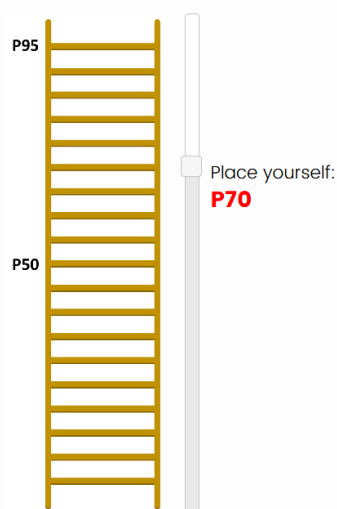
**Outcomes block.** This sequence asks about views on fairness of inequality within the respondent's cohort and (large) reference groups, and also about the role of effort versus luck, and political view. For each reference group, we asked a standard question about fairness of inequality and a standard question about the role of effort versus luck. We only asked two questions to avoid increasing the length of the survey too much and selected questions that could be applied with the same formulation across all reference groups. However, we also

Figure 2.1: Example Survey Pages

(A) Eliciting perceptions of position

Rank among all people **born in 1970**

You previously reported that you had a yearly income in 2017 of 400000 DKK before tax. We will now ask you to report where you think this income placed you on the income ladder in 2017 for people who were born in 1970. Use the slider to select your position. Later, we will inform you about your true position.



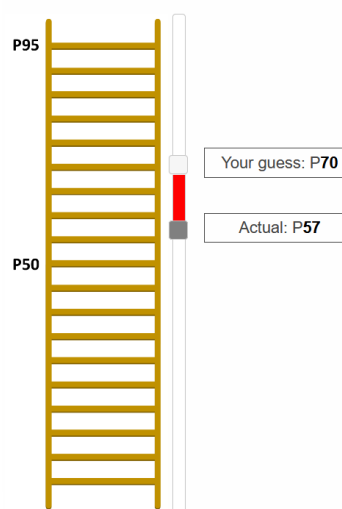
(B) Information treatment (example of cohort reference group)

**Rank among all people born in 1970**

You GUESSED that you were on position **P70**.

Based on the income you reported, your TRUE position is **P57**.

You are actually 13 positions lower on the ladder than you thought.



(C) Question on Unfairness of Inequality

On a scale from 1 to 7 where 1 is "Completely fair", 4 is "Neither fair or unfair" and 7 is "Completely unfair", indicate to what extent you think that is fair or unfair that there are differences in income among people born the same year as you **WITHIN** the following groups that you are yourself a part of?

	Completely fair		Neither fair or unfair			Completely unfair	
	1	2	3	4	5	6	7
Differences in income among people <b>born in 1970</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Differences in income among <b>men</b> born in 1970	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Differences in income among people, living in <b>København's municipality</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Differences in income among people with the educational level <b>Master or PhD program</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Differences in income among people working in the sector <b>Finance and insurance</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*Notes:* Panel A shows the question eliciting the respondent's perceived position in the income distribution. In this example the respondent is born in 1970, has an income of 400,000 DKK and thinks they are in position 70. The slider is initialized at P1. Panel B shows part of the information treatment this respondent receives. The bottom panel shows a screenshot of the fairness of inequality question and illustrates how the reference groups are adapted (in bold) based on the respondent's earlier answers so as to ask directly about their reference group.



asked a few questions related to general well-being and views on inequality and show results for these outcomes in Appendix 2.D.

The main questions are as follows:

*“On a scale from 1 to 7 where 1 is “Completely fair”, 4 is “Neither fair nor unfair” and 7 is “Completely unfair”, indicate to what extent you think that it is fair or unfair that there are differences in income among people born the same year as you WITHIN the following groups that you are yourself a part of?”*

[The screen then lists five reference groups, filling out their labels directly with the respondent’s information from the earlier block, as can be seen in Panel C of Figure 2.1.]

*“Now, think about people born the same year as you WITHIN these groups (indicated below). On a scale from 1 to 7 where 1 is “Only luck”, 4 is “Equally important”, and 7 is “Only effort”, indicate to what extent you think that differences in income are caused by differences in people’s efforts over their lifetime or rather by luck? By luck, we mean conditions, which you have no control over. By effort, we mean conditions, which you can control.”*

*“Which party would you vote for if there was a general election today? [11 parties; Other; Do not wish to answer]”*

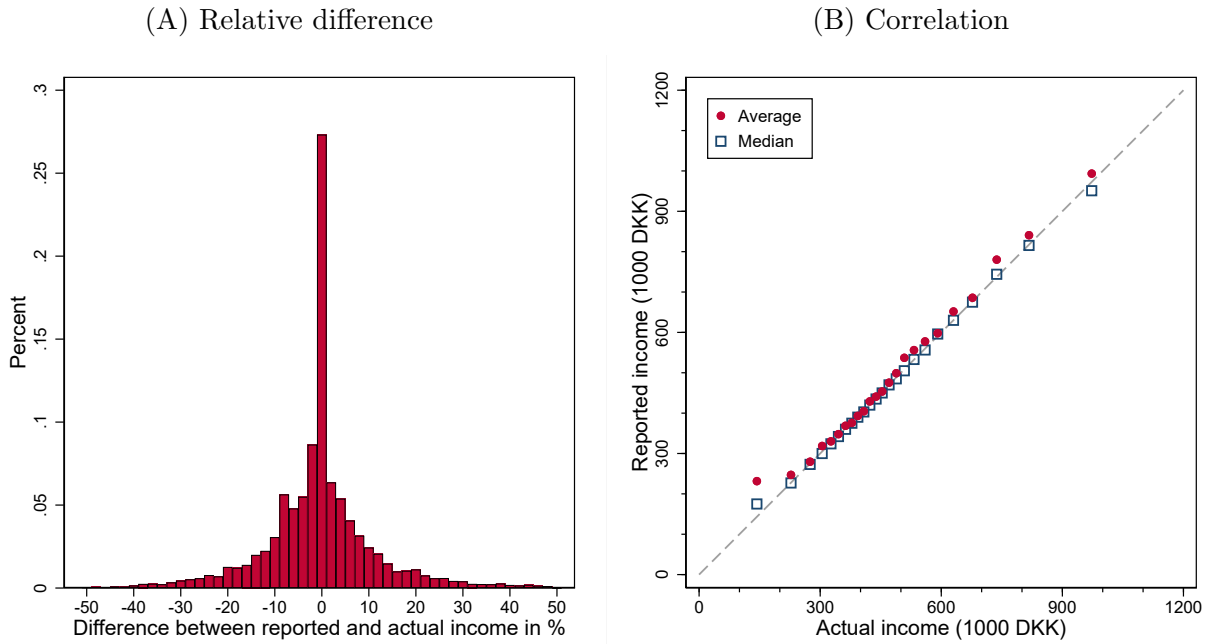
### 2.2.3 Response Quality

Figure 2.2 shows that the reported incomes in the survey match the actual incomes on tax returns well. Panel A depicts the distribution of the difference between reported and actual income. The average percentage difference is less than 0.5%. More than 25% of the respondents report an income that deviates less than 1% from their actual income and more than 70% of the respondents have a discrepancy that is less than 10%.<sup>3</sup> Panel B shows that across different levels of actual income both the average reported income and the median reported income within each bin are very close to the actual income.

---

<sup>3</sup>We see a small spike at a reported income 8-9% below actual income. Respondents are asked to report their income including labor-market contribution, which is 8% of income before taxes; a few respondents seem to report their income excluding these contributions.

Figure 2.2: Reported and actual income



*Notes:* The left panel shows a histogram of the relative difference between reported and actual income in %. The bin width is 2 and the plot is truncated at  $\pm 50$ . The right panel shows binned scatter plots of the average and median reported income against actual income (measured in 1000 DKK). The 25 bins have approximately the same number of respondents.

Next, we describe how the reference groups reported by the respondents align with the the official classifications in the administrative data and how we can account for possible discrepancies in the analysis. Appendix Table A-2.3 shows that information on gender and cohort are aligned and that 98% of the respondents report living in the correct municipality. Respondents are also relatively precise when they report educational level and sector of work, and the mismatches can be explained and dealt with in a consistent manner. Overall, for 74% of respondents, education levels reported in the survey match the register data and for 72% the sector reported matches. In fact, on the education level dimension, 93% of the respondents with a bachelor or master program as their highest level of education according to the registers report the correct level of education. Almost half of the respondents who report an incorrect educational level have a vocational education and training program as their highest level of education. The majority of these respondents report their highest level of education as either upper secondary school or short cycle higher education. The explanation for the first group is that many consider upper secondary school as a higher level than vocational education, but according to the standard education classification this is not

the case. For the second group, the majority have an education within *Office, commercial and business service*, and therefore plausibly think they have a short cycle higher education, but according to the education classification these are also categorized as vocational educations.

For the sector dimension, it is understandable that some respondents have difficulties in knowing the correct *label* of their sector, which is based on the standard classification and labelling of sectors as described in Table 2.2. For two of the large and well-defined sectors such as *Finance and insurance* and *Public administration, education, health and social activities* 89% and 91% of people in the sectors correctly report working in those sectors. In the smaller and less well-defined sectors such as *Culture, leisure and other services* and *Real estate activities* only 50% and 56% of people in the sectors correctly report working in those sectors. For instance, for *Real estate activities* more than half of those who report an incorrect sector, report working in either *Construction* or *Business service*. The discrepancies are thus likely attributable to genuine ambiguity, rather than to careless answers or misunderstandings.

The benchmark results we present use the reference groups respondents believe they belong to. Appendix 2.E shows that the conclusions are unchanged if we instead use their actual reference groups or only include respondents who perceive their reference group correctly. This reflects that, to start with, the differences are not large across the groups that are difficult for respondents to differentiate between.

## 2.3 Perceptions and Misperceptions about Social Positions

In this section, we describe people’s perceptions and misperceptions about the distribution of income and their own position in their cohort and their various reference groups.

### 2.3.1 Cohort Income Distributions and Within-Cohort Position

We start with people’s perceptions of the income distribution of their cohort. Our approach of asking about perceptions at the cohort level neutralizes life-cycle effects. This is both practically convenient and normatively more relevant. Arguably, large income variations due

to life cycle effects are considered normatively less important than large income differences across similar, working-age people. Respondents in our chosen target cohorts are close to the peak of their career paths and income trajectories, with much of their permanent income already realized.<sup>4</sup>

**Perception of the cohort income distribution.** Panel A of Figure 2.3 plots respondents' misperceptions about the P50 income level (red curve) and the P95 income level (blue curve) relative to the actual levels. For the P50 income level, errors are symmetric around zero and bell-shaped. 45% of respondents estimate the median with at most a 10% error; 75% estimate it with at most a 25% error. For comparison, the errors when people report their own income (black curve) are such that 70% (respectively, 90%) percent report correctly within a 10% (respectively, 25%) error band.<sup>5</sup> Against this benchmark, people seem reasonably well aware of the P50 income level of others in their cohort. As compared to the perceived P50, there is larger variance of the perceived P95 level and a small majority of people underestimating its level.

Panel B reveals an increasing relationship between the average perceived P50 income level and the respondent's own position in the distribution; the same holds for the median perceived P50. Higher-income respondents tend to over-estimate the P50 and lower-income people to underestimate it. Yet, except for respondents in the very top and the very bottom of the distribution, the average prediction errors are within 5% of the actual P50 value and the mean and median are very similar.<sup>6</sup> Similarly, there is a weak positive relationship between the perceived P95 level and the position of the individuals in Panel C. At most income levels up to percentile 95, the average perception error for the P95 is below 10%. By contrast, those in the very top of the distribution starkly overestimate the P95 by 50%. For most income positions, the median perception is slightly below the average perception.

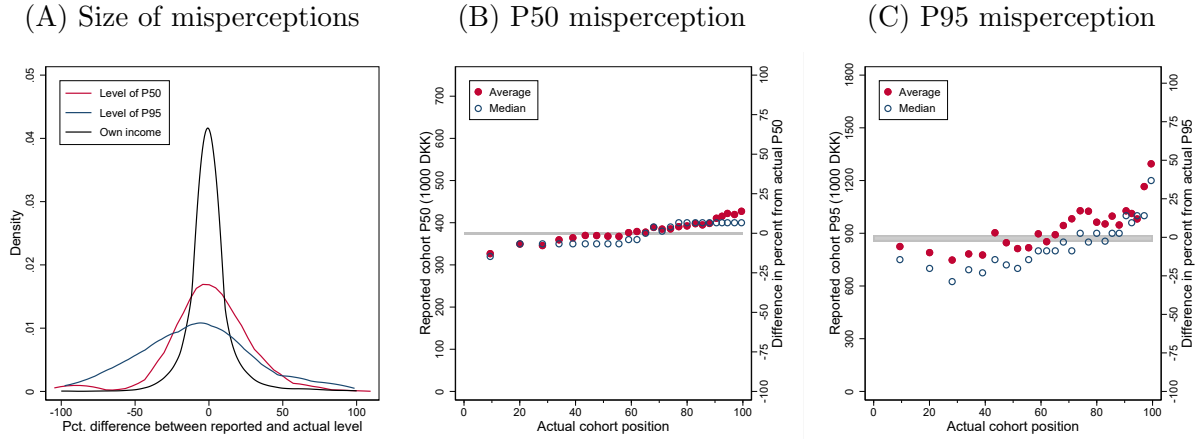
---

<sup>4</sup>Appendix 2.F highlights the pitfalls of asking about the full income distribution without specifying proper age limits. Both the P50 and P95 income by cohort varies drastically across different ages or by excluding or including some cohorts.

<sup>5</sup>Regarding this comparison, note that the spike at exactly zero in Panel A of Figure 2.2 suggests that many of the respondents have checked their actual income on the tax return when answering the survey. It is not possible to similarly easily find information about the median of a cohort.

<sup>6</sup>Equivalently, in terms of positions, an income level 5% below the actual median (DKK 350,000) corresponds to a percentile position of 44-45 within the cohorts and an income level 5% above (DKK 400,000) corresponds to a percentile position of 56-57.

Figure 2.3: Perceived P50 and P95 of cohort

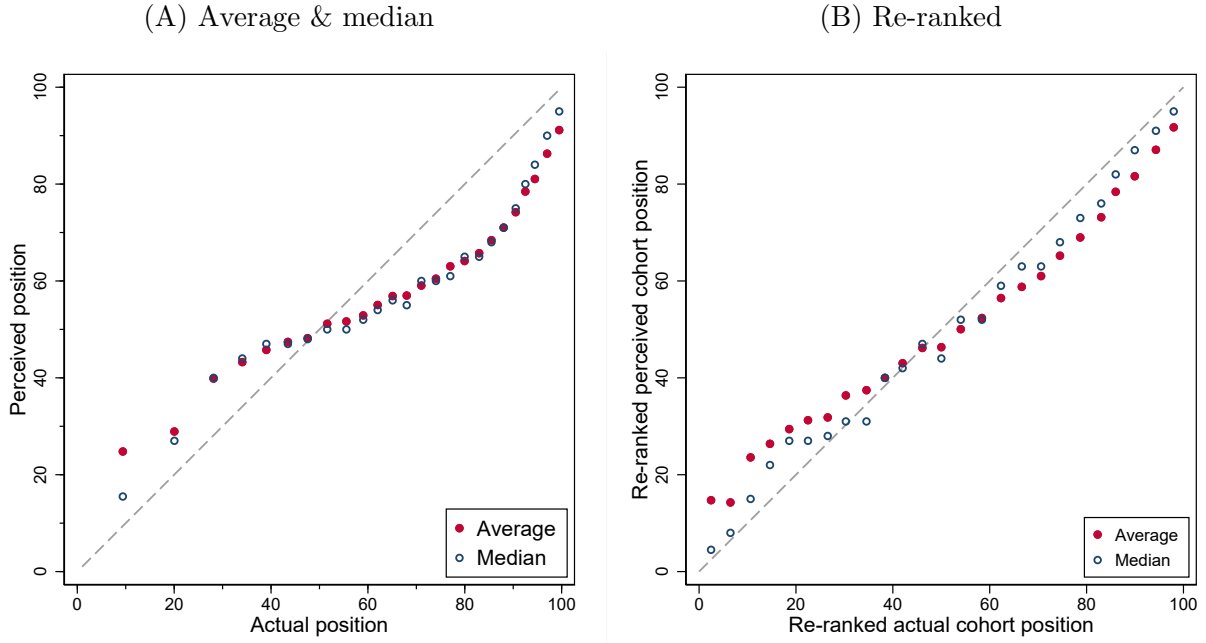


*Notes:* Panel A shows the distribution across respondents of the misperception in percent of the level of P50 and P95 (i.e., the percent difference between perceived and actual levels). For comparison, we plot the misperception of the respondent's own income as reported in the survey and their actual income on the tax return. The distributions are smoothed using epanechnikov kernels with a bandwidth of 5 for *Own income*, 10 for *Level of P50* and 15 for *Level of P95*. Panel B (respectively, Panel C) is a bin scatter with 25 bins of the average and median perceived P50 (respectively, P95) reported in DKK (left scale) and the corresponding misperception in percent (right scale) by actual position in the within-cohort income distribution. The perceived P50 and P95 are winsorized at the 5th and 95th percentiles within each bin.

**Perception of own position within the cohort.** Figure 2.4 shows the relation between respondents' actual and perceived positions in their cohort. Panel A plots average and median perceived position against actual position.<sup>7</sup> The two curves both have a horizontal, inverted S-shape, whereby people below the median income level tend to overestimate their position, while people above the median tend to underestimate it. The misperceptions seem large compared to the misperceptions of the moments of the distribution (P50 and P95), but to some extent such a pattern is mechanical because of a simple “center bias” logic: people at the highest position can only weakly underestimate their position, while people at the lowest position can only weakly overestimate it. Put differently, perceived income ranks are not uniformly distributed, unlike actual income ranks (see Appendix Figure A-2.8 for the distribution of perceived ranks). To filter out this center bias mechanism, we ask: Are those ranked at position  $x$  with respect to actual income also ranked at position  $x$  with respect to perceived position (i.e., is  $x$  also their rank in the *perceived* position distribution)? Panel B of Figure 2.4 shows that this is indeed the case, with all points lying close to the 45 degree line. Thus, the systematic misperception of own position almost disappears.

<sup>7</sup>Appendix Figure A-2.7, Panel A shows that if we compute the actual position in different ways using the average income on the tax return over the last three years or the reported income in the survey, the pattern is the same. Panel B depicts the 25th and the 75th percentile by actual position. It reveals more variation in perceived position of people in the very bottom of the distribution relative to those at the very top.

Figure 2.4: Perceived position within cohort



*Notes:* The left panel is a bin scatter of the average and median perceived position by actual position in 25 equally-sized bins. Actual position is based on the income from the tax return. In the right panel, we re-rank both actual and reported position, such that they are uniformly distributed from 1 to 100 in our sample, and plot the average and median perceived position by actual position.

**Which respondents are most accurate?** In addition to actual cohort position, what are the characteristics that drive respondents' perceptions and misperceptions? Table 2.3 regresses measures of accuracy and inaccuracy of perceived position, P50, and P95 on indicator variables capturing the respondents' gender (2 groups), political view (3), area of residence (5), education level (6), and sector of work (11) as well as cohort fixed effects and actual income position fixed effects. For conciseness, we only show the coefficients for the significant characteristics, with the full table in Table A-2.4. The dependent variable in Panel A is an indicator for whether the respondent is among the 20% of the sample with the largest misperception of their position, the P50, or the P95; by contrast, the dependent variable in Panel B is an indicator for whether the respondent is among the top 20% most accurate ones.<sup>8</sup> We find that more educated respondents and male respondents are significantly more likely to be accurate across all domains (own position, P50, and P95). Characteristics related to area of residence and sector of work are not strongly associated with misperceptions.

<sup>8</sup>The 20% most inaccurate respondents misperceive their position by more than 22 percentiles, the P50 by more than 33%, and P95 by more than 66%. The 20% most accurate ones misperceive their position by less than 5 percentiles, the P50 by less than 6%, and P95 by less than 10%.

Table 2.3: Accuracy of perceptions regressed on individual characteristics

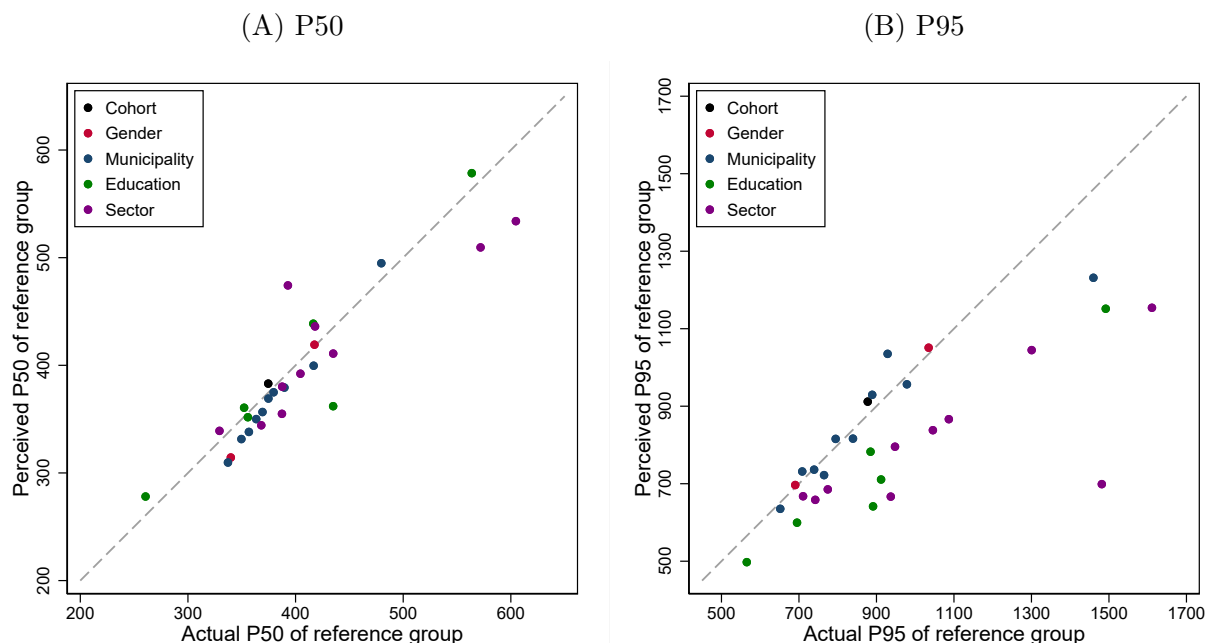
	A. Top 20% most inaccurate			B. Top 20% most accurate		
	Position	P50	P95	Position	P50	P95
Male	-0.04*** (0.01)	-0.06*** (0.01)	-0.04*** (0.01)	0.04*** (0.01)	0.02* (0.01)	0.05*** (0.01)
Left-wing	-0.02* (0.01)	0.00 (0.01)	0.03** (0.01)	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)
Right-wing	-0.03** (0.01)	-0.01 (0.01)	-0.00 (0.01)	0.03** (0.01)	0.01 (0.01)	0.00 (0.01)
Vocational education	-0.06** (0.02)	-0.04 (0.02)	-0.00 (0.02)	0.01 (0.02)	-0.00 (0.02)	0.04 (0.02)
Upper secondary edu.	-0.03* (0.02)	-0.03* (0.02)	-0.02 (0.02)	0.01 (0.02)	0.03 (0.02)	0.01 (0.02)
Short cycle higher edu.	-0.05** (0.02)	-0.07*** (0.02)	-0.04 (0.02)	0.02 (0.02)	0.02 (0.02)	0.06** (0.02)
Bachelor programs	-0.08*** (0.02)	-0.06** (0.02)	-0.02 (0.02)	0.04* (0.02)	0.02 (0.02)	0.04* (0.02)
Masters programs and PhD	-0.15*** (0.02)	-0.11*** (0.02)	-0.01 (0.02)	0.10*** (0.02)	0.05** (0.02)	0.06** (0.02)
<i>N</i>	9415	9415	9415	9415	9415	9415
<i>R</i> <sup>2</sup>	0.087	0.034	0.025	0.079	0.036	0.029
Cohort FE	✓	✓	✓	✓	✓	✓
Actual position FE	✓	✓	✓	✓	✓	✓
Regional FE	✓	✓	✓	✓	✓	✓
Sector FE	✓	✓	✓	✓	✓	✓

*Notes:* In Panel A (respectively, Panel B), the outcomes are indicator variables equal to 1 if the respondent is among the 20% of respondents with the largest (respectively, smallest) misperceptions for each variable. *Left-wing* and *Right-wing* are indicators that equal one if a respondent reported having either a very left-wing/left-wing or very right-wing/right-wing view on economic policy. For the educational level dummies, the baseline group is primary education. The *Actual position FE* is fixed effects for all 100 positions in the income distribution. Appendix Table A-2.4 shows an expanded table including coefficients on regions of residence and sectors of work. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

### 2.3.2 Income Distribution of and Position in Large Reference Groups

**Perceptions of income distribution.** Panel A of Figure 2.5 plots the average perceived P50 for different reference groups of respondents in those groups against the actual P50 of each group. Each point represents either the overall cohort, a gender group, an education group, a sector, or a set of municipalities. For example, the two red dots show how men perceive P50 of men and how women perceive P50 of women. Municipalities, which are too numerous to be plotted individually, are grouped into ten bins defined by median municipality income. In general, respondents estimate the P50 income level of their various

Figure 2.5: Perceived and actual P50 and P95 levels of large groups



*Notes:* For gender, we show one point for men and one for women. For municipality we divide the respondents into 10 similar sized groups based on the actual municipality P50 and P95 income and plot one point for each group. Each education level and sector are also represented by one point. The points show the means of the reported P50 or P95 by respondents in that group, winsorized at the 5th and 95th percentiles within the group.

reference groups relatively well. Misperceptions are largest for some sectors, particularly for the two sectors with the highest median income levels, namely “Finance & Insurance” and “Information & communication” where respondents tend to underestimate the median income.<sup>9</sup> The distributions of P50 misperceptions are similar across reference groups and mirror the distribution of misperceptions of the cohort P50 displayed in Figure 2.3 (see Appendix Figure A-2.9).

Panel B shows the perceived P95 levels for the different reference groups. Individuals are also well aware of this moment of the whole cohort distribution (black point), the gender-specific distribution (red points) and the municipality-specific distribution (blue points). Note, however, that they systematically underestimate P95 of their education group and, in particular, underestimate P95 in the sector where they work (the green and purple points are all below the 45 degree line). Thus, respondents are not well-aware of the degree of inequality

<sup>9</sup> “Information & communication” covers a wide range of industries, from computer programming to the publication of newspapers. It does not include advertising or marketing.



within their education groups and within their sectors.<sup>10</sup> Panel B and C of Appendix Figure A-2.9 shows these patterns also hold if we use group medians instead of averages.

How do respondents' perceived P50 levels for different reference groups vary with their own social positions? Figure 2.6 shows the perceived P50 levels for different large reference groups as a function of respondents' own actual positions, split by type into each panel (gender, municipality, education, and sector of work).

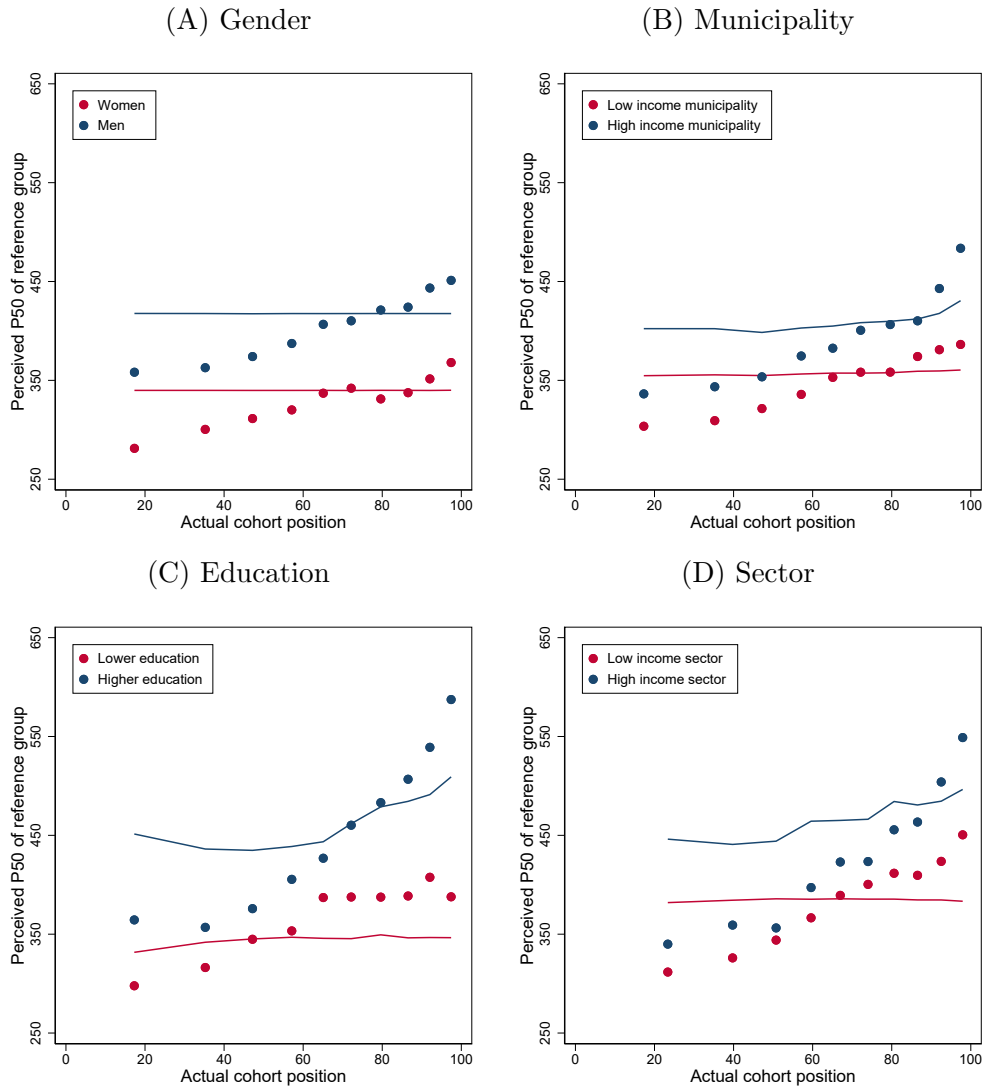
In Panel A, the horizontal lines represent the P50 income level of men and women. The dots show the average perceived P50 level of their gender group among men and women, by bins of actual cohort position. Conditional on being at the same income position in their cohort, men tend to correctly report a higher P50 income and women a lower P50 income for their own gender group. In fact, the vertical distance between the estimates of men and women at the same cohort position is very close to the actual difference between the P50 incomes. Yet, there is a systematic bias in perceptions: high-income men and women overestimate the P50 income level of their gender group, while lower-income men and women underestimate it. The belief that the median income of one's gender group is closer to one's own income is similar to the one observed for the cohort median income in Figure 2.3.

In Panel B, we split respondents into two roughly equally-sized groups, depending on whether they live in a low-income or high-income municipality. For each of these two municipality groups and by bins of actual cohort position, we plot the bin average perceived P50 income and the average actual P50 income of their municipality for respondents. We repeat the same procedure for education groups and sectors (Panel C and D). For all reference groups at all income levels, we observe that people belonging to a high income group consistently report a higher P50 for their group than people belonging to the corresponding low-income group. For the municipality and education reference groups, the differences in perceptions between low-income and high-income groups (vertical differences between the blue and red dots) tend to be somewhat smaller than the actual differences (difference between the red and blue lines) except towards the top of the distribution, while for the sector reference group these differences in perceptions are much smaller, i.e., much more compressed, than the

---

<sup>10</sup>The outlier in the lower-right corner is the sector "Agriculture, forestry and fishing." This is a small sector in Denmark measured by the number of employed people in the sector and we only have 80 respondents in our sample that work in this sector. Furthermore, it is a sector with large income inequality: the P50 income level is the lowest of the ten sectors, yet it has the second highest P95 income level.

Figure 2.6: Perceived and Actual P50 Levels by Cohort Position

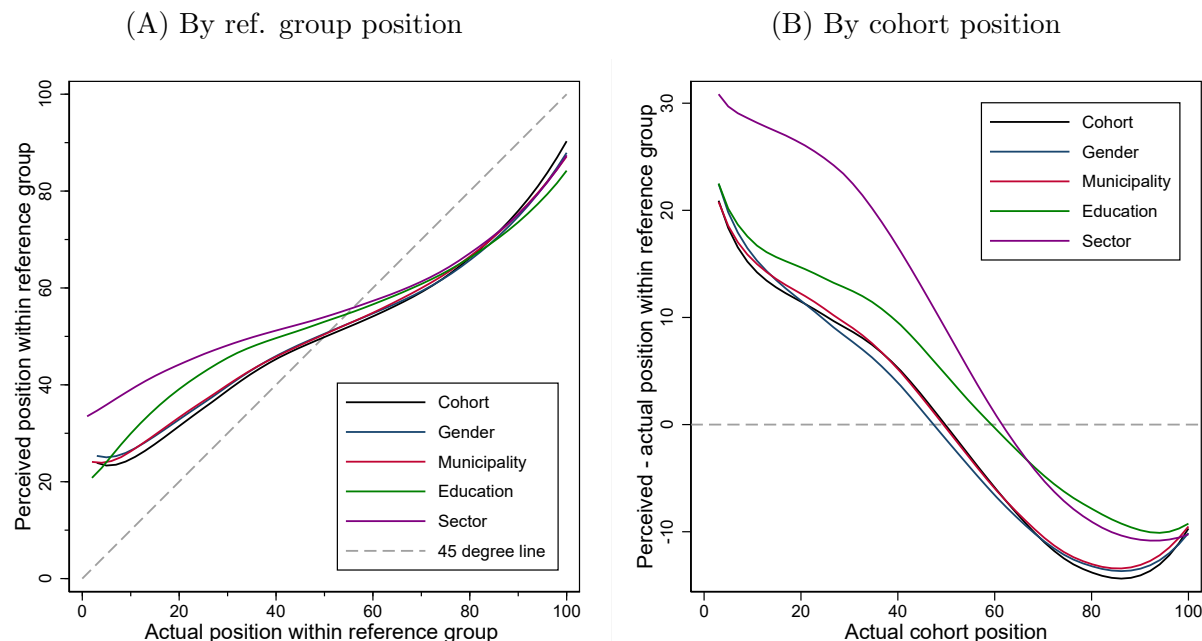


*Notes:* The solid lines indicate the actual average P50 income of their reference group for respondents grouped into bins by actual cohort positions. The dots indicate the bin average of perceived P50 income of the reference group. Municipalities and sectors are classified into two equally-sized groups based on their P50. The *Higher education* category includes short cycle higher education, bachelor programs, and master programs.

actual differences. To a large extent, this reflects the fact that people with low and middle income working in high-income sectors tend to significantly underestimate the P50 of the sector.

**Perceptions of own position.** Panel A of Figure 2.7 plots respondents' perceptions of their own position within each reference group as a function of their actual position within that group. To better compare the different reference group positions, we show local linear polynomials for each group in the same plot. Panel B recasts this information in a different way, by plotting respondents' misperception of their reference group positions for given overall

Figure 2.7: Perceived and actual position within large reference groups



Notes: Panel A plots perceived position within each reference group as a function of actual position in that reference group. Panel B plots misperception of position in reference group by cohort position. The local linear polynomials have a bandwidth of 10. We use reported reference groups both for actual and perceived positions.

position *in the cohort*. People in the lower part of the income distribution tend to overestimate their position within all reference groups; people in the higher part of the distribution tend to underestimate their positions. Especially within sector or education groups, the relation between actual and perceived position is quite weak. At all income levels, people tend to be most optimistic about their position within their education group and their sector. In particular, people with income below the median level greatly overestimate their position within the sector they work in. For example, people at the 20th percentile among their co-workers on average think they are well above the 40th percentile, while people at the 20th percentile in their municipality believe they are around the 30th percentile. This is in line with our previous conclusions that people in general tend to underestimate the inequality in their sector (as reflected by a negative misperception of P95), and that people in the lower part of the distribution significantly understate the P50 income level of their sector.

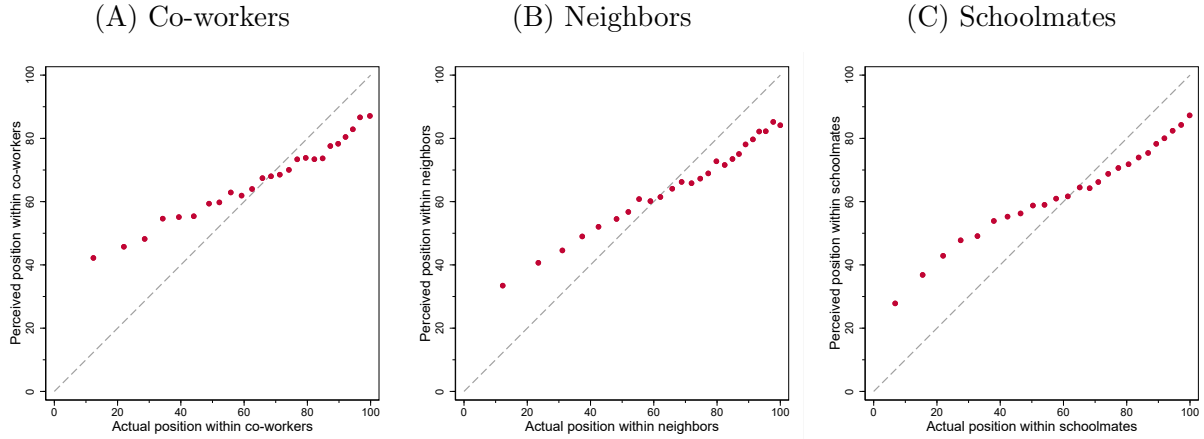
**Robustness and Extensions.** To what extent do people report very similar positions for all reference groups? If they do not really know, they may be tempted to simply state the same perceived position across the groups. Appendix Figure A-2.10 shows this not to be

the case. For each perceived position in the overall cohort distribution, we observe a lot of variation in perceived positions in the reference groups. In addition, we may wonder about the correlation between actual positions across the different reference groups. Appendix Figure A-2.11 shows respondents' actual position for each large reference group (top panels) and perceived position (bottom panels) as a function of their actual positions within the cohort. We split respondents into high-income reference groups (blue dots) and low-income reference groups (red dots), which both represent roughly half of the sample. This is reminiscent of Figure 2.6, which focuses on the P50 of the reference group, while the current figure focuses on the individual's own position. The top, left panel shows that men's position in the cohort is higher than their position in the distribution of men, while the women's position is lower than their position in the distribution of women. Furthermore, we can compare the link between actual position in the reference groups and in the cohort to the link between the perceived position in the reference group and actual cohort position, which is done in the second row of Appendix Figure A-2.11. The first panel shows that women's perceived position in the women distribution and men's perceived position in the men's distribution as a function of their overall position within the cohort have the by-now familiar inverted S-shapes because of the center bias logic described earlier. In addition, the differences in the perceptions of men and women about their positions in the respective reference groups are smaller than the actual differences. For the other reference groups, we see a similar pattern. The differences in perceptions between high- and low-income groups are most compressed for education and sector in line with the result for the perceived median. In general, the differences in perceptions are more compressed when looking at own position compared to P50, which can be explained by the center bias effect of misperceptions that moves perceptions of own position towards the mean.

### 2.3.3 Position in Small Reference Groups

In this section, we look at perceptions related to smaller reference groups that are perhaps easier to relate to and closer to a respondent's daily life. Figure 2.8 shows how people rank themselves among co-workers at the same workplace, among neighbors living on the same road (if living in a house) or stairwell (if living in an apartment), and among former

Figure 2.8: Perceived position within small reference groups

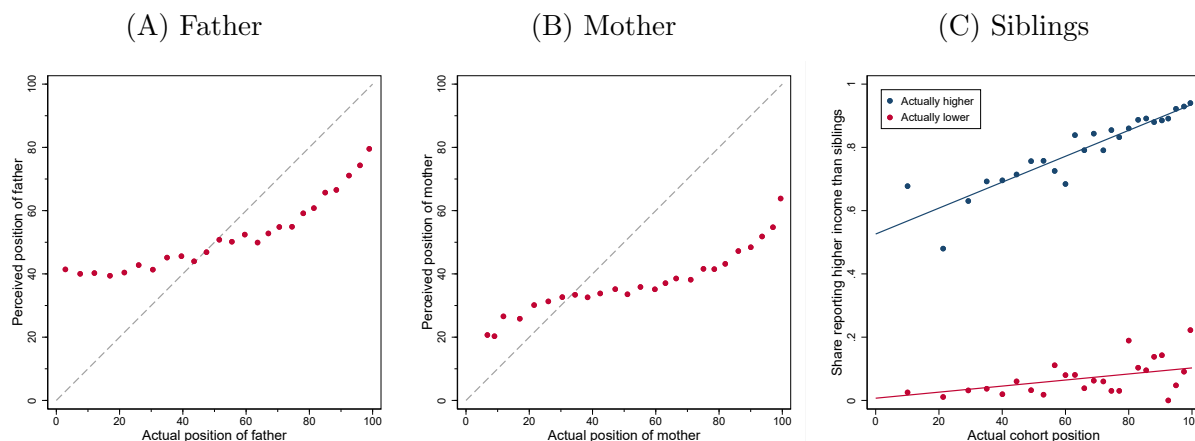


*Notes:* The panels depict the average perceived position of respondents among their co-workers in the same firm, their neighbors, and their former schoolmates, as a function of their actual position within these groups. There are 25 equally-sized bins in each panel.

schoolmates. Recall from Section 2.2 that for each group, we first asked the respondent about the perceived number of individuals in the group ( $N$ ) and then asked them to report their perceived income position ( $X$ ) on a horizontal slider going from 1 to  $N$ . Finally, to construct Figure 2.8, we compute the perceived percentile rank as  $\frac{X}{N} \cdot 100$  and the actual percentile rank using the true  $X$  and  $N$  from the register data.

The graph of the perceived position among co-workers at the same workplace in Panel A is very similar to the result for perceived position among co-workers in the same sector in Figure 2.7. In both cases, people who are in the bottom of the distribution believe that they are higher up than they truly are, e.g., respondents at the 20th percentile among co-workers in the same firm or sector on average believe that they are above the 40th percentile in those groups. In the upper part of the distribution people underestimate their positions, but the misperceptions are smaller than in the bottom. The graph of the perceived position among neighbors in Panel B is similar, but with smaller misperceptions at the lower part of the distribution. Actually, this graph is very similar to the graph for municipalities in Figure 2.7, which aligns with the conclusion that misperceptions at the bottom are larger when comparing yourself to co-workers than when comparing to people living in your area. The graph of the perceived position among schoolmates in Panel C has the same shape as the other graphs. One notable conclusion from this graph is that people tend to make smaller

Figure 2.9: Perceived position of family members



*Notes:* We asked about the respondents' perceived position of father's or mother's position when the respondent was 15 years old relative to parent's of other children from the same cohort as the respondent. We asked men about their father's position and we asked women about their mother's position. The right panel only includes responses from people reporting the correct number of siblings. *Actually higher* means that a respondent's income is at least 25 pct. higher than the siblings' mean income. *Actually lower* means that a respondent's income is more than 25 pct. below the siblings' mean income.

errors when ranking themselves among their former schoolmates during their teenage years than when ranking themselves among their current co-workers.<sup>11</sup>

Figure 2.9 shows results for family members. Panels A and B show the perceived positions of parents of the respondents, when the respondents were fifteen years old, as a function of the true position in the ranking of parents of children from the same cohort. The line is close to flat for the ranking of fathers in Panel A, except for fathers who were ranked in the top 25%, indicating that respondents have little idea about the historical income positions of their fathers. The same holds for mothers, where respondents starkly underestimate the ranking of mothers, including at the very top. Panel C shows the share of respondents who report having income higher than their siblings, as a function of their actual cohort position and split by whether they actually do have higher income than their sibling or not. In general, respondents are well aware whether they are making more or less than their siblings. The share of those who report making more and who in fact make less is very small, except for

<sup>11</sup>A deviation between perceived and actual rank could reflect that people misperceive the number of people belonging to their reference group rather than their own position within the group. In Appendix Figure A-2.12, we show that respondents are well aware of the size of their reference groups. The exception is a small share of respondents who have more than one hundred neighbors and underestimate that number. Appendix Figure A-2.13 shows that we obtain similar results if we restrict the analysis to respondents whose reported number of people in the small reference group matches the number observed in the register data within a 10% error band or if we use bin medians instead of bin averages. In addition, Appendix Figure A-2.13 shows similar patterns for co-workers and for neighbors if we split the respondents into people working in small firms versus large firms and into those living in apartments versus houses.

respondents who are ranked close to the top of their cohort. On the other hand, information is far from perfect even among siblings and, in line with evidence above, misperceptions correlate systematically with the respondent’s cohort position. People who are higher ranked than their siblings are less aware of this if they have a low cohort rank, and people who are lower ranked than their siblings are more likely to wrongly believe they are higher ranked if they belong to the upper part of the cohort distribution.

## 2.4 Relationship Between Social Positions and Fairness Views

This section analyzes the relationship between social position and views on fairness. We approach this question in four ways. First, we study the contemporaneous correlations between social position and fairness views within and across reference groups. We then leverage our data on respondents’ past income histories to relate fairness views to overall changes in social position over the lifetime and to specific changes in social position caused by major life events such as unemployment, health shocks, disability, and promotions. Finally, we analyze how fairness views are causally affected by changes in perceived social positions induced by our information treatment.

As described in Section 2.2.2, we study three types of outcomes: views on the fairness of inequality within each reference group, the belief that effort matters more than luck for differences in income within a reference group, and right-wing support represented by support for political parties that vary from left-wing to right-wing. The “unfairness (of inequality)” variable for each reference group ranges from 1 (inequality within the group is considered completely fair) to 7 (inequality within the group is considered completely unfair). The perceived “importance of effort (relative to luck)” variable measures to what extent a respondent believes that differences in income are caused by differences in people’s effort or rather by luck on a scale of 1 (only luck matters) to 7 (only effort matters).

The “Right-wing” variable is based on the party that the respondent plans to vote for if there was a general election today. They could choose among 11 parties and the two other options, “Other” and “Do not wish to answer” for a total of 13 options. To classify

parties into left- and right-wing in an objective, data-driven way, we take the control group respondents and use their answers to the economic policy view question to rank these 13 options based on the average economic policy views of respondents. The party ranked 1 has voters in the survey who on average report to be most left-wing while the party ranked 13 has voters in the survey who on average report to be most right-wing.<sup>12</sup>

In our benchmark regressions, we standardize the outcome variables into z-scores by subtracting the control group mean and dividing by the standard deviation. As a robustness check, we show the results are similar if we use indicator outcome variables throughout instead of categorical ones (see Appendix 2.G).

### 2.4.1 Fairness Views Within and Across Reference Groups

How fair do people consider income differences within the various reference groups to be and how does this vary with their own position in those reference groups? Panel A of Figure 2.10 plots the perceived unfairness of inequality for each reference group against the perceived position in that reference group. First, it shows that those who think they are ranked higher in their cohort believe inequality within the cohort to be less unfair. Similarly, those ranked higher within each reference group systematically find differences in income within that group to be less unfair. Second, views on how fair or unfair inequality is for the cohort, gender, and municipality reference groups are similar in magnitudes and exhibit an analogous relation to respondents' position. But compared to these groups, income inequalities within education and sector groups are considered considerably more unfair at all positions in the distribution. Panel B shows that those positioned higher within the different reference groups also tend to believe that income differences in these groups are the result of effort rather than luck. In this dimension, there is no major difference between reference groups either on the level or the slope.

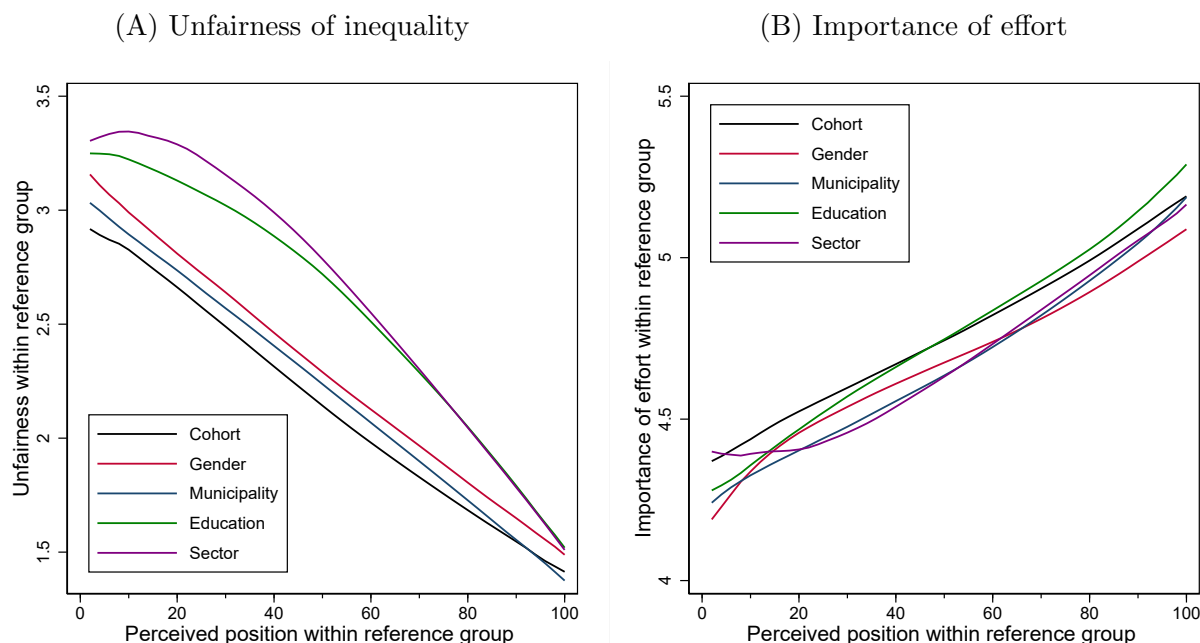
The same patterns hold if we instead plot the outcomes against actual (rather than perceived) position within each reference group or within cohort on the x-axis, as can be

---

<sup>12</sup>This data-driven ranking of parties aligns almost perfectly with the subjective ranking of the Danish parties by fourteen experts in Green-Pedersen and Kosiara-Pedersen (2020). For robustness, we verify that our results hold if we directly use the answers to the question on economic policy views rather than party support where this is possible (see Appendix 2.H for the full set of tables and figures).



Figure 2.10: Unfairness of inequality and importance of effort across large reference groups



*Notes:* The left panel plots the unfairness of inequality variable (locally linear polynomials with bandwidth 20), which is on a scale of 1 (completely fair) to 7 (completely unfair). The right panel plots the perceived importance of effort on a scale of 1 (only luck matters) to 7 (only effort matters). The sample is restricted to respondents in the control group only.

expected in light of the findings in Section 2.3 that perceived positions are in general in line with actual ones (see Appendix Figure A-2.14, which also shows the distribution of the outcomes by reference group).

Table 2.4 confirms the graphical analysis in Figure 2.10 without controls (Panel A) and with fixed effects for cohort, gender, municipality, educational level, sector of work, and employment status (Panel B). The last column shows the third main outcome, i.e., right-wing political views. For comparability across columns and figures, outcome variables are standardized into z-scores by subtracting the mean and dividing by the standard deviation of the control group sample for all figures and tables below. The “Outcome mean” row shows the average, non-standardized fairness views by reference group and highlights the different levels of perceived unfairness across reference groups. The precisely estimated means confirm that education and sector-level income differences are considered significantly more unfair than inequality overall within cohort or within other reference groups.

The inclusion of controls does not affect the results substantially. A higher position in the cohort and a higher positive misperception of it significantly correlate with weaker perceived

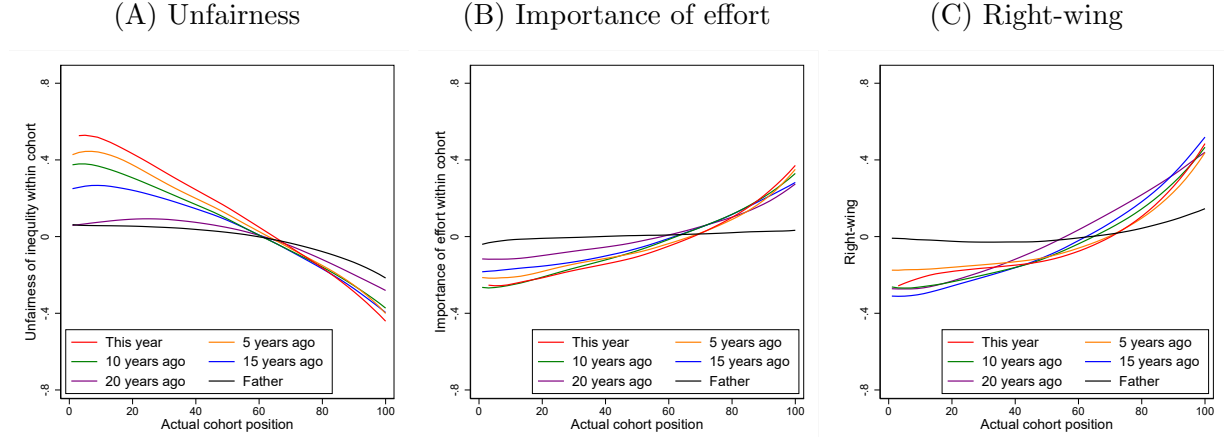
Table 2.4: Unfairness, importance of effort and political view by actual position and position misperception

	Unfairness of inequality					Importance	Right-
	Cohort	Gender	Mun.	Edu.	Sector	of effort	wing
<b>Panel A: No controls</b>							
Position	-1.18*** (0.07)	-1.13*** (0.07)	-1.27*** (0.07)	-1.36*** (0.07)	-1.42*** (0.07)	0.82*** (0.07)	0.90*** (0.07)
Misperception	-0.38*** (0.10)	-0.61*** (0.09)	-0.48*** (0.09)	-0.38*** (0.08)	-0.86*** (0.09)	0.44*** (0.10)	0.49*** (0.10)
<b>Panel B: With controls</b>							
Position	-1.09*** (0.09)	-1.01*** (0.08)	-1.09*** (0.08)	-1.03*** (0.08)	-1.19*** (0.08)	0.92*** (0.09)	0.99*** (0.08)
Misperception	-0.40*** (0.10)	-0.41*** (0.10)	-0.42*** (0.09)	-0.30*** (0.09)	-0.61*** (0.09)	0.47*** (0.10)	0.54*** (0.10)
<i>N</i>	4692	4692	4692	4692	4452	4692	4692
Outcome mean	2.01 (0.02)	2.16 (0.02)	2.09 (0.02)	2.54 (0.03)	2.53 (0.03)	4.81 (0.02)	7.10 (0.05)

*Notes:* The sample is restricted to control group respondents. All outcomes are as defined in the text, but standardized to z-scores. *Position* denotes the actual position within the reference group from percentile 1 to 100 divided by 100. A coefficient of 1 means that going from the bottom of the distribution to the top increases the outcome by one standard deviation. Similarly, *Misperception* is the difference between perceived and actual position within the reference group divided by 100. In the last two columns, the position and misperception are computed in the distribution of the cohort. *Controls* are cohort, municipality, education, gender, and sector fixed effects (including unemployed/not in workforce). *Outcome mean* is the mean of the non-standardized outcome variable. Standard errors on the estimates are reported in the parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

unfairness of inequality, a stronger belief in the role of effort over luck, and a higher likelihood of voting for a right-wing party. The precisely estimated coefficients on actual position show that, across reference groups, moving up by 10 positions in the income distribution is correlated with a 0.12-0.14 standard deviation increase in perceived unfairness. These are substantial effects, equal to around one third of the gap in perceived fairness between respondents who voted for left-wing parties (*Rød blok*) and those who voted for right-wing parties (in Danish *Blå blok*) in 2015. The effects of moving up in misperceived rank across the reference groups are smaller, i.e., around 30-60% of the effects of moving up in actual rank, but they are still significant and sizable.

Figure 2.11: History of past social positions, unfairness, importance of effort, and political views



*Notes:* Bandwidth for local linear polynomials is 20. For *Father*, the x-axis is the father's position among fathers when the respondent was 15 years old. The outcomes are standardized z-scores of the unfairness variable, perceived importance of effort relative to luck, and right-wing political views. Sample restricted to the control group.

## 2.4.2 Historical Variation in Social Positions

What is the relative importance of current social position and social position back in time in shaping views on fairness, the role of effort, and political affiliation? To address this question, we make use of the rich register data to reconstruct people's income and economic histories for the last twenty years and correlate them with their views today. In this section, we focus on their overall income path, while in the next section, we look at the effects of changes in social position due to specific negative and positive shocks.

Figure 2.11 plots respondents' views against their position in their cohort, measured at different points in time in five-year intervals, as well as against their father's position relative to other fathers in the cohort, measured when the individual was 15 years old. Panel A depicts the standardized (z-score) unfairness variable, i.e., the extent to which within-cohort income differences are considered unfair. In line with the aforementioned findings, there is a strong negative correlation between current social position and perceived unfairness. This association between current fairness views and social position becomes weaker when measuring social position at different points back in time. The father's social position is essentially uncorrelated with current fairness views. Recall from Section 2.3 that perceived and actual position are quite closely aligned for most respondents. Thus, although we do not

Table 2.5: History of past social positions, unfairness, importance of effort, and political view

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Position father	-0.059 (0.037)	-0.051 (0.037)	-0.068 (0.037)	-0.081* (0.036)	-0.075* (0.037)	0.025 (0.037)	0.147*** (0.035)
Position -20 yr.	-0.122** (0.042)	-0.097* (0.041)	-0.108** (0.041)	-0.113** (0.040)	-0.131** (0.041)	0.061 (0.042)	0.263*** (0.039)
Position -15 yr.	-0.203*** (0.048)	-0.169*** (0.048)	-0.148** (0.048)	-0.126** (0.047)	-0.139** (0.048)	0.131** (0.049)	0.268*** (0.046)
Position -10 yr.	-0.085 (0.058)	-0.100 (0.057)	-0.117* (0.057)	-0.186*** (0.056)	-0.162** (0.057)	0.206*** (0.058)	0.163** (0.055)
Position -5 yr.	-0.108 (0.070)	-0.096 (0.068)	-0.152* (0.069)	-0.108 (0.067)	-0.130 (0.069)	0.037 (0.070)	-0.015 (0.066)
Position this yr.	-0.655*** (0.074)	-0.618*** (0.073)	-0.647*** (0.073)	-0.678*** (0.071)	-0.761*** (0.074)	0.540*** (0.074)	0.479*** (0.070)
Observations	9046	9046	9046	9046	8575	9046	9046
Controls	✓	✓	✓	✓	✓	✓	✓

*Notes:* All outcomes are z-scores. *Position* denotes the cohort position from percentile 1 to 100 divided by 100. A coefficient of 1 means that going from the bottom of the distribution to the top increases the outcome by one standard deviation. *Importance of effort* is for income differences within cohort. *Position father* is the respondent's father's income rank when the respondent was 15 years old compared to other fathers of 15 year olds. *Controls* includes a treatment indicator, cohort, municipality, gender, education, and sector fixed effects (including unemployed/not in workforce). Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

have respondent's history of perceived positions (as opposed to actual positions), it is likely that these have co-moved to a significant extent over the life time.

Panels B and C show how the perceived importance of effort relative to luck and support for right-wing parties relate to social positions over time. These links appear more stable regardless of the time at which social position is considered, suggesting that these are stickier outcomes.<sup>13</sup>

Putting this all together and to understand whether one's history of social positions over time is correlated with fairness view conditional on current position, each column in Table 2.5 shows the regression coefficients of the variables in the columns on positions 20, 15, 10 and 5

<sup>13</sup>These results need to be interpreted in light of the degree of income mobility over different time spans. Appendix Tables A-2.6 and A-2.7 show that, naturally, the correlation between the current social position and past positions decreases as we move back in time. If we think of position as being composed of a permanent income component and a transitory component then the results are consistent with political views being mostly related to the permanent component, while fairness views are mainly related to the current position, i.e., the sum of the permanent component and the current transitory component.

years ago, as well as current position, and controlling for cohort, gender, education, sector, and treatment fixed effects. The table shows that the coefficients on current social position are large on views of unfairness, in particular when compared to political views. Some of the historical positions are also significantly, but less related to fairness views. Political views are related to current and past positions; and even the position of the father, conditional on the individual’s own historical positions. Views on the importance of effort are correlated most strongly with current position, and more weakly so with past positions. Appendix Tables A-2.8 and A-2.9 show that these findings are unchanged if we omit the controls, or use average positions over five-year intervals instead of yearly position 20, 15, 10 and 5 years ago.

### 2.4.3 Variation in Social Positions due to Major Life Events

We now consider four major life events – negative and positive – that could have affected people over the past ten years: having experienced an unemployment spell, having become disabled, having had a health condition that required hospitalization, and having received a promotion.

For each of these shocks, we perform the analysis on the subsample of individuals who did not experience this shock in any of the years 2008-11 (the “pre-shock period”) and assign a treatment indicator to individuals who experienced the shock sometime during the years 2012-2017 (“shock period”). We regress our fairness outcomes, perceptions of the role of effort, and being right-wing on the treatment indicator, detailed individual level controls including fixed effects for cohort, gender, municipality, education, sector and percentile cohort position prior to the shock.<sup>14</sup> Thus, the question we are asking is: conditional on starting

---

<sup>14</sup>The unemployment shock is defined as more than three months of unemployment in at least one year in the shock period. We estimate this regression on respondents who were not unemployed according to this definition at any time in the pre-shock period and who were in the workforce for the entire ten-year period. A disability shock is identified by a respondent receiving disability insurance benefits in one of the years in the shock period (according to the official Integrated Database for Labour Market Research, IDA, from Statistics Denmark). We only use respondents who were not on disability insurance in the pre-shock period. Hospitalization denotes at least one emergency room visit or hospital visit by referral from a general practitioner, but excluding visits due to congenital diseases, pregnancy, or routine checks, which do not reflect unexpected health shocks. In the regression we only use respondents who were not hospitalized according to this definition in the pre-shock period. Finally, promotion in the table indicates if a respondent switched from a job position as regular employee in the pre-shock period to a management position in the following period. Note that we do not include the vote in the last general election in the benchmark table, since it is from 2015 and therefore not measured in the pre-shock period. Since it can nevertheless serve as a proxy for past political views, we show the results controlling for party fixed effects in Table A-2.12.

from the same position ten years ago, and conditional on an array of personal characteristics, do those that experienced one of these four shocks that shifted their social position hold different views today than those who did not?

Table 2.6 shows the results. Column 1 shows the effect of the shocks on the current social position; columns 2 through 8 show the effects on our usual outcomes; column 9 reports the size of the subsamples of analysis for each type of shock; and column 10 indicates the share of respondents affected by the shock during the shock period. The unemployment and promotion shocks affect 5-7 percent of the sample and are related to large changes in social position (in opposite directions). Disability shocks hit less than 1 percent of the sample, but are related to a drop of around 22 percentiles in social position. By contrast, half of the sample visits the hospital during the shock period and this is associated with a smaller drop in social position of 2 percentiles.

Respondents who have experienced any of the negative shocks (unemployment, disability, and hospitalization) are significantly more likely to consider inequality within the different reference groups unfair. To the contrary, those who have been promoted are less likely to consider inequality unfair, especially within sector, which is the domain most closely related to work promotions. Consistent with our previous results, the effects of the shocks are weaker and less significant on the perceived importance of effort relative to luck and on right-wing views.

Of course, these results do not necessarily identify the causal effects of these shocks, as they are not random and may be correlated with other unobservable characteristics of the respondents that also affect their views. Still, the detailed controls, fixed effects, and accounting for the starting position ten years ago give confidence in the estimation. We actually obtain similar effects if we omit individual level controls except for starting position (see Appendix Table A-2.13), suggesting that there is not a highly systematic correlation between these individual characteristics and the likelihood of these shocks occurring. It is also informative that we are able to study four different types of shocks, with some that could be perceived as less influenced by individual choice, conditional on a detailed set of controls (e.g., hospitalization or disability), than others (e.g., promotion).

Table 2.6: Historic shocks, unfairness, importance of effort and political views

	Current position (1)	Unfairness of inequality					Importance of effort (7)	Right-wing (8)	N	Affected % (10)
		Cohort (2)	Gender (3)	Mun. (4)	Edu. (5)	Sector (6)				
Unemployment	-12.2*** (0.78)	0.20*** (0.051)	0.23*** (0.050)	0.21*** (0.050)	0.18*** (0.049)	0.20*** (0.052)	-0.11* (0.052)	-0.061 (0.049)	7537	5.03
Disability	-21.9*** (2.31)	0.30* (0.13)	0.54*** (0.13)	0.42** (0.13)	0.27* (0.13)		-0.31* (0.13)	-0.25* (0.13)	9246	0.61
Hospitalization	-1.83*** (0.49)	0.093** (0.029)	0.079** (0.028)	0.089** (0.028)	0.060* (0.028)	0.039 (0.028)	-0.010 (0.029)	-0.018 (0.028)	4749	55.5
Promotion	8.51*** (0.74)	-0.12** (0.045)	-0.11** (0.044)	-0.12** (0.045)	-0.17*** (0.044)	-0.21*** (0.044)	0.14** (0.045)	0.19*** (0.043)	7970	6.66
Pre-shock position	✓	✓	✓	✓	✓	✓	✓	✓		
Controls	✓	✓	✓	✓	✓	✓	✓	✓		

*Notes:* All outcomes z-scores. Each cell in the table is a separate regression of the column outcome on the row regressor and the controls indicated in the bottom part of the table. The explaining variables are all indicators that equal 1 if the respondent experienced the shock between 2012 and 2017. In each row, we exclude respondents who already experienced this type of shock in the pre-period (2008-2011). For *Unemployment*, we only use respondents who were in the workforce in the entire period. For *Disability*, we do not estimate the effect on fairness within sector, because very few disabled people work. *Controls* included in all regressions are a treatment indicator, cohort, gender, municipality, education and sector fixed effects (incl. unemployed/not in workforce), all measured in 2008. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Note also that the shocks that move social position the most are also those that have the largest effects on fairness views. As we have four different shocks, we can tentatively exploit the fact that they move social positions to different extents and estimate the implied impact of social position on fairness views under the assumption that the only channel through which the shocks acted on views was through social position. Put differently, we can do a suggestive IV-type analysis, in which we instrument for current position using the occurrence of these shocks. Appendix Table A-2.14 shows that the pooled IV results that use all four shocks at once are close in magnitude to the baseline correlations from Table 2.4. Using the shocks as instruments one by one yields broadly consistent effects as well, except for hospitalization, which is both very frequent and shifts income position only by a little (the “first stage” is weak). The magnitude of the effects of social position on fairness views from this suggestive IV is thus very consistent with the simple OLS correlations. Of course, the exclusion restriction is not necessarily satisfied here. Indeed, while these shocks do shift social position, there could be other channels through which they affect views too, over and above social position.

#### **2.4.4 Survey Experimental Variation in Social Position: Correction of Misperceptions**

In this section, we analyze the causal impact of changes in perceived position on fairness views. As described in Section 2.2.2, we informed a random half of the sample (the treatment group) about their true social positions in the five large reference groups before asking about their views on fairness, effort versus luck, and political attitudes.<sup>15</sup>

Table 2.7 shows the main experimental results. The outcome variables are the same as in Tables 2.4-2.6. The rows here show the coefficients on an indicator variable for having a positive misperception of one’s position within the given reference group (i.e., overestimating one’s position), and on the interactions of treatment status with having a positive or a negative misperception for the reference group under consideration. In line with the correlations in Table 2.4, respondents with positive misperceptions of their position in any given reference

---

<sup>15</sup>Due to the wish to treat all respondents fairly, the other half of the sample (the control group) was informed about their true positions only *after* these outcome questions, with no possibility to go back and change their answers. Hence, their answers to the outcome questions cannot be affected by this information.



Table 2.7: Survey information experiment and unfairness, importance of effort and political view

	Unfairness of inequality					Importance of effort	Right-wing
	Cohort	Gender	Mun.	Edu.	Sector		
Positive misperception	-0.132*** (0.034)	-0.131*** (0.033)	-0.133*** (0.033)	-0.108*** (0.031)	-0.267*** (0.034)	0.112** (0.034)	0.117*** (0.034)
$T \times \text{Positive}$	0.154*** (0.037)	0.082* (0.036)	0.099** (0.035)	0.060* (0.030)	0.089** (0.031)	-0.010 (0.038)	-0.013 (0.037)
$T \times \text{Negative}$	0.023 (0.024)	0.022 (0.024)	0.028 (0.025)	0.044 (0.027)	0.005 (0.028)	0.006 (0.024)	-0.027 (0.024)
$N$	9331	9331	9331	9331	8854	9331	9331
Position FE	✓	✓	✓	✓	✓	✓	✓

Notes: All outcomes are z-scores. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position within the reported reference group specified in each column. For the *Importance of effort* and *Right-wing* outcomes, we use cohort misperception.  $T \times \text{Positive}$  is an interaction of the treatment indicator and the *Positive misperception* indicator.  $T \times \text{Negative}$  is an interaction of the treatment indicator and an indicator for having a misperception  $\leq 0$ . In the regressions we also include a constant term. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

group in general believe inequality to be less unfair, conditional on position fixed effects. When they are informed about their misperception, their views on inequality revert back to being more aligned with those of respondents at the same income position who did not initially overestimate their position (i.e., the coefficients on treatment status interacted with positive misperceptions tend to cancel out those on having a positive misperception).

Since the treatment is showing the positions within all reference groups, the effect is a composite of learning about all these positions. Therefore, Appendix Table A-2.15 shows that, if we restrict the analysis to respondents who had either only positive misperceptions (i.e., overestimated their position) or only negative misperceptions for all reference groups, the coefficients are even larger and more significant. This makes sense given that the treatment corrected the position misperceptions for these respondents consistently in the same direction for all reference groups.

Respondents who start with a negative misperception for a given reference group do not change their fairness views, even after they are informed that they are, in fact, ranked higher than they thought. Appendix Table A-2.15 confirms that this remains the case even if we restrict the sample to respondents who had either only negative or only positive misperceptions across all reference groups. The effect of correcting misperceptions is thus asymmetric, with “bad news” weighing more heavily than “good news” on respondents.

The last two columns show that the treatment has no significant effects on the role of effort versus luck and on political preferences. Similarly, we do not detect any significant effects on other views and attitudes (see Appendix Table A-2.16). Overall, these results suggest that changes in perceived social position have stronger effects on fairness than on other normative views, and that these effects are asymmetric.

The Appendix reports several robustness checks: The results are unaffected if we include controls (see Appendix Table A-2.17). The treatment information is based on people's reported income, not on their actual income, so one may potentially worry that the treatment effects could be affected by errors in people's reported income. However, this is not the case. As Appendix Table A-2.18 shows, restricting the sample to respondents who reported their own income accurately gives similar results. We also check that the heterogeneous treatment effect is indeed driven by the misperception, rather than by income per se. Appendix Table A-2.19 shows that the results are unchanged if we also include treatment-income interactions in the regressions.

## 2.5 Conclusion

Our results provide new answers to some of the long-standing questions asked in the Introduction. First, we find that people are overall well informed about social positions, although they have some systematic misperceptions. They tend to believe others are closer to themselves than they really are and, in this sense, perceive inequalities to be smaller than they are. This finding could be explained, for instance, by anchoring to one's own income or Bayesian shrinkage. Nevertheless, the main conclusion is that, with some exceptions, misperceptions are not large. They may seem large when assessing people's own position, as normally done in the literature, but this is in many cases due to a mechanical center bias when asking about ranks.

Second, we provide a number of results suggesting that people's fairness views are strongly related to their social position and change when their positions change. Fairness views correlate more strongly with current social positions than with past positions, and movements up or down in social positions caused by real-life shocks are also related to corresponding

differences in people’s fairness views. When informing people about their true positions in their reference groups, we observe changes in fairness views across all the reference groups, but with asymmetric effects: Those that are told that they are lower-ranked than they thought perceive inequality as more unfair. By contrast, those who are informed they are ranked higher than they thought do not change their views on fairness. Thus, when it comes to correcting misperceptions, bad news weigh more heavily than good news in people’s minds.

Third, people consider inequalities conditional on the same level of education or work sector as most unfair. Exactly in these dimensions, where it matters most, people are least informed about inequality and lower-income people strongly overestimate their positions. It is an open question what drives this observed pattern, but it could have important implications for wage setting and career dynamics. One reason could be that people see education and type of work as choices of individuals that are key determinants of income. People from the same cohort, gender, or municipality can have very different education levels and types of work, and so it may be more expected that they earn different levels of income. To the contrary, inequality conditional on the key determinants of income (education and sector of work) may appear to be due to factors outside of people’s control and, therefore, perceived as more unfair. Another reason could be that people have different aspirations across reference groups and that, hence, admitting that they have a low position within education group or sector could lead to more resentment. In any case, these different views on the fairness of inequality in various reference groups highlights the importance of decomposing changes in inequality into those happening within sectors or firms and those happening across, as done by, e.g., Card et al. (2013) and Song et al. (2019).

Are our results portable to other countries? Clearly, Denmark is one of the most equal countries in the world (Atkinson and Sørensen, 2016; Boserup et al., 2016; Jakobsen et al., 2020) and attitudes vary across countries (Alesina et al., 2001, 2018; Almås et al., 2020). Yet, because we analyze rank positions, relative inequality perceptions (e.g., differences between perceived and actual P95 levels compared to corresponding P50 levels) and relative fairness views across reference groups, it is not *a priori* clear that our results should be different in one direction or the other.

Key to our analysis and findings is the linking of large-scale survey data on perceived social positions and fairness to administrative records on actual social positions across time, life events and reference groups. We see this combination of subjective and objective information as a promising avenue to learn more about the determinants of perceptions and attitudes.

## References

- Alesina, A. and G.-M. Angeletos (2005). Fairness and redistribution. *American Economic Review* 95(4), 960–980.
- Alesina, A., E. Glaeser, and B. Sacerdote (2001). Why Doesn't the United States Have a European-Style Welfare State? *Brookings Papers on Economic Activity* 2001(2), 187–277.
- Alesina, A., S. Stantcheva, and E. Teso (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review* 108(2), 521–54.
- Almås, I., A. W. Cappelen, K. G. Salvanes, E. Ø. Sørensen, and B. Tungodden (2017). Fairness and family background. *Politics, Philosophy & Economics* 16(2), 117–131.
- Almås, I., A. W. Cappelen, and B. Tungodden (2020). Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians? *Journal of Political Economy* 128(5), 1753–1788.
- Andersen, H. Y. and S. Leth-Petersen (2020). Housing Wealth or Collateral: How Home Value Shocks Drive Home Equity Extraction and Spending. *Journal of the European Economic Association* 0(0), 1–38.
- Atkinson, A. B. and J. E. Søgaaard (2016). The Long-Run History of Income Inequality in Denmark. *The Scandinavian Journal of Economics* 118(2), 264–291.
- Baker, M., Y. Halberstam, K. Kroft, A. Mas, and D. Messacar (2019). Pay transparency and the gender gap. NBER Working Paper 25834. *National Bureau of Economic Research*.
- Blanchflower, D. G. and A. J. Oswald (2004). Well-being Over Time in Britain and the USA. *Journal of Public Economics* 88(7-8), 1359–1386.
- Bénabou, R. and E. A. Ok (2001). Social Mobility and the Demand for Redistribution: The Poum Hypothesis. *The Quarterly Journal of Economics* 116(2), 447–487.
- Boserup, S. H., W. Kopczuk, and C. T. Kreiner (2016, May). The Role of Bequests in Shaping Wealth Inequality: Evidence from Danish Wealth Records. *American Economic Review* 106(5), 656–61.
- Boskin, M. J. and E. Sheshinski (1978). Optimal redistributive taxation when individual welfare depends upon relative income. *The Quarterly Journal of Economics* 92(4), 589–601.
- Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics* 128(3), 967–1015.
- Card, D., A. Mas, E. Moretti, and E. Saez (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review* 102(6), 2981–3003.
- Charité, J., R. Fisman, and I. Kuziemko (2015, March). Reference points and redistributive preferences: experimental evidence. NBER Working Paper 21009. *National Bureau of Economic Research*.
- Clark, A. E. and A. J. Oswald (1996). Satisfaction and comparison income. *Journal of Public Economics* 61(3), 359–381.
- Cruces, G., R. Perez-Truglia, and M. Tetaz (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics* 98, 100–112.
- Cullen, Z. and R. Perez-Truglia (2018a). How Much Does Your Boss Make? The Effects of Salary Comparisons. NBER Working Paper 24841. *National Bureau of Economic Research*.

- Cullen, Z. B. and R. Perez-Truglia (2018b). The salary taboo: Privacy norms and the diffusion of information. NBER Working Paper 25145. *National Bureau of Economic Research*.
- Duesenberry, J. (1949). *Income, saving, and the theory of consumer behavior*. Economic studies : No. 87. Harvard University Press.
- Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. A. David and M. W. Reder (Eds.), *Nations and households in economic growth*, pp. 89 – 125. Academic Press.
- Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization* 27(1), 35–47.
- Easterlin, R. A. (2001). Income and happiness: Towards a unified theory. *The Economic Journal* 111(473), 465–484.
- Easterlin, R. A., L. A. McVey, M. Switek, O. Sawangfa, and J. S. Zweig (2010). The happiness-income paradox revisited. *Proceedings of the National Academy of Sciences of the United States of America* 107(52), 22463–22468.
- Epper, T., E. Fehr, H. Fehr-Duda, C. T. Kreiner, D. D. Lassen, S. Leth-Petersen, and G. N. Rasmussen (2020). Time discounting and wealth inequality. *American Economic Review* 110(4), 1177–1205.
- Fehr, D., J. Mollerstrom, and R. Perez-Truglia (2019). Your place in the world: The demand for national and global redistribution. NBER Working Paper 26555. *National Bureau of Economic Research*.
- Ferrer-i Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect. *Journal of Public Economics* 89(5-6), 997–1019.
- Fisman, R., I. Kuziemko, and S. Vannutelli (2020). Distributional Preferences in Larger Groups: Keeping up with the Joneses and Keeping Track of the Tails. *Journal of the European Economic Association* 0(0), 1–32.
- Giuliano, P. and A. Spilimbergo (2014). Growing up in a recession. *Review of Economic Studies* 81(2), 787–817.
- Green-Pedersen, C. and K. Kosiara-Pedersen (2020). Det danske partisystem mellem kontinuitet og forandring. *Politica - Tidsskrift for politisk videnskab* 52(3), 215–233.
- Jakobsen, K., K. Jakobsen, H. Kleven, and G. Zucman (2020). Wealth taxation and wealth accumulation: Theory and evidence from Denmark. *The Quarterly Journal of Economics* 135(1), 329–388.
- Karadja, M., J. Mollerstrom, and D. Seim (2017). Richer (and holier) than thou? The effect of relative income improvements on demand for redistribution. *Review of Economics and Statistics* 99(2), 201–212.
- Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica* 79(3), 651–692.
- Kreiner, C. T., D. D. Lassen, and S. Leth-Petersen (2019). Liquidity constraint tightness and consumer responses to fiscal stimulus policy. *American Economic Journal: Economic Policy* 11(1), 351–379.
- Kuziemko, I., R. W. Buell, T. Reich, and M. I. Norton (2014). 'Last-place aversion': Evidence and redistributive implications. *The Quarterly Journal of Economics* 129(1), 105–149.

- Kuziemko, I., M. I. Norton, E. Saez, and S. Stantcheva (2015). How elastic are preferences for redistribution? Evidence from randomized survey experiments. *American Economic Review* 105(4), 1478–1508.
- Luttmer, E. F. P. (2005). Neighbors as Negatives: Relative Earnings and Well-Being. *The Quarterly Journal of Economics* 120(3), 963–1002.
- Malmendier, U. and S. Nagel (2011, February). Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *The Quarterly Journal of Economics* 126(1), 373–416.
- Meltzer, A. H. and S. F. Richard (1981). A rational theory of the size of government. *Journal of Political Economy* 89(5), 914–927.
- Perez-Truglia, R. (2020). The effects of income transparency on well-being: Evidence from a natural experiment. *American Economic Review* 110, 1019–1054.
- Roth, C. and J. Wohlfart (2018). Experienced inequality and preferences for redistribution. *Journal of Public Economics* 167, 251–262.
- Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019). Firming up inequality. *The Quarterly Journal of Economics* 134(1), 1–50.
- Weinzierl, M. (2014). The promise of positive optimal taxation: Normative diversity and a role for equal sacrifice. *Journal of Public Economics* 118, 128–142.

## 2.A Survey

### 2.A.1 Survey link and questions in English

**Link:** [https://cebi.eu.qualtrics.com/jfe/form/SV\\_6PcXP1t0Mw89iqp](https://cebi.eu.qualtrics.com/jfe/form/SV_6PcXP1t0Mw89iqp)

#### Background and Political views

1. What is your birth year?  
*Dropdown menu with years. Only 1969-1973 accepted.*
2. What is your gender?  
*Male; Female*
3. How many siblings do you have with the same biological mother and father as you?  
*0; 1; 2 or more*
4. Which municipality did you live in at the beginning of 2017? Note that in the following options, some of the municipalities are grouped together.  
*Dropdown menu with Danish municipalities*
5. Which of the following categories best describes your highest educational level?  
*Primary education; Upper secondary education; Vocational education and training; Short cycle higher education; Bachelor program or vocational bachelor education; Master program or PhD program*
6. What was your employment status at the beginning of 2017?  
*Full-time employment; Part-time employment; Self-employed; Unemployed; Not in the workforce*
7. Which sector did you work in at the beginning of 2017? Note that we mean the sector which your workplace belongs to. For example, if you work with PR in a bank you should choose the sector “Finance and insurance” and not the sector “Information and communication”.  
*Construction; Real estate activities; Business services; Finance and insurance; Trade and transport; Manufacturing, raw material extraction and utilities; Information and communication; Culture, leisure and other services; Agriculture, forestry and fishing; Public administration, education, health and social work activities*
8. Which party did you vote for in the last general election (in 2015)?  
*Socialdemokratiet; Venstre, Danmarks Liberale Parti; Radikale Venstre; Enhedslisten - De Rød-Grønne; Det Konservative Folkeparti; Alternativet; SF - Socialistisk Folkeparti; Liberal Alliance; Kristendemokraterne; Dansk Folkeparti; Other; Did not vote; Do not wish to answer*
9. How would you describe your attitude on economic policy?  
*Very left-wing; Left-wing; Moderate; Right-wing; Very right-wing*

#### Income

1. We will now ask you about your total income BEFORE tax in 2017. You should NOT include contributions to employer-managed pension schemes or mandatory pension contributions. When we later will inform you about your own position, it is important



that you state your total income as precisely as possible. If you are in doubt about the amounts, you can view them on your annual statement for 2017 from SKAT under *Opgørelse af indkomst* below *Før AM-bidrag*. You can also see a description of the different categories below. Note: In the scheme below we ask you to please state the yearly amounts in entire thousand DKK. If you enter 1 this corresponds to 1,000 DKK. *Salary and fees; Net profit from self-employment; Unemployment benefits, social assistance, study grants and pension payments*

## Perceptions

1. Instruction video
2. We will now ask you a question to see if you have understood the video's explanation of the ladder's different positions. Think about a person with an income where 73 out of 100 people have an income that is the same as or lower than this person's income. 27 out of 100 people have an income which is higher than this person's income. Select this person's position on the income ladder using the slider below.
3. What do you think the income for P50 was in 2017 for individuals born in [PIPED BIRTH YEAR]? Remember that P50 is the income, where half have an income that is the same as or lower than this income, and half have an income that is higher than this income. Remember also that income is before tax for the whole of 2017 and consists of salary, net profit from self-employment, other business income, unemployment benefits, transfers and payments from private and public pensions. Note: Please state your answer in entire thousand DKK. If you enter 1 it corresponds to 1,000 DKK
4. We will now ask you what you think the before tax income for P50 was in 2017 for the groups below which you are a part of. The first slider shows your answer from the previous question. You can use the other sliders to select what you think the income was for P50 for the different groups of people who were born the same year as you. *One horizontal slider for each reference group. The slider for cohort is locked at the amount entered in the previous question.*
5. What do you think the income for P95 was in 2017 for individuals born in [PIPED BIRTH YEAR]? Remember that P95 is the income where 95 out of 100 have an income that is the same as or lower than this income, and 5 out of 100 have an income that is higher than this income. Please state your answer in entire thousand DKK. If you enter 1 it corresponds to 1,000 DKK
6. We will now ask you what you think the before tax income for P95 was in 2017 for the groups below which you are a part of. The first slider shows your answer from the previous question. You can use the other sliders to select what you think the income was for P95 for the different groups of people who were born the same year as you. *One horizontal slider for each reference group. The slider for cohort is locked at the amount entered in the previous question.*
7. Rank among all people born in [PIPED BIRTH YEAR]. You previously reported that you had a yearly income in 2017 of [PIPED INCOME] DKK before tax. We will now ask you to report where you think this income placed you on the income ladder in 2017 for people who were born in [PIPED BIRTH YEAR]. Use the slider to select your position. Later, we will inform you about your true position.

8. Rank among [PIPED GENDER]. Now, think about all [PIPED GENDER] born in [PIPED BIRTH YEAR]. Use the slider to select where you think you were placed on the income ladder in 2017 for this group of people. Later, we will inform you about your true position.
9. Rank within [PIPED MUNICIPALITY] municipality. Now, think about people who also lived in [PIPED MUNICIPALITY] municipality at the beginning of 2017 and were born in [PIPED BIRTH YEAR]. Use the slider to select where you think you were placed on the income ladder in 2017 for this group of people. Later, we will inform you about your true position.
10. Rank within the educational level [PIPED EDUCATION]. Now, think about people whose educational level also was [PIPED EDUCATION] at the beginning of 2017 and were born in [PIPED BIRTH YEAR]. Use the slider to select where you think you were placed on the income ladder in 2017 for this group of people. Later, we will inform you about your true position.
11. Rank within the sector [PIPED SECTOR]. Now, think about people who also worked in the sector [PIPED SECTOR] at the beginning of 2017 and were born in [PIPED BIRTH YEAR]. Use the slider to select where you think you were placed on the income ladder in 2017 for this group of people. Later, we will inform you about your true position.
12. Think about your [FOR WOMEN: mother's. FOR MEN: father's] total income in the year where you turned 15. Compared to [FOR WOMEN: mothers. FOR MEN: fathers] of children, who were also born in [PIPED BIRTH YEAR], where do you think your [FOR WOMEN: mother. FOR MEN: father] was placed on the income ladder in the year where you turned 15?
13. Is your income higher or lower than [FOR RESPONDENTS WITH ONE SIBLING: your brother's/sister's income? FOR RESPONDENTS WITH 2 OR MORE SIBLINGS: the average income of your siblings?]  
*Higher; The same; Lower*
14. Think about your co-workers at the beginning of 2017. By co-workers, we mean the people who had the same workplace as you at the beginning of 2017. A workplace usually has the same address so if you for instance worked in a chain store then your co-workers are those who worked in the same store as you and not all the people, who were employed in the same firm. How many people worked at your workplace at the beginning of 2017 incl. you? If you do not remember the exact number then report your best guess.
15. Imagine that we rank you and your colleagues by your income in 2017 such that the person with the lowest income is number 1 and the person with the highest income is number [PIPED # COWORKERS]. What do you think your position was in this rank in 2017?
16. Think about your neighbors at the beginning of 2017. By neighbors, we mean the people who lived on the same road as you, if you lived in a house, or the people living on the same stairwell as you, if you lived in an apartment. Think only about the people, who were between 25 and 65 years old. How many people lived on the same road or on the same stairwell as you, including your own household, at the beginning of 2017? If you do not remember the exact number then report your best guess.

17. Imagine that we rank you and your neighbors by your income in 2017 such that the person with the lowest income is number 1 and the person with the highest income is number [PIPED # NEIGHBORS]. What do you think your position was in this rank in 2017?
18. Think about your schoolmates when you were 15 years old. By schoolmates, we mean everybody at your school who was born in [PIPED BIRTH YEAR], and not just the people in your class. How many schoolmates were you including yourself? If you do not remember the exact number then report your best guess.
19. Imagine that we rank you and your schoolmates by your income in 2017, such that the person with the lowest income is number 1 and the person with the highest income is number [PIPED # SCHOOLMATES]. What do you think your position was in this rank in 2017?

## Treatment

For the treatment group this block appears here. For the control group it appears after the block "Outcomes".

For each reference group, cohort/gender/municipality/educational level/sector, we provide the following information on separate pages along with a visualization of the difference:

You GUESSED that you were on position PXX.

Based on the income you reported, your TRUE position is PXX.

You are actually X positions higher/lower on the ladder than you thought.

## Outcomes

1. On a scale from 1 to 7 where 1 is "Completely fair", 4 is "Neither fair nor unfair" and 7 is "Completely unfair", indicate to what extent you think that it is fair or unfair that there are differences in income among people born the same year as you WITHIN the following groups that you are yourself a part of?
  - (a) Differences in income among people born in [PIPED BIRTH YEAR]
  - (b) Differences in income among [PIPED GENDER] born in [PIPED BIRTH YEAR]
  - (c) Differences in income among people living in [PIPED MUNICIPALITY] municipality
  - (d) Differences in income among people with the educational level [PIPED EDUCATION]
  - (e) Differences in income among people working in the sector [PIPED SECTOR]
2. Now, think about people born the same year as you WITHIN these groups (indicated below). On a scale from 1 to 7 where 1 is "Only luck", 4 is "Equally important", and 7 is "Only effort", indicate to what extent you think that differences in income are caused by differences in peoples' efforts over their lifetime or rather by luck? By luck, we mean conditions, which you have no control over. By effort, we mean conditions, which you can control.
  - (a) Reason for different incomes among people born in [PIPED BIRTH YEAR]?

- (b) Reason for different incomes among [PIPED GENDER] born in [PIPED BIRTH YEAR]?
  - (c) Reason for different incomes among people living in [PIPED MUNICIPALITY] municipality?
  - (d) Reason for different incomes among people with the educational level [PIPED EDUCATION]?
  - (e) Reason for different incomes among people working in the sector [PIPED SECTOR]?
3. Which party would you vote for if there was a general election today?  
*Socialdemokratiet; Venstre, Danmarks Liberale Parti; Radikale Venstre; Enhedslisten - De Rød-Grønne; Det Konservative Folkeparti; Alternativet; SF - Socialistisk Folkeparti; Liberal Alliance; Kristendemokraterne; Dansk Folkeparti; Nye Borgerlige; Other; Do not wish to answer*
4. Below, you see six statements which you can agree or disagree with. On a scale from 1 to 7 where 1 is "Completely agree", 4 is "Neither agree nor disagree", and 7 is "Completely disagree", indicate to what extent you agree or disagree with each statement.
- (a) Income inequality is a problem in Denmark
  - (b) The government should increase redistribution of income by increasing taxes and transfers to reduce inequality
  - (c) I am generally satisfied with my life
  - (d) My work has generally paid off
  - (e) People with high incomes have worked hard for their income and deserve it
  - (f) If a person is poor this is mainly due to lack of effort from his or her side

## Outro

1. It is important for our study that we only use responses from people, who have given the survey their full attention. You will automatically participate in the lottery no matter what you answer, but we would like to know how much attention you have given the survey.  
*1 I barely gave the survey any attention; ... ; 7 I gave the survey my full attention*
2. Do you think that the survey was biased?  
*Yes, it was right-winged; Yes, it was left-winged; No, it was neutral*
3. If you have any comments about the survey, then you are welcome to write them here:

## 2.A.2 Instruction video link and script

**Link:** <https://www.dropbox.com/s/ya1z0nlmii5tkpo/Instruktionsvideo.m4v?dl=0>

We will now ask you some questions regarding the distribution of income between Danes born the same year as you.

It may be difficult to answer, but we ask you to try your best.

There are differences between peoples' incomes. Some people have a high income, others have a low income.

The ladder to the left illustrates **how the incomes are distributed between Danes** born the same year as you.

This is also called the income distribution.

Think of 100 people born the same year as you.

They are ranked according to their income such that the person with the lowest income is at the bottom of the ladder and the person with the highest income is at the top of the ladder.

Look at the person next to the first rung of the ladder.

5 out of 100 people (i.e. 5 %) have an income which is the same as or **lower** than the income of this person.

We call this P5, because the person has position 5 on the income ladder.

The person on the middle rung has position 50.

**Exactly half** of all people (i.e. 50 %) born the same year as you have an income which is the same as or **lower** than the income of this person and **exactly half** have an income which is **higher** than the income of this person.

We call the position in the middle for P50.

Remember that P50 is the position in the middle since we will use this several times in the following questions.

The person next to the top rung has position 95.

95 out of 100 (i.e. 95 %) have an income which is the same as or **lower** than the income of this person and only 5 out of 100 people born the same year as you (i.e. 5 %) have an income which is higher than the income of this person.

Remember what P95 indicates since we are going to use this several times.

Shortly, we will now ask you what you think the income is for P50 and P95, respectively, for Danes born the same year as you

Next, we will ask you what you think **your position** is on the ladder.

You are welcome to watch the video again if you are not sure of the meaning of the different positions.

## 2.A.3 Survey screenshots

Figure A-2.1: Income question

We will now ask you about your **total income BEFORE tax in 2017**. You should **NOT** include contributions to employer-managed pension schemes or mandatory pension contributions. When we later will inform you about your own position, it is important that you state your total income as precisely as possible. If you are in doubt about the amounts, you can view them on your annual statement for 2017 from SKAT under *Opgørelse af indkomst* below **Før AM-bidrag**. You can also see a description of the different categories below.

**Note:** In the scheme below we ask you to please state the **yearly** amounts in **entire thousand DKK**. If you enter 1 this corresponds to 1,000 DKK.

Salary and fees	<input type="text"/>	thousand DKK
Net profit from self-employment	<input type="text"/>	thousand DKK
Unemployment benefits, social assistance, study grants and pension payments	<input type="text"/>	thousand DKK
Total	<input type="text" value="0"/>	thousand DKK

### Examples

**Salary and fees:** Taxable wage income before tax and before labor market contribution and fees. You should include:

- Value of fringe benefits
- Taxable foreign wage
- Wage during sickness and maternity/paternity leave
- Fees from board duties, consultancy work, talks etc.
- Value of stock options, severance pay and anniversary bonus

On your tax statement this corresponds to box 11 + 12 + 14.

**Net profit from self-employment:** Net profit from self-employment after capital income and expenses. On your tax statement this corresponds to box 111 minus box 112.

**Unemployment benefits, social assistance, study grants and pension payments:** Unemployment benefits, cash benefits, sickness benefits, maternity/paternity benefits, study grants, payments from private pensions, public pensions and disability pensions. On your tax statement this corresponds to box 16.

Figure A-2.2: Elicitation of cohort P50 perception

What do you think the income for **P50** was in 2017 for individuals born in 1970?

Remember that P50 is the income, where half have an income that is the same as or lower than this income, and half have an income that is higher than this income.

Remember also that income is before tax for the whole of 2017 and consists of salary, net profit from self-employment, other business income, unemployment benefits, transfers and payments from private and public pensions.

**Note:** Please state your answer in **entire thousand DKKs**. If you enter 1 it corresponds to 1,000 DKK.



*Notes:* The figure shows a screenshot from the survey for a person who reported being born in 1971.

Figure A-2.3: Elicitation of large reference group P50 perceptions

We will now ask you what you think the before tax income for **P50** was in 2017 for the groups below, which you are a part of. The first slider shows your answer from the previous question. You can use the other sliders to select, what you think the income was for P50 for the different groups of **people who were born the same year as you**.

P50 for people **born in 1970**

400.000

P50 for **men** born in 1970

20.000

P50 for people who also lived in **Københavns municipality**

20.000

P50 for people who also had the educational level **Master or PhD program**

20.000

P50 for people who also worked in the sector **Finance and insurance**

20.000

*Notes:* The top slider shows the piped answer to the question in figure A-2.2 and cannot be moved. The sliders go from 20,000 to 8,069,000 in 200 steps according to  $Y = 20000 * EXP(0.03 * Step)$ . In the middle position the slider has the value 402,000.

Figure A-2.4: Elicitation of number of co-workers and position among co-workers

(A) Number

Think about your co-workers in the beginning of 2017. By co-workers we mean the people who had the same workplace as you in the beginning of 2017. A workplace usually has the same address so if you for instance worked in a chain store then your co-workers are those who worked in the same store as you and **not** all the people who were employed in the same firm.

How many people worked in your workplace at the beginning of 2017 incl. you? If you do not remember the exact number then report your best guess.

50

(B) Position

Imagine that we rank you and your co-workers by your incomes in 2017 such that the person with the lowest income is number 1 and the person with the highest income is number 50. What do you think your position was in this rank in 2017?

50

1

Place yourself:

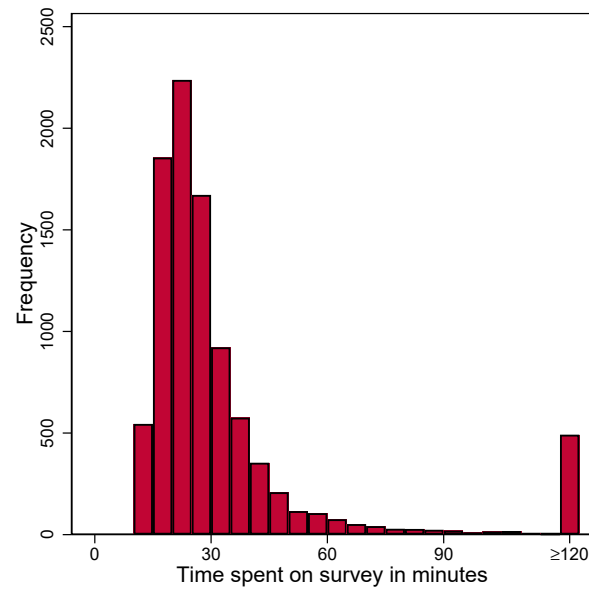
**Number 1 out of 50 in my workplace.**

*Notes:* The panels show screenshots from two pages in the survey. On the first page in this example, the respondent reports having 50 co-workers (the box is empty as default). On the second page, this number is piped as the max of the slider, and when the respondent moves the slider with the cursor the red position number changes accordingly.



## 2.B Data Collection, Survey Design and Administrative Data

Figure A-2.5: Full distribution of time spent on answering the survey



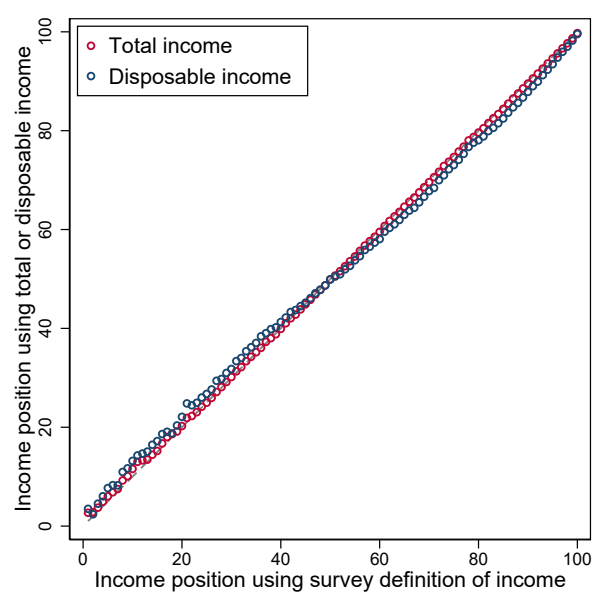
*Notes:* The figure shows the distribution of time spent on the survey for the analysis sample. The bin width is 5 minutes and the distribution is censored above 120 minutes.

Table A-2.1: Attrition analysis

	Not in sample	
<b>Panel A</b>		
Treatment	0.011	(0.008)
Male	-0.083***	(0.008)
Age	0.001	(0.003)
Married	-0.021**	(0.008)
Ref.: Middle 40%		
Bottom 50 %	0.149***	(0.009)
Top 10 %	-0.060***	(0.012)
Ref.: Master programs		
Primary education	0.157***	(0.017)
Upper secondary edu.	0.017	(0.019)
Vocational education	0.086***	(0.012)
Short cycle higher edu.	0.014	(0.017)
Bachelor programs	0.026*	(0.012)
Ref.: Nothern Jutland		
Copenhagen	0.016	(0.015)
Sealand	-0.000	(0.016)
Southern Denmark	0.007	(0.015)
Middle Jutland	-0.014	(0.015)
Observations	13667	
<b>Panel B</b>		Share
Not in the final sample		0.312
Drop out at consent question		0.010
Drop out at income question		0.102
Drop out before treatment		0.242
Drop out after treatment		0.012
Screened out		0.049

*Notes:* Respondents who dropped out before the treatment, were not assigned to either the treatment or control group. We randomly assign these individuals to one of the groups. The number of observations in the regression in Panel A is 19 lower than total number of people who started the survey. This is because we miss educational information for these individuals. The sum of *Drop out before treatment*, *Drop out after treatment* and *Screened out* sum to 30.3%. The last 0.9% are people who are assigned to the control but do not complete the survey. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Figure A-2.6: Comparison of position using different income definitions



*Notes:* The figure uses all individuals born from 1969 to 1973 observed in the income register data. N=389,759. We use total income and disposable income as defined by Statistics Denmark.

Table A-2.2: Treatment balancing

	Control	Treatment	Difference	
Actual cohort position	64.003	64.370	-0.367	(0.513)
Treatment information	-5.812	-6.048	0.237	(0.335)
Cohort misperception	-5.767	-6.064	0.297	(0.353)
Left-wing	0.219	0.222	-0.003	(0.009)
Right-wing	0.236	0.241	-0.004	(0.009)
Male	0.511	0.518	-0.007	(0.010)
Age	47.058	46.998	0.060*	(0.029)
Primary education	0.077	0.075	0.001	(0.005)
Upper secondary education	0.061	0.054	0.007	(0.005)
Vocational education	0.317	0.312	0.005	(0.010)
Short cycle higher education	0.090	0.091	-0.001	(0.006)
Bachelor programs	0.264	0.274	-0.010	(0.009)
Master programs	0.190	0.193	-0.003	(0.008)
Self-employed	0.038	0.036	0.003	(0.004)
Employee	0.902	0.904	-0.001	(0.006)
Unemployed	0.011	0.013	-0.001	(0.002)
Private sector	0.660	0.657	0.003	(0.010)
Not in work force	0.048	0.048	-0.000	(0.004)
Copenhagen	0.087	0.086	0.001	(0.006)
Sealand	0.237	0.229	0.008	(0.009)
Southern Denmark	0.199	0.215	-0.016	(0.008)
Middle Jutland	0.312	0.308	0.004	(0.010)
Nothern Jutland	0.164	0.161	0.003	(0.008)
<i>N</i>			9415	

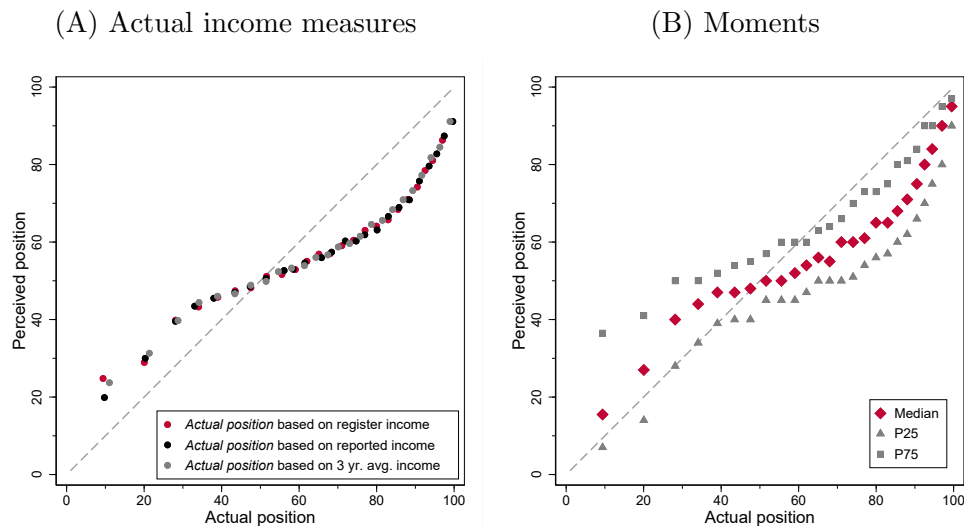
*Notes:* Column 1 and 2 show the group means of the variables. Column 3 shows the difference. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.3: Match between survey response and register data

	N	Share
Correct cohort	9,415	1.00
Correct gender	9,415	1.00
Correct municipality	9,239	0.98
Correct level of education	6,958	0.74
Correct sector	6,768	0.72
All correct	4,952	0.53

## 2.C Perceptions and Misperceptions about Social Positions

Figure A-2.7: Perceived Position Within Cohort



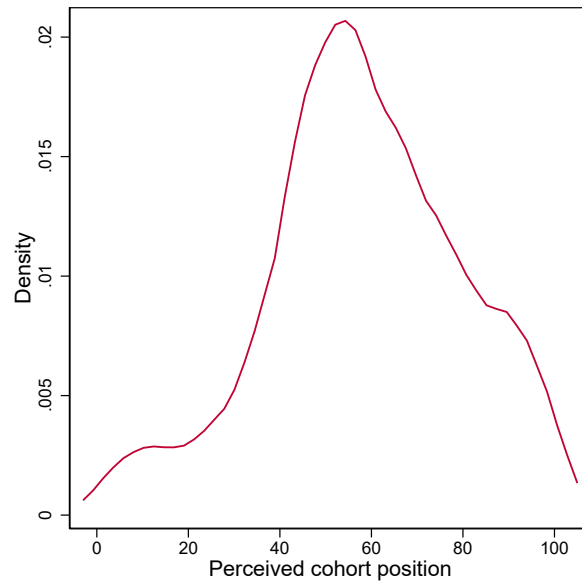
*Notes:* The left panel is a bin scatter of the average perceived position by actual position (in 25 equally-sized bins). Actual position is either based on the actual income observed on the tax return, the income reported in the survey, or a three-year average of actual income. The right panel shows the 25th, 50th and 75th percentile of perceived position by bins of actual income position.

Table A-2.4: Accuracy of perceptions regressed on individual characteristics

	A. Top 20% most inaccurate			B. Top 20% most accurate		
	Position	P50	P95	Position	P50	P95
Male	-0.04*** (0.01)	-0.06*** (0.01)	-0.04*** (0.01)	0.04*** (0.01)	0.02* (0.01)	0.05*** (0.01)
Left-wing	-0.02* (0.01)	0.00 (0.01)	0.03** (0.01)	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)
Right-wing	-0.03** (0.01)	-0.01 (0.01)	-0.00 (0.01)	0.03** (0.01)	0.01 (0.01)	0.00 (0.01)
Middle Jutland	-0.02 (0.02)	0.01 (0.02)	0.03 (0.02)	-0.01 (0.02)	0.00 (0.02)	0.00 (0.02)
Southern Denmark	-0.03 (0.02)	0.01 (0.02)	0.03 (0.02)	0.00 (0.02)	-0.01 (0.02)	-0.02 (0.02)
Sealand	-0.01 (0.02)	0.04** (0.02)	0.02 (0.02)	0.01 (0.02)	-0.00 (0.02)	0.01 (0.02)
Copenhagen Area	-0.02 (0.02)	0.02 (0.02)	0.01 (0.02)	-0.01 (0.02)	-0.01 (0.02)	0.01 (0.02)
Vocational education	-0.06** (0.02)	-0.04 (0.02)	-0.00 (0.02)	0.01 (0.02)	-0.00 (0.02)	0.04 (0.02)
Upper secondary edu.	-0.03* (0.02)	-0.03* (0.02)	-0.02 (0.02)	0.01 (0.02)	0.03 (0.02)	0.01 (0.02)
Short cycle higher edu.	-0.05** (0.02)	-0.07*** (0.02)	-0.04 (0.02)	0.02 (0.02)	0.02 (0.02)	0.06** (0.02)
Bachelor programs	-0.08*** (0.02)	-0.06** (0.02)	-0.02 (0.02)	0.04* (0.02)	0.02 (0.02)	0.04* (0.02)
Masters programs and PhD	-0.15*** (0.02)	-0.11*** (0.02)	-0.01 (0.02)	0.10*** (0.02)	0.05** (0.02)	0.06** (0.02)
Construction	0.03 (0.03)	-0.01 (0.03)	-0.01 (0.03)	0.00 (0.03)	-0.04 (0.03)	0.01 (0.03)
Real estate activities	0.04 (0.04)	-0.03 (0.04)	-0.07 (0.04)	-0.00 (0.04)	-0.06 (0.04)	0.02 (0.04)
Business service	0.06* (0.02)	0.01 (0.02)	-0.01 (0.02)	-0.04 (0.03)	-0.02 (0.03)	0.01 (0.03)
Finance and insurance	-0.01 (0.03)	-0.02 (0.03)	-0.06* (0.03)	0.05 (0.03)	-0.05 (0.03)	0.04 (0.03)
Trade and transport	0.06* (0.02)	0.01 (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.03 (0.02)	-0.01 (0.02)
Manufacturing	0.04 (0.02)	0.00 (0.02)	-0.03 (0.02)	-0.02 (0.02)	-0.03 (0.03)	0.00 (0.02)
Information and communication	-0.02 (0.03)	-0.02 (0.03)	-0.03 (0.03)	0.00 (0.03)	-0.02 (0.03)	0.00 (0.03)
Culture and leisure	0.01 (0.03)	-0.03 (0.03)	-0.04 (0.03)	0.07* (0.03)	0.02 (0.03)	0.02 (0.03)
Agriculture	0.08 (0.05)	-0.10* (0.05)	-0.06 (0.05)	-0.04 (0.05)	0.04 (0.05)	0.07 (0.05)
Public adm., edu. and health	0.03 (0.02)	-0.02 (0.02)	-0.03 (0.02)	-0.01 (0.02)	-0.01 (0.02)	0.02 (0.02)
<i>N</i>	9415	9415	9415	9415	9415	9415
<i>R</i> <sup>2</sup>	0.087	0.034	0.025	0.079	0.036	0.029
Cohort FE	✓	✓	✓	✓	✓	✓
Actual position FE	✓	✓	✓	✓	✓	✓

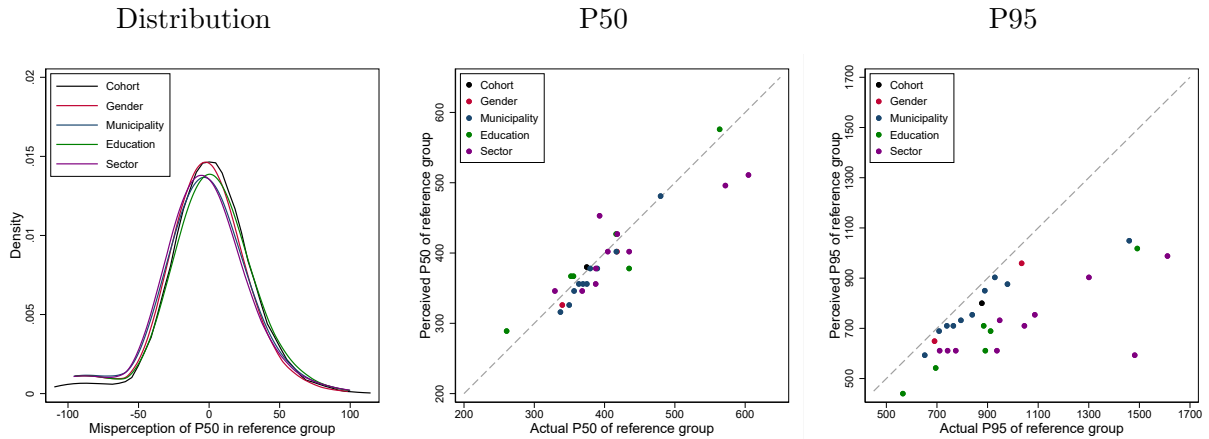
*Notes:* In Panel A (respectively, Panel B), the outcomes are indicator variables equal to 1 if the respondent is among the 20% of respondents with the largest (respectively, smallest) misperceptions for each variable. The *Actual position FE* is fixed effects for all 100 positions in the income distribution. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Figure A-2.8: Distribution of perceived cohort position



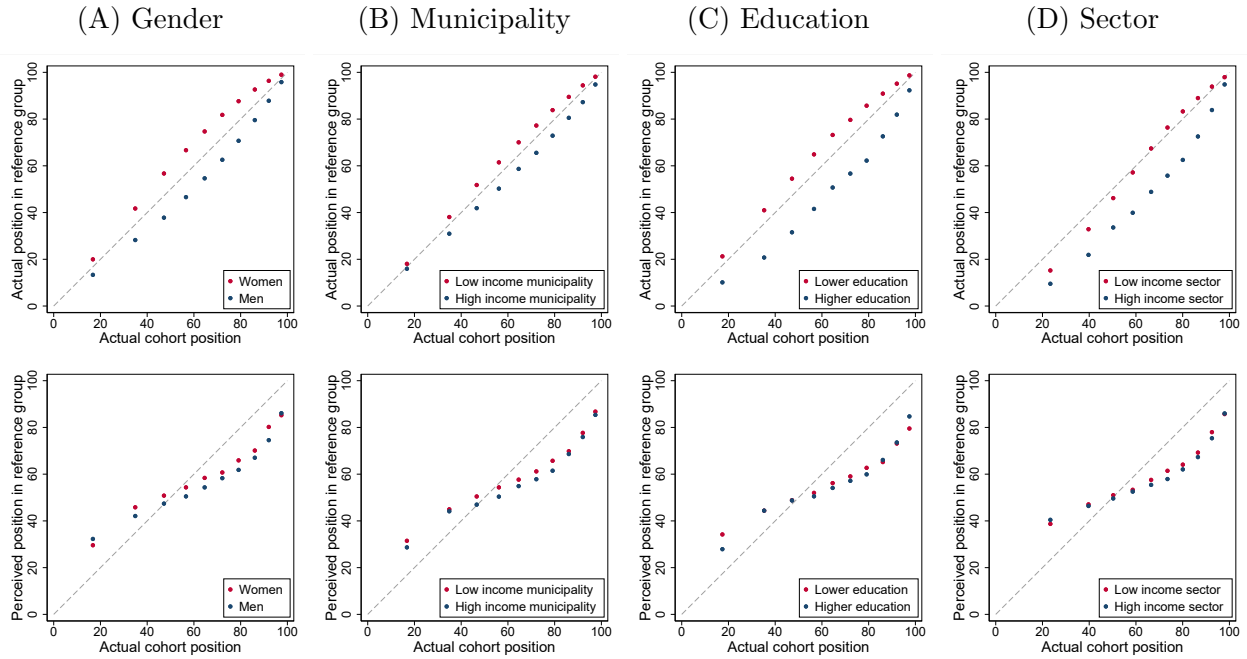
*Notes:* The panel shows a density plot and is constructed using Epanechnikov kernels with a bandwidth of 5.

Figure A-2.9: Distribution of P50 misperceptions and median perceived P50 and P95 incomes for large reference groups



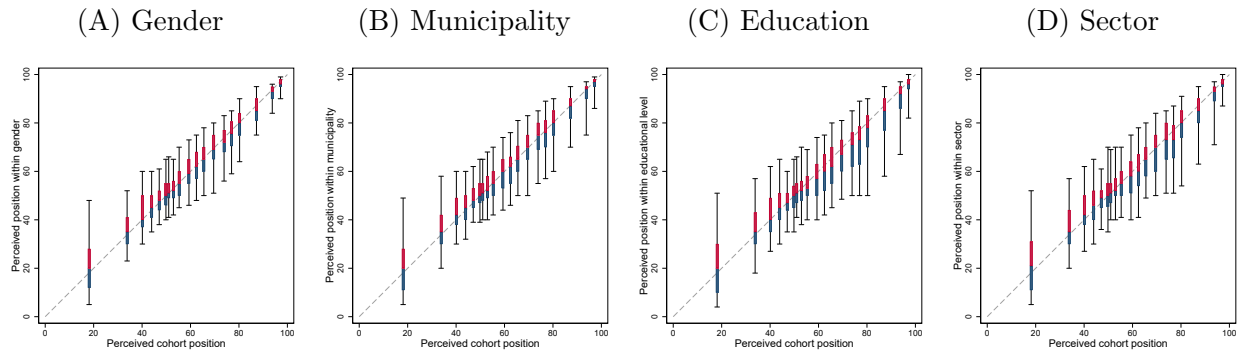
*Notes:* In the left panel, we show the distributions of P50 misperceptions in the large reference groups. The distributions are smoothed using epanechnikov kernels with a bandwidth of 15. In the middle and right panels, we show bin medians instead of bin means using the same sample as in figure 2.5. For gender, we show one scatter for men and one for women. For municipality we divide the respondents into 10 similar sized groups based on the actual municipality P50 and P95 income and plot one scatter for each group. For education and sector we show one scatter for each educational level or sector.

Figure A-2.11: Correlation between Actual Cohort Position and Actual versus Perceived Position within Reference Groups



*Notes:* The top panels plot *actual* reference group positions by actual cohort positions, while the bottom panels plot *perceived* reference group positions by actual cohort positions. The high/low income split for municipality is based on the median of within cohort in sample actual municipality P50 income. This is also the case for sector. For education, *Higher education* is short cycle higher education, bachelor programs and master programs.

Figure A-2.10: Variation in perceived position across large reference groups

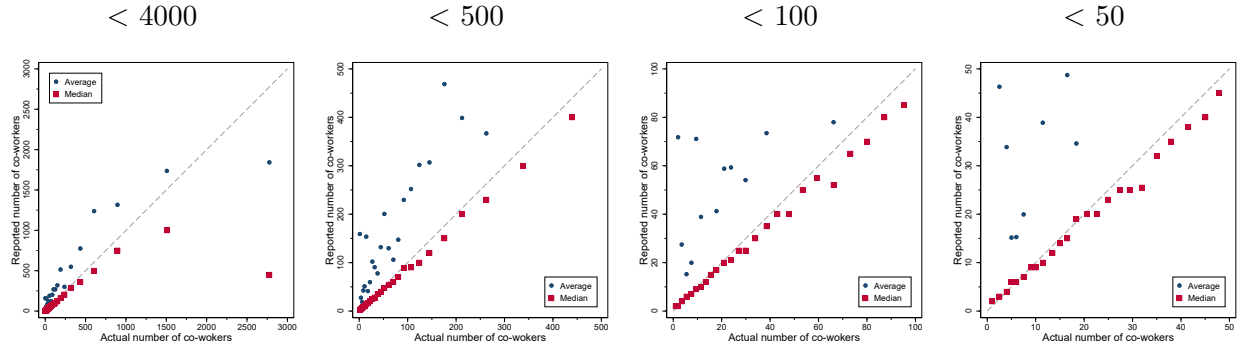


*Notes:* This figure shows 5th, 25th, 50th, 75th and 95th percentile of reported position within the large reference group by bins of perceived cohort position.

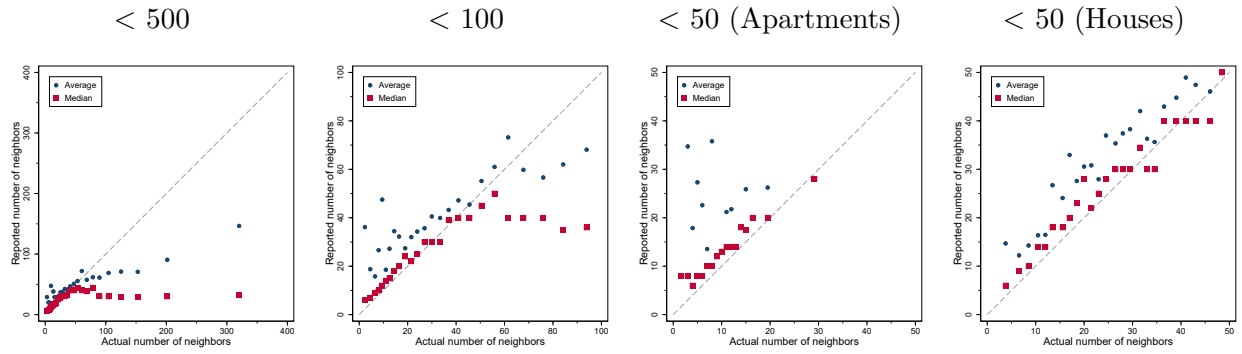


Figure A-2.12: Actual and reported number of people in small reference groups

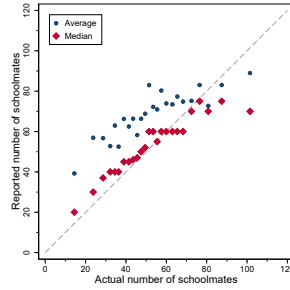
(A) Co-corkers



(B) Neighbors

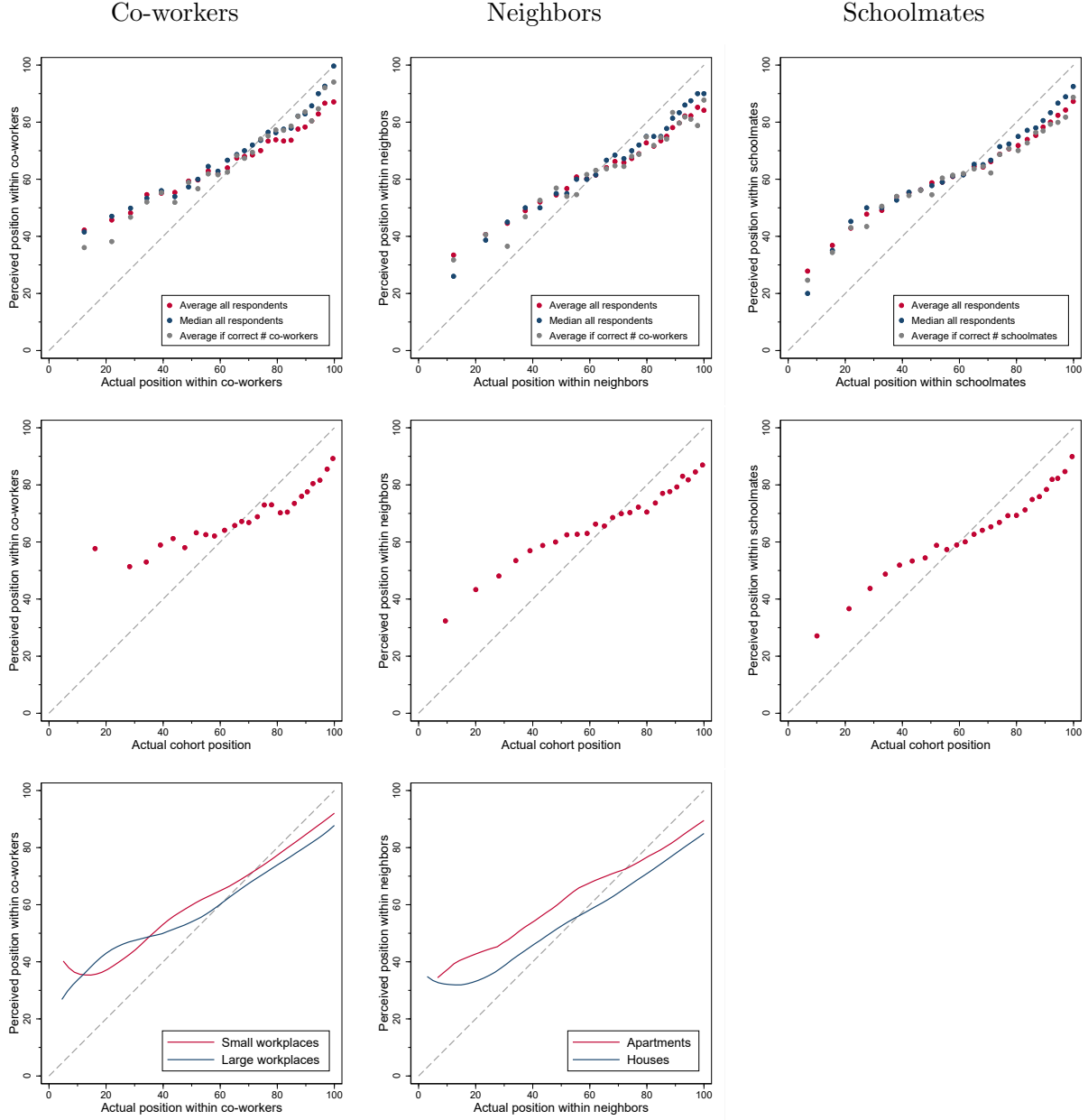


(C) Schoolmates



*Notes:* The figures show bin scatters of the reported number of co-workers by the actual number of co-workers. In each panel, the sample is restricted to observations where the *Actual number of co-workers* is below a certain threshold. All observations are used to calculate the bin averages but the panels only show the averages if they are smaller than the threshold. There are 25 bins in each panel and there is the same number of observations behind each bin. The bin averages are only plotted if they are lower than the maximum actual number. For *Schoolmates*, the figure is based on respondents enrolled in "Grundskole" (Basic School) at age 15. The figure excludes observations from one very large school.

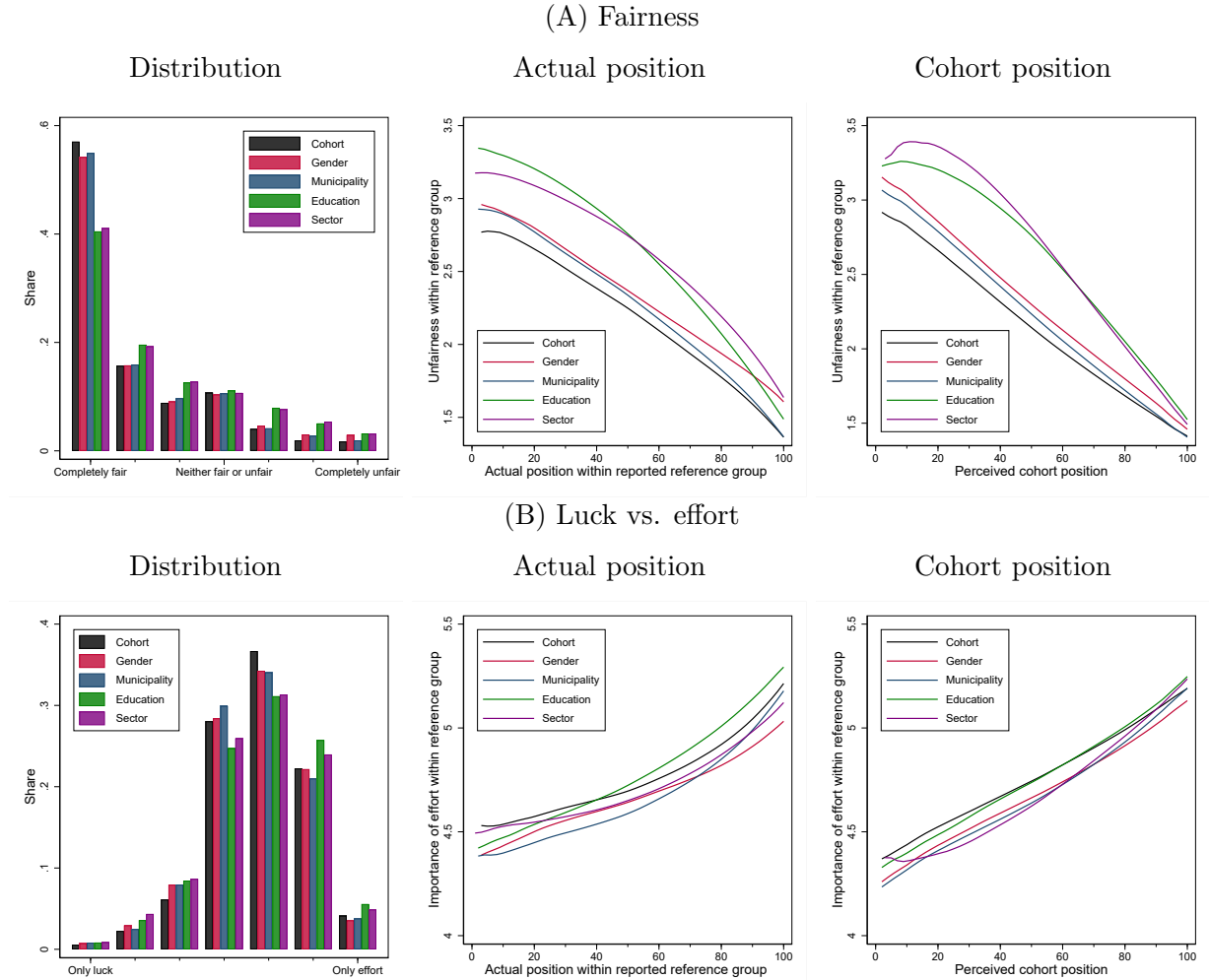
Figure A-2.13: Perceived position in small reference groups



*Notes:* There are 25 bins in each panel. They are of equal size, except the top bin for co-workers and neighbors in the top panels, which have more observations. The top panels show similar patterns as in figure 2.8 using medians instead of averages or restricting the sample to respondents who reports a number of people in the small reference group that matches the number observed in the register data  $\pm 10\%$ . In the middle panels, we use actual cohort position instead actual position within the small reference group. Again, we see a pattern similar to figure 2.8. In the bottom panels, the local linear polynomials have a bandwidth of 10 and are based on the respondents who report the correct number of people in the reference group  $\pm 10\%$ . Small workplaces have 10 to 100 employees. Large workplaces have more than 100 employees.

## 2.D Relationship Between Social Positions and Fairness Views

Figure A-2.14: Unfairness of inequality and importance of effort within large reference groups: Distribution and by position



*Notes:* The figure only uses responses from the control group. The local linear polynomials have a bandwidth of 20. The left column panels show the raw distribution of the answers on the 7 point scale. In the middle column panels, we use actual position within each reference group instead of perceived position as in figure 2.10. In the right column panels, we use perceived cohort position.

Table A-2.5: Inequality views by actual position and position misperception

	Ineq. not prob.	Less redist.	Less satis.	Work not paid off	Rich not deserve	Poor's own fault
<b>Panel A: No controls</b>						
Position	2.32*** (0.12)	2.07*** (0.13)	-1.22*** (0.11)	-2.20*** (0.11)	-1.45*** (0.11)	-1.08*** (0.11)
Misperception	0.95*** (0.18)	0.78*** (0.19)	-0.38* (0.16)	-1.08*** (0.16)	-0.83*** (0.16)	-0.69*** (0.16)
<b>Panel B: With controls</b>						
Position	1.17*** (0.08)	1.15*** (0.09)	-0.83*** (0.09)	-1.36*** (0.09)	-1.07*** (0.09)	-0.57*** (0.09)
Misperception	0.50*** (0.10)	0.46*** (0.10)	-0.32** (0.10)	-0.70*** (0.10)	-0.60*** (0.10)	-0.44*** (0.10)
<i>N</i>	4690	4690	4690	4690	4690	4690
Outcome mean	4.08 (0.03)	4.52 (0.03)	2.32 (0.02)	2.53 (0.02)	3.19 (0.02)	4.78 (0.02)

*Notes:* In the table we only use the control group respondents. All outcomes are z-scores. *Position* denotes the actual cohort position from percentile 1 to 100 divided by 100. A coefficient of 1 means that going from the bottom of the distribution to the top increases the outcome by one standard deviation. Similarly, *Misperception* is the difference between perceived and actual cohort position divided by 100. *Controls* includes cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects. *Outcome mean* is the mean of the non-standardized outcome variable. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.6: Pairwise correlations of historic income positions

	-20 yr.	-15 yr.	-10 yr.	-5 yr.	This yr.
-20 yr.	1.00	0.41	0.29	0.22	0.22
-15 yr.	0.41	1.00	0.62	0.54	0.51
-10 yr.	0.29	0.62	1.00	0.70	0.65
-5 yr.	0.22	0.54	0.70	1.00	0.78
This yr.	0.22	0.51	0.65	0.78	1.00

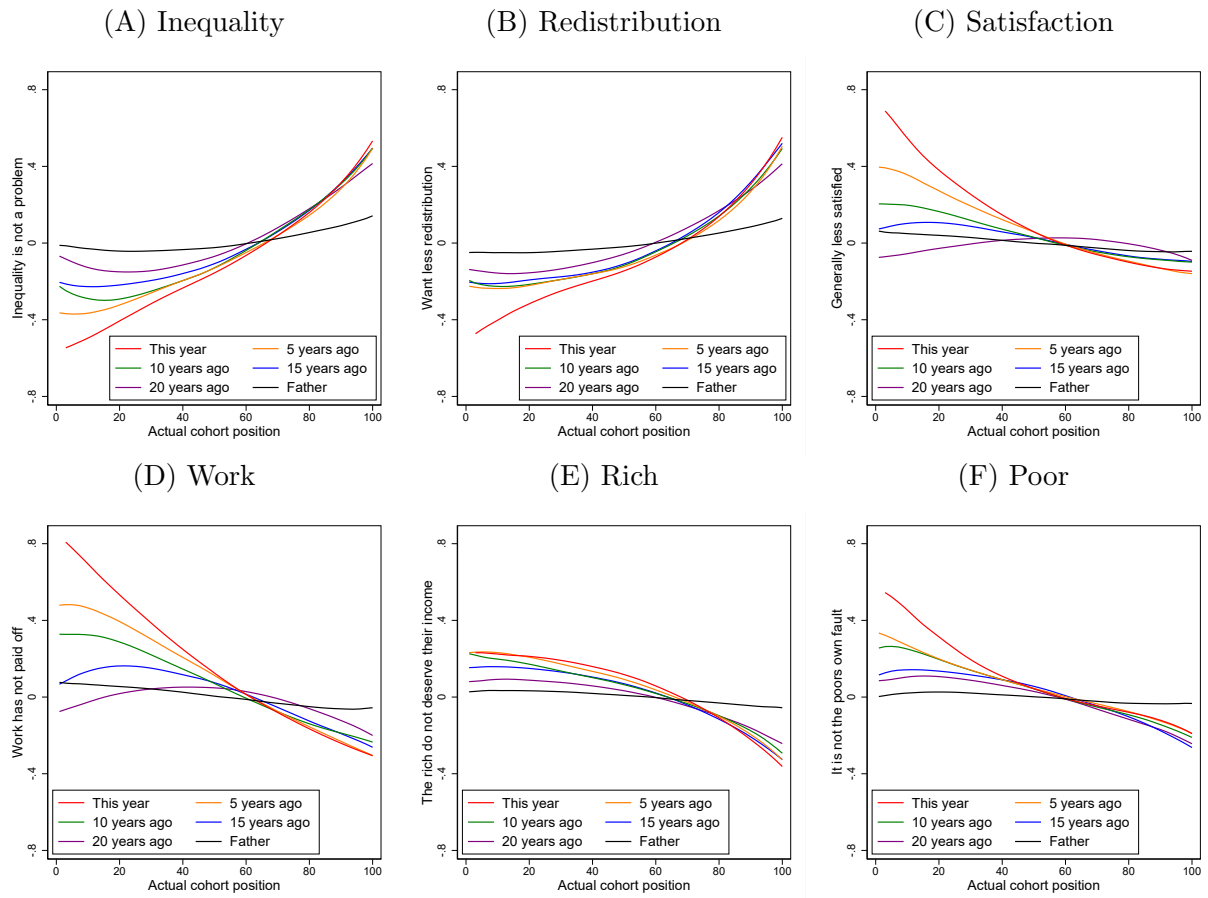
*Notes:* Based on the full cohorts born between 1969 and 1973. We only include individuals we observe in all years. N=356,556.

Table A-2.7: Transition matrix of income position 20 years ago and today

	Quintile today					Total
	1	2	3	4	5	
Quintile 20 years ago						
1	5.2	3.2	3.1	3.9	4.4	19.8
2	5.2	4.7	3.5	3.3	3.2	19.9
3	3.4	6.0	4.9	3.4	2.3	20.1
4	2.0	4.0	5.5	5.0	3.5	20.1
5	1.5	2.1	3.8	5.4	7.3	20.1
Quintile 10 years ago						
1	9.5	4.4	2.0	1.5	1.0	18.3
2	4.0	8.7	4.8	1.9	0.7	20.1
3	1.9	4.4	8.0	4.9	1.3	20.5
4	1.2	1.8	4.7	8.6	4.2	20.6
5	0.8	0.7	1.3	4.2	13.4	20.5
Total	17.4	20.1	20.8	21.1	20.7	100.0

*Notes:* Based on the full cohorts born between 1969 and 1973. We only include individuals we observe in all years. The columns do not sum to 20% each because immigrants are included when the income percentiles are generated but not in the table, since we do not observe them historically. N=356,556.

Figure A-2.15: History of past social positions and inequality views



*Notes:* Bandwidth for local linear polynomials is 20. For *Father*, the x-axis is the father's position among fathers when the respondent was 15 years old. In all panels, the y-axis is the z-score for the survey answers. Sample restricted to the control group.

Table A-2.8: History of past social positions, unfairness, importance of effort and political view without controls

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Position father	-0.050 (0.036)	-0.027 (0.036)	-0.063 (0.036)	-0.040 (0.035)	-0.047 (0.036)	-0.023 (0.036)	0.052 (0.035)
Position -20 yr.	-0.153*** (0.037)	-0.183*** (0.037)	-0.150*** (0.037)	-0.308*** (0.036)	-0.265*** (0.037)	0.169*** (0.037)	0.557*** (0.036)
Position -15 yr.	-0.228*** (0.048)	-0.217*** (0.047)	-0.182*** (0.047)	-0.228*** (0.047)	-0.236*** (0.048)	0.171*** (0.048)	0.376*** (0.047)
Position -10 yr.	-0.099 (0.057)	-0.155** (0.057)	-0.145* (0.057)	-0.290*** (0.056)	-0.269*** (0.057)	0.205*** (0.058)	0.197*** (0.056)
Position -5 yr.	-0.110 (0.069)	-0.093 (0.068)	-0.159* (0.068)	-0.077 (0.068)	-0.113 (0.070)	-0.009 (0.069)	-0.094 (0.068)
Position this yr.	-0.739*** (0.068)	-0.725*** (0.068)	-0.764*** (0.068)	-0.716*** (0.067)	-0.858*** (0.071)	0.527*** (0.069)	0.406*** (0.067)
Observations	9046	9046	9046	9046	8575	9046	9046
Controls							

*Notes:* All outcomes are z-scores. *Importance of effort* is in the cohort dimension. *Position father* is the respondent's father's income rank when the respondent was 15 years old compared to other fathers of 15 year olds. All positions used as explaining variables have been re-scaled to go from 0.1 to 1. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.9: History of past social positions, unfairness, importance of effort and political view using 5-year average income positions

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Pos. -20 to -16	-0.224*** (0.058)	-0.188** (0.057)	-0.187** (0.058)	-0.205*** (0.056)	-0.198*** (0.057)	0.197*** (0.059)	0.566*** (0.055)
Pos. -15 to -11	-0.143 (0.073)	-0.145* (0.072)	-0.131 (0.072)	-0.132 (0.070)	-0.121 (0.072)	0.143 (0.074)	0.130 (0.070)
Pos. -10 to -6	-0.097 (0.091)	-0.096 (0.089)	-0.137 (0.089)	-0.161 (0.087)	-0.145 (0.089)	0.084 (0.091)	0.034 (0.086)
Pos. -5 to -1	-0.264* (0.113)	-0.205 (0.112)	-0.266* (0.112)	-0.263* (0.109)	-0.259* (0.113)	0.129 (0.114)	-0.000 (0.108)
Position this yr.	-0.515*** (0.092)	-0.509*** (0.090)	-0.513*** (0.090)	-0.528*** (0.088)	-0.634*** (0.092)	0.460*** (0.092)	0.470*** (0.087)
Observations	9388	9388	9388	9388	8895	9388	9388
Controls	✓	✓	✓	✓	✓	✓	✓

*Notes:* All outcomes are z-scores. *Importance of effort* is in the cohort dimension. *Pos. -20 to -16*, *Pos. -15 to -11*, *Pos. -10 to -6* and *Pos. -5 to -1* are five year average cohort positions. All positions used as explaining variables have been re-scaled to go from 0.1 to 1. *Controls* includes a treatment indicator, cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .



Table A-2.10: History of past social positions and inequality views

	Inequal. not prob.	Less redist.	Less satis.	Work not paid off	Rich don't deserve	Not poors own fault
Position father	0.087* (0.035)	0.115** (0.036)	-0.006 (0.038)	0.009 (0.037)	-0.029 (0.037)	-0.019 (0.037)
Position -20 yr.	0.142*** (0.040)	0.136*** (0.041)	0.000 (0.042)	-0.017 (0.041)	-0.147*** (0.042)	-0.110** (0.041)
Position -15 yr.	0.087 (0.046)	0.190*** (0.047)	-0.103* (0.049)	-0.114* (0.048)	-0.169*** (0.049)	-0.077 (0.048)
Position -10 yr.	0.112* (0.055)	0.040 (0.057)	0.061 (0.059)	-0.021 (0.058)	0.023 (0.058)	0.070 (0.058)
Position -5 yr.	0.062 (0.066)	0.012 (0.068)	-0.206** (0.071)	-0.282*** (0.069)	-0.208** (0.070)	-0.051 (0.069)
Position this yr.	0.694*** (0.070)	0.700*** (0.072)	-0.455*** (0.075)	-0.714*** (0.073)	-0.551*** (0.074)	-0.261*** (0.073)
Observations	9036	9036	9036	9036	9036	9036
Controls	✓	✓	✓	✓	✓	✓

*Notes:* All outcomes are z-scores. *Importance of effort* is in the cohort dimension. *Position father* is the respondent's father's income rank when the respondent was 15 years old compared to other fathers of 15 year olds. All positions used as explaining variables have been re-scaled to go from 0.1 to 1. *Controls* includes a treatment indicator, cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.11: Historic shocks and inequality views

	Inequal. not prob. (1)	Less redist. (2)	Less satis. (3)	Work not paid off (4)	Rich don't deserve (5)	Not poors own fault (6)	N (7)	Affected % (8)
Unemployment	-0.16** (0.050)	-0.12* (0.050)	0.15** (0.052)	0.17*** (0.051)	0.11* (0.052)	-0.029 (0.051)	7531	5.03
Disability	-0.29* (0.13)	-0.44*** (0.13)	0.29* (0.14)	0.37** (0.13)	0.26 (0.14)	0.50*** (0.13)	9238	0.60
Hospitalization	-0.035 (0.028)	-0.012 (0.028)	0.076** (0.029)	0.075** (0.029)	0.057* (0.029)	0.066* (0.028)	4746	55.6
Promotion	0.19*** (0.043)	0.20*** (0.044)	-0.13** (0.045)	-0.10* (0.045)	-0.17*** (0.045)	-0.052 (0.045)	7964	6.65
Pre-shock position FE	✓	✓	✓	✓	✓	✓		
Controls	✓	✓	✓	✓	✓	✓		

*Notes:* All outcomes are z-scores. Each cell in the table is a separate regression of the column outcome on the row regressor and the controls indicated in the bottom part of the table. The explaining variables are all indicators that equal 1 if the respondent experienced the shock between 2012 and 2017. In each row, we exclude respondents who experienced the shock between 2008 and 2011. For *Unemployment*, we only use respondents who were in the workforce in the entire period. *Controls* includes cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects, all measured in 2008, and a treatment indicator. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.12: Historic shocks, unfairness, importance of effort and political views with last vote fixed effects

	Unfairness of inequality					Importance of effort (6)	Right-wing (7)	N	Affected %
	Cohort (1)	Gender (2)	Mun. (3)	Edu. (4)	Sector (5)				
Unemployment	0.17*** (0.050)	0.20*** (0.049)	0.18*** (0.049)	0.15** (0.048)	0.18*** (0.050)	-0.088 (0.051)	-0.036 (0.034)	7537	5.03
Disability	0.15 (0.13)	0.40** (0.13)	0.28* (0.13)	0.13 (0.12)		-0.19 (0.13)	0.030 (0.087)	9246	0.61
Hospitalization	0.090** (0.028)	0.076** (0.028)	0.086** (0.028)	0.059* (0.027)	0.039 (0.027)	-0.0093 (0.029)	-0.021 (0.019)	4749	55.5
Promotion	-0.068 (0.044)	-0.066 (0.044)	-0.066 (0.044)	-0.11** (0.043)	-0.16*** (0.043)	0.098* (0.045)	0.072* (0.030)	7970	6.66
Pre-shock position FE	✓	✓	✓	✓	✓	✓	✓		
Controls	✓	✓	✓	✓	✓	✓	✓		
Last vote FE	✓	✓	✓	✓	✓	✓	✓		

*Notes:* All outcomes z-scores. Each cell in the table is a separate regression of the column outcome on the row regressor and the controls indicated in the bottom part of the table. The explaining variables are all indicators that equal 1 if the respondent experienced the shock between 2012 and 2017. In each row, we exclude respondents who already experienced this type of shock in the pre-period (2008-2011). For *Unemployment*, we only use respondents who were in the workforce in the entire period. For *Disability*, we do not estimate the effect on fairness within sector, because very few disabled people work. *Controls* included in all regressions are cohort, gender, municipality, education and sector fixed effects (incl. unemployed/not in workforce), all measured in 2008, and a treatment indicator. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.13: Historic shocks, unfairness, importance of effort and political view without controls

	Current		Unfairness of inequality				Importance		Right-wing	N	Affected %
	position	Cohort	Gender	Mun.	Edu.	Sector	of effort				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
Unemployment	-12.3*** (0.81)	0.17** (0.051)	0.19*** (0.050)	0.18*** (0.050)	0.098 (0.050)	0.13* (0.052)	-0.069 (0.051)	0.024 (0.051)	7537	5.03	
Disability	-24.8*** (2.40)	0.32* (0.13)	0.56*** (0.13)	0.44*** (0.13)	0.26* (0.13)		-0.31* (0.13)	-0.17 (0.13)	9246	0.61	
Hospitalization	-2.20*** (0.51)	0.11*** (0.029)	0.098*** (0.028)	0.11*** (0.028)	0.080** (0.028)	0.065* (0.028)	-0.022 (0.029)	-0.046 (0.029)	4749	55.5	
Promotion	8.86*** (0.77)	-0.13** (0.045)	-0.12** (0.044)	-0.13** (0.045)	-0.16*** (0.044)	-0.22*** (0.045)	0.15** (0.045)	0.20*** (0.045)	7970	6.66	
Pre-shock position FE	✓	✓	✓	✓	✓	✓	✓	✓			
Controls									✓		

*Notes:* All outcomes are z-scores. Each cell in the table is a separate regression of the column outcome on the row regressor and the controls indicated in the bottom part of the table. The explaining variables are all indicators that equal 1 if the respondent experienced the shock between 2012 and 2017. In each row, we exclude respondents who experienced the shock between 2008 and 2011. For *Unemployment*, we only use respondents who were in the workforce in the entire period. For *Disability*, we do not estimate the effect on fairness within sector, because very few disabled people work. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.14: Historic shocks, unfairness, importance of effort and political view using 2SLS

	Unfairness of inequality				Importance of effort	Right-wing	N	Affected %
	Cohort (1)	Gender (2)	Mun. (3)	Edu. (4)	Sector (5)			
Unemployment	-1.60*** (0.41)	-1.77*** (0.39)	-1.71*** (0.41)	-1.34*** (0.37)	-2.34*** (0.60)	0.87* (0.42)	0.50 (0.39)	7537 5.03
Disability	-1.35* (0.60)	-2.31*** (0.58)	-1.84** (0.57)	-1.27* (0.59)		1.42* (0.61)	1.14* (0.57)	9246 0.61
Hospitalization	-5.11** (1.93)	-3.84* (1.55)	-4.81* (1.87)	-3.45 (1.77)	-2.64 (2.32)	0.55 (1.56)	0.97 (1.48)	4749 55.5
Promotion	-1.45** (0.52)	-1.31** (0.50)	-1.36** (0.50)	-1.69*** (0.43)	-1.99*** (0.43)	1.69** (0.53)	2.28*** (0.52)	7970 6.66
Pooled	-1.61*** (0.34)	-1.93*** (0.33)	-1.75*** (0.33)	-1.49*** (0.30)	-2.15*** (0.35)	1.25*** (0.31)	1.08*** (0.31)	29502 12.2
Pre-shock position FE	✓	✓	✓	✓	✓	✓	✓	
Controls	✓	✓	✓	✓	✓	✓	✓	

*Notes:* All outcomes are z-scores. Each cell in the table is a separate 2SLS regression of the column outcome on current position instrumented using the row regressor and the controls indicated in the bottom part of the table. The instruments are all indicators that equal 1 if the respondent experienced the shock between 2012 and 2017. In each row, we exclude respondents who experienced the shock between 2008 and 2011. For *Unemployment*, we only use respondents who were in the workforce in the entire period. For *Disability*, we do not estimate the effect on fairness within sector, because very few disabled people work. *Controls* includes cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects, all measured in 2008, and a treatment indicator. Standard errors in parentheses. In the pooled regression we cluster the standard errors at the individual level. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.15: Survey information experiment and unfairness, importance of effort and political view with same treatment direction across all reference groups

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Positive misperception	-0.149** (0.053)	-0.065 (0.052)	-0.134** (0.052)	-0.229*** (0.051)	-0.186*** (0.056)	0.146** (0.054)	0.120* (0.052)
$T \times \text{Positive}$	0.205*** (0.050)	0.112* (0.050)	0.159** (0.050)	0.134** (0.048)	0.150** (0.051)	-0.027 (0.051)	-0.027 (0.049)
$T \times \text{Negative}$	0.038 (0.035)	0.013 (0.035)	0.032 (0.035)	0.037 (0.034)	0.010 (0.035)	0.029 (0.036)	-0.044 (0.035)
$N$	4701	4701	4701	4701	4385	4701	4701
Position FE	✓	✓	✓	✓	✓	✓	✓

Notes: All outcomes are z-scores. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position.  $T \times \text{Positive}$  and  $T \times \text{Negative}$  are interactions of the treatment indicator and the positive and 0 or negative misperception indicators. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.16: Survey information experiment and inequality views

	Inequal. not prob.	Less redist.	Less satis.	Work not paid off	Rich don't deserve	Not poors own fault
Positive misperception	0.121*** (0.033)	0.113*** (0.034)	-0.082* (0.035)	-0.155*** (0.034)	-0.148*** (0.034)	-0.125*** (0.034)
$T \times \text{Positive}$	-0.021 (0.036)	0.027 (0.037)	-0.007 (0.038)	0.031 (0.037)	0.047 (0.038)	0.016 (0.037)
$T \times \text{Negative}$	-0.012 (0.023)	0.013 (0.024)	0.016 (0.024)	0.015 (0.024)	-0.008 (0.024)	0.020 (0.024)
$N$	9323	9323	9323	9323	9323	9323
Position FE	✓	✓	✓	✓	✓	✓

Notes: All outcomes are z-scores. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position.  $T \times \text{Positive}$  and  $T \times \text{Negative}$  are interactions of the treatment indicator and the positive and 0 or negative misperception indicators. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.17: Survey information experiment and unfairness, importance of effort and political view with controls

	Unfairness of inequality					Importance of effort	Right-wing
	Cohort	Gender	Mun.	Edu.	Sector		
Positive misperception	-0.134*** (0.034)	-0.070* (0.033)	-0.107** (0.033)	-0.081** (0.031)	-0.184*** (0.033)	0.118*** (0.034)	0.122*** (0.033)
$T \times \text{Positive}$	0.170*** (0.037)	0.082* (0.035)	0.103** (0.035)	0.065* (0.029)	0.091** (0.030)	-0.020 (0.037)	-0.036 (0.035)
$T \times \text{Negative}$	0.022 (0.024)	0.029 (0.024)	0.025 (0.024)	0.041 (0.026)	0.005 (0.027)	0.007 (0.024)	-0.023 (0.023)
$N$	9331	9331	9331	9331	8854	9331	9331
Position FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓

*Notes:* All outcomes are z-scores. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position.  $T \times \text{Positive}$  and  $T \times \text{Negative}$  are interactions of the treatment indicator and the positive and 0 or negative misperception indicators. *Controls* includes cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.18: Survey information experiment and unfairness, importance of effort and political view if actual and reported income match

	Unfairness of inequality					Importance of effort	Right-wing
	Cohort	Gender	Mun.	Edu.	Sector		
Positive misperception	-0.065 (0.041)	-0.111** (0.040)	-0.120** (0.039)	-0.077* (0.037)	-0.221*** (0.041)	0.016 (0.042)	0.057 (0.042)
$T \times \text{Positive}$	0.137** (0.046)	0.137** (0.045)	0.106* (0.043)	0.036 (0.036)	0.078* (0.038)	0.053 (0.046)	-0.055 (0.046)
$T \times \text{Negative}$	0.027 (0.027)	0.001 (0.028)	0.023 (0.027)	0.013 (0.031)	-0.010 (0.032)	0.017 (0.027)	-0.030 (0.027)
$N$	6660	6537	6539	6272	5873	6660	6660
Position FE	✓	✓	✓	✓	✓	✓	✓

*Notes:* In this table, we only use respondents whose reported income generate treatment information that is at most five positions from the information they would have received if the reported and actual income exactly matched. All outcomes are z-scores. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position.  $T \times \text{Positive}$  and  $T \times \text{Negative}$  are interactions of the treatment indicator and the positive and 0 or negative misperception indicators. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.19: Survey information experiment and unfairness, importance of effort and political view with interaction of treatment and high income

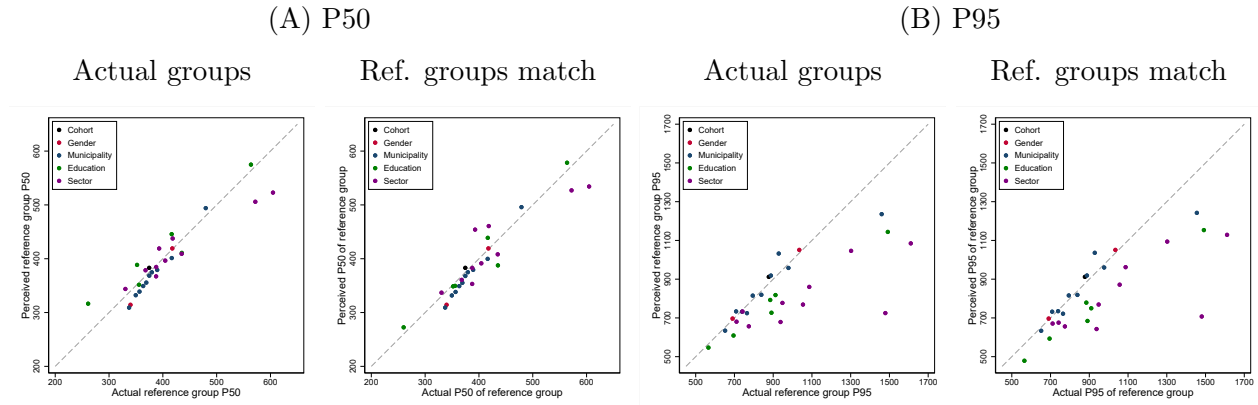
	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Positive misperception	-0.130*** (0.036)	-0.134*** (0.035)	-0.121*** (0.034)	-0.091** (0.034)	-0.253*** (0.037)	0.070 (0.036)	0.112** (0.036)
$T \times \text{Positive}$	0.157*** (0.042)	0.076 (0.041)	0.123** (0.041)	0.082* (0.033)	0.105** (0.035)	-0.076 (0.042)	-0.021 (0.042)
$T \times \text{Negative}$	0.029 (0.051)	0.008 (0.051)	0.077 (0.050)	0.101* (0.049)	0.051 (0.055)	-0.150** (0.051)	-0.047 (0.051)
$T \times \text{High income}$	-0.006 (0.052)	0.016 (0.052)	-0.058 (0.050)	-0.068 (0.048)	-0.051 (0.053)	0.181*** (0.052)	0.023 (0.052)
$N$	9331	9331	9331	9331	8854	9331	9331
Position FE	✓	✓	✓	✓	✓	✓	✓

*Notes:* All outcomes are z-scores. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position.  $T \times \text{Positive}$  and  $T \times \text{Negative}$  are interactions of the treatment indicator and the positive and 0 or negative misperception indicators.  $T \times \text{High income}$  is an interaction of the treatment indicator and an indicator that equals 1 if the individual has a position in the group above 50. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .



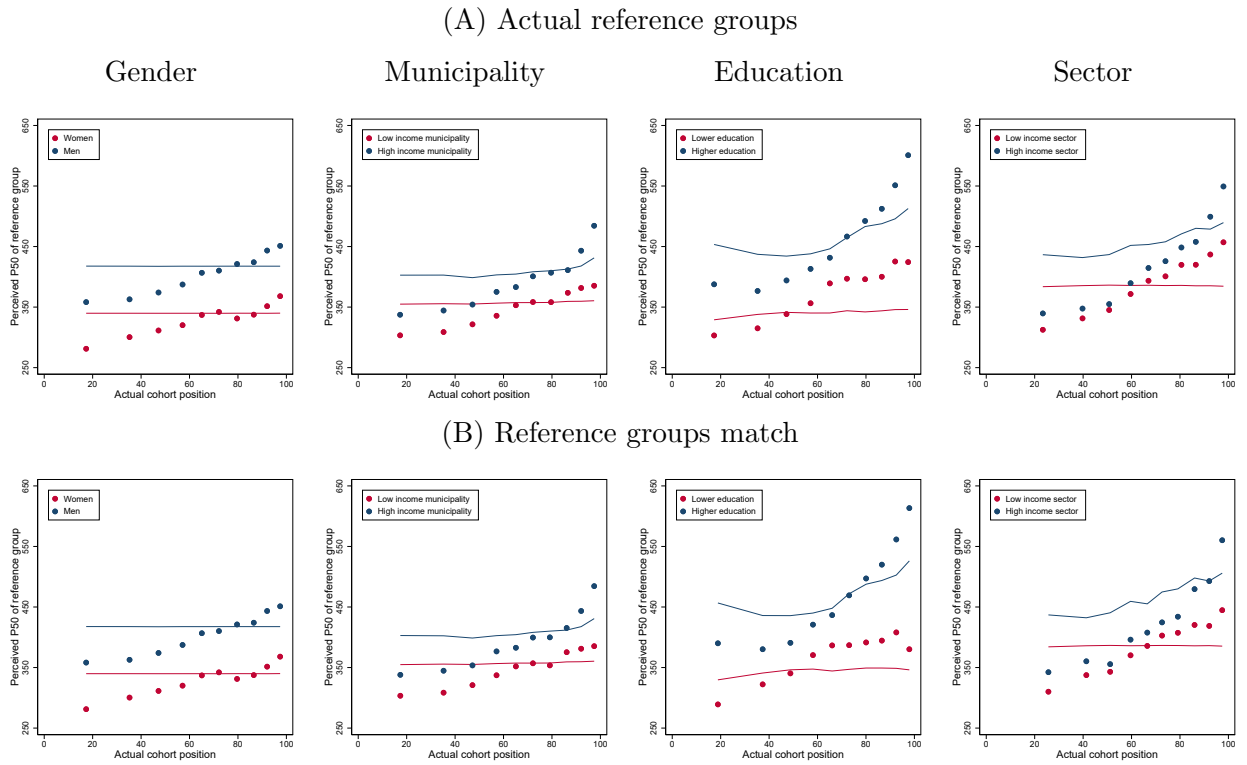
## 2.E Large Reference Groups: Using Actual Groups and Restricted Sample

Figure A-2.16: Perceived P50 and P95 Incomes for large reference groups using actual groups and restricted sample



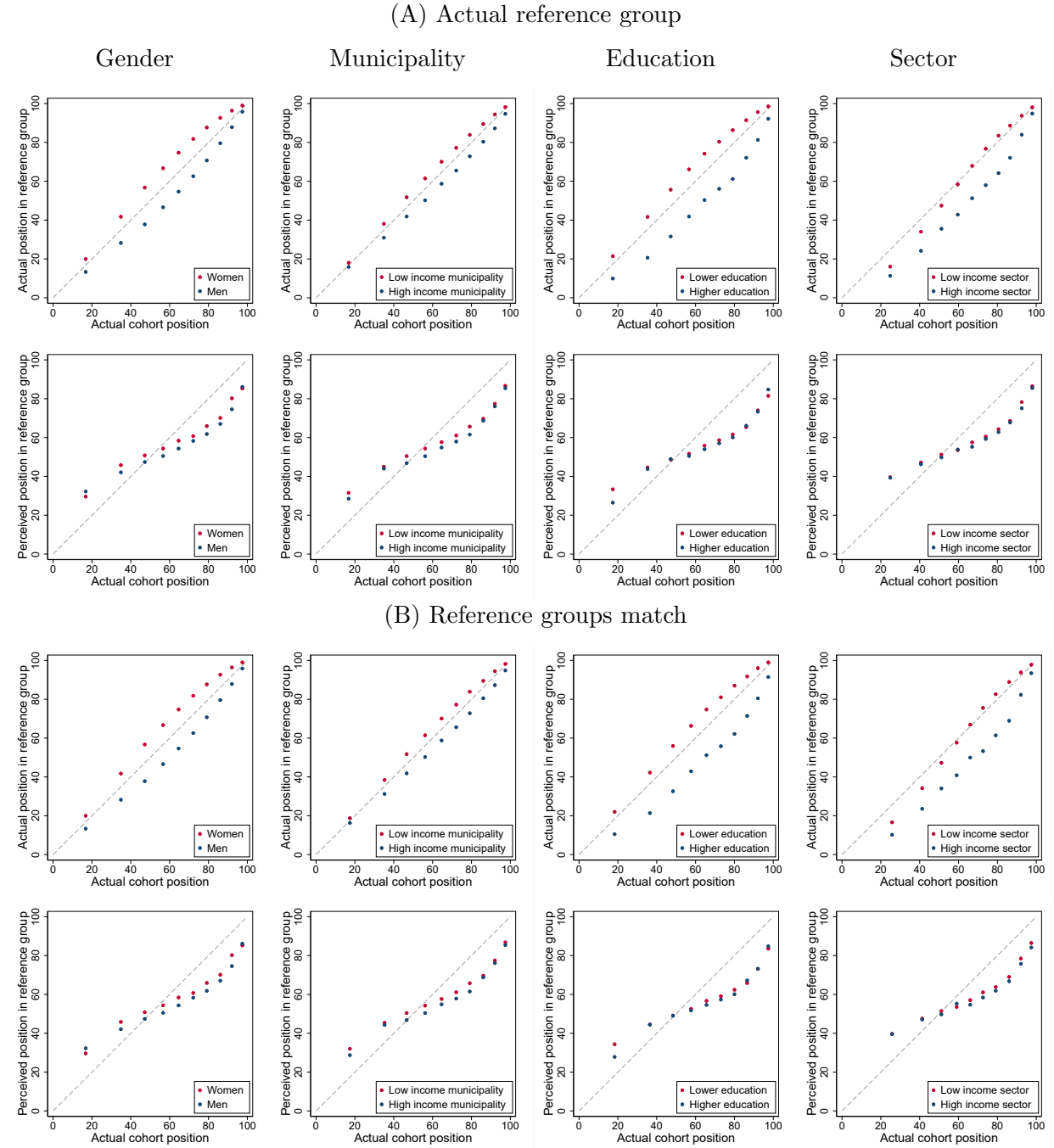
*Notes:* In the first and third panels, we use actual reference group instead of reported reference group. In the second and fourth panels, we restrict the sample and only include respondents in each reference group if the reported group matches the group observed in the register data. For gender, we show one scatter for men and one for women. For municipality we divide the respondents into 10 similar sized groups based on the actual municipality P50 and P95 income and plot one scatter for each group. For education and sector we show one scatter for each educational level or sector. The scatters show the means of the reported P50 or P95 winzorized at the 5th and 95th percentile within the group.

Figure A-2.17: Perceived P50 incomes for large reference groups using actual groups and restricted sample



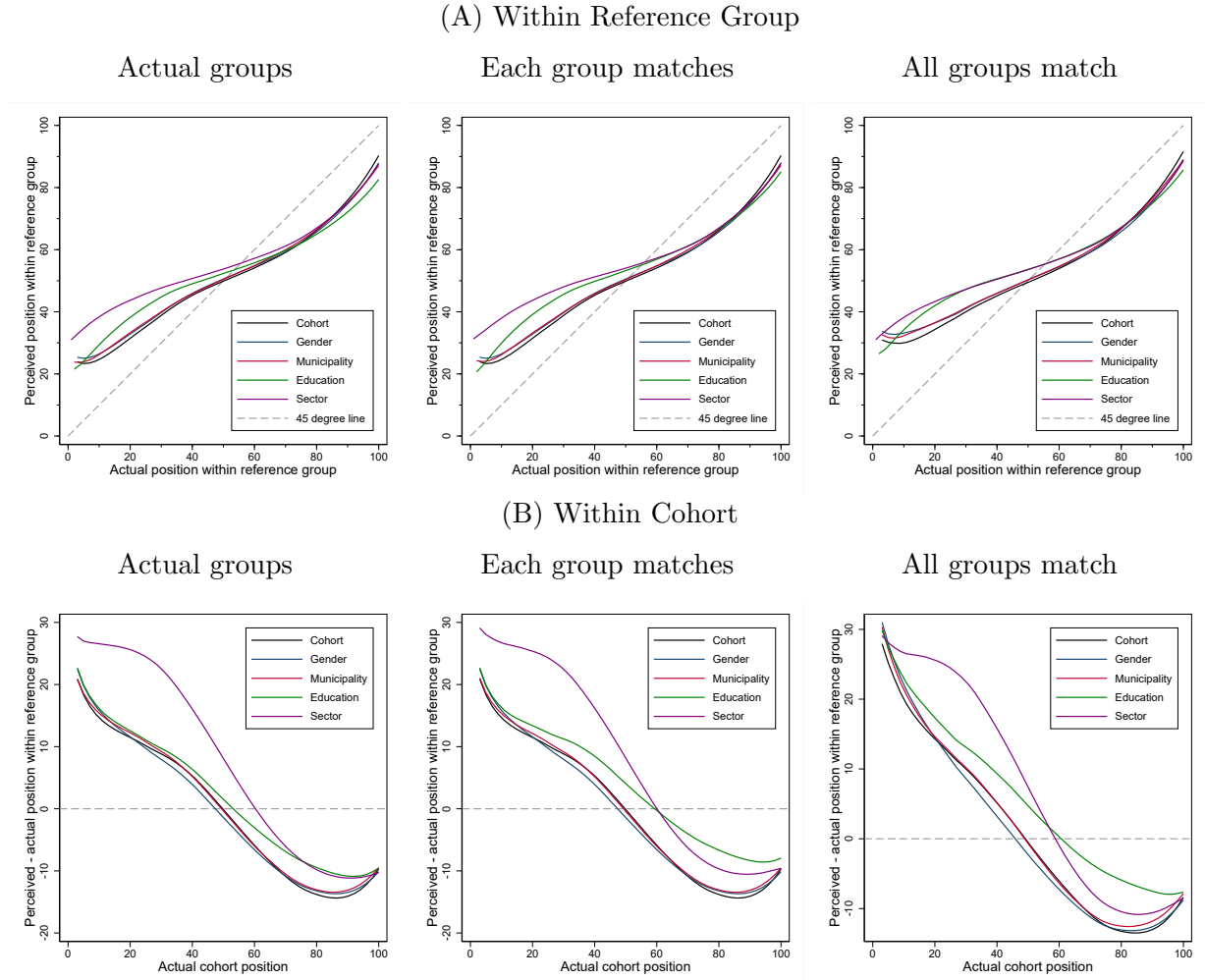
*Notes:* In the top panels, we use actual reference group instead of reported reference group. For each reference group, the bottom panels only include respondents whose reported reference group matches the actual reference group. The solid lines indicate the actual average P50 for each group within the bin. The high/low income split for municipality is based on the median of within cohort in sample actual municipality P50 income. This is also the case for sector. For education, *Higher education* is short cycle higher education, bachelor programs and master programs.

Figure A-2.18: Correlation between actual cohort position and within large reference groups using actual groups and restricted sample



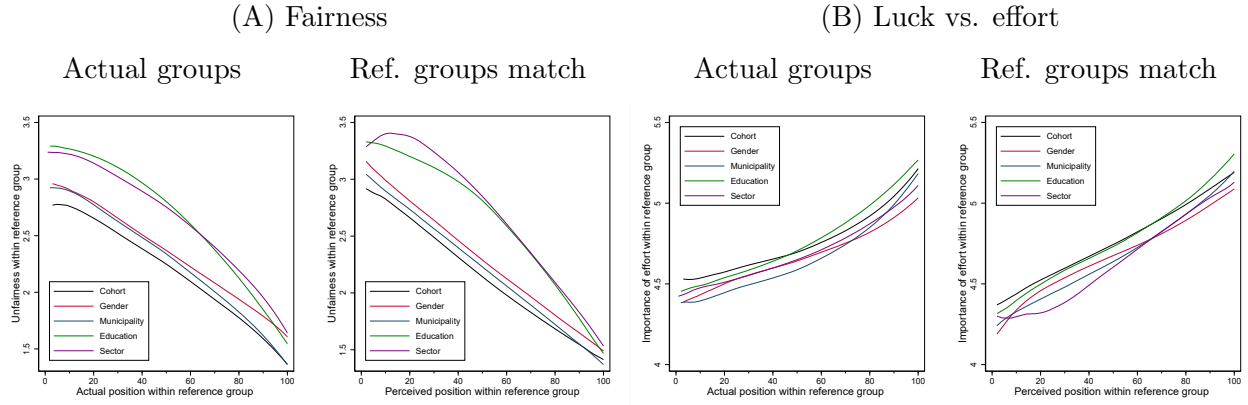
*Notes:* In Panel A, we use actual reference groups instead of reported reference groups. For each reference group, the figures in Panel B only include respondents whose reported reference group matches the actual reference group. The high/low income split for municipality is based on the median of within cohort in sample actual municipality P50 income. This is also the case for sector. For education, *Higher education* is short cycle higher education, bachelor programs and master programs.

Figure A-2.19: Perceived and actual position for large reference groups using actual groups and restricted sample



*Notes:* In the left panels, we use actual reference groups instead of reported reference groups. In the middle panels, we only use respondents in each reference group if the reported group matches the group observed in the register data. In the right panels, only respondents where all reported groups match the groups in the register data are included and the sample is the same across groups. The local linear polynomials have a bandwidth of 10.

Figure A-2.20: View on fairness and effort vs. luck within large reference groups using actual groups and restricted sample



*Notes:* The figure only uses responses from the control group. The local linear polynomials have a bandwidth of 20. In the first and third panels, we use actual reference groups instead of reported reference groups. In the second and fourth panels, we only include respondents for each group if the reported group matches the group observed in the register data.

Table A-2.20: Unfairness, importance of effort and political view by actual position and position misperception using actual reference groups

	Unfairness of inequality					Importance of effort	Right-wing
	Cohort	Gender	Mun.	Edu.	Sector		
<b>Panel A: No controls</b>							
Position	-1.18*** (0.07)	-1.13*** (0.07)	-1.27*** (0.07)	-1.38*** (0.07)	-1.44*** (0.07)	0.82*** (0.07)	0.90*** (0.07)
Misperception	-0.38*** (0.10)	-0.61*** (0.09)	-0.48*** (0.09)	-0.43*** (0.08)	-0.83*** (0.09)	0.44*** (0.10)	0.49*** (0.10)
<b>Panel B: With controls</b>							
Position	-1.09*** (0.09)	-1.01*** (0.08)	-1.10*** (0.08)	-1.04*** (0.08)	-1.22*** (0.08)	0.92*** (0.09)	0.99*** (0.08)
Misperception	-0.40*** (0.10)	-0.41*** (0.10)	-0.41*** (0.09)	-0.28*** (0.09)	-0.54*** (0.09)	0.47*** (0.10)	0.54*** (0.10)
<i>N</i>	4692	4692	4692	4692	4332	4692	4692

*Notes:* In the table, we only use the control group respondents and use actual reference groups instead of reported groups. All outcomes are z-scores. *Position* denotes the actual position within reference group from percentile 1 to 100 divided by 100. Similarly, *Misperception* is the difference between perceived and actual position within the reference group divided by 100. *Controls* includes cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.21: Unfairness, importance of effort and political view by actual position and position misperception using restricted sample

	Unfairness of inequality					Importance	Right-
	Cohort	Gender	Mun.	Edu.	Sector	of effort	wing
<b>Panel A: No controls</b>							
Position	-1.18*** (0.07)	-1.13*** (0.07)	-1.28*** (0.07)	-1.45*** (0.08)	-1.46*** (0.09)	0.82*** (0.07)	0.90*** (0.07)
Misperception	-0.38*** (0.10)	-0.61*** (0.09)	-0.48*** (0.09)	-0.52*** (0.10)	-0.93*** (0.11)	0.44*** (0.10)	0.49*** (0.10)
<b>Panel B: With controls</b>							
Position	-1.09*** (0.09)	-1.01*** (0.08)	-1.10*** (0.09)	-1.12*** (0.09)	-1.19*** (0.10)	0.92*** (0.09)	0.99*** (0.08)
Misperception	-0.40*** (0.10)	-0.41*** (0.10)	-0.41*** (0.09)	-0.39*** (0.10)	-0.58*** (0.11)	0.47*** (0.10)	0.54*** (0.10)
<i>N</i>	4692	4692	4600	3453	3218	4692	4692
Outcome mean	2.01 (0.02)	2.16 (0.02)	2.08 (0.02)	2.60 (0.03)	2.58 (0.03)	4.81 (0.02)	7.10 (0.05)

*Notes:* In the table we only use the control group respondents and for each reference group, we only include respondents whose reported group matches the group reported in the register data. All outcomes are z-scores. *Position* denotes the actual position within reference group from percentile 1 to 100 divided by 100. Similarly, *Misperception* is the difference between perceived and actual position within the reference group divided by 100. *Controls* includes cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects. *Outcome mean* is the mean of the non-standardized outcome variable. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.22: History of past social positions, unfairness, importance of effort and political view using restricted sample

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Position father	-0.059 (0.037)	-0.051 (0.037)	-0.070 (0.037)	-0.112** (0.042)	-0.025 (0.044)	0.025 (0.037)	0.147*** (0.035)
Position -20 yr.	-0.122** (0.042)	-0.097* (0.041)	-0.118** (0.041)	-0.103* (0.046)	-0.148** (0.049)	0.061 (0.042)	0.263*** (0.039)
Position -15 yr.	-0.203*** (0.048)	-0.169*** (0.048)	-0.147** (0.048)	-0.143** (0.053)	-0.081 (0.058)	0.131** (0.049)	0.268*** (0.046)
Position -10 yr.	-0.085 (0.058)	-0.100 (0.057)	-0.119* (0.058)	-0.195** (0.066)	-0.177* (0.071)	0.206*** (0.058)	0.163** (0.055)
Position -5 yr.	-0.108 (0.070)	-0.096 (0.068)	-0.143* (0.069)	-0.109 (0.078)	-0.205* (0.086)	0.037 (0.070)	-0.015 (0.066)
Position this yr.	-0.655*** (0.074)	-0.618*** (0.073)	-0.646*** (0.073)	-0.683*** (0.085)	-0.729*** (0.094)	0.540*** (0.074)	0.479*** (0.070)
Observations	9046	9046	8878	6698	6164	9046	9046
Controls	✓	✓	✓	✓	✓	✓	✓

*Notes:* For each reference group, we only include respondents if the reported group matches the group observed in the register data. All outcomes are z-scores. *Position father* is the respondent's father's income rank when the respondent was 15 years old compared to other fathers of 15 year olds. All positions used as explaining variables have been re-scaled to go from 0.1 to 1. *Controls* includes a treatment indicator, cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.23: Historic shocks, unfairness, importance of effort and political view using restricted sample

	Unfairness of inequality					Importance	Right-
	Cohort (1)	Gender (2)	Mun. (3)	Edu. (4)	Sector (5)	of effort (6)	wing (7)
Unemployment	0.20*** (0.051)	0.23*** (0.050)	0.21*** (0.051)	0.18** (0.059)	0.24*** (0.069)	-0.11* (0.052)	-0.061 (0.049)
Observations	7537	7537	7397	5599	5498	7537	7537
Affected %	5.03	5.03	4.96	4.80	3.62	5.03	5.03
Disability	0.30* (0.13)	0.54*** (0.13)	0.42** (0.14)	0.36* (0.15)		-0.31* (0.13)	-0.25* (0.13)
Observations	9246	9246	9073	6841		9246	9246
Affected %	0.61	0.61	0.57	0.61		0.61	0.61
Hospitalization	0.093** (0.029)	0.079** (0.028)	0.091** (0.029)	0.054 (0.033)	0.0084 (0.034)	-0.010 (0.029)	-0.018 (0.028)
Observations	4749	4749	4662	3543	3257	4749	4749
Affected %	55.5	55.5	55.5	55.5	54.3	55.5	55.5
Promotion	-0.12** (0.045)	-0.11** (0.044)	-0.12** (0.045)	-0.19*** (0.051)	-0.21*** (0.054)	0.14** (0.045)	0.19*** (0.043)
Observations	7970	7970	7832	5950	5622	7970	7970
Affected %	6.7	6.7	6.7	6.7	6.4	6.7	6.7
Pre-shock position FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓

*Notes:* For each reference group, we only include respondents if the reported group matches the group observed in the register data. All outcomes are z-scores. Each cell in the table is a separate regression of the column outcome on the row regressor and the controls indicated in the bottom part of the table. The explaining variables are all indicators that equal 1 if the respondent experienced the shock between 2012 and 2017. In each row, we exclude respondents who experienced the shock between 2008 and 2011. For *Unemployment*, we only use respondents who were in the workforce in the entire period. For *Disability*, we do not estimate the effect on fairness within sector, because very few disabled people work. *Controls* includes cohort fixed effects, an indicator for men, municipality fixed effects, educational level fixed effects and sector (incl. unemployed/not in workforce) fixed effects, all measured in 2008, and a treatment indicator. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .



Table A-2.24: Survey information experiment and unfairness, importance of effort and political view using actual reference groups

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Positive misperception	-0.132*** (0.034)	-0.131*** (0.033)	-0.131*** (0.033)	-0.140*** (0.032)	-0.249*** (0.034)	0.112** (0.034)	0.117*** (0.034)
T × Positive	0.154*** (0.037)	0.082* (0.036)	0.097** (0.035)	0.079* (0.031)	0.087** (0.032)	-0.010 (0.038)	-0.013 (0.037)
T × Negative	0.023 (0.024)	0.022 (0.024)	0.030 (0.025)	0.033 (0.026)	0.013 (0.027)	0.006 (0.024)	-0.027 (0.024)
N	9331	9331	9331	9331	8647	9331	9331
Position FE	✓	✓	✓	✓	✓	✓	✓

Notes: In the table, we use actual reference groups observed in the register data instead of reported reference groups. All outcomes are z-scores. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position. *T × Positive* and *T × Negative* are interactions of the treatment indicator and the positive and 0 or negative misperception indicators. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

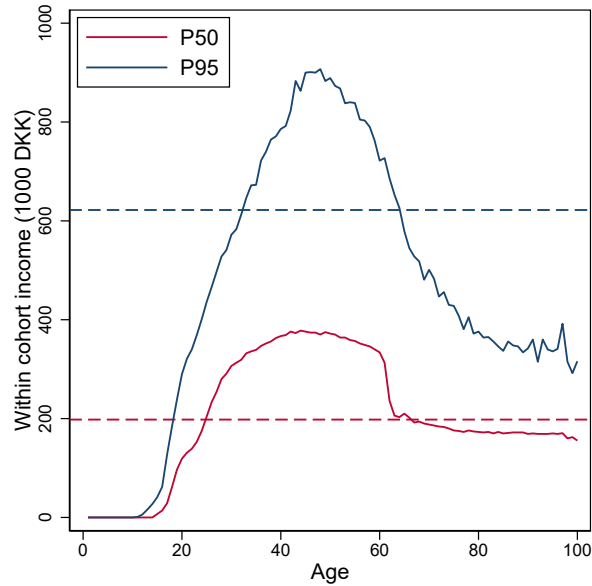
Table A-2.25: Survey information experiment and unfairness, importance of effort and political view using restricted sample

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Positive misperception	-0.132*** (0.034)	-0.131*** (0.033)	-0.131*** (0.033)	-0.168*** (0.036)	-0.277*** (0.040)	0.112** (0.034)	0.117*** (0.034)
T × Positive	0.154*** (0.037)	0.082* (0.036)	0.100** (0.035)	0.058 (0.034)	0.098** (0.038)	-0.010 (0.038)	-0.013 (0.037)
T × Negative	0.023 (0.024)	0.022 (0.024)	0.034 (0.025)	0.019 (0.031)	0.001 (0.033)	0.006 (0.024)	-0.027 (0.024)
N	9331	9331	9156	6901	6356	9331	9331
Position FE	✓	✓	✓	✓	✓	✓	✓

Notes: For each reference group, we only use respondents if the reported group matches the group observed in the register data. All outcomes are z-scores. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position. *T × Positive* and *T × Negative* are interactions of the treatment indicator and the positive and 0 or negative misperception indicators. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

## 2.F Moments in the Income Distribution

Figure A-2.21: Within cohort P50 and P95 by age



*Notes:* This figure shows the within cohort P50 and P95 income based on a 10% sample of the full population in Denmark. We use the same income definition as in the survey which excludes early retirement benefits, since the cohorts surveyed are not yet eligible for this benefit. The age cut-off for early retirement benefits is 60 and therefore we see a sharp drop at this age. We include pension payments, since we cannot disentangle old age pension and disability pension.

Table A-2.26: Moments in the full income distribution

	Income distribution percentiles				
	P5	P25	P50	P75	P95
Full population	0	57	198	358	622
Adult population	36	158	261	394	670
Working age population	39	217	333	447	751
45-50 year olds	112	262	373	502	896

*Notes:* This table shows different moments of the income distribution in 1,000 DKK based of different definitions of the population. The moments are based on a 10% sample of the full population in Denmark in 2017. *Adult population* are individuals from age 18 and up. *Working age population* are from age 25 to 65.

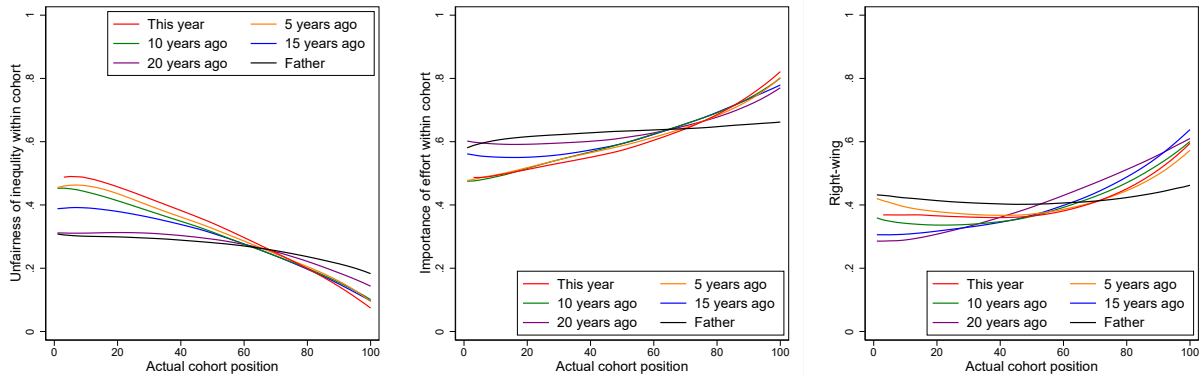
## 2.G Using Indicator Outcome Variables

Table A-2.27: Unfairness, importance of effort and political view by actual position and position misperception using indicator outcomes

	Unfairness of inequality					Importance of effort	Right-wing
	Cohort	Gender	Mun.	Edu.	Sector		
<b>Panel A: No controls</b>							
Position	-0.52*** (0.03)	-0.51*** (0.03)	-0.57*** (0.03)	-0.65*** (0.03)	-0.69*** (0.04)	0.45*** (0.03)	0.32*** (0.03)
Misperception	-0.18*** (0.04)	-0.25*** (0.04)	-0.22*** (0.04)	-0.20*** (0.04)	-0.42*** (0.04)	0.24*** (0.05)	0.18*** (0.05)
<b>Panel B: With controls</b>							
Position	-0.51*** (0.04)	-0.48*** (0.04)	-0.52*** (0.04)	-0.52*** (0.04)	-0.59*** (0.04)	0.45*** (0.04)	0.35*** (0.04)
Misperception	-0.19*** (0.04)	-0.18*** (0.04)	-0.19*** (0.04)	-0.16*** (0.04)	-0.30*** (0.04)	0.25*** (0.05)	0.19*** (0.05)
<i>N</i>	4692	4692	4692	4692	4452	4692	4692
Outcome mean	.27 (0.01)	.30 (0.01)	.29 (0.01)	.40 (0.01)	.40 (0.01)	.63 (0.01)	.41 (0.01)

*Notes:* The sample is restricted to control group respondents. All outcomes are indicators that equal 1 if the outcome z-score is larger than 0. *Position* denotes the actual position within the reference group from percentile 1 to 100 divided by 100. Similarly, *Misperception* is the difference between perceived and actual position within the reference group divided by 100. *Controls* are cohort, municipality, education, gender, and sector fixed effects (including unemployed/not in workforce). *Outcome mean* is the mean of the indicator outcome variable. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Figure A-2.22: History of past social positions, unfairness, importance of effort, and political views using indicator outcomes



*Notes:* Bandwidth for local linear polynomials is 20. For *Father*, the x-axis is the father's position among fathers when the respondent was 15 years old. All outcomes are indicators that equal 1 if the outcome z-score is larger than 0. Sample restricted to the control group.

Table A-2.28: History of past social positions, unfairness, importance of effort, and political view using indicator outcomes

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Position father	-0.023 (0.016)	-0.021 (0.017)	-0.028 (0.017)	-0.051** (0.018)	-0.047** (0.018)	0.021 (0.018)	0.060*** (0.018)
Position -20 yr.	-0.038* (0.018)	-0.035 (0.019)	-0.028 (0.019)	-0.062** (0.020)	-0.062** (0.020)	0.027 (0.020)	0.132*** (0.020)
Position -15 yr.	-0.089*** (0.021)	-0.075*** (0.022)	-0.068** (0.022)	-0.059** (0.023)	-0.061** (0.023)	0.064** (0.023)	0.093*** (0.023)
Position -10 yr.	-0.045 (0.026)	-0.054* (0.026)	-0.072** (0.026)	-0.080** (0.028)	-0.088** (0.028)	0.069* (0.028)	0.075** (0.028)
Position -5 yr.	-0.035 (0.031)	-0.027 (0.032)	-0.038 (0.031)	-0.032 (0.033)	-0.036 (0.034)	0.056 (0.034)	-0.053 (0.033)
Position this yr.	-0.311*** (0.033)	-0.314*** (0.034)	-0.322*** (0.033)	-0.356*** (0.035)	-0.397*** (0.036)	0.246*** (0.036)	0.202*** (0.035)
Observations	9046	9046	9046	9046	8575	9046	9046
Controls	✓	✓	✓	✓	✓	✓	✓

*Notes:* All outcomes are indicators that equal 1 if the outcome z-score is larger than 0. *Position* denotes the cohort position from percentile 1 to 100 divided by 100. A coefficient of 1 means that going from the bottom of the distribution to the top increases the outcome by one standard deviation. *Importance of effort* is for income differences within cohort. *Position father* is the respondent's father's income rank when the respondent was 15 years old compared to other fathers of 15 year olds. *Controls* includes a treatment indicator, cohort, municipality, gender, education, and sector fixed effects (including unemployed/not in workforce). Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.29: Historic shocks, unfairness, importance of effort and political views using indicator outcomes

	Unfairness of inequality				Importance of effort	Right-wing	N	Affected %
	Cohort (1)	Gender (2)	Mun. (3)	Edu. (4)	Sector (5)			
Unemployment	0.082*** (0.023)	0.093*** (0.023)	0.074** (0.023)	0.073** (0.025)	0.070** (0.026)	-0.068** (0.025)	7537 (8)	5.03 (9)
Disability	0.090 (0.059)	0.15* (0.061)	0.17** (0.060)	0.11 (0.063)		-0.16* (0.065)	9246 (8)	0.61 (9)
Hospitalization	0.036** (0.013)	0.031* (0.013)	0.033* (0.013)	0.028* (0.014)	0.017 (0.014)	-0.0080 (0.014)	4749 (8)	55.5 (9)
Promotion	-0.057** (0.020)	-0.056** (0.021)	-0.057** (0.021)	-0.089*** (0.022)	-0.13*** (0.022)	0.072** (0.022)	7970 (8)	6.66 (9)
Pre-shock position	✓	✓	✓	✓	✓	✓	✓	
Controls	✓	✓	✓	✓	✓	✓	✓	

Notes: All outcomes are indicators that equal 1 if the outcome z-score is larger than 0. Each cell in the table is a separate regression of the column outcome on the row regressor and the controls indicated in the bottom part of the table. The explaining variables are all indicators that equal 1 if the respondent experienced the shock between 2012 and 2017. In each row, we exclude respondents who experienced the shock between 2008 and 2011. For *Unemployment*, we only use respondents who were in the workforce in the entire period. For *Disability*, we do not estimate the effect on fairness within sector, because very few disabled people work. *Controls* included in all regressions are cohort, gender, municipality, education and sector fixed effects (incl. unemployed/not in workforce), all measured in 2008, and a treatment indicator. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.30: Survey information experiment and unfairness, importance of effort and political view using indicator outcomes

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Positive misperception	-0.056*** (0.015)	-0.049** (0.015)	-0.069*** (0.015)	-0.052*** (0.015)	-0.135*** (0.016)	0.051** (0.017)	0.040* (0.017)
$T \times \text{Positive}$	0.061*** (0.017)	0.030 (0.017)	0.036* (0.016)	0.027 (0.015)	0.039** (0.015)	-0.001 (0.018)	-0.003 (0.019)
$T \times \text{Negative}$	0.001 (0.011)	0.010 (0.011)	0.001 (0.011)	0.035** (0.013)	0.007 (0.013)	0.003 (0.012)	-0.001 (0.012)
$N$	9331	9331	9331	9331	8854	9331	9331
Position FE	✓	✓	✓	✓	✓	✓	✓

*Notes:* All outcomes are indicators that equal 1 if the outcome z-score is larger than 0. *Positive misperception* is an indicator that equals 1 if the perceived position is larger than the actual position.  $T \times \text{Positive}$  and  $T \times \text{Negative}$  are interactions of the treatment indicator and the positive and 0 or negative misperception indicators. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

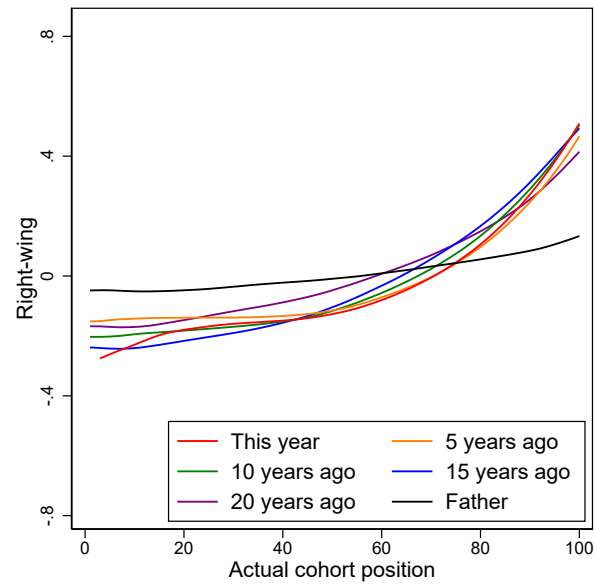
## 2.H Using Economic Policy View as the “Right-wing” Outcome

Table A-2.31: Unfairness, importance of effort and political view by actual position and position misperception using economic policy view as right-wing outcome

	Unfairness of inequality					Importance	Right-
	Cohort	Gender	Mun.	Edu.	Sector	of effort	wing
<b>Panel A: No controls</b>							
Position	-1.18*** (0.07)	-1.13*** (0.07)	-1.27*** (0.07)	-1.36*** (0.07)	-1.42*** (0.07)	0.82*** (0.07)	0.91*** (0.07)
Misperception	-0.38*** (0.10)	-0.61*** (0.09)	-0.48*** (0.09)	-0.38*** (0.08)	-0.86*** (0.09)	0.44*** (0.10)	0.45*** (0.10)
<b>Panel B: With controls</b>							
Position	-1.09*** (0.09)	-1.01*** (0.08)	-1.09*** (0.08)	-1.03*** (0.08)	-1.19*** (0.08)	0.92*** (0.09)	0.86*** (0.09)
Misperception	-0.40*** (0.10)	-0.41*** (0.10)	-0.42*** (0.09)	-0.30*** (0.09)	-0.61*** (0.09)	0.47*** (0.10)	0.43*** (0.10)
<i>N</i>	4692	4692	4692	4692	4452	4692	4692
Outcome mean	2.01 (0.02)	2.16 (0.02)	2.09 (0.02)	2.54 (0.03)	2.53 (0.03)	4.81 (0.02)	3.01 (0.01)

*Notes:* The sample is restricted to control group respondents. All outcomes are z-scores. *Position* denotes the actual position within the reference group from percentile 1 to 100 divided by 100. Similarly, *Misperception* is the difference between perceived and actual position within the reference group divided by 100. *Controls* are cohort, municipality, education, gender, and sector fixed effects (including unemployed/not in workforce). *Outcome mean* is the mean of the non-standardized outcome variable. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Figure A-2.23: History of past social positions, unfairness, importance of effort, and political views using economic policy view as right-wing outcome



*Notes:* Bandwidth for local linear polynomials is 20. For *Father*, the x-axis is the father's position among fathers when the respondent was 15 years old. The outcomes is the standardized z-score of the economic policy view variable. Sample restricted to the control group.



Table A-2.32: History of past social positions, unfairness, importance of effort, and political view using economic policy view as right-wing outcome

	Unfairness of inequality					Importance of effort	Right- wing
	Cohort	Gender	Mun.	Edu.	Sector		
Position father	-0.059 (0.037)	-0.051 (0.037)	-0.068 (0.037)	-0.081* (0.036)	-0.075* (0.037)	0.025 (0.037)	0.122*** (0.036)
Position -20 yr.	-0.122** (0.042)	-0.097* (0.041)	-0.108** (0.041)	-0.113** (0.040)	-0.131** (0.041)	0.061 (0.042)	0.141*** (0.041)
Position -15 yr.	-0.203*** (0.048)	-0.169*** (0.048)	-0.148** (0.048)	-0.126** (0.047)	-0.139** (0.048)	0.131** (0.049)	0.268*** (0.047)
Position -10 yr.	-0.085 (0.058)	-0.100 (0.057)	-0.117* (0.057)	-0.186*** (0.056)	-0.162** (0.057)	0.206*** (0.058)	0.197*** (0.057)
Position -5 yr.	-0.108 (0.070)	-0.096 (0.068)	-0.152* (0.069)	-0.108 (0.067)	-0.130 (0.069)	0.037 (0.070)	-0.030 (0.068)
Position this yr.	-0.655*** (0.074)	-0.618*** (0.073)	-0.647*** (0.073)	-0.678*** (0.071)	-0.761*** (0.074)	0.540*** (0.074)	0.358*** (0.072)
Observations	9046	9046	9046	9046	8575	9046	9046
Controls	✓	✓	✓	✓	✓	✓	✓

*Notes:* All outcomes are z-scores. *Position* denotes the cohort position from percentile 1 to 100 divided by 100. *Position father* is the respondent's father's income rank when the respondent was 15 years old compared to other fathers of 15 year olds. *Controls* includes a treatment indicator, cohort, municipality, gender, education, and sector fixed effects (including unemployed/not in workforce). Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-2.33: Historic shocks, unfairness, importance of effort and political views using economic policy view as right-wing outcome

	Current position (1)	Unfairness of inequality					Importance of effort (7)	Right-wing (8)	N	Affected % (10)
		Cohort (2)	Gender (3)	Mun. (4)	Edu. (5)	Sector (6)				
Unemployment	-12.2*** (0.78)	0.20*** (0.051)	0.23*** (0.050)	0.21*** (0.050)	0.18*** (0.049)	0.20*** (0.052)	-0.11* (0.052)	-0.049 (0.050)	7537	5.03
Disability	-21.9*** (2.31)	0.30* (0.13)	0.54*** (0.13)	0.42** (0.13)	0.27* (0.13)		-0.31* (0.13)	-0.26* (0.13)	9246	0.61
Hospitalization	-1.83*** (0.49)	0.093*** (0.029)	0.079** (0.028)	0.089** (0.028)	0.060* (0.028)	0.039 (0.028)	-0.010 (0.029)	-0.0082 (0.028)	4749	55.5
Promotion	8.51*** (0.74)	-0.12** (0.045)	-0.11** (0.044)	-0.12** (0.045)	-0.17*** (0.044)	-0.21*** (0.044)	0.14** (0.045)	0.17*** (0.044)	7970	6.66
Pre-shock position FE	✓	✓	✓	✓	✓	✓	✓	✓		
Controls	✓	✓	✓	✓	✓	✓	✓	✓		

Notes: All outcomes z-scores. Each cell in the table is a separate regression of the column outcome on the row regressor and the controls indicated in the bottom part of the table. The explaining variables are all indicators that equal 1 if the respondent experienced the shock between 2012 and 2017. In each row, we exclude respondents who experienced the shock between 2008 and 2011. For *Unemployment*, we only use respondents who were in the workforce in the entire period. For *Disability*, we do not estimate the effect on fairness within sector, because very few disabled people work. *Controls* included in all regressions are cohort, gender, municipality, education and sector fixed effects (incl. unemployed/not in workforce), all measured in 2008, and a treatment indicator. Standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

# Chapter 3

## **Preferences predict who commits crime among young men**

with Thomas Epper, Ernst Fehr, Claus Thustrup Kreiner, Søren Leth-Petersen and Gregers  
Nytoft Rasmussen

## 3.1 Introduction

In choice theory of crime, individuals trade-off their benefits from criminal activity against the opportunity costs of legitimate activity and the risks of future costs due to apprehension and punishment (Becker, 1968; Ehrlich, 1973; Freeman, 1999). People can face different trade-offs, for instance due to differences in how much they can otherwise earn in the labor market, or they can evaluate the trade-offs differently due to differences in preferences. A large empirical literature documents that variation in the trade-off people face predicts who commit crime (Freeman, 1999; Chalfin and McCrary, 2017; Nagin, 2013; Draca and Machin, 2015). In contrast, little is known about the role of preferences.

Key preference parameters are risk tolerance and impatience. Intuitively, crime provides a benefit now but at the risk of a cost in the future which makes crime less attractive for people who dislike risk and for people who care more about future well-being (Freeman, 1999). Thus, more risk tolerant and more impatient people are more prone to commit crime than others (see Appendix Section 3.A for a formal derivation in a basic two-period model of criminal behavior). In standard choice theory, people are entirely driven by self-interest, but the theory can be extended to allow for altruistic motives or, more sophisticated, other-regarding preferences (Andreoni, 1990; Fehr and Schmidt, 1999; Garoupa, 2003). More altruistic people will commit less crime because they care about the costs they inflict on others.

In this paper, we ask whether preferences predict who commits crime among young men? We focus on young men (age 18-19) who are known to have much higher crime rates than women and older people (Grogger, 1995; Freeman, 1999; Levitt and Lochner, 2001). To answer the question, we leverage a unique combination of data where we use state-of-the-art methods to elicit preference parameters of young men in Denmark and link this experimental data to administrative records with information about all criminal offences and a large set of individual background characteristics.

Preferences are typically considered to be relatively fixed personal characteristics and it is basically impossible to randomly assign preference-characteristics to people (Epper et al., 2020; Kreiner et al., 2020). To analyze the role of preferences, we regress observed criminal behavior on elicited preferences while controlling thoroughly for other characteristics of the individuals.

The link between experimental data and administrative records enables us to include an extraordinary rich set of relevant control variables. This includes school performance, area of residence, immigrant status, family size, birth order, parental socio-economic status, criminal records of parents, and family stress as measured by parental divorce or unemployment. Moreover, when we collected the experimental data on preferences, we also asked about self-control which is known to be a strong predictor of crime (Gottfredson and Hirschi, 1990; Pratt and Cullen, 2000; Moffitt et al., 2011). We use this information to control for behavioral factors other than risk, time and social preferences.

We also investigate the relation between preferences and different types of criminal offenses. Arguably, the hypothesis that cost-benefit considerations help explain crime seems more appropriate for property crimes than violent, drug and sexual offences where lack of self-control may be a more important driver.

To our knowledge, this study is the first to link experimentally-elicited risk, time, and social preferences to actual criminal behavior. We follow a large literature in experimental economics that elicits people's preferences using incentivized choice experiments, where participants receive payments according to their decisions in the experiment. This literature documents pervasive variation in preference parameters including that some people are much more risk tolerant, impatient and motivated by self-interest than others (Frederick et al., 2002; Bruhin et al., 2010; Fehr and Schmidt, 2006). Studies also show that these parameters are correlated with real-life behavior and outcomes in accordance with theoretical predictions, for example savings, high school graduation, disciplinary referrals of school children, entrepreneurship, BMI and smoking (Barsky et al., 1997; Chabris et al., 2008; Castillo et al., 2011; Sutter et al., 2013; Noussair et al., 2014; Kerr et al., 2019; Castillo et al., 2019; Epper et al., 2020). The number of participants in our elicitation-experiment is large compared to previous studies and is sampled randomly from the population in contrast to many previous studies based on samples of students (Falk et al., 2013). A large population sample is important for our agenda where accurate estimates are a challenge because crime offences are a low-frequency outcome, and one where students differ significantly from non-students.

Most closely related to our work is a recent study that finds a significant relationship between patience and criminal offences (Åkerlund et al., 2016). The study uses a non-

incentivized survey question to measure patience and demonstrates that this predicts crime. Interestingly, the authors find that the predictive power of patience is about 1/3 of the power of cognitive skills and that patience is most predictive of property crime, which is similar to our findings. Our study differs (*i*) by providing a more comprehensive and behaviorally elicited set of preference measures that include risk preferences and social preferences and (*ii*) by including a self-control measure in the set of predictors. This enables us to (*i*) show that risk tolerance is the strongest predictor of crime among the preference parameters and (*ii*) document that risk and time preferences are specifically predictive for property crime where self-control is not a significant predictor whereas in the domain of violent, drug and sexual offenses self-control is a key predictor while risk and time preferences are not predictive.

## 3.2 Materials and Methods

Based on a random sample provided from population registries by Statistics Denmark, we conducted an online incentivized preference-elicitation experiment among 18-19 y-old individuals in Denmark in 2018. We invited 13,799 individuals to participate, who all received a personalized letter from the University of Copenhagen inviting them to participate on a customized internet platform.

The invitations were distributed through an electronic mailbox (E-boks), which is the default way to receive mail from public authorities in Denmark. Previous research suggests that response rates are higher when contacting a random sample of potential respondents in this way compared to sending out physical invitation letters (Hvidberg et al., 2020). In our case, 39% of the invited individuals logged on to the platform.

For the elicitation of each preference parameter, the participants were presented a series of choice situations, resembling methods previously used in the literature as described below. Before making decisions in these choice situations, participants watched an animated instruction video and completed a tutorial session. The choice situations were presented in random order. Each choice situation involved a monetary trade-off, and towards the end of the session one choice situation was randomly selected to be paid out. The average payment to participants was DKK 250 (USD 40). After the random selection, participants typed their

Figure 3.1: Screenshots of online experiment eliciting time and risk preferences



cell number and the money was then transferred through Mobile Pay, a Danish app used for fast transfer of money between individuals, paying bills and paying in stores.

The elicitation of time preferences used a money-earlier-or-later experiment (Epper et al., 2020) with 16 choice situations. Panel A of Figure 3.1 shows a screen shot of one of the choice situations in the time preference task. In this choice situation, an individual could choose to get DKK 250 paid out in 8 weeks or to save all or some of the money for later and receive the savings plus an interest rate of 2.4% in 16 weeks. In this example, the individual chose to save DKK 100 corresponding to a savings rate of 40% ( $100/250$ ). This gave a payout of DKK 150 in 8 weeks and a payout of DKK 102.4 in 16 weeks. The rate of return and the time profile varied across the choice situations as described in Appendix Table A-3.7.

We compute the mean savings rate across the choice situations of each individual and then use this measure to rank people on a 1-100 scale, corresponding to their percentile positions in the distribution of elicited impatience. The degree of impatience may be computed in more sophisticated ways, e.g. by estimating a structural model, but as we show in Appendix Table A-3.4 this does not change the results. A reason is that the rank position of an individual is a quite robust measure. For example, it is unaffected by other ways of computing impatience from the experiment if this keeps the relative positions of people unchanged.

The elicitation of risk preferences is based on an investment experiment (Gneezy and Potters, 1997) with 15 choice situations. Panel B of Figure 3.1 shows a screen shot for one of the situations. Here, an individual could choose to get DKK 250 with certainty or invest

some or all of the money in a lottery which yielded an average rate of return of 6%, but including a risk of a significant loss. In this example, the individual chose to keep DKK 50 and invest DKK 200 in the lottery, corresponding to an investment of 80% of the initial DKK 250 endowment. The investment gave DKK 80 (a loss) with a chance of 40% and DKK 300 (a win) with a chance of 60%. The outcome of the lottery and the sum of money earned were displayed on the screen afterwards. If this situation was selected for payment then the individual would receive the money within 24 hours. The probability of winning and the expected rate of return varied across the choice situations as described in Appendix Table A-3.8. Similar to the measurement of impatience, we compute the mean investment of each individual across the choice situations and then use this measure to rank people, thereby obtaining their percentile positions in the distribution of risk tolerance.

The elicitation of altruism uses 20 different dictator games (Forsythe et al., 1994) (described in Appendix Table A-3.9) that systematically vary the costs and benefits of giving as well as the resulting inequality (similar to (Fehr et al., 2020)). These dictator games enable us to construct an overall measure of altruism as well as decomposing this measure into behindness aversion and aheadness aversion. This elicitation included 20 choice situations with different payment profiles. A simple measure of altruism is the money given to the other person as a share of what the individual could potentially give to the other person (the giving rate). As with the other preference measures, we compute the mean giving rate across the 20 choice situations and then use this measure to rank the individual relative to other participants. Appendix Figure A-3.1 shows a screenshot for one of the choice situations.

Conducting the experimental tasks was time consuming and we, therefore, only included a few survey questions. We asked individuals about age and gender, which can be verified in the registers, and about their school GPA, which is used in a few special cases where grades are not included in the register data. Most importantly, we also included a question about the self-assessed capacity to exercise self-control, which is known to be a strong predictor of crime. On average, participants spent 45 minutes from start to end.

We link the experimental data to administrative records that contain all criminal offences as well as detailed background information about the individuals. Since crime is a low-frequency outcome, we define a person as criminal if he or she is convicted of a criminal



offence committed during the age span 15-20. We exclude traffic offences, which are common but mostly minor offences that only involve small fines. We mainly focus on the males in our sample, which are known to have much higher crime rates than females. The crime rate is 19% for the invited males compared to 3% for the invited females. More details on the crime data and the many background variables included are provided in Appendix Section 3.B. The appendix also includes summary statistics on the random sample of invited men and on the analysis sample of men who completed all experiments (Appendix Table A-3.1). In line with previous research inviting people to participate in surveys/experiments (Epper et al., 2020; Hvidberg et al., 2020), those who selected into the platform have somewhat different outcomes, in particular lower crime rates. If we account for the differences between the participants and the population of 18 y-old men by reweighting the observations based on the propensity score estimated on observable characteristics, preferences play an even stronger role than in the main analysis, c.f. column (6) of Appendix Table A-3.6.<sup>1</sup>

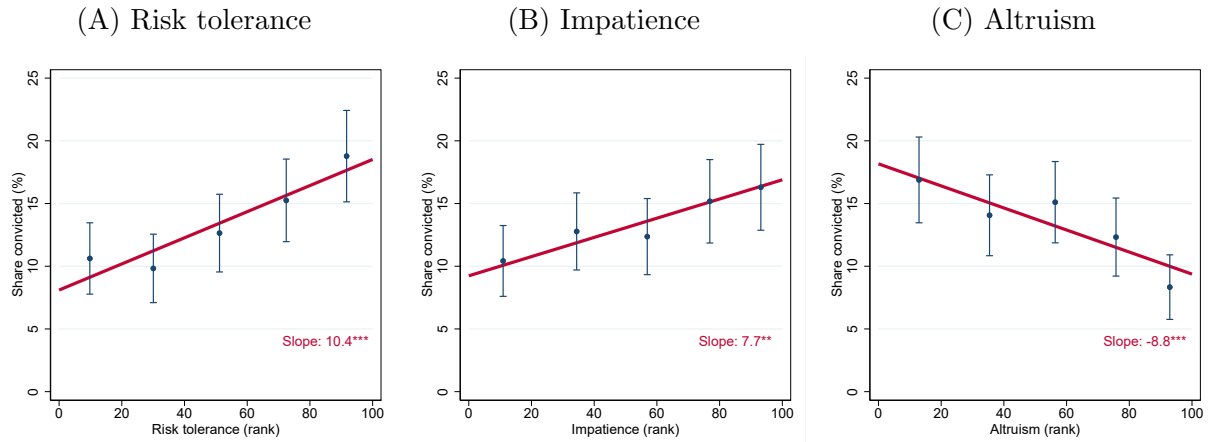
### 3.3 Preferences and Criminal Offences

Figure 3.2 illustrates the relationship between individuals' preference parameters and their propensity to commit crimes. In all diagrams, the vertical axis shows the probability of being convicted of a crime. On the horizontal axis in panel A, we rank individuals according to their level of risk tolerance going from percentile 1 to 100 in the distribution of risk tolerance. Similarly, Panel B ranks individuals according to their level of impatience and panel (c) ranks individuals according to their level of altruism. The diagrams show a strong and almost linear relationship between each of the preference parameters and the propensity to commit crime. All relationships are statistically significant at the 1 percent level of significance.

---

<sup>1</sup>The project was approved by the Danish Data Protection Agency under Agreement 2015-57-0125-0008 and was also approved by Statistics Denmark and the Internal Review Board at the Department of Economics, University of Copenhagen. Data and programs are stored in a separate directory at Statistics Denmark with project number 704856. All empirical analyses were carried out with the software Stata/MP 16.1 using the secure internet interface of Statistics Denmark. Individual-level data is subject to the European Union's General Data Protection Regulation (GDPR). Due to security considerations, the data may not be transferred to computers outside Statistics Denmark. Researchers interested in obtaining access to the data employed in this paper are required to submit a written application to gain approval from Statistics Denmark. Applications can be submitted by researchers who are affiliated with Danish institutions accepted by Statistics Denmark, or by researchers outside of Denmark who collaborate with researchers affiliated with these institutions. We will assist in any way we can with this procedure.

Figure 3.2: Association between preferences and criminal offences



Notes: 95% CIs based on robust standard errors. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Going from the most risk averse individuals to the most risk tolerant individuals is associated with a change in the crime propensity from 8% to 18%, and moving up 10 percentiles in the risk tolerance distribution is associated with a 1 percentage point increase in the crime propensity. The association with impatience is slightly weaker. In this case, moving up 10 percentiles is associated with a 0.8 percentage point increase in the crime propensity. Moving up 10 percentiles in the distribution of altruism is associated with a 0.9 percentage point decrease in the the crime propensity.

Table 3.1 shows results from estimations of multivariate probit models. All three preference parameters are still strongly associated with crime when we move from the bivariate analysis in the graphs to the multivariate analysis in column (1). The table reports the estimated marginal effects of a change in each of the preferences parameters, for given values of the other preference parameters, on the probability of having committed an offence. The marginal effects are of the same magnitude as the slopes in Figure 3.2.

In column (2), we move beyond the strong assumption of full rationality in decision-making by including a measure of self-control. Self-control is known to be strongly associated with crime propensities (Gottfredson and Hirschi, 1990; Pratt and Cullen, 2000; Moffitt et al., 2011; Grasmick et al., 1993). This is also the case in column (2). Importantly, the effects of the preference parameters are nearly unchanged. The effect of impatience is most affected which is consistent with the intuitive idea that impatience and self-control are related concepts.

It is well-known that cognitive skills and socio-economic background are among the strongest predictors of criminal behavior (Herrnstein and Murray, 1994; Hjalmarsson and

Table 3.1: Preferences and probability of being convicted of an offence

	Probability of having a conviction					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk tolerance	10.9*** (2.5)	10.1*** (2.4)	8.3*** (2.4)	8.3*** (2.4)	8.1*** (2.4)	7.9*** (2.4)
Impatience	8.9*** (2.5)	7.6** (2.5)	5.9* (2.5)	5.2* (2.5)	5.2* (2.5)	5.0* (2.4)
Altruism	-7.0** (2.4)	-6.9** (2.4)	-2.7 (2.5)	-2.5 (2.4)	-2.3 (2.4)	-2.5 (2.4)
Self-control		-13.2*** (2.5)	-11.4*** (2.5)	-10.4*** (2.5)	-10.2*** (2.5)	-10.1*** (2.4)
GPA			-16.7*** (2.6)	-14.2*** (2.7)	-12.9*** (2.7)	-13.7*** (2.8)
Parental income					-1.4 (2.7)	-1.6 (3.4)
Convicted parent (=1)					6.3** (2.1)	5.7** (2.2)
Observations	2254	2254	2254	2254	2254	2254
Individual controls				✓	✓	✓
Parental controls						✓

*Notes:* The table reports the marginal effects on the percentage share of respondents committing crime from estimated probit models. *Risk tolerance*, *Impatience*, *Altruism*, *Self-control*, *GPA*, and *Parental income* are all within cohort in sample rank. *Convicted parent* is an indicator. *Individual controls* include regional FE, large city indicator, immigrant and descendant status, a living with both parents indicator, an only child indicator, a first born indicator and an indicator for misreported age or gender in survey. *Parental controls* include educational level, age at child's birth, employment status and unemployment history. Robust standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Lindquist, 2012; Gregory, 2004; Schwartz et al., 2015). One likely reason is that low-skilled individuals and individuals with a poor family background obtain lower wages in the labor market and, therefore, face lower opportunity costs of crime (Chalfin and McCrary, 2017). It can also work through preferences, which are known to be correlated with cognitive skills and socio-economic background (Dohmen et al., 2010). We show in Appendix Table A-3.2 that all three preferences measures are correlated with cognitive skills measured in terms of individuals' GPA at the end of compulsory schooling (age 15-16) but the correlation between altruism and GPA is particularly pronounced: Less skilled individuals are on average less altruistic. This might relate to Strain Theory in Psychology, which hypothesizes that

frustration among lower-class youths leads to increased crime (Levitt and Lochner, 2001; Cloward and Ohlin, 1960).

A key question is whether the preference parameters have any explanatory power beyond cognitive skills and socio-economic background. In column (3) of Table 3.1, we include the individual's percentile rank position in the GPA distribution in the regression. As expected, this is a strong predictor of crime. Moving up ten percentiles in the grade distribution is associated with a 1.67 percentage point decrease in the crime propensity conditional on the other characteristics. Now, altruism is no longer significant, but the effects of risk tolerance and impatience are still large and strongly significant. The effect of moving up one percentile in the risk tolerance distribution or in the impatience distribution corresponds to about a half and one-third of the effect, respectively, of moving up one percentile in the GPA distribution.

In column (4), we include a large set of additional variables that are likely predictive of crime, including region of residence, living in a large city, divorce of parents, immigrant status, having siblings, and birth order (Glaeser and Sacerdote, 1999; Kling et al., 2005; Dustmann and Landersø, 2005; Glaeser et al., 1996, 2003). This reduces the effect of impatience somewhat, but has no impact on the effect of risk tolerance. In column (5), we include the percentile position of parents in the distribution of parental income. We also include information on whether parents have committed crime, which is known to be a strong predictor of criminal offences of sons (Glueck and Glueck, 1950; Gregory, 2004; Hjalmarsson and Lindquist, 2012). The position in the income distribution is, in isolation, strongly correlated with crime (see Appendix Table A-3.2), but not when it is included together with the other variables in column (5). As expected, parental crime is strongly associated with the criminal propensity of sons. Most importantly, the effects of risk tolerance and impatience are almost unchanged when going from column (4) to column (5).

In column (6), we also account for a large set of characteristics of each parent that can be important, including educational background, age at child birth, recent employment status and unemployment history (Lochner and Moretti, 2004). This has nearly no impact on the estimated effects of interest. In total, column (6) includes 55 control variables beyond the three preference parameters of interest (see Appendix Table A-3.3).

Table 3.2: Preferences and probability of being convicted of different offences

	Property offence		Violent, drug or sexual offence	
Risk tolerance	3.21**	(1.06)	2.34	(1.31)
Impatience	2.16*	(0.96)	0.87	(1.38)
Altruism	-1.28	(0.98)	-1.33	(1.36)
Self-control	-1.47	(0.95)	-6.47***	(1.50)
GPA	-5.09***	(1.45)	-5.73***	(1.70)
Parental income	-1.19	(1.42)	-0.53	(1.77)
Convicted parent (=1)	1.18	(0.78)	1.62	(1.26)
Observations	2254		2254	
Mean outcome (%)	2.40		4.21	
Individual controls	✓		✓	
Parental controls	✓		✓	

*Notes:* The results correspond to column (6) in Table 1, but with two different outcomes: probability of having committed a property offence and probability of having committed a violent, drug or sexual offence. Robust standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Overall, school GPA is the strongest predictor of criminal behavior. Across all the specifications that include school GPA, the effect of risk tolerance is significant at the 0.1% level and its size is about a half of the effect of school GPA. The effect of impatience is significant at the 5% level and is about one third of the effect of school GPA.

Table 3.2 reports results from running the analysis separately on property offences and violent, sexual and drug offences using the specification with all controls as in column (6) of Table 3.1. The estimates show that risk tolerance and impatience significantly predict property offences, but do not predict violent, drug and sexual offences. Conversely, self-control does not significantly predict property offences, but is a very strong predictor of violent, drug and sexual offences. These results suggest an intuitive separation of crime with property offences explained well by preferences as hypothesized in the basic choice theory framework while violent, drug and sexual offences are more related to self-control problems.

### 3.4 Additional Results and Robustness Checks

Here we report the conclusions from additional results and robustness checks shown in Appendix Tables A-3.4, A-3.5 and A-3.6. In Appendix Table A-3.4, we include additional behavioral measures that can be computed from the experiments. For example, we compute a measure that indicates whether respondents are present-biased (Laibson, 1997), which relates to the broader measure of self-control problems included in the analysis. This additional measure is insignificant. We also use the behavioral data from our social preference experiment to decompose our measure of altruism into a measure of “behindness aversion” and a measure of “aheadness aversion”. A behindness averse person is willing to give up money in the experiment to reduce the amount the other person receives if this reduces disadvantageous inequality, i.e., the payoff advantage of the other person. An aheadness averse person is willing to increase the other persons’ payoff in the domain of advantageous inequality, i.e., when the decision-maker has a payoff advantage over the other person. An individual can be behindness averse in the domain of disadvantageous inequality while simultaneously being aheadness averse in the domain of advantageous inequality (Fehr and Schmidt, 1999). When we substitute these measures for the overall altruism measure, we find that crime is strongly associated with behindness aversion, but not with aheadness aversion. As in the case with altruism, behindness aversion becomes insignificant when we control for GPA. Finally, we redo the main analysis but rank the participants in the preference distributions according to structural estimates of their preference parameters as often done in the experimental literature. (Fisman et al., 2007; Bruhin et al., 2019). This does not change the results.

Since crime is a low-frequency outcome, we define an individual as criminal if having any criminal offences over a five-year age span. In Appendix Table A-3.5, we look separately at crime offences after the elicitation of preferences and before the elicitation of preferences. We also look at the sub-sample of individuals who have no criminal offences before the elicitation of preferences and then look at whether preferences predict who will commit crime after the elicitation of preferences. The effects become smaller since the crime propensity by construction is lower, but the effects are quite stable across the different samples. Importantly,

the effect of risk tolerance is significant in all cases, larger than the effect of impatience and at least half as large as the effect of GPA.

In Appendix Table A-3.6, we report sensitivity results with respect to changes in variable measurement, sample selection and empirical specification. It shows that the effects of the key preference parameters are unchanged if we allow for more flexible regression specifications with decile indicators for GPA and parental income (column 2) or if we only include math grades in GPA (column 3), which might be a better proxy for differences in cognitive skills. Similarly, the conclusions are the same if we estimate a linear probability model instead of a probit model (column 4) or substitute the percentile rank variables with the corresponding z-scores (column 5). The effects of interest are a little higher if we account for selection into the experiment by using propensity score weighted regressions that account for observable differences between those who participated in the experiment and those who did not (column 6). If we include women in the sample (column 7), who have a crime rate of only 3% compared to 19% for men, then the effect of risk preference falls somewhat, but both risk and time preferences are still significant. In the main analysis, if individual information does not exist on an explanatory variable then this is captured by an indicator variable. If we instead remove respondents altogether when information is missing on one or more variables then this reduces all the main effects of interest somewhat, but the relative magnitudes of risk, impatience, self-control and GPA are unchanged (column 8).

### 3.5 Concluding Remarks

Our results show that differences in preferences predict who commits crime. Risk tolerance, impatience and altruism are all associated with the crime propensity. Impatience and, in particular, risk tolerance strongly predicts crime when we control for an extensive set of background characteristics. The most risk tolerant individuals have a crime propensity that is 8-10 percentage points higher than the least risk tolerant individuals. This effect is half the size of the effect of cognitive skills which is the best predictor of crime.

Evidence on crime levels across countries do not place Denmark as very different from other countries (Enzmann et al., 2010). In that respect, there is no reason to believe that our results on criminal behavior should be unique to Denmark.

One of the criminal justice system's key functions is deterrence of crime. The choice theory of crime implies that increases in the certainty and severity of sanctions deters crime. However, our evidence on the importance of differences in preferences, which are typically assumed to be stable personal traits, suggests that those who commit crimes are also those who are least responsive to policy initiatives that increase certainty and severity of sanctions. This might help explain the somewhat mixed evidence on the effectiveness of sanctions on deterrence (Chalfin and McCrary, 2017).

We also find that variation in preferences in particular predict property offences, but not "crimes of passion" such as violent, sexual and drug offences. Conversely, we find that self-control problems predict crimes of passion, but not property crime. At a broader level, these findings, together with previous results (Kleven et al., 2011; Åkerlund et al., 2016), suggest that the choice theory framework is most appropriate to help understand certain types of crime such as white-collar crime, but not for understanding other types of crime such as violence and sexual assaults where other behavioral parameters appear to be more appropriate.



## References

- Åkerlund, D., B. H. H. Golsteyn, H. Grönqvist, and L. Lindahl (2016). Time discounting and criminal behavior. *Proceedings of the National Academy of Sciences* 113(22), 6160–6165.
- Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal* 100(401), 464–477.
- Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *The Quarterly Journal of Economics* 112(2), 537–579.
- Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76(2), 50.
- Bruhin, A., E. Fehr, and D. Schunk (2019). The many Faces of Human Sociality: Uncovering the Distribution and Stability of Social Preferences. *Journal of the European Economic Association* 17(4), 1025–1069.
- Bruhin, A., H. Fehr-Duda, and T. Epper (2010). Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. *Econometrica* 78(4), 1375–1412.
- Castillo, M., P. J. Ferraro, J. L. Jordan, and R. Petrie (2011). The today and tomorrow of kids: Time preferences and educational outcomes of children. *Journal of Public Economics* 95(11–12), 1377–1385.
- Castillo, M., J. L. Jordan, and R. Petrie (2019). Discount Rates of Children and High School Graduation. *The Economic Journal* 129(619), 1153–1181.
- Chabris, C. F., D. Laibson, C. L. Morris, J. P. Schuldt, and D. Taubinsky (2008). Individual laboratory-measured discount rates predict field behavior. *Journal of Risk and Uncertainty* 37(2–3), 237–269.
- Chalfin, A. and J. McCrary (2017, March). Criminal Deterrence: A Review of the Literature. *Journal of Economic Literature* 55(1), 5–48.
- Cloward, R. A. and L. E. Ohlin (1960). *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York, NY, US: Free Press.
- Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review* 100(3), 1238–1260.
- Draca, M. and S. Machin (2015, August). Crime and Economic Incentives. *Annual Review of Economics* 7(1), 389–408.
- Dustmann, C. and R. Landersø (Forthcoming). Child’s Gender, Young Father’s Crime, and Spillover Effects in Criminal Behavior. *Journal of Political Economy* X(X), X.
- Ehrlich, I. (1973). Participation in Illegitimate Activities: A Theoretical and Empirical Investigation. *Journal of Political Economy* 81(3), 521–565.
- Enzmann, D., I. H. Marshall, M. Killias, J. Junger-Tas, M. Steketee, and B. Gruszczynska (2010, March). Self-reported youth delinquency in Europe and beyond: First results of the Second International Self-Report Delinquency Study in the context of police and victimization data. *European Journal of Criminology* 7(2), 159–183.
- Epper, T., E. Fehr, H. Fehr-Duda, C. T. Kreiner, D. D. Lassen, S. Leth-Petersen, and G. N. Rasmussen (2020). Time discounting and wealth inequality. *American Economic Review* 110(4), 1177–1205.

- Falk, A., S. Meier, and C. Zehnder (2013). Do Lab Experiments Misrepresent Social Preferences? The Case of Self-Selected Student Samples. *Journal of the European Economic Association* 11(4), 839–852.
- Fehr, E., T. Epper, and J. Senn (2020). Other-regarding Preferences and Redistributive Politics. Working Paper No. 339, University of Zurich, Department of Economics, Zurich.
- Fehr, E. and K. M. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics* 114(3), 817–868.
- Fehr, E. and K. M. Schmidt (2006). Chapter 8 The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories. In S.-C. Kolm and J. M. Ythier (Eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*, Volume 1. Elsevier.
- Fisman, R., S. Kariv, and D. Markovits (2007). Individual Preferences for Giving. *American Economic Review* 97(5), 1858–1876.
- Forsythe, R., J. L. Horowitz, N. E. Savin, and M. Sefton (1994). Fairness in Simple Bargaining Experiments. *Games and Economic Behavior* 6(3), 347–369.
- Frederick, S., G. Loewenstein, and T. O'Donoghue (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature* 40(2), 351–401.
- Freeman, R. B. (1999). The economics of crime. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, pp. 3529–3571. Elsevier.
- Garoupa, N. (2003). Behavioral Economic Analysis of Crime: A Critical Review. *European Journal of Law and Economics* 15(1), 5–15.
- Glaeser, E. L. and B. Sacerdote (1999, December). Why is There More Crime in Cities? *Journal of Political Economy* 107(S6), S225–S258.
- Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996). Crime and Social Interactions. *The Quarterly Journal of Economics* 111(2), 506–548.
- Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (2003). The Social Multiplier. *Journal of the European Economic Association* 1(2-3), 345–353.
- Glueck, E. and S. Glueck (1950). *Unraveling Juvenile Delinquency*. Cambridge, MA: Harvard University Press.
- Gneezy, U. and J. Potters (1997). An Experiment on Risk Taking and Evaluation Periods. *The Quarterly Journal of Economics* 112(2), 631–645.
- Gottfredson, M. R. and T. Hirschi (1990). *A General Theory of Crime*. A General Theory of Crime. Stanford University Press.
- Grasmick, H. G., C. R. Tittle, R. J. Bursik Jr., and B. J. Arneklev (1993). Testing the Core Empirical Implications of Gottfredson and Hirschi's General Theory of Crime. *Journal of Research in Crime and Delinquency* 30(1), 5–29.
- Gregory, N. (2004). Crime and the Family: Like Grandfather, Like Father, Like Son? *The British Journal of Forensic Practice* 6(4), 32–36.
- Grogger, J. (1995). The Effect of Arrests on the Employment and Earnings of Young Men. *The Quarterly Journal of Economics* 110(1), 51–71.
- Herrnstein, R. J. and C. Murray (1994). *The Bell Curve: Intelligence and Class Struture in American Life*. New York: Free Press.

- Hjalmarsson, R. and M. J. Lindquist (2012). Like Godfather, Like Son: Exploring the Intergenerational Nature of Crime. *Journal of Human Resources* 47(2), 550–578.
- Hvidberg, K. B., C. T. Kreiner, and S. Stantcheva (2020). Social Position and Fairness Views. NBER working paper 28099, National Bureau of Economic Research, Cambridge, MA, US.
- Kerr, S. P., W. R. Kerr, and M. Dalton (2019). Risk attitudes and personality traits of entrepreneurs and venture team members. *Proceedings of the National Academy of Sciences* 116(36), 17712–17716.
- Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica* 79(3), 651–692.
- Kling, J. R., J. Ludwig, and L. F. Katz (2005). Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment. *The Quarterly Journal of Economics* 120(1), 87–130.
- Kreiner, C. T., S. Leth-Petersen, and L. C. Willerslev-Olsen (2020). Financial Trouble Across Generations: Evidence from the Universe of Personal Loans in Denmark. *The Economic Journal* 130(625), 233–262.
- Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics* 112(2), 443–478.
- Levitt, S. D. and L. Lochner (2001). The Determinants of Juvenile Crime. In J. Gruber (Ed.), *Risky Behavior among Youth: An Economic Perspective*, pp. 327–374. University of Chicago Press.
- Lochner, L. and E. Moretti (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *American Economic Review* 94(1), 155–189.
- Moffitt, T. E., L. Arseneault, D. Belsky, N. Dickson, R. J. Hancox, H. Harrington, R. Houts, R. Poulton, B. W. Roberts, S. Ross, M. R. Sears, W. M. Thomson, and A. Caspi (2011, February). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences* 108(7), 2693–2698.
- Nagin, D. S. (2013, August). Deterrence: A Review of the Evidence by a Criminologist for Economists. *Annual Review of Economics* 5(1), 83–105.
- Noussair, C. N., S. T. Trautmann, and G. van de Kuilen (2014). Higher Order Risk Attitudes, Demographics, and Financial Decisions. *The Review of Economic Studies* 81(1), 325–355.
- Pratt, T. C. and F. T. Cullen (2000). The Empirical Status of Gottfredson and Hirschi’s General Theory of Crime: A Meta-Analysis. *Criminology* 38(3), 931–964.
- Schwartz, J. A., J. Savolainen, M. Aaltonen, M. Merikukka, R. Paananen, and M. Gissler (2015). Intelligence and criminal behavior in a total birth cohort: An examination of functional form, dimensions of intelligence, and the nature of offending. *Intelligence* 51, 109–118.
- Sutter, M., M. G. Kocher, D. Glätzle-Rützler, and S. T. Trautmann (2013). Impatience and Uncertainty: Experimental Decisions Predict Adolescents’ Field Behavior. *American Economic Review* 103(1), 510–531.

### 3.A A Basic Choice Model of Criminal Behavior

Here we illustrate the role of risk preferences and time preferences in a simple two-period model of criminal behavior. We assume the preferences of an individual can be represented by the objective function

$$\Omega(a) = u(c_1(a)) + \beta [pu(c_2^A(a)) + (1-p)u(c_2^B(a))],$$

where  $c_1$  denotes consumption in period 1, depending on whether the individual engage in crime ( $a = 1$ ) or not ( $a = 0$ ), whereas  $p$  is the probability of detection in which case consumption in period 2 becomes  $c_2^A$ . Otherwise consumption in period 2 equals  $c_2^B$ . The parameter  $\beta$  is the weight of an individual on future utility (patience) and  $u(\cdot)$  is a concave period utility function.

The consumption levels equal

$$\begin{aligned} c_1(a) &= (1 + a\gamma)y, \\ c_2^A(a) &= (1 - a\eta)y, \\ c_2^B(a) &= y, \end{aligned}$$

where  $y$  is a fixed income,  $\gamma$  is the gain from criminal activity measured in proportion to the fixed income, and  $\eta$  is the loss if detected measured in proportion to income. Gains and losses can be money/consumption but, more broadly, it can also include non-pecuniary effects that correspond to a given gain or loss in consumption.

It is optimal to engage in criminal activity iff  $\Omega(1) > \Omega(0)$ , which gives the condition

$$u((1 + \gamma)y) + \beta [pu((1 - \eta)y) + (1 - p)u(y)] > (1 + \beta)u(y). \quad (3.1)$$

If it is assumed that the consumption levels associated with criminal activity are well approximated by a second-order Taylor expansions around the non-crime consumption level,  $y$ , then we can write

$$\begin{aligned} u((1 + \gamma)y) &\approx u(y) + u'(y)\gamma y + \frac{1}{2}u''(y)(\gamma y)^2, \\ u((1 - \eta)y) &\approx u(y) - u'(y)\eta y + \frac{1}{2}u''(y)(\eta y)^2. \end{aligned}$$

By substituting these expressions into the inequality (3.1) and simplifying, we obtain the condition

$$\gamma y + \frac{1}{2} \frac{u''(y)}{u'(y)} (\gamma y)^2 - \beta p \eta y + \beta p \frac{1}{2} \frac{u''(y)}{u'(y)} (\eta y)^2 > 0.$$

Using the definition of the coefficient of relative risk aversion,  $\theta \equiv -\frac{u''(y)y}{u'(y)}$ , this may be rewritten as

$$\gamma \left(1 - \frac{1}{2}\theta\gamma\right) > \beta p \eta \left(1 + \frac{1}{2}\theta\eta\right). \quad (3.2)$$

The LHS is the gain of engaging in crime, which is decreasing in the risk aversion parameter  $\theta$  because marginal utility is decreasing. The RHS is the expected loss, which is increasing in the risk aversion parameter and in the patience parameter  $\beta$ .

If we impose two basic regularity conditions,  $0 \leq \theta \leq 2/\gamma$  and  $0 \leq \beta \leq \gamma/(p\eta)$ ,<sup>2</sup> then we have

**Proposition 1** (a) For given values of  $(\eta, \gamma, p, \theta)$  fulfilling  $0 < \theta < \frac{2}{\gamma}$  there exists  $\hat{\beta} \in \left(0, \frac{\gamma}{p\eta}\right)$  such that individuals with  $\beta < \hat{\beta}$  commit a crime, while individuals with  $\beta \geq \hat{\beta}$  do not commit a crime. (b) For given values of  $(\eta, \gamma, p, \beta)$  fulfilling  $0 < \beta < \frac{\gamma}{p\eta}$  there exists  $\hat{\theta} \in \left(0, \frac{2}{\gamma}\right)$  such that individuals with  $\theta < \hat{\theta}$  commit a crime, while individuals with  $\theta \geq \hat{\theta}$  do not commit a crime.

**Proof.** (a) The crime condition [3.2] applies with equality if  $\hat{\beta}(\theta) = \frac{\gamma(1-\frac{1}{2}\theta\gamma)}{p\eta(1+\frac{1}{2}\theta\eta)}$ , which is the  $\beta$ -threshold where an individual is indifferent between committing a crime or not. Since  $\hat{\beta}(0) = \frac{\gamma}{p\eta}$  and  $\hat{\beta}\left(\frac{2}{\gamma}\right) = 0$  it follows that  $\hat{\beta} \in \left(0, \frac{\gamma}{p\eta}\right)$ . (a) The condition [3.2] applies with equality if  $\hat{\theta}(\beta) = \frac{2}{\gamma} \frac{1-\beta p\eta/\gamma}{1+\beta p(\eta/\gamma)^2}$ , which is the  $\theta$ -threshold where an individual is indifferent between doing crime or not. Since  $\hat{\theta}(\beta) = 0$  and  $\hat{\theta}(0) = \frac{2}{\gamma}$  it follows that  $\hat{\theta} \in \left(0, \frac{2}{\gamma}\right)$ .

This result shows that individuals with sufficiently low degree of risk aversion  $\theta$  and patience  $\beta$  are predicted to commit crime.

### 3.B Administrative data

From Statistics Denmark we use the two crime registers for charges (KRSI) and convictions (KRAF). Entries in the two registers can be linked using a file number (`journr3`). From KRSI we use the data of the committed offence (`sig_ger1dto`). Not all convictions have a related charge. In these cases we assume that the offence was committed on the date of the conviction. The convictions in the KRAF register are categorized using seven digit codes in the variable `afg_ger7`. The first digit indicates whether the conviction is related to the criminal code (1), the traffic code (2) or other special laws (3). We discard convictions under the traffic code. The two first digits are used to divide the convictions into sexual offences (11), violent offences (12), property offences (13) and drug offences (32). We use this information to generate indicators for having a conviction for different types of crime. From the register UDFK we have information on grades in lower secondary school. We compute the grade point average across subjects in the final exams. For 63 participants we do not have information on any final exams in 9th or 10th grade. For 43 of these participants, we use self-reported GPA in the survey or their mark for general proficiency where possible. For the remaining 20 participants we assume that they have not passed any final exams and assign them a GPA of 0. Based on the grade point average, we compute the rank/percentile position of each participant in the grade point average distribution of all participants, `gpa_r`. We use ordinal ranking, which ensures a uniform distribution of percentile positions, also when some of the underlying observations are identical. We also compute the GPA only based on the subject math, `gpa_math_r`, which we use in the sensitivity analyses in column 3 of Table A-3.6.

We use the population register (BEF) to identify participants' gender, birth year, immigrant and descendant status, region of residence, and whether their municipality is characterized as a city (according to the "Kommunegrupper" characterization provided by Statistics

<sup>2</sup>If  $\theta > 2/\gamma$  then the (approximated) marginal utility would be negative. If  $\beta > \gamma/(p\eta)$  then criminal activity is never optimal independent of the degree of risk aversion  $\theta$ .

<sup>3</sup>Names written with monospaced typewriter typeface refer to variable names. For the variables from Statistics Denmark's records, these are the names Statistic Denmark has assigned (see definitions here: [www.dst.dk/da/Statistik/dokumentation/Times](http://www.dst.dk/da/Statistik/dokumentation/Times)).

Denmark). We also use information in the BEF register to identify the participants' parents and the variable `fm_mark` to identify whether they live with both parents or not, e.g. if the parents are divorced. Using municipality (`kom`) and address (`bopikom`) in the BEF register we link participants to the household register (HUST). Here we obtain information on the number of children in the household when the participant was 0 years old as an indicator of whether the participant is the first born child, and the number of children in the household when the participant was 10 years old as an indicator for having siblings.

For the parents, we also use the population register to compute their age when the participant was born and the crime register to compute an indicator of whether they have ever committed a crime, excl. traffic offences. Furthermore, we use the income register (IND) to compute the parents' average income in 2015 prices across the years when the participant was 17 to 19 years old. As for the grade point average, we compute the percentile position in the parent income distribution, `income_parents_r`. We use the employment register (IDAP) to compute indicators of parents employment status in 2017 (employed, self-employed, unemployed and not in the work force) and unemployment between 2008 and 2017. We first compute the average share of time they have been unemployed (based on `arledgr`) and compute indicators of not having been unemployed, having been unemployed from 0 to 10% of the time and from 10% to 100% of the time. Finally, we use the education register (UDDA) to compute indicators for the parents' levels of education based on the Danish ISCED classification (primary or lower secondary, upper secondary, short cycle tertiary, bachelor or equivalent, and master, doctoral or equivalent).

For each participant, we predict the probability of participation, i.e. the propensity score, based on a probit model where we include the register data that is available for both participants and non-participants as explanatory variables. We use this for the propensity score weighting in column 6 of Table A-3.6.

### 3.C Experimental data

Based on the data from our incentivized behavioral experiments we compute each individual's time, risk and social preferences. In addition, the experimental data enable us to identify individuals' present bias and their behindness and aheadness aversion which we use for sensitivity and robustness checks.

We use the time experiment to compute the impatience measure, which is the mean share kept across choice situations in Table A-3.7. Based on this, we compute the percentile positions in the impatience distribution for all participants, `impatience_r`, in the same way as we do for grade point average and parental income. Since participants were randomly assigned to either a high or low stake condition, we compute the percentiles separately for the two conditions. We also use the data from the time experiment to compute percentile positions within the distribution of present bias, `present_bi_r`. Here we first compute the difference in the share kept between two situations with the same interest rate but with different timing of payout, e.g. situation 1 and 9 in Table A-3.7. We then compute the mean across the 8 computed differences and rank the participants accordingly. Similarly, we use the risk experiment to compute each participant's risk tolerance percentile position, `risk_toler_r`. Finally, we use the social experiment to compute each participant's altruism percentile position, `altruism_r`. As seen in Table A-3.9, the cost of giving varies across situations, and in particular it is positive in some cases, i.e. it cost money for the participant to increase the others payoff, and negative in others, i.e. the participant gets money for increasing the others payoff. We compute a measure of Aheadness aversion from the situations with a positive cost of giving (i.e. negative slope in the trade-off) and a measure of Behindness aversion

from the situations with a negative cost of giving (i.e. a positive slope in the trade-off). Again, we compute each participant's positions in the distributions, `altruism_pos_r` and `altruism_neg_r`.

As an alternative to the non-parametric preference measures described above, we also estimate structural preference parameters for the three choice domains and use these in the sensitivity analysis in column 5 of Table A-3.4. To estimate the discrete choice models we assume random utility with choice-domain-specific error parameters. More specifically, we back out individual-level parameters from mixed logit models with normally distributed behavioral parameters. As for the non-parametric preference measures, we then compute percentile positions based on the estimated parameter distribution. We implement a hierarchical Bayesian procedure to estimate the mixed logit models. This procedure is described in detail elsewhere (Allenby, 1997; Train, 2001).<sup>4</sup> In the following models,  $i$  denotes the individual and  $j$  denotes the choice situation.

Based on the risk task, we define  $R_{ij}$  as the binary lottery the individual faces as a consequence of the chosen allocation (see section 3.2) and estimate an expected utility model (Morgenstern and Neumann, 1953), which assigns the value  $V$  to the alternative  $R_{ij}$ :

$$V(R_{ij}) = p_j u_i(w_{1ij}) + (1 - p_j) u_i(w_{2ij}),$$

with  $u_i(w_{ij}) = w_{ij}^{1-\rho_i}$  where  $p_j$  is the probability that the good state occurs,  $w_{1ij}$  and  $w_{2ij}$  are the payoffs in the good and the bad state, and  $\rho_i$  denotes the (Arrow-Pratt) coefficient of relative risk aversion. Higher values of  $\rho_i$  indicates comparatively more risk aversion.

Based on the time task, we define  $T_{ij}$  as the income stream the individual faces as a consequence of the chosen allocation and estimate a quasi-hyperbolic discounted utility model (Laibson, 1997),<sup>5</sup> which assigns the value  $V$  to the alternative  $T_{ij}$ :

$$V(T_{ij}) = d(t_{1j}) v_i(w_{1ij}) + d(t_{2j}) v_i(w_{2ij}),$$

with  $v_i(w_{ij}) = w_{ij}^{1-\gamma_i}$ , where  $t_{1j}$  and  $t_{2j}$  denote the payment delays in months relative to the time of the experiment ( $0 \leq t_{1j} < t_{2j}$ ),  $w_{1ij}$  and  $w_{2ij}$  are the payoffs at the earlier and later point in time, and  $\gamma$  denotes an (Arrow-Pratt-type) coefficient of relative aversion towards income fluctuations over time. Finally,  $d(t_{.j}) = 1$  if  $t_{.j} = 0$  and  $d(t_{.j}) = \beta_i e^{-\eta_i \frac{t_{.j}}{12}}$  otherwise, where  $\eta_i \geq 0$  denotes the (annualized) rate of time preference and  $\beta_i$  present bias (for  $\beta_i < 1$ ) or future bias (for  $\beta_i > 1$ ).

Based on the social task, we define  $S_{ij}$  as the interpersonal distribution the individual faces as a consequence of the chosen allocation and estimate an inequality aversion model (Fehr and Schmidt, 1999), which assigns the value  $V$  to the alternative  $S_{ij}$ :

$$V(S_{ij}) = w_{\text{own}j} - \alpha_i (w_{\text{other}j} - w_{\text{own}j}) \mathbb{1}[w_{\text{own}j} < w_{\text{other}j}] - \beta_i (w_{\text{own}j} - w_{\text{other}j}) \mathbb{1}[w_{\text{own}j} > w_{\text{other}j}],$$

where  $w_{\text{own}j}$  is the payoff to the participant,  $w_{\text{other}j}$  is the payoff to the other person and  $\alpha_i$  and  $\beta_i$  denote behindness aversion and aheadness aversion, respectively.

Participants in the experiment also responded to a few survey questions: self-reported gender, self-reported birth year, self-reported GPA and self-reported self-control. For the last variable the question asked was “I am good at exercising self-control in my actions and

<sup>4</sup>We assume uninformative prior distributions. In addition, we adopt appropriate transformations of the distributions to ensure that the estimated parameters lie within the range of their theoretical support.

<sup>5</sup>Since the individuals were assigned to a high or low stake condition, we estimate separate models for the two conditions and obtain condition-specific rankings of individuals to avoid the ordering being confounded by the magnitude effect (see e.g. Frederick et al. (2002)).

decisions” and participants answered this question on a 7 point Likert scale where 1 was “strongly disagree” and 7 was “strongly agree”. Using the same procedure as for the other explanatory variables, we compute the percentile position in the self-control distribution, **selfcontrol\_r**. For all the percentile position variables we also compute the z-scores based on the raw variables where we subtract the mean across all participant and divide by the standard deviation. We use this in Table A-3.6.



### 3.D Figures

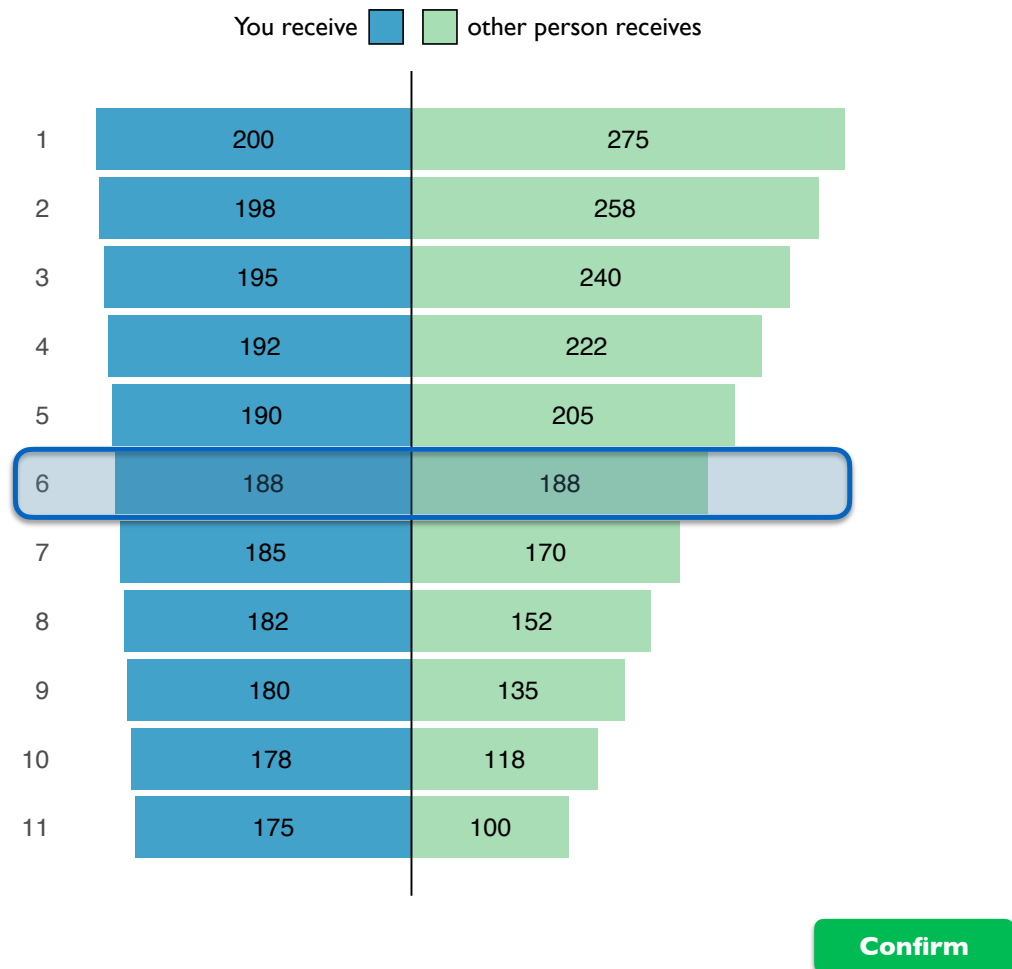


Figure A-3.1: Screenshot of online experiment eliciting social preferences

*Notes:* This is an example of one of the choice situations (Situation 9 in Table A-3.9). The subject can choose one out of eleven payoff allocations where the blue bars to the left show the money received by the subject, while the green bars to the right show the money received by a randomly assigned other person. The individual has chosen option number six, which gives DKK 188 to each person.

### 3.E Tables

Table A-3.1: Summary statistics

	Sample (1)	Logins (2)	Invited (3)	Difference (1) - (3) (4)	P-value (%) (5)
Crime age 15-20 (%)	13.4	14.0	19.0	-5.6	0.0
GPA	7.5	7.4	6.6	0.9	0.0
Immigrant (%)	3.2	3.4	5.2	-2.0	0.0
Descendant (%)	6.6	7.0	7.9	-1.3	0.3
Northern Jutland (%)	9.8	9.9	10.5	-0.7	16.3
Middle Jutland (%)	24.6	24.3	23.7	0.9	21.0
Southern Denmark (%)	25.4	25.2	23.1	2.3	0.2
Copenhagen (%)	26.4	26.2	28.0	-1.7	3.1
Sealand (%)	13.8	14.4	14.7	-0.9	13.1
Capitol municipality (%)	21.9	21.7	23.9	-2.0	0.6
Large city municipality (%)	11.9	11.8	11.0	0.9	12.8
Small city municipality (%)	25.7	25.2	24.3	1.4	6.4
Hinterland municipality (%)	19.0	19.1	17.8	1.2	8.1
Rural municipality (%)	21.5	22.2	23.0	-1.5	3.9
First born (%)	37.9	37.7	36.4	1.5	7.6
Only child (%)	10.8	10.3	10.2	0.6	26.4
Lives w. both parents (%)	60.1	58.9	55.0	5.1	0.0
Father's inc. (1.000 DKK)	565.9	552.4	542.3	23.6	5.0
Mother's inc. (1.000 DKK)	399.4	394.1	386.8	12.6	0.7
Mother's age at birth	30.2	30.1	29.8	0.4	0.0
Father's age at birth	32.7	32.6	32.5	0.2	8.0
Missing father info (%)	2.3	2.6	3.3	-1.0	0.0
Missing mother info (%)	0.6	0.8	1.7	-1.1	0.0
Observations	2254	2650	7054		

*Notes:* In column (1), the *Sample* consists of the respondents who completed all the experiments on the online platform and are used in the analysis. In column (2), *Logins* are everyone who logged into the online platform. Column (3) shows descriptives for the random sample of 18 year old men who were invited to participate. Column (4) shows the difference between participants used in the analysis and everyone who was invited (incl. participants). Column (5) shows the P-values of the differences in column (4). The P-values are calculated using partially overlapping samples t-test with Welch's degrees of freedom (Derrick et al., 2017).

Table A-3.2: Pairwise correlations of regressors and outcome

	Convicted	Risk	Impatience	Altruism	Self-control	GPA	Income	Conv. parent
Convicted	1.00	0.09	0.07	-0.07	-0.12	-0.17	-0.10	0.10
Risk tolerance	0.09	1.00	-0.12	-0.11	-0.04	-0.13	-0.05	0.04
Impatience	0.07	-0.12	1.00	-0.07	-0.11	-0.12	-0.07	0.01
Altruism	-0.07	-0.11	-0.07	1.00	0.01	0.27	0.08	-0.06
Self-control	-0.12	-0.04	-0.11	0.01	1.00	0.12	0.08	-0.04
GPA	-0.17	-0.13	-0.12	0.27	0.12	1.00	0.34	-0.12
Parental income	-0.10	-0.05	-0.07	0.08	0.08	0.34	1.00	-0.08
Convicted parent (=1)	0.10	0.04	0.01	-0.06	-0.04	-0.12	-0.08	1.00

*Notes:* The table shows pairwise correlations between the different key regressors in the analysis and between each regressor and the outcome (convicted).

Table A-3.3: Economic preferences and probability of being convicted of an offence

	Probability of having a conviction	
Risk tolerance	7.87***	(2.37)
Impatience	4.97*	(2.44)
Altruism	-2.46	(2.42)
Self-control	-10.06***	(2.43)
GPA	-13.75***	(2.76)
Parental income	-1.57	(3.37)
Convicted parent (=1)	5.69**	(2.21)
Geography (=1)		
Middle Jutland	1.33	(2.39)
Southern Denmark	0.21	(2.35)
Copenhagen	3.26	(2.64)
Sealand	1.19	(2.74)
Urban area	-1.42	(1.75)
Family and background (=1)		
Immigrant	-2.50	(5.06)
Descendant	5.29*	(2.69)
Lives with both parents	-3.59*	(1.49)
Only child	1.67	(2.23)
First born	-4.62**	(1.60)
Misreport age or gender (=1)	3.53	(4.14)
Missing father information	-8.85	(9.05)
Father's age at birth (=1)		
24-25	-2.06	(4.87)
26-27	-2.30	(4.40)
28-29	-4.95	(4.46)
30-31	-2.35	(4.38)
32-33	-4.37	(4.43)
34-35	-6.65	(4.58)
36-37	-8.51	(4.76)
38-39	-6.96	(4.98)
>39	-8.70	(4.91)
Father's educational level (=1)		
Upper secondary	-0.30	(2.04)
Short cycle tertiary	0.64	(3.33)
Bachelor or equivalent	2.75	(2.71)
Master, Doctoral or equivalent	2.19	(3.02)
Missing	-3.82	(5.66)
Father's employment status (=1)		
Self-employed	2.23	(2.64)
Unemployed	-12.13	(6.90)
Not in the workforce	4.46*	(2.27)
Missing	1.49	(3.30)
Father's avg. unemployment 10 yr. (=1)		
1-10%	0.06	(1.90)
11-100%	-0.24	(2.44)
Missing mother information	1.89	(11.44)
Mother's age at birth (=1)		
24-25	1.78	(3.55)
26-27	2.49	(3.23)
28-29	4.70	(3.25)
30-31	3.97	(3.41)
32-33	2.59	(3.52)
34-35	2.76	(3.63)
36-37	1.82	(4.18)
38-39	5.19	(4.94)
>39	8.51	(5.62)
Upper secondary	-1.58	(2.18)
Mother's educational level (=1)		
Short cycle tertiary	-4.25	(4.12)
Bachelor or equivalent	-0.40	(2.49)
Master, Doctoral or equivalent	2.44	(3.17)
Missing	4.27	(5.93)
Mother's employment status (=1)		
Self-employed	3.32	(3.30)
Unemployed	-10.22	(5.89)
Not in the workforce	1.73	(2.06)
Missing	0.59	(5.43)
Mother's avg. unemployment 10 yr. (=1)		
1-10%	1.42	(1.84)
11-100%	4.83*	(2.08)
Observations	2254	

Notes: The table reports the marginal effects in p.p. from the same probit model in column (6) of Table 3.1. Robust standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-3.4: Economic preferences and probability of being convicted of an offence

	Probability of having a conviction				
	(1)	(2)	(3)	(4)	(5)
Present bias	2.0 (2.4)			0.8 (2.3)	-3.5 (2.7)
Aheadness aversion		-1.8 (2.4)		-1.8 (2.3)	-2.3 (2.6)
Behindness aversion			10.9*** (2.5)	1.4 (2.7)	3.0 (2.7)
Risk tolerance				7.9*** (2.4)	7.9** (2.4)
Impatience				4.9* (2.4)	6.2* (2.7)
Self-control				-10.0*** (2.4)	-10.3*** (2.4)
GPA				-13.9*** (2.8)	-13.4*** (2.8)
Parental income				-1.5 (3.4)	-2.1 (3.4)
Convicted parent (=1)				5.7** (2.2)	5.5* (2.2)
Observations	2254	2254	2254	2254	2254
Individual controls				✓	✓
Parental controls				✓	✓
Structural prefs.					✓

*Notes:* The table reports the marginal effects in p.p. from estimated probit models. *Present bias*, *Behindness aversion*, *Aheadness aversion*, *Risk tolerance*, *Impatience*, *Self-control*, *GPA*, and *Parental income* are all within cohort in sample ranks. *Convicted parent* is an indicator. *Individual controls* include regional FE, an urban area indicator, immigrant and descendant status, a living with both parents indicator, an only child indicator, a first born indicator and an indicator for misreported age or gender in survey. *Parental controls* include educational level, age at child's birth, employment status and unemployment history. *Structural prefs.* indicates that we use structurally estimated preference measures in column (5) instead of the non-parametric measures. Robust standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-3.5: Economic preferences and probability of being convicted of an offence committed before or after the preference elicitation

	Age 15-18		Age 19-20		19-20 (no crime 15-18)	
	(1)	(2)	(3)	(4)	(5)	(6)
Risk tolerance	5.8*** (1.7)	4.6** (1.6)	7.4*** (2.2)	4.9* (2.0)	6.0** (2.1)	4.5* (2.0)
Impatience	5.3** (1.7)	3.6* (1.7)	5.5* (2.1)	2.2 (2.0)	4.4* (2.1)	1.9 (2.0)
Altruism	-4.1* (1.7)	-1.6 (1.6)	-5.3* (2.1)	-2.4 (2.1)	-3.5 (2.0)	-1.4 (2.0)
Self-control		-2.8 (1.6)		-8.6*** (2.1)		-8.1*** (2.1)
GPA		-9.0*** (2.0)		-9.7*** (2.4)		-6.6** (2.3)
Parental income		-2.7 (2.4)		-0.3 (2.8)		0.8 (2.7)
Convicted parent (=1)		1.5 (1.4)		4.7* (1.9)		4.8** (1.8)
<i>N</i>	2254	2254	2254	2254	2125	2125
Mean outcome (%)	5.7	5.7	9.3	9.3	8.1	8.1
Controls		✓		✓		✓

*Notes:* The table reports the marginal effects in p.p. from estimated probit models. In columns 1-2, we consider whether individuals committed a crime at age 15-18. In columns 3-4, we consider whether individuals committed a crime at age 19-20. In columns 5-6, we consider whether individuals committed a crime at age 19-20 conditional on not having committed a crime at age 15-18. *Risk tolerance*, *Impatience*, *Altruism*, *Self-control*, *GPA*, and *Parental income* are all within cohort in sample rank. *Convicted parent* is an indicator. *Controls* include regional FE, an urban area indicator, immigrant and descendant status, a living with both parents indicator, an only child indicator, a first born indicator, an indicator for misreported age or gender in survey, parents' educational level, parents' age at child's birth, parents' employment status and unemployment history. Robust standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-3.6: Economic preferences and probability of being convicted of an offence committed at age 15-20

	Baseline	Flexible	Math GPA	LPM	z-scores	Weighted	Incl. women	All information
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Risk tolerance	7.9*** (2.4)	7.4** (2.4)	7.9*** (2.4)	8.1** (2.5)	2.3*** (0.7)	9.1** (3.0)	4.7*** (1.4)	6.8** (2.5)
Impatience	5.0* (2.4)	5.3* (2.4)	5.1* (2.4)	4.7* (2.4)	1.6* (0.7)	7.0* (3.0)	5.4*** (1.5)	4.5 (2.5)
Altruism	-2.5 (2.4)	-2.5 (2.4)	-2.3 (2.4)	-2.4 (2.5)	-0.9 (0.7)	-1.0 (3.0)	0.2 (1.4)	-2.8 (2.5)
Self-control	-10.1*** (2.4)	Category indicators	-9.9*** (2.5)	-10.2*** (2.6)	-2.9*** (0.6)	-11.3*** (3.1)	-7.0*** (1.4)	-8.0** (2.6)
GPA	-13.7*** (2.8)	Decile indicators	-12.5*** (2.6)	-13.1*** (2.7)	-3.8*** (0.7)	-17.2*** (3.7)	-12.2*** (1.6)	-12.0*** (2.9)
Parental income	-1.6 (3.4)	Decile indicators	-1.7 (3.4)	-2.3 (3.5)	-0.5 (0.5)	-3.5 (4.0)	0.0 (1.9)	-1.5 (3.6)
Convicted parent (=1)	5.7** (2.2)	6.1** (2.2)	5.6* (2.2)	7.8* (3.1)	5.5* (2.2)	7.2** (2.7)	3.4** (1.2)	5.0* (2.3)
Observations	2254	2254	2254	2254	2254	2254	4503	2011
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Parental controls	✓	✓	✓	✓	✓	✓	✓	✓

*Notes:* Column (1) is the baseline result from the Table 3.1 in the main text. Column (2) includes flexible controls for GPA and parental income, in the form of dummies for each decile, instead of continuous measures. Column (3) uses GPA only for math. Column (4) shows OLS estimates in p.p. from a linear probability model. Column (5) uses z-scores for the explanatory variables instead of ranks. Column (6) weighs the observations with the inverse probability of being in the sample using all explanatory variables that are also available for non-participants. Column (7) includes women born in 1999. Column (8) excludes participants for whom we do not observe all information in the administrative data, most importantly GPA and parental information. *Risk tolerance*, *Impatience*, *Altruism*, *Self-control*, *GPA*, and *Parental income* are all within cohort in sample ranks. *Convicted parent* is an indicator. *Controls* include regional FE, an urban area indicator, immigrant and descendant status, a living with both parents indicator, an only child indicator, a first born indicator, an indicator for misreported age or gender in survey, parents' educational level, parents' age at child's birth, parents' employment status and unemployment history. Robust standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Table A-3.7: Choice situations in the time experiment

Situation	Low stakes		High stakes		$t_1$	$t_2$	Rate
	$x_1$	$x_2$	$x_1$	$x_2$			
1	250	251	7500	7530	0	2	0.024
2	250	256	7500	7680	0	2	0.153
3	250	261	7500	7830	0	2	0.295
4	250	266	7500	7980	0	2	0.451
5	250	271	7500	8130	0	2	0.622
6	250	276	7500	8280	0	2	0.811
7	250	281	7500	8430	0	2	1.016
8	250	286	7500	8580	0	2	1.242
9	250	251	7500	7530	2	4	0.024
10	250	256	7500	7680	2	4	0.153
11	250	261	7500	7830	2	4	0.295
12	250	266	7500	7980	2	4	0.451
13	250	271	7500	8130	2	4	0.622
14	250	276	7500	8280	2	4	0.811
15	250	281	7500	8430	2	4	1.016
16	250	286	7500	8580	2	4	1.242

*Notes:*  $x_1$  is the amount the participant can get paid out sooner and  $x_2$  is the amount the participant can get paid out later. They differ by whether the participant was assigned to the low or high stake treatment.  $t_1$  indicates the sooner payout time (either within 24 hours (0) or in 2 months) while  $t_2$  indicates the later payout time (either in 2 months or in 4 months). The user interface displayed the delays in weeks to avoid confounds by payments at different weekdays. Rate is the annualized interest rate the participant gets on the amount saved for two months. For instance, in Situation 1, Rate=0.024 refers to a yearly interest rate of 2.4%. Panel (a) of Figure 1 in the main text illustrates situation 10 with low stakes.



Table A-3.8: Choice situations in the risk experiment

Situation	DKK	$p$	Good	Bad
1	250	0.5	1.21	0.81
2	250	0.2	1.41	0.91
3	250	0.8	1.11	0.61
4	250	0.5	1.31	0.71
5	250	0.2	1.61	0.86
6	250	0.8	1.16	0.41
7	250	0.5	1.35	0.75
8	250	0.2	1.65	0.90
9	250	0.8	1.20	0.45
10	250	0.6	1.50	0.40
11	250	0.4	1.72	0.62
12	250	0.6	1.45	0.35
13	250	0.4	1.67	0.57
14	250	0.5	1.51	0.50
15	250	0.5	1.61	0.60

*Notes:* *DKK* is the amount the participant can keep or invest in the lottery.  $p$  is the probability that the lottery will give the good state. *Good* is the multiplier of the investment in the good state while *Bad* is the multiplier of the investment in the bad state. Panel (b) of Figure 1 in the main text illustrates situation 10.

Table A-3.9: Choice situations in the social experiment

Situation	$own_1$	$other_1$	$own_2$	$other_2$	Cost of giving
1	262.5	137.5	112.5	237.5	1.500
2	250.0	125.0	125.0	250.0	1.000
3	237.5	112.5	137.5	262.5	0.667
4	225.0	112.5	150.0	262.5	0.500
5	212.5	112.5	162.5	262.5	0.333
6	212.5	100.0	162.5	275.0	0.286
7	200.0	100.0	175.0	275.0	0.143
8	187.5	100.0	187.5	275.0	-0.000
9	175.0	100.0	200.0	275.0	-0.143
10	175.0	112.5	200.0	262.5	-0.167
11	162.5	100.0	212.5	275.0	-0.286
12	162.5	112.5	212.5	262.5	-0.333
13	187.5	187.5	200.0	250.0	-0.200
14	212.5	112.5	187.5	187.5	0.333
15	187.5	187.5	212.5	262.5	-0.333
16	250.0	125.0	187.5	187.5	1.000
17	187.5	187.5	225.0	275.0	-0.429
18	262.5	162.5	187.5	187.5	3.000
19	187.5	187.5	192.5	292.5	-0.048
20	192.5	92.5	187.5	187.5	0.053

*Notes:*  $own_1$  is the amount the participant gets if he/she gives the smallest possible amount to the other person.  $other_1$  is the smallest possible amount to give.  $own_2$  is the amount the participant gets if he/she gives the largest possible amount to the other person.  $other_2$  is the largest possible amount to give. Thus, for a given choice situation ( $other_1$ ,  $other_1$ ) and ( $other_2$ ,  $other_2$ ) represent the most extreme allocations in the set of feasible payoff allocations. In every choice situation there were 9 further feasible payoff allocations located between the extremes in an equi-distant way. *Cost of giving* denotes the cost of the participant per DKK given to the other person. A negative value means that the participant benefit per DKK given to the other person.

## References

- Allenby, G. (1997). An Introduction to Hierarchical Bayesian Modeling. Tutorial Notes, Advanced Research Technique Forum, American Marketing Association.
- Derrick, B., B. Russ, D. Toher, and P. White (2017). Test statistics for the comparison of means for two samples that include both paired and independent observations. *Journal of Modern Applied Statistical Methods* 16(1), 137–157.
- Fehr, E. and K. M. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics* 114(3), 817–868.
- Frederick, S., G. Loewenstein, and T. O'Donoghue (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature* 40(2), 351–401.
- Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics* 112(2), 443–478.
- Morgenstern, O. and J. V. Neumann (1953). *Theory of Games and Economic Behavior*. Princeton university press.
- Train, K. (2001). A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit. Working Paper, Department of Economics, University of California, Berkeley.