Essays in Health and Labor Economics

DISSERTATION
of the University of St.Gallen
School of Management,
Economics, Law, Social Sciences,
International Affairs and Computer Science,
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Patrick Michel Chuard-Keller

from

Oberbüren (St.Gallen)

Approved on the application of

Prof. Dr. Dr. hc. Monika Bütler

and

Prof. Dr. Rafael Lalive

Dissertation no. 5208

D-Druck Spescha, St. Gallen 2022

Essays in Health and Labor Economics

DISSERTATION
of the University of St.Gallen
School of Management,
Economics, Law, Social Sciences,
International Affairs and Computer Science,
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Patrick Michel Chuard-Keller

from

Oberbüren (St.Gallen)

Approved on the application of

Prof. Dr. Dr. hc. Monika Bütler

and

Prof. Dr. Rafael Lalive
Prof. Dr. Michael Lechner

Dissertation no. 5208

D-Druck Spescha, St. Gallen 2022

The University of St.Gallen, School of Management, Economics, Law, Social Sciences, International Affairs and Computer Science, hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St.Gallen, November 8, 2021

The President:

Prof. Dr. Bernhard Ehrenzeller

Acknowledgments

Frankly, writing a dissertation is an «ego project». It involves many hours in front of a screen: thinking, coding, writing—mostly alone. This is even more true when writing a dissertation during a pandemic. However, I would not have written a single word without the help, wisdom, and understanding of others.

Foremost, I am deeply indebted to my supervisor, Monika Bütler. This is not only because you gave me the time and freedom to conduct research on any topic I was interested in, but also because you inspired me with your attitude towards identifying, tackling, and explaining relevant social problems. I am extremely honored to have been supervised by such a renowned economist.

I would like to extend my sincere thanks to my external supervisor, Rafael Lalive from the University of Lausanne, for reviewing my thesis and for his useful suggestions and supportive feedback at my pre-defense and public defense. Also, I wish to thank Beatrix Eugster, head of the thesis committee, for her supportive feedback and for the private lessons in «how to keep calm in academia».

I would also like to thank some of my former employers for paving the way towards a PhD. Therefore, special thanks to Manuel Wälti, my former superior at the Swiss National Bank. Your excitement and dedication to all the different facets of economics was highly contagious. My time at «Iconomix » was insightful and fun, even though the prize for the best picture on the Flipchart of Fame went to my friend Arno Gartmann... Admittedly, it is hard to top a picture of dancing with the president. Also, thanks to my superior at the Zurich University of Applied Sciences, Simon Wieser, for introducing me to applied research and teaching me the importance of having a «signature color codes» in graph.

Although COVID-19 put severe restrictions on time spent with peers, I always enjoyed the scientific exchange and—let's be honest—foremost the distraction, when available. Many thanks to my top-of-the-world buddies, Gabriel Okasa, aka Mr. Stylo, and Georgi

Lautliev, aka Mr. Legend. Also, many thanks to Veronica Schmiedgen-Grassi from Lugano, my co-author of the intergenerational mobility paper—not only, but also—for sticking to my color scheme. I very much enjoyed working with you. I am still impressed by how well structured a do-file can be, and still a bit a jealous of the palm tree behind you on your webcam. Thanks also to my chair and office companion, Nadia Myohl, for always having a lot of chitchats ready and for successfully distracting me with exceptionally boring topics before my public defense. Also, I would like to thank Aurélien Sallin for philosophical inspiration and for giving me «Arbi junior» the sourdough—sadly, I must confess now, he did not make it. RIP Arbi junior.

Thanks also to my co-author Giuseppe Sorrenti for putting me on the multigrading research project when all other research was put on hold. Although that paper is not in this thesis—mainly to spare my thesis committee another paper—I have learned a great deal while working with you. Apart from research related stuff... also that one can confidently wear a Mickey Mouse shirt as a professor.

With a little help from my friends, is not only a great song but especially true when doing a PhD. Many thanks to my talented long-term friend and singer Katrin Corazza. Thanks not only for the delicious food in Käti's Mittagsstübli next to my office, but also for letting my guitar accompany your voice in the last two decades and allowing me to forget about identification strategies when performing with you. Also, a special thanks to my friend Remo Messerli, aka Dr. Louis Lustig, for the weekly summary of the «Making Sense» podcast, the great discussions at the Table of Truth, and mostly for the delicious Quittengelee from Belchenhof (In case you read this, it is empty again. Just saying...). And a big Dankeschön to my friend Stefan Rauber, aka Paul Boss, for the legendary times in the Fischerstube, and for giving me the privilege of being a godfather to your daughter.

Also, many thanks to my family, especially to my parents Irene and Jörg Keller. You always supported me in my academic endeavors, even though you rarely knew what I was doing—admittedly, neither did I sometimes. Thanks also to my sister Alexandra Keller for letting me improve my teaching skills by playing «Schüeleris» as children—even if you did not want to play it. But you know, that's school.

One of my papers analyzes grandparental effects on grandchild outcomes, and although there is no Control-Patrick, I am pretty sure that the Treatment-Patrick benefitted greatly from his grandparents. Thus, I am highly grateful to Marianne and Robert Weber, who originally convinced me—with a little help from Dolcetto—to go to high school, and helped to increase overall educational mobility in Switzerland. Many thanks to my grandmother, Lore Weber, for teaching me perseverance, and to Jda Keller for her cheerful mood and for encouraging me to persist.

Thanks also to my «in-laws»: Merci beaucoup Marc Chuard, for making space for this book on your shelf and for driving me, at brief notice and scarcely dressed, to Bern. And thank you, Nathalie Chuard, aka PSIL, for your contagious super mood and for the (hopefully ongoing) supply of chocolate. And thank you Marion Chuard, for your positive and inspiring energy during my PhD and the delicious food that you always bring when visiting us.

The biggest thanks of all goes to my amazing wife and co-author, Caroline Chuard. Thank you for not only encouraging me to do a PhD, but also for the intellectual support during the booms—and even more so for the mental support during the busts—of the business cycle of motivation. I do not want to imagine how the experience would have been without your help, understanding, and wisdom. Asking you to marry me was by far the best idea I have ever had in my PhD years. Thank you for saying yes.

Winterthur, December 2021
Patrick Chuard-Keller

Contents

Abstract

This thesis conducts empirical analyses in the field of labor and health economics. It can be divided into two parts: The first part studies the effect of policies on health outcomes; the second part analyzes social mobility in Switzerland. More specifically, the first part investigates the effect of a birth allowance on newborn health, fertility, and birth scheduling. We find that introducing such a «baby bonus» leads to a sizable but only temporary increase in the fertility rate and a small, but permanent increase in the birth weight of the newborn. However, we do not find that introducing such a birth allowance leads to birth scheduling. The other chapter on health effects of policies studies the effect of introducing early public retirement on male mortality. There is a strong and significant increase in the mortality rate for the cohort that is eligible for two years of early retirement. Looking at the underlying mortality causes of those men shows that there is a discontinuous jump in deaths related to alcohol dependence. Also, the effect is highest for single men. This points at potential issues with a loss of structure associated with early retirement. The first chapter in the social mobility part studies intergenerational income and educational mobility in Switzerland. The results show that income mobility is high, while educational mobility is low. We argue that this divergence is linked to the vocational education and training (VET) system, which provides good wage outcomes at relatively little formal education. We also find that educational tracks starting with VET and adding additional education contribute a lot to upward mobility. This is likely because VET reduces the impact of parental credit constraints on a child's human capital acquisition. The next chapter in the social mobility part analyzes income and educational mobility over three generations. Linking income and education of a child to the one of parents and grandparents shows that economic advantages in terms of income decay at a geometric rate. Thus, even strong family privileges disappear relatively quickly. However, the persistence of educational inequality over generations is much stronger. The last chapter combines income social mobility with policy analysis. It studies the effect of cross-border immigration on income mobility of natives. The results show that the removal of restrictions for cross-border workers in Switzerland led to a decrease in the wage for children from low-income parents in regions more affected by cross-border immigration. This happens mechanically because children from low-income parents more often learn occupations and choose occupational tracks more affected by the influx of immigrants.

Zusammenfassung

Diese Dissertation beinhaltet empirische Analysen im Bereich der Arbeits- und Gesundheitsökonomie. Die Arbeit lässt sich in zwei Teile gliedern: Der erste Teil befasst sich mit gesundheitlichen Auswirkungen von Politikmassnahmen. Der zweite Teil analysiert die soziale Mobilität in der Schweiz. Das erste Kapitel untersucht, ob eine Geburtszulage die Fertilität, die Gesundheit von Neugeborenen oder den Geburtstermin beeinflusst. Wir stellen fest, dass die Einführung zu einer beträchtlichen, aber nur vorübergehenden Erhöhung der Geburtenrate und einer kleinen, aber dauerhaften Erhöhung des Geburtsgewichts von Neugeborenen führt. Das zweite Kapitel untersucht, ob die Einführung von Frühpensionierung in der öffentlichen Altersvorsorge die Mortalität von Männern beeinflusst. Die Resultate zeigen, dass die Kohorte mit der Möglichkeit sich zwei Jahre früher pensionieren zu lassen eine höhere Mortalität aufweist. Ein Blick auf deren Todesursachen zeigt, dass Todesfälle in Zusammenhang mit Alkoholabhängigkeit ansteigen. Zudem ist der Effekt bei alleinstehenden Männern am ausgeprägtesten. Dies weist auf mögliche Probleme mit einem Strukturverlust im Zusammenhang mit der Frühpensionierung hin. Das erste Kapitel des zweiten Teils beschreibt die intergenerationelle Einkommens- und Bildungsmobilität in der Schweiz. Die Einkommensmobilität ist hoch, obwohl die Bildungsmobilität tief ist. Diese Diskrepanz hängt wahrscheinlich mit dem Berufsbildungssystem zusammen, welches gute Lohnergebnisse bei wenig formaler Bildung bietet. Wir stellen fest, dass Bildungswege, die mit einer Berufsbildung beginnen und einer zusätzlichen Ausbildung enden, die Aufwärtsmobilität in der Schweiz treiben. Dies ist darauf zurückzuführen, dass die Berufsbildung die Auswirkungen von finanziellen Restriktionen der Eltern auf den Humankapitalerwerb eines Kindes verringert. Das nächste Kapitel im Teil der sozialen Mobilität analysiert die Einkommens- und Bildungsmobilität über drei Generationen. Es zeigt sich, dass Vorteile in Bezug auf Einkommen mit geometrischer Geschwindigkeit abnehmen und so nach einigen Generation verschwinden. Vorteile in Bezug auf Bildung verschwinden aber wesentlich langsamer. Das letzte Kapitel untersucht die Auswirkungen von Grenzgängern auf die Einkommensmobilität von Einheimischen. Die Ergebnisse zeigen, dass die Aufhebung der Beschränkungen für Grenzgänger in der Schweiz zu einem Lohnrückgang für Kinder von einkommensschwachen Eltern führte. Dies geschieht deshalb, weil Kinder von einkommensschwachen Eltern häufiger Berufe erlernen und Bildungswege wählen, die stärker vom Zuzug von Einwanderern betroffen sind.

Chapter 1

Introduction

This thesis contributes to two main strands in the empirical health and labor economics literature. The first part analyzes how economic policy reforms can affect health. Does a birth allowance increase fertility or increase newborn health? Does allowing for early retirement affect life expectancy of men? The second part analyzes equality of opportunity. More specifically, it sheds light on the state of intergenerational income and education mobility in Switzerland over one and two generations. Furthermore, it analyzes how a large policy change, the abolition of restriction for cross-border workers, affects intergenerational mobility of natives.

Chapter 2, written jointly with Caroline Chuard, analyzes the impact of birth allowances on fertility, newborn health, and birth scheduling in Switzerland. We exploit variation in time and space of introducing such birth allowances: 11 out of 26 cantons introduced a baby bonus during the last 50 years at different points in time. In addition, we study whether the introduction led to birth scheduling, that is, shifting of the birth with the intention to receive the birth allowance. We do not find evidence for birth-scheduling. However, we discover a sizable but only temporary increase in the fertility rate of 5.5 percent and a permanent but diminishing increase in the birth weight of 2.8 percent. Interestingly, we document substantial heterogeneity by citizenships of mothers.

Chapter 3 studies how early retirement affects health of men. Studying the effect of retirement on health is difficult since the retirement decision itself is likely influenced by the health status. To circumvent such endogeneity issues, I use exogenous variation in the eligibility of early public pension withdrawal in a regression discontinuity design. Specifically, I exploit the 10th AHV revision in Switzerland, which gave men born as of a certain date the option to retire one and two years earlier. As on outcome, I use mortality. This variable has the advantage that reliable data exists and that it is more reliable than survey data. To estimate the effect, I draw from two full sample administrative data sets: the mortality and the retirement register. The results show a significant increase in mortality for the reform that allowed men to retire two years earlier.

Looking at the drivers of this effect reveals some interesting insights. First, there is a strong and significant increase in deaths related to alcohol dependence. While this does not mean that men died because of alcohol, it tells us something about their unhealthy lifestyle. Second, the effect is mostly driven by men living in the German-speaking area, which points at cultural effects. Previous research shows that individuals in the German-speaking regions show stronger norm towards work and might therefore suffer more from a loss of structure coming with early retirement. Third, the effect is strongest for single men, which could point at the previously documented negative effect of loneliness on health. Forth, the effect is strongest in the middle of the income distribution, which suggests that poverty might not be the principal cause. Several of those points resemble the «deaths of despair» that have been documented by Case and Deaton (2015) in the US. It tells us that providing men the option to retire early can be dangerous. A similar study from Austria titles this with the apt words: Fatal attraction (Kuhn et al, 2020).

Chapter 4, 5, and 6 study equality of opportunity. The idea that all children should have the same chance to succeed, the idea of the American Dream, is widespread in most societies. One way to measure equality of opportunity is to compare the relationship between child and parent income, the so called intergenerational income mobility. While this is quite an important proxy for the fluidity of a society, there are only few countries with reliable estimates because data requirements are challenging.

Chapter 4, written jointly with Veronica Schmiedgen-Grassi, documents intergenerational income and education mobility for Switzerland. We use a large administrative data set that covers the universe of labor income since 1982 which is linked over generations and matched to census and survey data. The results show that Switzerland has a high intergenerational income mobility. This is surprising because we also find that educational mobility is low. So far, economic models predict that low educational mobility would translate into low-income mobility. This does not seem to be the case in Switzerland. The reason for this puzzle might lie in the idiosyncratic education system in Switzerland, in which most children opt for vocational education and training after compulsory school. Indeed, many children that achieve the American Dream start with vocational education and training and add some sort of higher education after it. VET comes at almost no cost for parents and still provides ample options for further education after the apprenticeship. Thus, credit constraints of parents—that in the economic models are the reason for persistence in income inequality over generations—are less important in the acquisition of children's human capital. This finding that the VET system could drive upward mobility could be vital for other countries that struggle with low income mobility.

Chapter 5 goes one step further, or rather one generation backwards, and looks at how the income of grandparents is related to the income of their grandchildren. One of the core questions in such multigenerational models is whether one can predict the persistence in inequality from two adjacent generations to more remote generations. Usually, one assumes that persistence in inequality decays geometrically. For example, if the intergenerational income elasticity is 0.5 between two generations, it would be 0.25 within 3 generations and 0.125 within 4 generations. Thus, even with large intergenerational elasticities, economic status gained by one generation would disappear relatively quickly. This is where the early 20th century proverbial saying comes from: «From shirtsleeves to shirtsleeves in three generations».

However, recent studies challenge this geometric decay and find that the persistence over more than two generations is stronger than the extrapolation from the two generation prediction would suggest. I also test those predictions for Switzerland with linking labor income and education between three generations. The results show that there is indeed a geometric decay when looking at labor income. However, the decay is much slower when looking at educational attainment. This points to the fact that intergenerational mobility might be domain specific. One interpretation is that it is easier to influence the education of children and grandchildren than the income. Another explanation might also be that this is idiosyncratic to Switzerland with its strong vocational education system which provides good wages outcomes with relatively little formal education.

In this last chapter 6, I analyze the effect of a large policy change in the labor market of Switzerland: The abolition of the restriction for cross-border workers in certain areas close to the border. Cross-border workers make up a large share of the workforce in Switzerland Thus, it is natural to ask how this policy change influenced equality of opportunity for natives. Thereby, I use the same data set that links incomes over generation and exploit the policy reform in a difference-in-difference framework. The results show that the policy reform decreases intergenerational income mobility. The income of children from the bottom quintile decreases, while it slightly increases for children with parents at the top quintile. The effect is large because it is not alleviated by changes in education or learned occupation followed by the influx of cross-border immigrants.

Although the topics are heterogeneous per se, they all have two common themes: the use of large administrative data and the analysis of policies or economic circumstances in Switzerland. This is to say that most of those chapters would not have been possible several years ago, when governments were more restrictive with giving researchers access to individual data. To test hypotheses and inform policymakers, it is crucial to have access to such data. Furthermore, Switzerland is an interesting country to study, despite its small size. Switzerland's data might not be as extensive as the prominent data from the Nordic countries. However, its federal organization with canton specific policy changes, its multi-linguistic nature, its prominent vocational education system, its peculiar status outside of the European Union and its specific immigration policy, its flexible labor market, somewhere between the US and Europe, gives plenty of opportunity for interesting research.

Chapter 2

Baby Bonus in Switzerland: Effects on Fertility, Newborn Health, and Birth Scheduling

joint with Caroline Chuard

Abstract This paper studies the effect of birth allowances (so-called baby bonus) on fertility, newborn health, and birth-scheduling in Switzerland. Switzerland provides an optimal quasi-experiment: 11 out of 26 cantons introduced a baby bonus during the last 50 years at different points in time. To identify the effect of changes in the baby bonus, we employ an event study with control groups using several administrative data sets on births, stillbirths, and infant deaths in Switzerland from 1969 to 2017. While there is no evidence for birth-scheduling, we find, however, a sizable but only temporary increase in the fertility rate of 5.5% and a permanent but diminishing increase in the birth weight of 2.8%. The latter effect is particularly strong at the lower end of the birth weight distribution. Furthermore, we document substantial heterogeneity by citizenship of mothers.

Published 2021 in «Health Economics» Journal, 30(9), 2092– 2123
doi: https://doi.org/10.1002/hec.4366

Keywords: birth allowances, birth scheduling, fertility;
JEL classification: H31, J13

We are very grateful to Matt Sutton and two anonymous referees for extensive and supporting feedback. Furthermore, we want to thank Mirjam Bächli, Monika Bütler, Marc Chuard, Michael Lechner, Hannes Schwandt, Josef Zweimüller, and seminar participants at the University of Zurich and the University of St. Gallen for comments and suggestions. The Swiss Federal Statistical Office has kindly provided us with the data. Any errors are our own.

2.1 Introduction

Children are expensive. Therefore, many governments introduced policies to ease financial pressure on families. Among them are birth allowances that incorporate a lump-sum transfer at the event of birth. Birth allowances — also called baby bonuses — are designed for the vulnerable transition from being a couple without a child to becoming parents. Providing financial support in this critical period can affect parental behavior in the medium- to long- and short-run.

In the medium- to long-run, birth payments could incentivize couples to become parents and consequently boost fertility. This is a crucial topic for countries with an aging population and with fertility rates below the replacement level of 2.1 children per woman. Due to the improved financial situation and the decline in financial stress, birth payments may also improve newborn health.

In the short-run, expecting parents might (re-)schedule births when a new baby bonus policy is introduced and thereby affect newborn health. Specifically, financial incentives may motivate parents to shift a birth forward or backward around the expected date of delivery. This can have severe long-run consequences for the unborn child because advancing or postponing a birth affects newborn health by giving the fetus less or more time to grow within the maternal womb (Gans & Leigh, 2009; Tamm, 2013; Neugart & Ohlsson, 2013; Brunner & Kuhn, 2014; Borra et al., 2019).

In this paper, we study the effect of introducing, increasing, or abolishing birth payments on fertility, newborn health, and birth-scheduling. For the empirical analysis, we draw on several administrative data sets from 1969 to 2017. We build outcome variables based on the Swiss birth register, the universe of stillbirths, and the statistics on infant deaths. Combining these outcome variables with cantonal information on birth allowances allows us to study the causal impact of birth payments in a unique quasi-experimental setting. Nevertheless, authorities leave cantons a certain degree of freedom for birth allowances: Cantons are free to implement birth payments and free to set the amount. Based on this, we implement an event study difference-in-difference estimation.

Our results show that introducing a baby bonus affects fertility and newborn health. The fertility rate increases by around 5.5% at the mean in the first year of the post-treatment period, but fades out quickly. Newborn weight increases by around 2.8% at the mean, and the effect diminishes over time. To study heterogeneous effects across the socio-economic spectrum, we approximate socio-economic status of the mother by her country of origin. We find that the fertility effect is driven by mothers with citizenship from LMICs. Birth weight, however, significantly increases for mothers from a high-income country.

In contrast to previous studies, we do not find evidence for birth-scheduling around the policy changes. We argue that this results from several features in the Swiss setting. First, changes are in absolute terms smaller than in other countries with birth allowances. Second, the daily birth number per canton is likely too small to detect significant results.

This paper contributes to the literature on impacts of cash transfers on fertility behavior and newborn health. Furthermore, it adds to the literature on policy announcement effects on birth-scheduling and newborn health.

Studies, such as Gans & Leigh (2009), Tamm (2013), Neugart & Ohlsson (2013), Brunner & Kuhn (2014), or Borra et al. (2019), have found evidence of birth scheduling and fertility adjustments .[1] Due to the quasi-experimental setting in the Swiss context, we are the first to introduce a plausible control group: Cantons that do not change their birth allowance system at a given point in time and cantons that never introduced such a policy. Previous studies almost exclusively analyzed national policy changes instead of cantonal policy changes. Thus, they had to rely on regularities in the data before or after the policy change to predict an alternative outcome in the absence of the policy change to which the actual number of births per specific day in the year could be compared to. There is an other issue when only using data from before the policy change to predict the alternative outcome: observations at the boundary of the sample period have a strong impact on the estimated time trend in case of a nonlinear trend. Furthermore, we can analyze introductions, increases, and abolition of the baby bonus within one country. This setting allows us to study asymmetries as the parental choice of delaying or scheduling a birth early is different.

More generally, there is a large and growing strand of literature analyzing the impact of cash transfers on fertility (Kearney, 2004; Milligan, 2005; Cohen et al., 2013; Laroque & Salanié, 2014; González, 2013; González & Trommlerová, 2021). Several of these studies find an impact on fertility when parents face financial support. Most closely related to our study is Milligan's (2005) analysis of a policy reform in birth allowances in Quebec. This policy led to transfers up to CAD 8,000 (roughly CHF 6,000)[2] for the third child. The author finds a strong effect on fertility. While the absolute amount is significantly higher in the Canadian study (depending on the canton multiplied with a factor from 3–6), the Swiss transfers are already being paid for the first child. Therefore, the Swiss case allows to study fertility effects at both the intensive and extensive margin. Second, while many works exclusively focus on fertility and labor market outcomes, we extend the analysis

[1] A related strand of literature investigates tax incentives and birth scheduling (Dickert-Conlin & Chandra, 1999; Schulkind & Shapiro, 2014; LaLumia et al., 2015). For the US, these papers show that the tax scheme incentivizes parents to schedule births in late December instead of early January.

[2] 1 CHF equals roughly 1 USD.

to newborn health outcomes giving a broader picture of cash transfers. Third, with the panel structure of the data and the long history of Swiss family allowances, we can study the impact of baby bonuses over time and show a fading out of the effect with every year after the implementation.

Finally, based on the staggered implementation across cantons and time, we can base our estimates on several policy changes which increases the external validity of our results. Due to the concerns raised recently by Goodman-Bacon (2021), we do not choose a two-way fixed effects (2WFE) difference-in-difference model. Goodman-Bacon (2021) show that the treatment estimate of 2WFE regression can be severely biased if effects change over time. Furthermore, it is difficult to assess the parallel trends assumption with this model. While several authors, among them De Chaisemartin & d'Haultfoeuille (2020), Callaway & Sant'Anna (2020), and Athey & Imbens (2018) propose solutions to the described problem, we chose to use an event study design with control group that incorporates dummies for every year relative to the introduction of the treatment. We will refer to this strategy by *event study DiD*. This setting allows us to study effects over time instead of a single coefficient under relatively mild assumptions which will be described more specifically in the empirical part.

We organize the remainder of the paper as follows. Section 2.2 describes the institutional background. Section 2.3 describes the data. Section 2.4 introduces the empirical strategy. We present various results, sensitivity analyzes, and a discussion of the results on fertility and newborn health in Section 2.5 and on birth-scheduling in Section 2.6. Section 2.7 concludes.

2.2 The Swiss Baby Bonus

Switzerland has a decentralized federal political system with three interdependent governmental levels (federal, cantonal, and municipal). Family allowances are regulated on the federal level. However, each canton has the authority to adjust the local payments individually. Family allowances are financed via contributions to the family compensation office and not via taxes. Therefore, they are detached from other regulatory decisions or tax incentives designed for families. Depending on the canton, expecting parents have to collect their family allowances directly from their employer or the family compensation office.

The Swiss political system is characterized by a direct democracy and a decentralized federalism. Each governmental layer is entitled to decide about all political issues in its sphere of influence. Furthermore, each important new constitutional amendment needs the consent of the voting population, which results both in lengthy processes of implementing new policies and in a tremendous variation of different policies. Thus, even if other family policies exist — such as incentives in tax systems or in child care — it is unlikely that they systematically interfere with the family allowances which we study.[3] Specifically, introductions of the baby bonus mostly occurred prior to other social policies supporting child care.

There are three different family allowances: (1) child allowances, which by federal law since 2009 have to be at least CHF 200 per month,[4] (2) education allowances, which by federal law have to be at least CHF 250 per month, and (3) the one-time birth payment with no federal minimum payout. Thus, cantons are free to implement a baby bonus and to define the amount paid. They may also increase the baby bonus or abolish it at any point in time. This gives rise to large variation across cantons.

An important difference between these benefits is that child and education allowances are monthly money transfers, while the baby bonus is a onetime payment. While the different allowances may change at the same point in time, eligibility to collect one type of allowance varies. All mothers living in a specific canton and giving birth after the implementation date of the baby bonus are eligible for the baby bonus. Thus, there is a sharp cutoff from one day to the next. For child and education allowances, every child eligible in a month can enjoy higher payments after a policy change.[5] This is to clarify that in practice the baby bonus and the child allowance paid in the first month after birth never offset each other.

In this paper, we will only focus on birth payments because the unique setup of this benefit allows us to analyze birth-scheduling, newborn health, and fertility effects. The birth payment is a unique payment to a woman who had a living birth (or a stillbirth after at least 23 weeks of gestation). The birth payment is per newborn. For a multiple birth, a mother can collect the baby bonus for each child.

The baby bonus may affect two outcome margins. Already pregnant mothers may want to shift their birth a few days to become or stay eligible for the birth payment. This is what we call the short-run margin. This short-run margin may affect newborn health via birth-scheduling. Importantly, this effect is not diluted by a change in the fertility behavior because all policy changes were announced less than seven months before the implementation.[6] Therefore, mothers were already pregnant at the time of the policy announcement. Thus, mothers can only schedule the birth in a short period around the expected date of birth.

[3]In a robustness check, we do, however, control for changes in child allowances and find similar results.

[4]The evolution of child allowances over time are depicted in Figure A.1 in the Appendix.

[5]Eligibility for child and education allowances depends on the age of the child, the educational track of the child, and the employment status of the parent.

[6]See for further information on announcement and implementation dates Table A.1 in the Appendix.

Birth-scheduling results from financial incentives of introductions, increases, and abolition of birth payments. On the one hand, births may be delayed beyond the date of implementation when a baby bonus is introduced or substantially increased. On the other hand, births may be brought forward when a baby bonus is abolished. It is more difficult to delay a birth than to schedule early, due to the natural end of every pregnancy. There are several ways to delay labor (Coomarasamy et al., 2003; Shapiro et al., 2013; Lima et al., 2018): One is to avoid stress or to take medication to delay labor by up to 48 hours. Another one is to postpone an already planned Cesarean section. Through a delay, a newborn is expected to have a higher weight, since the unborn had more time to grow in the mother's womb.

In the case of an abolition of the policy, mothers may want to speed up the pregnancy. Mothers can schedule a birth early via a Cesarean section or induce labor medically. These choices will lead to an earlier birth and a lighter newborn. As a result, mothers must weigh up their financial gain against the potentially harmful effect on their newborns' health.

In the medium- to long-run, mothers might also adjust their fertility behavior. Thus, higher birth allowances can increase fertility. This may be the result of explicitly choosing to have a(n additional) child, having children earlier or choosing not to abort. Furthermore, the payment may also improve newborn health. Either because different types of mothers choose to have a(n additional) child (i.e. a selection / composition effect) or because more money directly impacts newborn health via, for example, better maternal health or a change in maternal behavior (an extensive overview by Almond et al. (2018) documents various of these effects).

The latter channel is expected to be especially strong for parents with low-socio-economic status based on the findings of the literature on the fetal origins hypothesis. Financially distressed parents may benefit from this extra payment and negative pregnancy outcomes might be prevented and positive birth outcomes promoted. US welfare programs targeted toward low-socio-economic groups such as the Earned Income Tax Credit (EITC) studied by Hoynes et al. (2015) and the Food Stamp Program (FSP) studied by Almond et al. (2011) show substantial beneficial impacts on newborn health. However, the impact of additional income above a certain threshold is much less studied and not documented so far. Furthermore, it is also less clear whether a benefit paid in the future might affect health outcomes today. One argument could work via a reduction in maternal stress based on the knowledge of receiving a transfer in the very near future. Several studies have shown that stress due to various reasons affects newborn health (Camacho, 2008; Aizer, 2011; Currie & Rossin-Slater, 2013; Lee, 2014; Black et al., 2016; Quintana-Domeque & Ródenas-Serrano, 2017).

2.3 Data

We base the empirical analysis on several data sources. The data on all outcome measures such as newborn health, birth-scheduling, and fertility is coming from the Swiss Vital Statistics and Annual Population Statistics provided by the Federal Statistical Office (FSO). Information on the amount and the date of implementation of all birth allowances per canton is recorded by the Federal Social Insurance Office (FSIO).

2.3.1 Data Sources

Swiss Birth Register. The birth register covers all births from 1969 to 2017. It contains information on date of birth, sex, and beginning from 1979 birth outcomes, such as weight and length at birth. Based on the information about birth weight, we create a dummy for low birth weight defined as less than 2,500 grams to understand which part of the distribution is mostly affected. The latter outcome measures are linked to later life outcomes (Almond et al., 2018). Furthermore, the birth register provides information on birth order and birth interval in months to a preceding birth.

Using the information on the gender of the child, we generate the sex ratio. There are several arguments for how socio-economic and maternal health conditions during pregnancy can affect the sex ratio as summarized by Scalone & Rettaroli (2015). Improving socio-economic conditions can, based on a biological and evolutionary argument, favor boys in the maternal womb. This is because the male fetus is frailer. Thus, we follow the lines of Sanders & Stoecker (2015) and use the sex ratio in live births as a proxy for the miscarriage rate.

To calculate the crude birth rate per 1,000 people and the total fertility rate per woman, we merge the data on a canton-year level with the Annual Population Statistics.[7] The Annual Population Statistics is available from 1971 with detailed information on age-specific population starting in 1981. Thus, the crude birth rate can be reported from 1971 onward, while the total fertility rate is only available after 1981.

The birth register also contains information about the mother, such as her age, marital status, citizenship, municipality, and canton of residence. Maternal age, though, serves as an additional outcome measure (also in combination with the birth interval between two consecutive births) to study mechanisms that explain overall fertility and child health.

Stillbirths and Deaths. For the determination of more severe health outcomes, we rely on information provided in the statistics of stillbirths and deaths. As in the birth register,

[7] We follow conventional definition to measure these two rates. The crude birth rate is the total number of births divided by the total population multiplied by 1,000. The total fertility rate results from dividing the total number of births by the total number of fertile women aged 15–49 multiplied by 35, the total age range.

these data sets provide information on the date, municipality, and canton of occurrence. Based on these two measures, we calculate the stillbirth and infant (< 1 year) death rate per 1,000 births.

Birth Allowances. We have collected the full history of birth allowances per canton from several publications. From 1969 to 1992 the data were published in *Zeitschrift für die Ausgleichskassen* (Bundesamt für Sozialversicherungen, 1969–1992). The publication *AHI-Praxis* covers the period from 1993 to 2004 (Bundesamt für Sozialversicherungen, 1993–2004). Starting from 2005, the data are published online on the website of the FSIO (Bundesamt für Sozialversicherungen, 2005–2020). These publications record information on the date of implementation and the amount of the allowance per canton. Additionally to the date of implementation also the date of announcement is recorded. All health policy reforms were announced not more than seven months prior to their implementation. This guarantees that around the implementation date, the only adjustable margin is birth-scheduling and no fertility adjustment, as mothers were already pregnant by that time. In the long run, however, fertility can be affected.

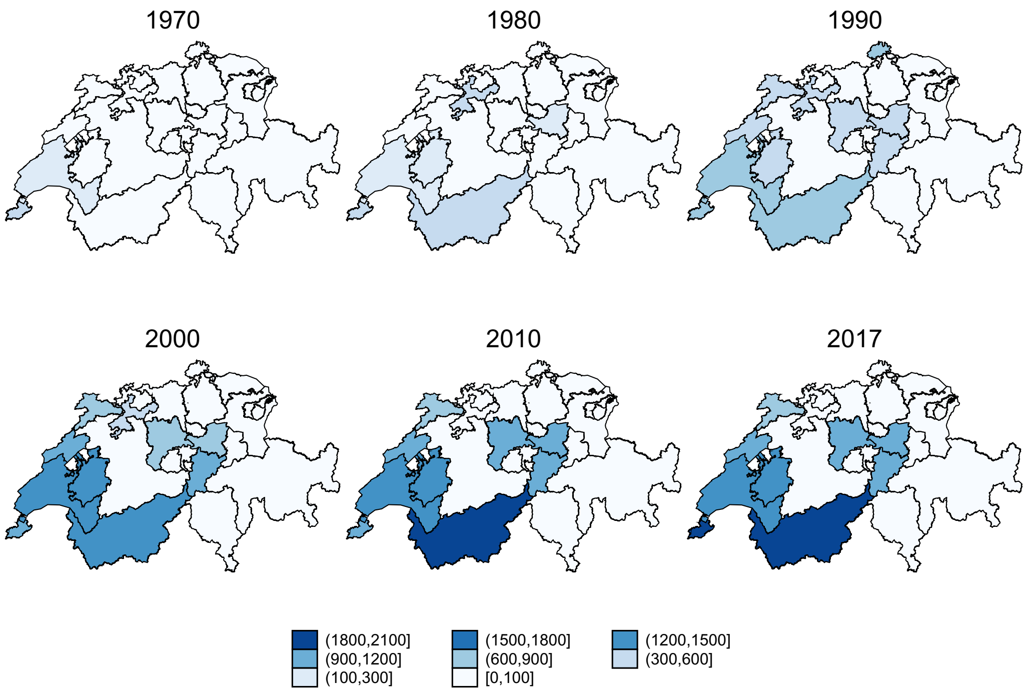
2.3.2 Descriptive Statistics

We show descriptive statistics for birth measures, child characteristics, and maternal characteristics for the overall sample in Table 2.1. Over the entire time period from 1969 to 2017, we observe on average around 81,000 births per year. The crude birth rate per 1,000 people in Switzerland is 11.8 and the total fertility rate per woman 1.6. On average, 5.0 fetuses out of 1,000 births die in a mother's womb and 7.6 infants out of 1,000 births die within the first year. At birth there are slightly more males (0.514) which directly translates into a sex ratio of 0.946 girls per 1 boy. The average Swiss family has a birth interval of slightly over 3 years between children and the average birth weight of a newborn is 3,334 grams. 5.7% of children are born with a birth weight of less than 2,500 grams. Mothers are on average 29 years old when giving birth, mostly married (91 percent) and 74 percent of them are Swiss.

Figure 2.1 shows the geographic variation in birth allowances for six different years. Cantons with birth allowances are mostly concentrated in the French speaking part and in the region of Central Switzerland. Figure A.2 shows the time variation in birth allowances for all cantons that introduced the baby bonus at some point in time. Three cantons (Geneva, Vaud, and Fribourg) have already put baby allowances in place before 1969.[8] Several cantons adjust the amount of baby allowances over time. Two cantons (Solothurn and Schaffhausen) abolish the baby bonus after some years again.

[8] As the data set on the outcome measures only starts in 1969, the variation before 1969 cannot be exploited in this paper and is therefore not shown in the graphs.

FIGURE 2.1: Geographic Variation of Birth Allowances



Notes: This figure shows the amount of birth allowances provided per child per canton in current year values. The focus is on geographical variation so that birth allowances are drawn for all cantons every 10 years up to the most recent year available.

TABLE 2.1: Descriptive Statistics

	Mean	Std. Dev.	Min	Max
Overall birth measures				
Total yearly births	80,578	7,083	71,375	102,520
Fertility rate	1.621	0.256	1.146	2.914
Crude birth rate (per 1,000 people)	11.762	2.170	7.531	19.170
Stillbirth rate (per 1,000 births)	5.083	2.841	0	22.422
Infant death rate (per 1,000 births)	7.581	4.939	0	36.082
Child characteristics				
Male	0.514	0.016	0.421	0.631
Birth interval (months)	37.550	3.102	27.366	48.301
Birth weight	3,333.551	49.468	3216.849	3902.548
Share Low Birth Weight	0.057	0.012	0.016	0.104
Maternal characteristics				
Age of the mother	28.985	1.586	25.952	32.631
Married at birth	0.905	0.070	0.635	0.991
Swiss at birth	0.735	0.119	0.408	0.977
N (Canton x Years)	1,274			

Notes: The full sample covers all births, stillbirths, and infant deaths from 1969 to 2017.

2.4 Empirical Strategy

To answer our research questions, we employ two types of event studies. We aim to identify the causal effect of the baby bonus on health and fertility outcomes in an *event study DiD* where we show how effects evolve over time, and specifically after the treatment. To test for birth-scheduling behavior, we use a *time-series event study* design where we predict the total number of births per day and canton in absence of the treatment and look for manipulation around the cutoff by displaying the residuals from the prediction.

2.4.1 Fertility and Newborn Health

To estimate the causal effect of the baby bonus on health and fertility, we compare cantons with baby bonus to cantons without baby bonus before and after the introduction.[9] In our main specification, we use an event study DiD model. Event studies differ from a standard difference-in-difference design in that the treatment is no longer uniquely characterized by a binary indicator, but a set of dummies indicating the time relative to the introduction. Thereby, we can analyze the evolution over event time. Furthermore, it allows to control for canton fixed effects, year fixed effects, and cantonal trends. Specifically, we estimate the following regression:

$$y_{ct} = \gamma_c + \lambda_t + \sum_{\tau=2}^{m} \delta_{-\tau} D_{c,t-\tau} + \sum_{\tau=0}^{q} \delta_{+\tau} D_{c,t+\tau} + \eta_c \times t + \epsilon_{ct}. \tag{2.1}$$

The dependent variable, y_{ct} represents the total fertility rate, the crude birth rate, the birth weight, share of low weight births, the stillbirth rate, the infant death rate, and the sex ratio as well as maternal age, and the birth interval in canton c and year t. γ_c are canton fixed effects, λ_t denotes (calender) year fixed effects, and η_c allows for canton specific linear time trends. ϵ_{ct} is an error term.

The variables $D_{c,t-\tau}$ and $D_{c,t+\tau}$ equal 1 in the m pre-treatment periods and in the q post-treatment periods, respectively. We omit category $\tau = -1$, which is the event-year before the introduction. Thus, the set of coefficients δ_τ for $\tau \in [-m, q]$ shows the change in outcomes in cantons with a baby bonus compared to cantons without a baby bonus relative to the event year $\tau - 1$.

We weight the estimates in the canton-year cell differently, depending on the outcome. The fertility rate is weighted by the number of fertile women, the crude birth rate is weighted by population size, and health measures are weighted by the number of births. Robust standard errors are clustered at the cantonal level.

To ensure that we have the same amount of pre-treatment years for all baby bonus cantons, we choose the pre-treatment period to be 5 years. The pre-treatment periods yield important insights to ensure that our identification strategy is valid. The key identifying assumption is the *parallel trends assumption*. This assumption states that in the absence of treatment, treated units would have experienced the same trends in average

[9] For this part of the analysis we focus solely on the introduction and not on increases of the baby bonus. With our empirical approach at hand, a study of increases is not straightforward and possibly even problematic if future changes in bonus size are endogenous, i.e. depend on the success of the introduction.

outcomes as the control units (i.e. those which never introduced the baby bonus). If the underlying parallel trends assumptions holds, pre-treatment coefficients should not significantly differ from 0. At the same time, pre-treatment period coefficients serve as a Granger causality test. If the policy is responsible for a change in the post-treatment periods, we would expect to have zero effects in the pre-treatment periods and a non-zero effect afterwards. As most introductions take place relatively early, we can use 9 post-treatment periods.[10] The long post-treatment periods allow us to show how the effect evolves over time.

Another assumption to identify the causal effect is the *stable unit treatment value assumption (SUTVA)*. This assumption would be violated if other policies were introduced that interfere with the introduction of the baby bonus policy. In a robustness analysis we, thus, control for child allowances — one of the other policies in the family allowances package, which, however, are much more regulated due to minimal payment amounts based on the federal law.

2.4.2 Birth-Scheduling

To test for birth-scheduling behavior, we use a time-series event study design. First, we collapse our individual level data on the daily cantonal level. Next, we regress the following equation:

$$y_{imyc} = \alpha + \beta_c + \gamma_y + \delta_m + \zeta_i + \epsilon_{imyc}, \tag{2.2}$$

where y_{imyc} is the total (log) count of births per day i, in month m, in year y, and canton c. With β_c we include canton fixed effects, and with γ_y and δ_m year and month fixed effects, respectively. ζ_i are, depending on the specification, day-of-week or day-of-year fixed effects. Day-of-week fixed effects can be more precisely estimated and root on the idea that daily births vary across the day of the week due to, for example, relatively few planned births via Cesarean sections on the weekend. Day-of-year fixed effects, instead, control for specific dates unrelated to the day of the week such as day-specific holidays or the first day of a month in case parents have a preference or aversion for any of these dates. Based on the fact, that our sample includes control cantons, the different coefficients (γ_y, δ_m, and ζ_i) on time fixed effects can be identified on top of a treatment effect in a given year and canton. Finally, we calculate residuals from a linear prediction and plot these residuals for the 60 days around the policy change. Thereby, we pool over the same event across cantons and time. Robust standard errors are clustered at the cantonal level.

Our identification of the birth-scheduling effect relies again on the assumptions of parallel trends. As other untreated cantons serve as control, these cantons must provide an

[10]Thus, $q = 8$ as we count the event year $\tau = 0$ as part of the post-treatment period.

appropriate counterfactual so that they describe the trend treated cantons would have followed in absence of the treatment. Furthermore, depending on the specification, the day-of-week or day-of-year fixed effects should not vary over the included time frame on top of the included year and month fixed effects. Visual inspection of the residual graphs, show that the residual approach results in a *noisy pattern around zero* further away from the cutoff date and, therefore, seems to appropriately de-trend the data.

We show the birth-scheduling effect individually for introductions, increases of above CHF 200, and the abolition of the policy. Given the specific event, we would expect a certain pattern accentuating the closer the cutoff. For example, in case of an introduction of a baby bonus we would expect parents to shift the birth after the introduction time. Thus the prediction in absence of the policy change would be too high before the policy's implementation and therefore the residuals below zero. While after the introduction, we would expect to observe a discontinuous jump in the residuals. An increase of the baby bonus would lead to the same pattern, while an abolition should lead to the opposite picture i.e. a negative jump around the cutoff.

2.5 Results: Fertility and Newborn Health

We show our main results of Equation (2.1) in Table 2.2 and Figure 2.2. Figure 2.2 plots the coefficients δ_τ relative to the time of introduction. The omitted category is event time $\tau = -1$, directly before the introduction in event time 0. Negative event times indicate pre-treatment periods.

We see significant changes in the post-treatment period for the fertility rate, age of mother at birth, birth weight, and the share of children being born with low birth weight. Looking at the coefficients of those four outcomes in the pre-treatment period, we see that they are small and in almost all cases not significantly different from 0. This suggests that the parallel trends assumption is plausible and that they pass the Granger causality test, which states that the effect of the treatment cannot occur before the treatment happened. Note that the low birth weight and birth weight coefficients show some significance in event time -5 and -3. However, the coefficients in pre-treatment are small and considerably, as well as significantly, different from the coefficients in the post-treatment periods. Furthermore, if anything the coefficients on low birth weight trended upwards so that the negative effect after the reform can be interpreted as a conservative estimate.

Looking at how the effect of the baby bonus on the fertility rate, age of mother at birth, birth weight, and low birth weight develops over time, reveals some additional insights. Over time, the effect tends to fade out — this is especially true for the fertility rate: After merely 4 years, the coefficient is close to 0 and insignificant. For age of mother at birth,

birth weight and low birth weight, this process is slightly slower, and the coefficient tends to stay marginally significant.

Table 2.2 shows the treatment coefficients that are plotted in Figure 2.2. Comparing the effects of δ_τ in the introduction period ($\tau = 0$) to the mean of the dependent variable, allows to interpret the effects better. The fertility rate increases by 0.089 which corresponds to an increase at the mean of 5.5%. This fertility effect is therefore large — although only transitory in nature. The effect on age of mother at birth is -0.474 corresponding to a reduction of roughly 6 months in age at birth. While we argue that maternal age at birth can give insights on the effect on fertility (i.e. the earlier a mother gives birth the less likely she is to be affected of negative shocks such as a health shock or a divorce), we discuss this more in detail in the next section on the potential mechanisms.

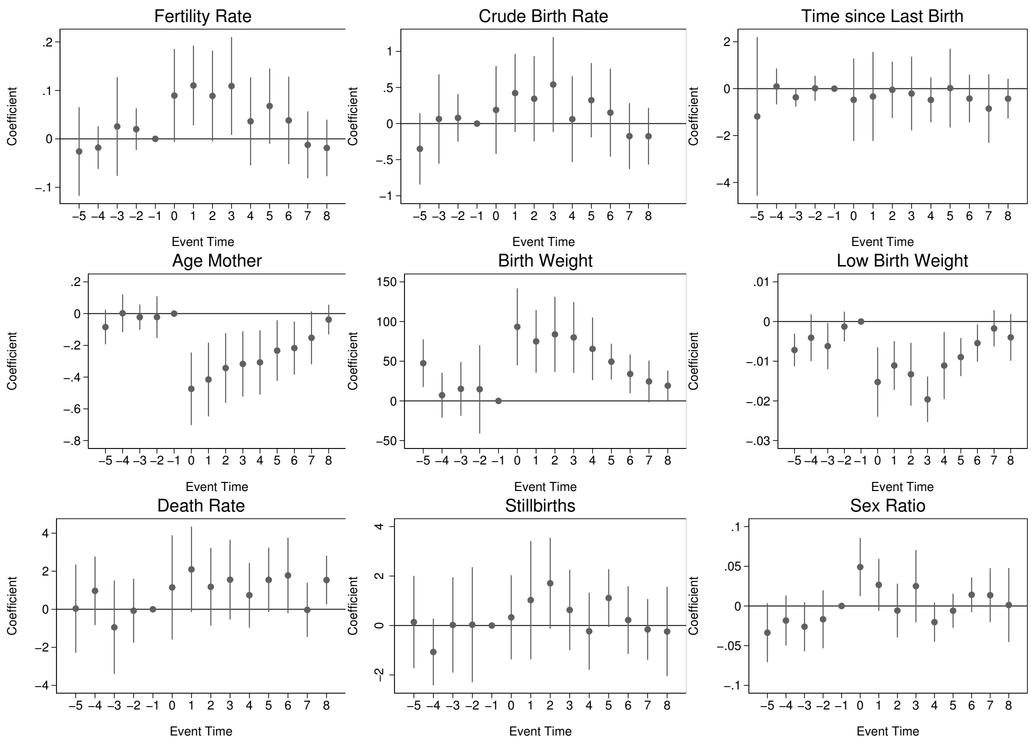
Furthermore, birth weight increases by 93 grams in the introduction period. This corresponds to an increase of 2.8% evaluated at the mean. The share of children being born with low birth weight reduces by 0.015 in the introduction period or 1.5 percentage points. In relative terms, this effect is quite large and corresponds to a decrease of 28% evaluated at the mean.

Opposite to these clear patterns, we do not find conclusive evidence of the effect on the crude birth rate, infant deaths, stillbirths, and the sex ratio. The crude birth rate shows positive coefficients in the post-treatment period, which would be in line with the findings on the fertility rate. However, the coefficients are not significant. For infant deaths, and stillbirths we cannot detect a significant change at the time of introduction, nor can we confirm that pre-treatment periods are different from post-treatment periods. While the coefficients mostly oscillate around 0, they tend to increase but stay insignificant in the post-event period. The coefficient on the sex ratio is significant in the post-treatment period 0. However, the effect is small and disappears after one period. Given the direction of the effect, this would indicate a small increase in the miscarriage rate. While both the average birth weight and the incidence of low birth weight improve — thus, an overall shift of the birth weight distribution to the right — this stands in contrast to the deteriorating severe health outcomes. The only way to explain such a pattern is through heterogeneous effects by maternal characteristics. However as the severe health measures are far from significant, we are cautious in interpreting this effect any further.

Other explanations for those small or non-significant changes at the time of introduction are multifold. On the one hand, the crude birth rate might be a too noisy measure. Population dynamics, such as immigration or changes in the age pyramid, make this rate an unreliable measure. Especially in Switzerland, with large immigration flows and an aging population, these influencing factors should not be neglected. On the other hand, the small but non-significant effects in the infant deaths, stillbirths, and the sex ratio can also be explained by the fact, that those are very severe and negative infant health measures, that are unlikely to be affected by a cash transfer in a developed economy.

We conduct several robustness checks. Table A.2 in the appendix shows the event study estimates without cantons that abolish the baby bonus after some time. Those are two cantons that abolished the baby bonus near the end of our data period. The overall pattern is unchanged to our baseline estimation. In Table A.3, we control for child allowances. Child allowances could potentially interfere with the baby bonus. However, as Figure A.1, depicting cantonal changes over time, shows, increases are mostly small and are, thus, unlikely to change fertility or newborn health. Table A.3 confirms this as the results are quantitatively and qualitatively similar to our baseline estimation.

FIGURE 2.2: Event Study DiD Results



Notes: This figure shows how the effect of the introduction of the baby bonus on the respective outcomes changes over time. It depicts δ_τ from Equation (2.1) for $\tau \in [-5, 8]$ for each specific outcome variable. The event year represents the year relative to the introduction of the baby bonus. The dots show the point estimates per event time, while the line corresponds to the 95% confidence interval.

Taken together, the results suggest that the introduction of the baby bonus had a sizable, but transitory effect on fertility. While we show with our robustness analysis that this dissipating effect is not driven by cantons that later abolish the baby bonus, we argue

TABLE 2.2: Event Study DiD Estimates

	(1) Fertility Rate	(2) Crude Birth Rate	(3) Birth Interval	(4) Age Mother	(5) Birth Weight	(6) Share Low Birth Weight	(7) Death Rate	(8) Stillbirth Rate	(9) Sex Ratio
Event Time Estimate:									
-5	-0.026	-0.350	-1.184	-0.085	47.483**	-0.007**	0.037	0.140	-0.034
	(0.044)	(0.238)	(1.632)	(0.053)	(14.490)	(0.002)	(1.118)	(0.901)	(0.018)
-4	-0.018	-0.254	0.094	0.002	7.167	-0.004	0.971	-1.071	-0.018
	(0.021)	(0.154)	(0.369)	(0.057)	(13.578)	(0.003)	(0.870)	(0.651)	(0.015)
-3	0.026	0.064	-0.371	-0.023	15.172	-0.006*	-0.949	0.021	-0.026
	(0.049)	(0.298)	(0.192)	(0.038)	(16.289)	(0.003)	(1.182)	(0.929)	(0.015)
-2	0.020	0.079	0.015	-0.022	14.659	-0.001	-0.073	0.031	-0.017
	(0.021)	(0.159)	(0.257)	(0.063)	(26.821)	(0.002)	(0.812)	(1.122)	(0.018)
0	0.089	0.189	-0.475	-0.474***	93.340***	-0.015**	1.151	0.331	0.049*
	(0.046)	(0.293)	(0.848)	(0.110)	(23.412)	(0.004)	(1.320)	(0.821)	(0.018)
1	0.110*	0.423	-0.332	-0.415**	74.973***	-0.011**	2.099	1.027	0.027
	(0.040)	(0.260)	(0.916)	(0.112)	(19.035)	(0.003)	(1.084)	(1.152)	(0.016)
2	0.089	0.344	-0.048	-0.343**	83.869**	-0.013**	1.180	1.708	-0.006
	(0.045)	(0.284)	(0.581)	(0.106)	(22.822)	(0.004)	(0.986)	(0.888)	(0.016)
3	0.109*	0.541	-0.207	-0.317**	79.970**	-0.020***	1.558	0.628	0.025
	(0.049)	(0.317)	(0.757)	(0.099)	(21.578)	(0.003)	(1.012)	(0.785)	(0.022)
4	0.036	0.061	-0.479	-0.307**	65.582**	-0.011*	0.739	-0.232	-0.020
	(0.044)	(0.287)	(0.461)	(0.098)	(19.020)	(0.004)	(0.820)	(0.750)	(0.012)
5	0.068	0.324	0.024	-0.233*	49.411***	-0.009***	1.549	1.112	-0.006
	(0.038)	(0.248)	(0.810)	(0.092)	(10.923)	(0.002)	(0.814)	(0.564)	(0.010)
6	0.038	0.151	-0.421	-0.217*	33.970**	-0.005*	1.780	0.222	0.014
	(0.044)	(0.294)	(0.490)	(0.081)	(11.793)	(0.002)	(0.957)	(0.660)	(0.011)
7	-0.013	-0.174	-0.848	-0.153	24.522	-0.002	-0.031	-0.159	0.014
	(0.033)	(0.220)	(0.706)	(0.080)	(12.638)	(0.002)	(0.688)	(0.593)	(0.016)
8	-0.019	-0.175	-0.428	-0.038	19.108*	-0.004	1.536*	-0.242	0.001
	(0.028)	(0.189)	(0.409)	(0.045)	(9.128)	(0.003)	(0.616)	(0.872)	(0.022)
Canton FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LinTrends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N (Canton x Years)	636	636	678	862	678	678	862	862	862
R2	0.952	0.937	0.857	0.993	0.719	0.694	0.79	0.672	0.157
Mean Dependent	1.63	11.87	37.15	28.86	3,342	0.053	7.84	5.13	0.94

Notes: This table shows coefficients for δ_τ from Equation (2.1) for $\tau \in [-5, 8]$ for each specific outcome variable. The event year represents the year relative to the introduction of the baby bonus. The omitted category is event time $\tau = -1$. Estimates in the canton-year cell are weighted corresponding to the following structure: fertility rate with number of fertile women; crude birth rate with total population; birth interval, age mother, birth weight, share low birth weight, death rate, stillbirth rate, and sex ratio with number of births. Robust standard errors (shown in parentheses) are clustered at the cantonal level and significance levels are indicated by * 0.05 ** 0.01 *** 0.001.

that this might be the result of a behavioral feature, the so called reference dependence (Kahneman & Tversky, 1979). As the birth payment becomes normal (i.e. the new reference point) to parents it might not affect fertility any longer. Furthermore we find a significant, but declining effect on birth weight. This declining effect seems especially driven by the lower end of the birth weight distribution as depicted with the effect on the share of newborns being born with low birth weight. Further, we find no evidence that more severe outcomes, such as infant deaths, stillbirths, or miscarriages as approximated by the sex ratio are affected.

We can compare these effects with other studies. In case of the effect on fertility, we consult the most closely related study of Milligan (2005). Based on the differential design of the birth allowances for Quebec, however, one must be careful when directly confronting the fertility effects. Overall Milligan (2005) documents a fertility effect ranging from 5 to 10%. As the large change in payments, though, almost exclusively occurs for third born children, the effect of fertility on third children increases to up to 25%. Multiplying our effect on fertility of 5.5% with the difference in the size of the baby bonus ranging from 3 to 6 depending on the canton, we find a similar effect overall equaling 16.5 to 33% and thus matching the 25% of Milligan (2005).

Comparing the effect on newborn health with other studies is more difficult, because of the very specific nature of the programs. Hoynes et al. (2015), for example, document a 2 to 3% reduction in the occurrence of low birth weight because of the EITC, while Almond et al. (2011) find a 0.5% increase in birth weight and a 10% decline in low birth weight due to the FSP. While both the EITC and the FSP are large transfer programs of roughly 2.7 of GDP and 2.2 of GDP in 2004, respectively, the Swiss baby bonus equaling 0.07 of GDP in 2017 is of much smaller size. As the effect on birth weight and especially low birth weight is declining over time, it is probably much more appropriate to compare the effect sizes in the last event year where birth weight significantly increases by 0.6% at the mean and the incidence of low birth weight declines by 7.5% at the mean, though this estimate is not statistically significant. These estimates are thus almost the same as in Almond et al. (2011). However, in Switzerland this result was achieved with much lower costs.

2.5.1 Mechanisms and Further Analyses

Switzerland is among the countries with the highest share of immigrants in Europe. Over 27% of the births in our sample from 1969 to 2017 are given by mothers without a Swiss citizenship. Thus, we can exploit the citizenship of the mother as a proxy for socio-economic status. To do so, we use the World Bank database which categorizes countries into four income levels: High-income, upper middle-income, lower middle-income, and low-income (The World Bank, 2020). To facilitate the analysis and increase group sample size, we generate two groups: High-income countries, consisting only of World Bank's *high-income* countries including Switzerland and *low- and middle-income* countries (LMICs), which comprises the other three categories. Roughly 90 percent of births are given by mothers from high-income countries and approximately 10 percent by mothers of LMICs. Most prevalent across the LMICs are Serbia, Turkey, North Macedonia, Sri Lanka, Bosnia and Herzegovina, and Brazil. Those countries account for 65 percent of births in this group. Certainly, the country of origin does not perfectly predict socio-economic status: even within nationality, socio-economic status might differ strongly.[11] Nevertheless, we argue that it serves well as a proxy.

Figures 2.3–2.4 and Table A.4 show the effect on the fertility rate and birth weight for mothers with a country's citizenship either in the high-income or LMIC category. We see a strong discrepancy between those two groups. While the fertility rate for mothers with a high-income country citizenship does not significantly react, we see a strong increase in the fertility rate of mothers with a citizenship from a LMIC. Thus, the average fertility effect that we find is almost solely driven by mothers from LMICs. The finding that mothers with a LMIC background select into giving birth due to the policy, raises the question whether those mothers or their children differ in other characteristics (such as newborn health) from the high-income country mothers. Looking at birth weight, we see that the birth weight for children from mothers from LMICs is higher than from mothers from high-income countries (3364 versus 3325 grams).[12] This difference might at least partially explain the increase in birth weight of the policy. To see whether this could be the case, we run the event study on birth weight for mothers from LMICs and high-income countries separately.[13] We show the results in Figure 2.4. The results indicate that the birth weight effect is driven by mothers from high-income countries. Although the effect is almost twice as large in absolute terms for LMIC mothers, it barely reaches significance. Contrary to that, the effect is strongly significant for high-income mothers. While this does not prove that the policy had a direct effect on birth weight, it rules out that LMIC mothers are the sole driver of the birth weight effect.

More generally, the seminal paper by Becker & Lewis (1973) highlights the quantity and quality trade off of having children. While a higher income might increase the overall fertility, it could also increase parents' investment in these children and thus their

[11] A large body of literature (see, for example, McDonald & Kennedy (2005); Antecol & Bedard (2006); Biddle et al. (2007); Chiswick et al. (2008); Constant et al. (2018) and many more) shows the so called *healthy immigrant effect* stating that immigrants tend to be healthier than comparing native populations. However, the same strand of literature also states that the health advantage of immigrants declines with time spent in the host country. As we do not know anything about the time spent in Switzerland, we argue that if anything, we estimate lower bounds as mothers from LMICs might be positively selected.

[12] The birth weight for the four income country groups is: High-income 3325 grams, upper middle-income 3393 grams, lower middle 3279 grams, and low-income 3308 grams.

[13] The group specific fertility rates are calculated by dividing the number of births from a specific group by the number of fertile women in the whole country.

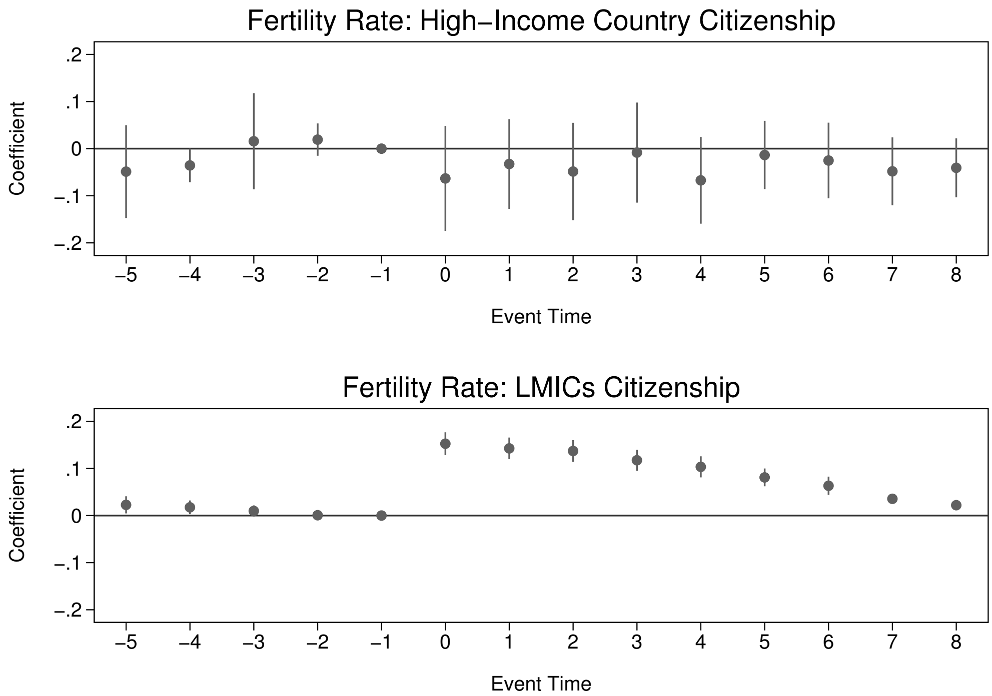
quality keeping the number of children constant. As such, a priori it is not clear which effect will dominate. Combining this takeaway with the fact that high-socioeconomic parents have shown to adapt more health promoting behavior in various settings (as extensively summarized by Almond et al. (2018) when it comes to smoking, drinking, healthy diets, doing exercises, going to the doctor, etc.) it is probably not surprising, that in this context the quantity effect (i.e overall fertility) dominates the quality effect (i.e. birth weight) for LMIC mothers and vice versa for high-income mothers.

We also analyze how the age of the mother and the time between two births change with the introduction of the policy. Subfigure 3 and 4 in Figure 2.2 show a relatively strong and significant reduction in the age of mothers giving birth and a not statistically significant decline in the birth interval.[14] More precisely, column 4 in Table 2.2 states that the age of the mother reduces by 0.474 in the introduction period, which corresponds to roughly 6 months. In light of this result, it is at first surprising that we do not find a fertility effect that is long lasting. Mothers who are deciding to have a planned child earlier are less likely to be affected of negative shocks to a partnership or their own health. However, in our setting the decision to have a child earlier does not translate into a long-run increase in fertility. This is very likely the result of the differential impact of fertility on high-income and LMIC mothers. The latter tend to be younger at birth (27.9 versus 29.2 years) so that opposite to the effect on birth weight, the effect on maternal age seems to be driven by a change in composition of mothers (see also Table A.5).

Finally, we also study the intensive and extensive margin of having a first child and having more children in Figure 2.5 and Table A.6. One might suspect that the intensive margin (i.e. having an additional child) would react more to a financial incentive as the marginal costs for children are decreasing. However, looking at point estimates, we see suggestive evidence that the fertility rate for the first child is slightly higher than for the second or third child. The latter is even totally unaffected by the policy. Though, none of these differences are statistically significant.

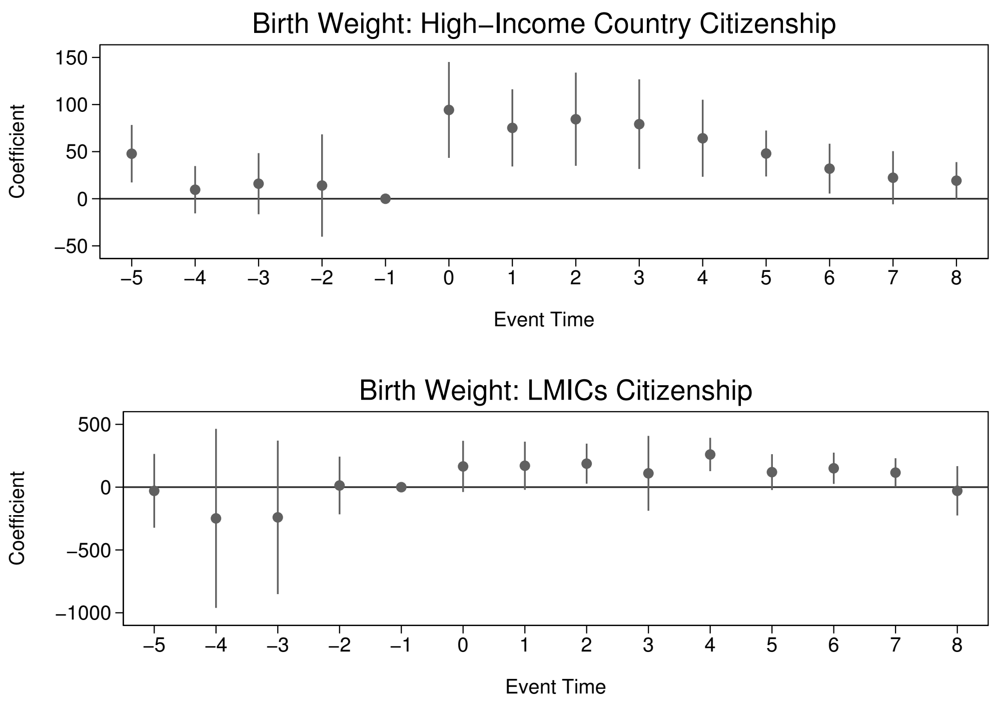
[14]Keep in mind that the measure of birth interval can only be calculated for higher order births excluding first births. This makes it harder to document a statistically significant effect.

FIGURE 2.3: Fertility Rate and Citizenship of Mother



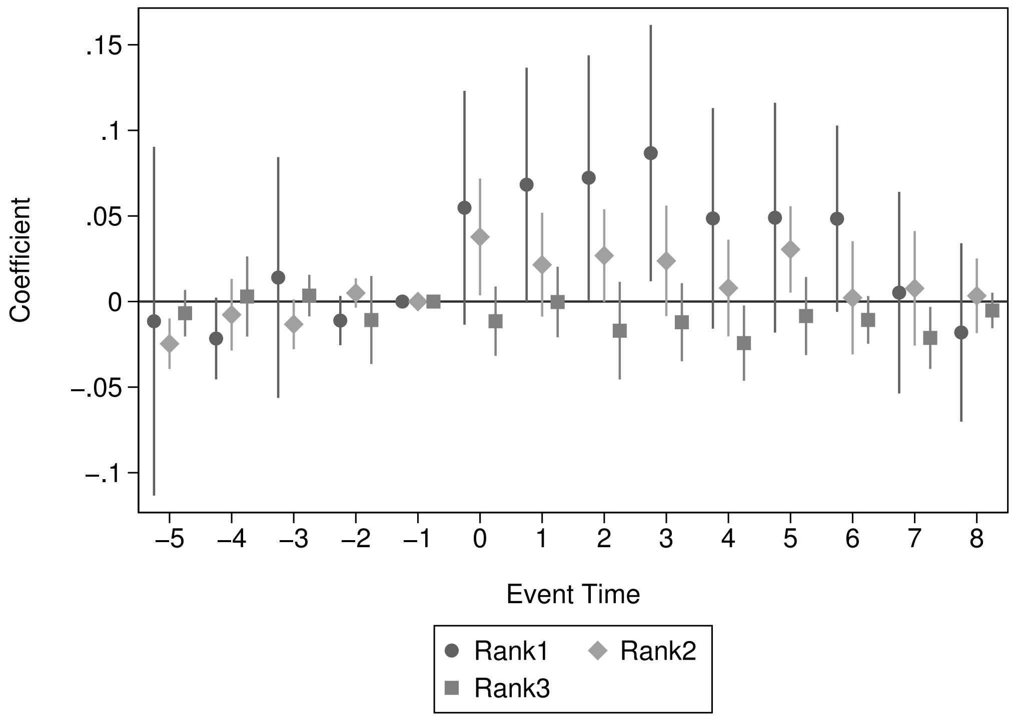
Notes: This figure shows how the effect of the introduction of the baby bonus on fertility by income country group changes over time. It depicts δ_τ from Equation (2.1) for $\tau \in [-5, 8]$ for the respective fertility rate. The event year represents the year relative to the introduction of the baby bonus. The dot shows the point estimate per event time, while the line corresponds to the 95% confidence interval.

FIGURE 2.4: Birth Weight and Citizenship of Mother



Notes: This figure shows how the effect of the introduction of the baby bonus on birth weight by income country group changes over time. It depicts δ_τ from Equation (2.1) for $\tau \in [-5, 8]$ for the respective birth weight. The event year represents the year relative to the introduction of the baby bonus. The dot shows the point estimate per event time, while the line corresponds to the 95% confidence interval.

FIGURE 2.5: Child Rank Specific Fertility Effect



Notes: This figure shows how the effect of the introduction of the baby bonus on fertility by child rank changes over time. It depicts δ_τ from Equation (2.1) for $\tau \in [-5, 8]$ for the respective fertility rate. The event year represents the year relative to the introduction of the baby bonus. The dot shows the point estimate per event time, while the line corresponds to the 95% confidence interval.

2.6 Results: Birth-Scheduling

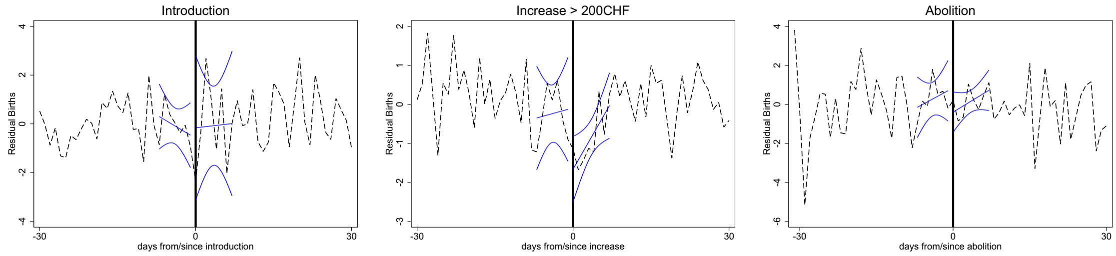
In Figure 2.6, we report the graphical results of our birth-scheduling analysis.

For none of the three events (introduction, increase, and abolition), there is a clear pattern around the policy change. This holds true for both specifications reported in Panel A (day-of-week fixed effects) and Panel B (day-of-year fixed effects), as well as for both the total count of births per day as shown in Figure 2.6 and the log of the outcome variable as shown in Figure A.3. If anything, there is a slight decline in births after an increase in the baby bonus of more than CHF 200.

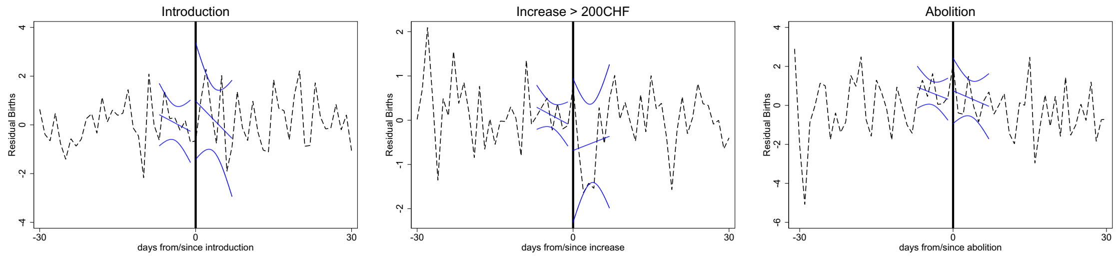
The are several possible reasons for the absence of birth-scheduling. First, daily birth counts per canton are small. This makes it hard to discover a statistically significant effect. While we study cantonal policy changes, previous literature analyzed national programs and thus national birth counts. Second, increases in birth allowances are much smaller as in other countries. Therefore, parents may not be willing to risk their child's health. Third, we also check newspaper articles for media coverage of the baby bonus policy changes. We search for articles about birth allowances on Factiva, one of the most

FIGURE 2.6: Birth-Scheduling Event Study

Panel A: Controlling for day-of-week fixed effects



Panel B: Controlling for day-of-year fixed effects



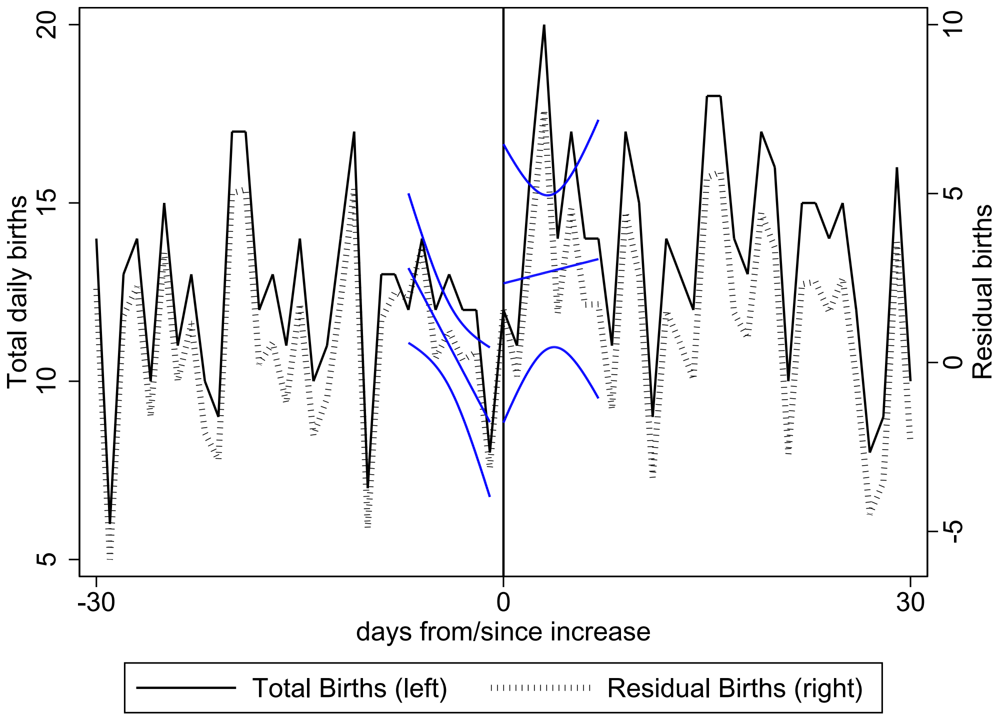
Notes: This figure shows the residuals (in dashed black line) from a linear prediction of estimating Equation (2.2) of total births per day and a linear fit including a 95% confidence interval (in blue) for the week before and after the policy change. Panel A reports the residual when ζ_i controls for day-of-week fixed effects and Panel B when ζ_i controls for day-of-year fixed effects. The three event studies combine either all introductions, increases above CHF 200, or abolition across cantons and time.

important database for press, company, and business information and the archives of several newspapers.[15] The results of this search are depicted graphically in Figure A.4. There was substantial media coverage on the Swiss baby bonus. Most of these articles describe differences across cantons or recently implemented changes in the payment structure. However, only roughly 25 percent of all relevant articles cover changes in the baby bonus scheme that are going to be implemented in the near future. This is especially true for increases. Thus, it is not surprising that birth-scheduling — a short-term behavioral change — is not taking place.

There is is one exception. In 2012, the canton of Geneva doubled the amount from CHF 1,000 to CHF 2,000. This increase was the result of a cantonal referendum initiated by the political left and led to a lot of discussion and widespread information exchange. Thus, we look in more detail at this specific increase, which is depicted in Figure 2.7. Both the total number of births and the residuals demonstrate that there is an increase in births happening after the policy change. However, the baseline daily count of births in Geneva is low and therefore the results are not significant.

[15]These newspapers include *Tages Anzeiger*, *Neue Zürcher Zeitung*, *Blick*, *St. Galler Tagblatt*, *24heures*, and *Le Temps*.

FIGURE 2.7: Birth-Scheduling Geneva: Policy Change on January 1st 2012



Notes: This figure shows the daily count of total births (left axis) and the residuals (right axis) from a linear prediction of estimating Equation (2.2) of total births per day. Additionally, a linear fit is added including a 95% confidence interval (in blue) for the week before and after the policy change. ζ_i controls for day-of-week fixed effects for the canton of Geneva. The 60 day window reports the 30 days pre- and post policy change (black vertical line) on January 1st 2012, where the baby bonus got doubled from CHF 1,000 to CHF 2,000 due to a public initiative.

Summarizing the analysis of birth-scheduling, there seem to be only marginal behavioral effects. This is likely the result of low daily birth counts and small changes in the payment structure. In the specific case of Geneva in 2012, where the baby bonus was doubled to CHF 2,000, graphical analysis suggests a postponement of births. But even in this case, we might not have enough power to document a significant effect because of the small amount of daily births.

2.7 Conclusion

We exploit a unique quasi-experimental setting in Switzerland that allows us to study the effect of birth allowances on fertility, newborn health, and birth-scheduling. In Switzerland, cantons are free to implement birth allowances. This gives rise to a lot of cantonal variation over time, which we use in an event study setting. Based on administrative data, we analyze various outcome measures—fertility rate, crude birth rate, birth weight, share of low-birth weight, infant deaths, stillbirths and the sex-ratio as a proxy for miscarriages.

To study birth-scheduling, we use a graphical time-series event study analysis based on daily birth counts. We base the birth-scheduling analysis on the fact that children were already in the womb when the policy introduction was announced. Thus, couples cannot react by becoming new parents, but they might shift the birth date to receive the baby bonus.

We do not find evidence for birth-scheduling. We argue that this results from several features, such as only minor changes in the amount and a small sample size due to a low cantonal birth count per day.

Looking at the effect of the baby bonus on fertility and newborn health, we find that the fertility rate and the birth weight increase. All significant effects are declining over time. While the effect on birth weight fades out but stays significantly different from zero, the effect on fertility is transitory. Furthermore, we find that mothers with nationalities from LMICs drive the fertility effect. We argue that this is an approximation for socio-economic status. The often fairly small baby bonus might still be an incentive for mothers with a low socio-economic status. Surprisingly, children from these mothers have a slightly higher birth weight than children from high-income country mothers. One might therefore speculate that the birth weight effect is driven by the former mothers' selection into birth. However, the subgroup analysis shows that birth weight does not significantly increase for women from LMICs, but for women with citizenship from high-income countries. While we cannot provide conclusive evidence that the allowance itself leads directly to better newborn health, this finding rules out that compositional changes are the sole driver of the effect on birth weight. Finally, we find that the effect on birth weight is especially strong at the lower end of the distribution documented by a significant decline in the occurrence of low birth weight.

Other health outcomes, such as infant deaths, stillbirths and the sex ratio do not show significant changes. We argue that those are severe measures and unlikely to be affected by a small financial bonus in a rich economy.

In terms of the fertility rate, we do not find significantly different effects at the intensive or extensive margin. However, we see suggestive evidence that the point estimate for the fertility rate of the first child is slightly higher than for the second child, while the effect of the fertility rate of the second child is significantly different from 0. The fertility rate for three or more children does not react to the policy.

Compared to other countries, the Swiss baby bonus is a cheap intervention to temporarily increase fertility and permanently improve newborn health. Importantly, other studies (Almond et al., 2018; Schwandt, 2018; Carneiro et al., 2015) suggest that health measures at birth translate into meaningful later life outcomes such as higher earnings or lower welfare dependence. Thus, the efficiency of the program might be underestimated by only studying outcomes visible at birth.

Our results entail important policy advice. Providing a birth allowance can increase fertility, even when the effect is not long lasting. Finally, an unintended, yet beneficial side-effect of this policy is the positive impact on newborn health.

Chapter 3

With Booze, you Lose: The Mortality Effects of Early Retirement

Abstract This study analyzes the effect of introducing early public pension schemes on male mortality. I exploit two reforms in Switzerland, which allowed men as of a certain cohort to retire one and two years before the statutory retirement age. This generates two sharp eligibility cutoff dates, which I use in a regression discontinuity design. I draw from two full sample administrative data sets: the mortality and the old age insurance register. Allowing men to retire two years before the statutory retirement age increases overall mortality by roughly 1.1 percentage points. The mortality effect is mainly driven by lifestyle diseases such as alcohol dependence and respiratory diseases related to smoking. The effects are largest for unmarried men and for men living in the German-speaking part of Switzerland who generally exhibit a stronger social norm towards work than men in the Latin-speaking part. Also, the effect is most pronounced in the middle of the income distribution. The results favor the lifestyle hypothesis, suggesting that retirement can increases mortality due to a loss of structure and a concomitant unhealthy lifestyle.

Keywords: early retirement, health behavior;
JEL classification: I18, J26

I am grateful for advice and inspiration by Monika Bütler, Stefan Staubli, Benedikt Lennart, Aurélien Sallin, Nathan Hendren, Marie Beigelman, Caroline Chuard, and Guido Cozzi. I also like to thank the participants of the PhD Seminar in St. Gallen, the ESPE 2021 conference, the Scottish Economic Society Conference, the Swiss Society of Economics and Statistics, the 8th IZA Workshop on Environment, Health and Labor Markets, the 7th IRDES-DAUPHINE Workshop on Applied Health Economics and Policy Evaluation, and the IHEA 2021 conference. Further, I thank the Swiss Federal Office of Statistics and the Central Compensation Office for providing the data. Any errors are my own.

3.1 Introduction

Retirement brings a great chunk of free time. Time to fill with inspiring and gratifying activities beyond the vicissitudes of day-to-day work. But retirement might cut both ways: While it can relieve retirees from strenuous and potentially harmful work, it also bears the risk that retirees suffer from a loss of structure and take up an unhealthy lifestyle to cope with the void.

Knowing how retirement affects health is not only important for public health, but also for the pension system. Demographic change will require reforms, such as increasing the statutory retirement age, to ease financial pressure of the pension system. If retiring later in life increases life expectancy, reforms that increase the retirement age have to be even stronger to compensate for the longer annuity period. On the other hand, if retiring later deteriorates health, an increase in health care costs could offset savings in pension.

In recent years, several compelling studies greatly advanced the literature on the health effects of retirement (see Kuhn (2018) for a concise overview). The results vary—even when using an objective health measure such as mortality. For example, Kuhn et al. (2020), Fitzpatrick & Moore (2018), and Blake & Garrouste (2013) find an increase in mortality due to (early) retirement, while Hagen (2018) and Hernaes et al. (2013) find no effect, Bloemen et al. (2017) and Hallberg et al. (2015) and find a decrease.

An explanation of those diverging results might be that retirement is not a clean-cut treatment, but rather a package of changes—a package that can differ from person to person and from country to country. Retirement not only increases the amount of *leisure* time but also decreases available *income*. When detecting an effect of retirement on health, it is vital to know what part of the retirement package triggered the effect. Is the gained leisure time used for harmful activities, as advocates of the lifestyle hypothesis suggest? Or is it the loss in income that deteriorates health, as implied by the income hypothesis? Another issue might be that effects on mortality materialize after a certain time and that some studies are not (yet) able to make statements on the long-term effect.

In this paper, I study the causal effect of introducing early retirement policies on mortality of men. Moreover, I test for empirical support of the lifetime or income hypothesis. To do so, I analyze effect heterogeneity by causes of death, geography, civil status, and lifetime income to shed light on the underlying mechanisms. In addition, I analyze how the mortality effect develops over time, since the follow-up period is almost two decades.

I address endogeneity by implementing a regression discontinuity design (RDD) and exploiting two reforms in Switzerland that allowed men as of a certain cohort to withdraw public pension one and two years before the statutory retirement age. The first reform allowed men born in the year 1933 to opt for early pension withdrawal in the year 1997—one year before the statutory retirement age; at 64 instead of 65. The second reform allowed men born in the year 1938 to opt for early pension withdrawal in the year 2001, two years before the statutory retirement age, at 63 instead of 65. Those two reforms generate two sharp cutoffs at the end of the birth years 1932 and 1937. The assignment to the policy is therefore as good as random around the birth cutoff.

I link two full sample administrative data sets. The first data set is the death register covering the universe of deaths and its causes in Switzerland. The second data set is the social security earnings register, which includes information about retirement, life-time and public pension income.

The results show a strong and significant increase in mortality for the reform that granted men access to retire two years before the statutory retirement age: Overall mortality until the age of 81 increases by roughly 1.1 percentage point, the net effect of the policy. The mortality effect of the reform that allowed to retire one year earlier is similar in size, but the estimate is not significantly different from zero. One reason for this divergence is likely the low take-up rate for one year of early retirement (only around 2% of the cohort). Therefore, the focus lies on the reform that introduced two years of early retirement with a take-up rate of a little over 4%. Interestingly, the mortality effect already kicks in—at least to a small degree—in the years after the announcement of the policy, notably before men could actually draw early public pension. This could be because of anticipation effects or because the availability of early public pension withdrawal triggers men to retire even earlier, for example by private or occupational pension.

The results survive a battery of robustness tests. I do not find any effects at other year and random cutoffs, nor do I find an effect for women at the reform cutoff cohort. This is reassuring because women around the cohort cutoff were not affected by the policy. Also, there is no effect on mortality before the announcement of the policy. Furthermore, the density in births around the cutoff is smooth.

The mechanisms analysis behind the mortality effect reveals interesting parallels to the *deaths of despair* as mentioned by Case & Deaton (2015) in the US in terms of *who* is affected (white middle-class men), the *causes of death* (alcohol), *civil status* (single men) and *timing* (around the year 2000). Specifically, the findings suggest that early retirement decreases life expectancy due to a loss of structure, followed by unhealthy coping behavior.

Looking at mortality causes and concomitant diseases, lifestyle diseases, especially alcohol dependence and (marginally significant) chronic airways obstruction (COPD), reveals a severe increase at the eligibility cutoff. This does not imply that the policy led men to drink or smoke themselves to death. The ultimate cause of death is often opaque.

What it tells us is that those deaths are *associated* with diseases that mirror unhealthy lifestyle behavior. The finding that alcohol consumption plays an important role during retirement is not surprising. Several studies document an increase in (harmful) alcohol consumption around the retirement age (Zins et al., 2011; Zantinge et al., 2014; Wang et al., 2014; Halonen et al., 2017). Also, for Switzerland, the Swiss Federal Office of Statistics reports that alcohol consumption increases sharply around the age cutoff 65, which is the standard retirement cutoff for men (BFS, 2019). While «only» 19 % of the male respondents in the age group 55 to 64 report to drink alcohol daily, this share is 34% for the men in the age group 65 to 74. Furthermore, the share of men with harmful chronic alcohol consumption is highest for this age group shortly after retirement. At the same time, causes of deaths that are less likely to be affected by lifestyle behavior, such as accidents or infectious diseases, do not change at the cutoff. Furthermore, single men are more likely to have an increased mortality when retiring earlier. This is in line with Richard et al. (2017), who find that loneliness is related to unhealthy lifestyle and worse health in Switzerland. Culture seems to be another important factor: The mortality effect is mostly driven by men living in the German-speaking part of Switzerland. Previous investigations propose that the social norm towards work is stronger in the German than in the Latin culture, therefore, ending work is likely to have more negative consequences in the German area (Eugster et al., 2011). Although it is very hard to disentangle income from lifestyle effects, I do not find evidence that low lifetime income is the core driver of the increase in mortality. If anything, the effect is strongest in the middle of the income distribution, which indicates that losing income plays less of a role.

I structure the rest of this paper as follows: Section 2 describes the institutional setting in Switzerland and provides an overview of the data. Section 3 lays out the empirical strategy used to identify the effect. Section 4 shows the main results and provides several robustness checks. Section 5 looks at underlying mechanisms and the last section concludes.

3.2 Literature

Several studies investigate the health effects of retirement—which underpins the relevance of this research question. However, results are quite mixed, ranging from beneficial, to neutral, to harmful. Potential reasons for the contradictory results might be that retirement is not a homogeneous, standardized treatment but a bundle whose content depends on a myriad of factors, such as age of the retirement, social status, network, and many other factors. Also, the way how the dependent variable health is defined and the identification strategy might play a role.[1]

[1] Kuhn (2018) gives a concise overview of the literature and theory of the health effect of retirement.

From a methodological point of view, many studies use an instrumental variable approach to tackle endogeneity of the retirement decision with the predicted probability to retire. Several other studies use a fuzzy regression discontinuity design with age as the running variable and the statutory retirement age as the cutoff (Fitzpatrick & Moore, 2018; Eibich, 2015; Müller & Shaikh, 2018). Certainly, those RDDs provide a credible identification strategy for short-term effects of early retirement on health. However, it is also important to look at the long-term effects of early retirement, since it is conceivable that many health effects manifest after a certain time—especially when due to changes in lifestyle behavior in retirement.

The outcome variable *health* can be classified as objective or subjective: Objective measures, such as mortality, hospitalization or illnesses, and subjective measures for which individuals are surveyed on the perception of their health. Even when looking at objective measures of health, the results are ambiguous. Probably most related to this paper is the work by Kuhn et al. (2020). They study a policy change in Austria that allowed workers in some regions to exit the labor force three years earlier. They find that blue-collar men are more likely to die before the age of 67, but do not find an effect for blue-collar women. Further, they estimate that an additional year in early retirement increases the risk of death before age 73 by 1.47 percentage points. Different from my study, many workers were pushed involuntarily into retirement. Fitzpatrick & Moore (2018) look at short-term mortality effects in the US. Using a RDD with age as the running variable, they find a discontinuous change in mortality at the US social security eligibility age 62 of 2% for males. Some of their additional analyzes suggest that the increase in male mortality is connected to associated lifestyle changes. Hernaes et al. (2013) study several reforms in Norway and find no effect of retirement on mortality.

Several studies exploit reforms for certain parts of the population. Bloemen et al. (2017) look at targeted incentive for civil servants to retire early. Similarly, Hallberg et al. (2015) look at male military officers. Both studies analyzing subpopulations find that retirement decreases mortality. Blake & Garrouste (2013) study private sector employees in France and discover that retirement increases mortality. Hagen (2018) looks at the health consequences of a two-year increase in the statutory retirement age of local government workers in Sweden and finds that the reform had no impact on mortality. Certainly, those results are not automatically valid for the entire population.

The impact on other objective health measures, such as hospital visits or health behavior, remains unclear. Behncke (2012) uses non-parametric matching and instrumental variables approach to identify the effect of retirement on health measures in the UK. She finds that retirement increases the risk of being diagnosed with a chronic condition. Specifically, it raises the risk of severe cardiovascular disease and cancer. Also, retirees have a higher risk of developing the metabolic syndrome, which is considered as an important risk factor for both cardiovascular diseases and cancer. On the other hand, Insler (2014) and Eibich (2015) find that positive health behavior increases because of retirement. In a recent study, Rose (2020) finds no immediate effect of retirement on cognitive ability, health care utilization, or mortality. For Denmark, Nielsen (2019) shows that early retirement leads to decreases in GP visits and hospitalizations of 8–10% in the short run, but has no effect on mortality. Heller-Sahlgren (2017) looks at short and long run mental health effects using retirement age thresholds. The results show no short-term effects of retirement on mental health, but a large negative longer-term impact. Fé & Hollingsworth (2016) find that retirement opens the gate to a sedentary life with an impoverished social component.

Subjective health measures should be considered with care. The literature finds mostly beneficial effects of retirement on health when health is measured subjectively (Eibich, 2015; Johnston & Lee, 2009; Insler, 2014). Mazzonna & Peracchi (2017) document an increases the age-related decline of health and cognitive abilities for most workers and find evidence of a positive immediate effect of retirement for workers with physically demanding occupations. Müller & Shaikh (2018) states that retirement affects own health positively, while the own retirement affects the health of the spouse negatively. Heller-Sahlgren (2017) finds no negative effect on mental health in the short-run, but does so in the long-run. Subjective health measures are prone to justification bias—the tendency of humans to justify their decision by denying potential negative consequences. For example, Johnston & Lee (2009) find different effects for objective and subjective health measures, even when using the same identification strategy and the same data. Thus, subjective health status should be looked at carefully. As stated by Kuhn (2018), mortality and its causes are the preferred measures available. Mainly, because there are many potential channels through which retirement affects health, and thus broader measures such as mortality should be preferred.

I add to the literature in several ways. First, I use a credible identification strategy allowing to differentiate short- and long-term effects: a regression discontinuity design (RDD) around a random date of birth cutoff. Most studies that use credible regression discontinuity designs use the default retirement age as a cutoff. This only grants to investigate mortality effects in a window around the cutoff. While this is a plausible identification strategy within a short-term window, it misses mortality effects that acquire later in life. Given that lifestyle behaviors can have irremediable health consequences that only materialize after a certain time lag, I argue that this is an important feature. Importantly, my time period is long enough to identify long-term effects, as the reforms took place in 1997 and 2001 and my data ends in 2019. Second, I use an objective health measure: administrative mortality data. Compared to the often used survey data, this has the advantage that it does not suffer from justification bias, measurement error, or sampling issues. Third, the unique cultural setting of Switzerland, which inherits the border of the two largest cultural groups in Europe (German and Latin), allows to look at heterogeneous effects across cultures. Those groups are especially interesting because it has been documented that the norm towards work is considerably higher in the German-part of Switzerland. Thus, losing work could therefore play an important mitigator. If the effects are different within the same setting, the whole literature on health effects of retirement must be cautious when transferring findings from country to country. Fourth, analyzing two reforms with different dosages of early retirement within the same setting allows to shed light on heterogeneous effects regarding the length of early retirement. Sixth, I use a reform applicable for, at least, all men. Many other studies focus on specific groups, such as male army officers or civil servants (Bloemen et al., 2017; Hallberg et al., 2015; Blake & Garrouste, 2013).

3.3 Institutional Setting and Data

3.3.1 Public Pension in Switzerland

The Swiss Old Age Insurance System offers a full pension to anyone reaching the statutory retirement age (SRA). For men, the statutory retirement age (SRA) is set at 65, for women at 64. During work-life, people contribute to the pay-as-you go pension system by paying social security taxes of 8.4% of their wage. Both employees and employers are required to pay contributions. Employee contributions are deducted directly from the salary. This contribution requirement starts from the age of 20 until the SRA. One year without contribution leads to a reduction in pension of 2.3% (1/44). Individuals without gaps in their contribution history receive a pension between roughly 14,000 CHF if average earnings are lower than 14,000 CHF and a maximum of approximately 28,000 CHF, if average earnings are higher or equal 84,000 CHF.

In the year 1997 and 2001, two policy changes were introduced which were part of a larger reform (*10th AHV reform*). Those policy changes allowed men to draw early pension before the SRA. In 1997, men born after December 31, 1932 were allowed to withdraw public pension at age 64 instead of 65. In 2001, men born after December 31, 1937 were allowed to take early public pension at the age of 63. The reform was known to the public several years in advance. This is because the reform was subject to a public mandatory referendum. On June 25, 1995, 60.7% of the Swiss population voted in favor of the new law. Retiring one year before the statutory retirement age comes at an actuarial reduction of the pension benefit of 6.8% per year. In case of retiring two years earlier, this amounts to 13.6%.

The *10th AHV reform* was mainly known to increase the SRA for women. More specifically, the female SRA was increased from 62 to 63 for women born as of 1939 and to

64 for women born as of 1942.[2] This raises the question of whether male mortality is influenced by this part of the policy reform. However, control and treatment cohorts are different for the male and the female policy change. Specifically, the year of birth cutoffs are different. This justifies the use of women as a placebo test around the cutoff in the robustness section.

Nevertheless, spillovers from wife to husband might still be possible, for example, due to the lower pension income or the decreased leisure time. This would especially be true if each spouse was affected by a reform because of their age difference. However, there are some empirical facts that ease concerns about a strong influence of this channel. First, the mean (and mode) of the age difference between spouses is three years. If the female reform affected male mortality one would expect to see the strongest discontinuity in male mortality at other cutoff years.[3] Second, the mortality effect is most pronounced for unmarried men. If the increase in the female statutory retirement age indeed influenced men, it would have increased health of the husband. Since the major changes with the reform are less leisure time and lower income, it is hard to imagine how this would have had a beneficial effect on their husband's health. Third, previous research shows that there is no influence of the (female) reform on male labor supply. Lalive & Staubli (2015) study the effects of this increase in retirement age on female labor supply and mortality. They find a significant effect of female labor supply, but no significant effect on female mortality. Importantly, they also do not no find an effect on male labor supply, suggesting that the retirement decision is not driven by the increase in the statutory retirement age of women. Forth, looking at discontinuities for female mortality at end-of-year cutoffs does not show any significant effect (see Figure B.2). The channel that female health was affected and, in turn, influenced male health is thus rather unlikely.

Throughout the paper, I refer to retiring at 63 instead of 65, as retiring two years before the SRA. When retiring at 63 was introduced in 2001, men were already able to draw public pension at 64 and thus one year of early retirement. One could argue that the reform that allowed men to retire at 63 only decreased eligible retirement age for one year. This would be true if the treatments are considered increasing linearly in intensity. However, it is more sensible to look at the two reforms as two distinct treatments instead: *Retiring at 63 (ER 63)* and *Retiring at 64 (ER 64)*.

TABLE 3.1: Summary Statistics

	ER 64		ER 63	
Cohort:	***1932***	***1933***	***1937***	***1938***
	(1)	(2)	(3)	(4)
Single	0.099	0.098	0.112	0.126
Married	0.675	0.683	0.681	0.664
Widowed	0.138	0.130	0.086	0.078
Divorced	0.088	0.089	0.121	0.132
German Language	0.718	0.718	0.701	0.705
French Language	0.230	0.228	0.241	0.236
Italian Language	0.049	0.050	0.054	0.054
1 Year Early Retirement	0.000	0.023	0.043	0.031
2 Years Early Retirement	0.000	0.000	0.000	0.042
Real Relevant Income (2017 CHF)	57,984	58,893	55,829	55,827
Size of Monthly Public Pension (2017 CHF)	1,185	1,193	1,092	1,068
Observations	56,899	57,937	57,122	58,677

Notes: This table shows summary statistics for available characteristics for the cohorts around the cutoff. Columns (1) and (2) look at the reform cutoff ER 64, columns (3) and (4) at the reform cutoff ER 63.

3.3.2 Data

I link two full administrative data sets. The first data set is the retirement register. It contains information on income during work-life, public pension benefits, and early pension withdrawals. It is provided by the Central Compensation Office, which is Switzerland's central implementing body for first-pillar social security. The second data set is the mortality register. It covers all deaths in Switzerland since 1969, including information on cause of death (ICD-10 codes), date of birth and municipality of residence. For the underlying research question, cohorts from 1930 to 1955 and observations years from 1990 to 2019 have been kindly provided by the Federal Office of Statistics.

In addition, I use aggregated census data issued by the Federal Office of Statistics to measure the number of men alive in the year 1990 per day of birth. Since I want to have relative mortality measure, I include only men born in Switzerland in the sample. This is because the number of births from the 1990 census data is also by men born in Switzerland. Otherwise, I would include immigrants in the nominator, but not in the denominator, which would then artificially increase the mortality rate. This is especially important because Switzerland experienced a sizable immigration starting in the 1990s.

[2]Women were still allowed to retire earlier at a (reduced) yearly reduction of 3.4% of their pension benefits.

[3]For the «female 1939 reform» at around 1936/1937 and for the «female 1942 reform» at 1939/1940.

Table 6.1 shows a summary statistic of the men opting for early public retirement for both reforms and for the cohorts above and below the cutoff. It shows that for the ER 63 reform, over 4.2% of the male population in Switzerland opts for early retirement. For retirement at 64, the share is smaller with 2.3%. Figure B.4 in the Appendix shows the share of men opting for two and one years of early retirement as a function of date of birth. By the rule of law, no men take up early retirement before the eligibility cutoff. Thus, there is perfect non-compliance before the cutoff.

How well does drawing early public pension withdrawal coincide with early retirement? Although pension withdrawal does not force people to stop working, the incentives are such that it is unfavorable to continue working once early retirement benefits are drawn. If individuals continue working, they are forced to pay social security taxes—even though they will never profit.[4] Under the occupational pension scheme, it is possible to take early retirement (starting from the age of 58). This is the second pillar of the Swiss pension system.

Another question is, how well drawing early public pension coincides with reacting to the policy. The announcement of introducing early public pension can trigger other responses as well. For example, in 1995 men could have decided to retire two years earlier in 2001, already adapting their lifestyle, and in the end still decide not to opt for early retirement, maybe because their firm strongly needed them (unemployment was very low in 2001). Therefore, those men would not show up in the official «complier share» but could still affect mortality. In econometric terms, this would violate the exclusion restriction. The instrument (early retirement policy) would not only affect mortality exclusively via retiring two years earlier, but also via other channels. I will elaborate on this issue in the next section. This makes it difficult to estimate a treatment effect of early retirement itself and I will therefore estimate the net effect of the policy instead.

Nevertheless, if one assumes that at least a considerable part of the net effect of the policy is due to early retirement, it is still important to know who those compliers are. Thus, Table 3.2 provides an overview in terms of income. The cohorts (4 birth years) that are included are within the bandwidth that is roughly chosen by the data driven bandwidth selection in the results section. When looking at the mean lifetime income, it is interesting that those drawing early public pension have a higher incomes—even though they miss one or two years to accumulate income. Also, when looking at the income quintiles, those opting for early retirement are more likely to be in the top three quintiles. To sum up, in terms of socio-economic background, men opting for early retirement are more likely in the upper half of the income distribution.

[4]There is an amount of exception: Until 1,400 Swiss Francs per month; individuals are not required to pay taxes.

TABLE 3.2: Complier Description Income

	(1)	(2)	(3)	(4)
	ER 64		ER 63	
	Cohorts: 1933-1936		Cohorts: 1938-1941	
	Not Early	***Early 1 yr***	***Not Early***	***Early 2 yrs***
Dead until 80	0.258	0.305	0.238	0.315
Relevant Real Income (CHF)	56,339	63,356	56,120	59,720
Inc Q1 (low)	0.205	0.088	0.206	0.123
Inc Q2	0.203	0.174	0.204	0.191
Inc Q3	0.198	0.217	0.197	0.222
Inc Q4	0.198	0.270	0.195	0.244
Inc Q5 (high)	0.195	0.251	0.197	0.219
Size of Monthly Pension Income (CHF / month)	1,147	1,291	1,089	1,052
Observations	176,350	3,527	187,963	8,420

Notes: This table describes the socio-economic differences of men retiring early vs men retiring not early. It does so for the cohorts that are within the bandwidth in the regression discontinuity design.

Table 3.2 also shows that those men opting for early retirement have a higher probability of being dead before they reach the age of 80, which is usually referred to as premature death. Men taking one year of early retirement have an 18 percent higher probability of being dead until the age of 80. Men taking two years of early retirement have a 32 percent higher probability. Certainly, this is not a causal effect. Men opting for early retirement are likely different from men retiring at the statutory retirement age. However, at least in terms of income, it is not those with the lowest income that retire most often early. If one assumes that health is positively correlated with income, this would mean that healthier men opt for early retirement.

3.4 Empirical Strategy

3.4.1 Identification

Simply regressing mortality on early retirement is likely to yield biased results. Reversed causality is one bias: Whether one retires early might itself depend on health status and subsequently on mortality. If unhealthier people are more likely to retire early, I would overestimate the effect of retirement on mortality. The results would also be biased because of omitted variables: Early retirees are likely to differ in characteristics that influence both health and the retirement decision.

To circumvent those biases, I use exogenous variation induced by two policy changes. Whether men were eligible for early retirement changed discontinuously at a certain date of birth. At the extreme, if a man was born a few seconds later at New Year's Eve 1932/33 or 1937/1938, he was eligible for early pension withdrawal.

In principle, two effects can be of interest. First, the effect of the policy introduction on overall mortality: Does the introduction of the policy affect mortality? This makes up a sharp regression discontinuity design (SRD). Second, the effect of early retirement on mortality: How does early retirement affect mortality? Here, the policy serves as an instrument for early retirement and the research design constitutes of a fuzzy regression discontinuity design (FRD). The two effects are closely linked. The difference is that in the FRD the effect of the policy on mortality is divided by the share of men reacting to the policy (i.e. the first stage). However, as I will explain below I will estimate a reduced form effect because it hard to estimate a precise LATE and because the policy relevance of a LATE is questionable in this setting with (most likely) strong effect heterogeneity.

The identifying assumptions are distinct for the SRD and the FRD, in the sense that FRD requires more assumptions than SRD. The first assumption is the *stable unit treatment value assumption* (SUTVA). It states that the potential outcomes for each person are unrelated to the treatment status of other individuals. If spillovers are present, this assumption can be violated. In my study this would be the case if the early retirement (eligibility) affected the mortality of the control group, for example due to general equilibrium effects. However, it is unlikely that large general equilibrium effects are present because around the treatment cutoff only one male cohort was eligible and only around 5% of men actually opted for early retirement.

The second assumption is the *continuity of conditional regression function*. This assumption implies that the running variable (date of birth) can be related to the outcome variable, but its association has to be smooth. This is the case when using mortality as an outcome and date of birth as a running variable, because mortality is on average a continuous function of age. It also means that absent the introduction of the policy, mortality would not have changed at the cutoff. Importantly, this requires that other determinants of mortality are not allowed to jump at the cutoff. Although I can never prove that there are not some unobserved determinants changing at the cutoff, there is to the best of my knowledge no other reform—for example in the realm of army, school, or health—related to this cutoff. Conveniently, I can use women to perform a placebo test at the cutoffs because they were not affected by the reform. I do not see an effect for women. Thus, if there are some unobserved determinants changing discontinuously at the cutoff, they would need to only affect men and not women. This certainly reduces the probability of a violation of the continuity assumption. Furthermore, graphical analysis shows that the outcome variable does not significantly jump at other end-of-year cutoffs and at random cutoffs (see Figure B.3 and Table B.1). Another way to challenge this assumption is to check whether other covariates are smooth (balanced) around the cutoff. However, covariates that are plausibly exogenous to the treatment are, such as schooling or income, are not available on a daily birth basis for the cohorts of interest.

The second assumption is the *continuity of conditional regression function*. This assumption implies that the running variable (date of birth) can be related to the outcome variable, but its association has to be smooth. This is the case when using mortality as an outcome and date of birth as a running variable, because mortality is on average a continuous function of age. It also means that absent the introduction of the policy, mortality would not have changed at the cutoff. Importantly, this requires that other determinants of mortality are not allowed to jump at the cutoff. Although I can never prove that there are not some unobserved determinants changing at the cutoff, there is to the best of my knowledge no other reform—for example in the realm of army, school, or health—related to this cutoff. Conveniently, I can use women to perform a placebo test at the cutoffs because they were not affected by the reform. I do not see an effect for women. Thus, if there are some unobserved determinants changing discontinuously at the cutoff, they would need to only affect men and not women. This certainly reduces the probability of a violation of the continuity assumption. Furthermore, graphical analysis shows that the outcome variable does not significantly jump at other end-of-year cutoffs and at random cutoffs (see Figure B.3 and Table B.1). Another way to challenge this assumption is to check whether other covariates are smooth (balanced) around the cutoff. However, covariates that are plausibly exogenous to the treatment, such as schooling or income, are not available on a daily birth basis for the cohorts of interest.

Another assumption is that the running variable cannot be *manipulated*. Here the running variable is date of birth. Thus, manipulation can be ruled out by construction. Nevertheless, Figure B.6 provides further evidence that the density is smooth around the cutoff.[5]

However, even if estimating a LATE was feasible, it is unclear what the benefits of such a result are. The resulting effect of the fuzzy regression discontinuity is a local average treatment effect (LATE). That is, the average effect of the policy on men that reacted to the policy. This does not necessarily translate to an average treatment effect (ATE), as the complier population is likely to differ in characteristics from the non-complier population (as shown empirically in the previous section). The LATE would only translate to an ATE if one assumes homogeneous effects—which would impose a very strong assumption. Therefore, for policy makers, it might be more useful to know what effect the

[5]Birth scheduling at the end-of-year cutoff has been documented. It is possible that parents at the end-of-year 1937 manipulated the date of birth of their offspring for some other reasons, e.g. school grade optimization. If this was the case, more people with date of birth in January should be observed. However, as my outcome variable is relative to the number of men with a certain day of birth, this bias is irrelevant in this context. Also, the density does not significantly change around the cutoff.

introduction of a certain policy has on the overall mortality in the population. So, the focus lies on the reduced form result.

3.4.2 Estimation

The primary goal of this study is to estimate the following reference model and retrieve the parameter τ which aims to capture the effect of early retirement on mortality.

$$Mortality_{itx} = \alpha + \tau EarlyRetirment_{itx} + u_{itx} \tag{3.1}$$

where i stands for individual, t for (calendar) time, and x for date of birth. The dependent variable $Mortality_{itx}$ measures the probability to die at calendar day t for each day of birth x. As explained above, τ is likely to be biased in this regression because the decision to retire early is likely $cov(EarlyRetirement_{itx}, u_{itx}) \neq 0$.

To circumvent endogeneity, I estimate the treatment effect within a regression discontinuity design (RDD). Let $\tilde{x}$ be the running variable *date of birth* centered around the cutoff dates x_0, precisely $\tilde{x} = x - x_0$. T_i indicates, if a man is born after the cutoff date and is therefore eligible for early retirement. This yields the following equation for the SRD:

$$Mortality_x = \beta_0 + \beta_1 \tilde{x} + \delta T_x + \beta_2 T_x \tilde{x} + \epsilon_x \tag{3.2}$$

where $Mortality_x$ measures the probability to be dead at the end of the data period per date of birth x. Specifically, for every date of birth, I calculate the sum of men dying between 1990 and 2019 and divide it by the number of men alive in 1990:

$$Mortality_x = \frac{1}{n_{x,1990}} \sum_{i \in x}^{t \leq 2019} \mathbb{1}[Death_i = 1]$$

Another measure for mortality would be the sum of deaths per day. However, I use this relative measure in the main specification to enhance robustness and facilitate interpretation.[6] It makes the measure even more robust to manipulation around the cutoff, such as birth scheduling, and helps against unlikely compositional changes (e.g. immigrants that due to administrative issues all got assigned the same date of birth). Also, it helps to interpret the effect, because the discontinuity in the outcome variable at the cutoff measures the increase in absolute risk of mortality. Importantly, the mortality measure I use differs from the traditional mortality measure, where the mortality rate is defined

[6] For simplicity, I refer to this measure as mortality rate. More precisely, it is at a «crude mortality rate» since the denominator is always relative to the year 1990

as the number of deaths scaled to the population size per unit of time. The mortality rate in this paper is always relative to the men alive in 1990. This is because precise population data is not available on a yearly basis for the relevant years in Switzerland. Further, to abstract from changes in population in-and outflow, I only use men born in Switzerland.

As suggested by Gelman & Imbens (2019), I use estimators based on local linear polynomials. To trade-off bias and variance in determining the bandwidth h, I use the data driven bandwidth selection method proposed by Calonico et al. (2014). Nevertheless, I will test whether the results are robust to shorter and longer bandwidths, as well as to higher order polynomials.

3.5 Results

3.5.1 Mortality Effect of the Policy

Figure 3.1 shows the men that died until 2019 per day of birth relative to the number of men alive in 1990. In Panel (a), the eligibility cutoff date is for two years early retirement, while Panel (b) looks at one year early retirement. The points are binned averages, while the line is a linear approximation within the data driven bandwidth.[7] Figure B.8 shows the same for different bin sizes and without visual guidance provided by the lines.

For two years of early retirement (ER 63), the discontinuity is clearly visible. At this end of birth year cutoff (1937/1938), mortality increases discontinuously. In Panel (b) (ER 64), the bins show almost no discontinuity, even though the global linear approximation detects a higher mortality after the cutoff.

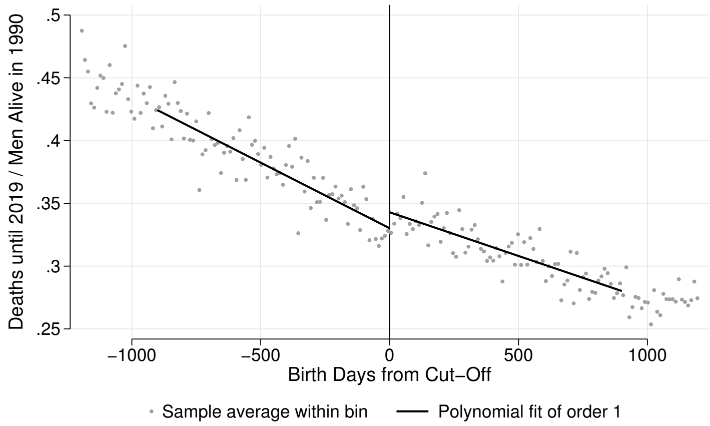
Table 3.3 provides estimates based on local linear regressions. Column (1) shows the main results of interest by looking at the relative mortality measures as specified in the previous chapter. Column (2) shows the absolute number of deaths to provide further insights on the raw data. The upper half of the Table shows the results for the ER 63 reform. The RD estimate states that the share of men dead until 2019 relative to men alive in 1990 increases by 1.1 percentage points. Figure B.5 in the Appendix shows that the result does not significantly change, when different bandwidths and orders of local polynomials are used. Baseline mortality right before the cutoff is 0.33. Thus, the relative net effect of the policy on male mortality is an increase of 3.3%. This is a strong increase —even though Switzerland has one of the highest life expectancies for men in the world (WorldBank, 2021).[8]

[7]The bins are determined by the mimicking variance evenly-spaced method.

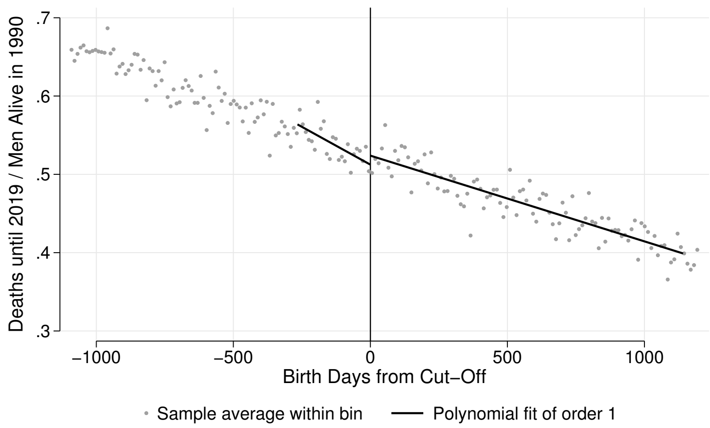
[8]If one would—despite the previously mentioned reservations—calculate a LATE with this estimate by assuming that 4.2% of men are compliers, one would estimate that two years of early retirement increases the risk to die before the age of 81 by 26 percentage points for those men that comply to the policy.

FIGURE 3.1: Regression Discontinuity Graphs

(A) ER 63



(B) ER 64



Notes: This figure shows the reduced form estimates for the two policy reforms: retiring at age 63 (a) and retiring at age 64 (b). The vertical axis measures the share of men dead until 2019 per day of birth x. The horizontal axis measures the days away from the cutoff. The linear approximation is based on the bandwidth chosen by the data driven bandwidth selection as stated in Table 3.3. Bin size is based on the mimicking-variance evenly spaced method. The line on the left side of the cutoff in Panel(b) is shorter because in the local linear approximation approach, a smaller bandwidth is selected since there is less data available before the cutoff.

As expected by visual inspection of Figure 3.1b, the discontinuity for the ER 64 is not significant. Nevertheless, the point estimate is positive and remarkably similar to the ER 63 reform which could indicate a discontinuity. This estimate might simply lack of power due to the relatively poor first stage.

Figure 3.2 shows how the effect for the ER 63 reform develops over time. The data period is split into five periods. The policy was announced in 1995 and introduced in 2001. Importantly, there is a precise zero in the pre-announcement period which can also be seen as placebo check for the identification strategy. The figure also shows that the point estimate during the anticipation period is positive, although not significantly different from zero. Bear in mind that sample size is reduced in this analysis, and the positive point estimate could indeed point towards anticipation effects. Interestingly, the mortality effect already kicks in in the first third (2001 to 2006) in the post period, but is highest—and significant—in the last third, when men are between 76 and 81 years old. The early kick in of the mortality effect around retirement is also something that has been observed in the US (Fitzpatrick & Moore, 2018).

Figure 3.2 shows how the mortality effect changes over time. Mortality is significantly higher within six years of the policy reform. Within the next six years, the point estimate becomes smaller and the effect insignificant. Within the last seven years of the sample period, the effect increases again. While it is surprising that the effect already manifests within a few years, it is consistent with Fitzpatrick & Moore (2018) who find increased mortality within a short period of time—after the default retirement age. The figure also shows how important it is to have a long follow up period to detect a (significant) mortality effect. If one had only data until 2006, one would not be able to find a significant effect.

3.5.2 Robustness

I perform several checks to assess the robustness of the increase in mortality due to the ER 63 reform. I refrain from analyzing the robustness of the ER 64 because—as explained above—there is likely not enough power to detect any effect. I begin with checking whether there is a discontinuity at the male policy eligibility cutoff for women. This serves as a placebo test because the mortality of women should not jump at this cutoff. Figure B.1 shows that there is no discontinuity around the 1937/1938 cutoff for women. Thus, if any other policy generated the jump at the cutoff, that policy can only have affected men.

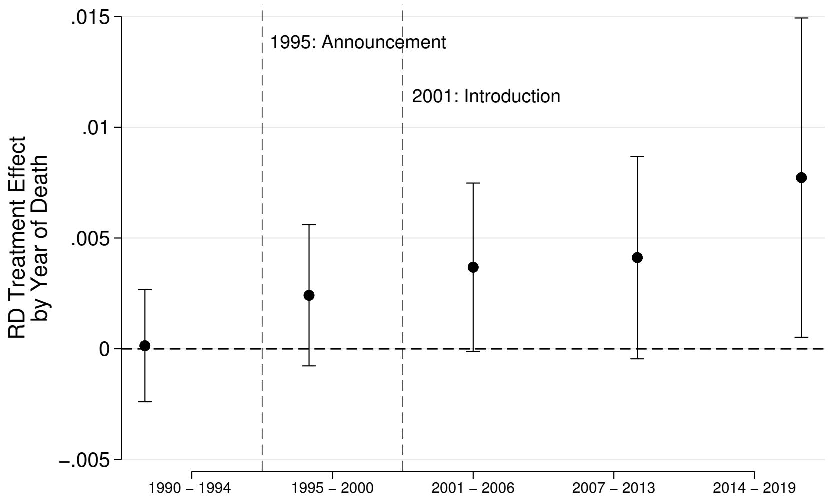
Then, I test if there is a discontinuity at other cutoffs. First, I focus on end of birth year cutoffs. The results are shown in Figure B.3. Apart from the end-of-year cutoff 1937/1938 (2 years of early retirement eligibility cutoff), there is no positive, significant

TABLE 3.3: Mortality Effects of the Policy

	(1)	(2)
Dep.Var.:	Mortality Rate (relative)	Number of Deaths per Day of Birth
ER 63		
RD Estimate	0.011**	3.219***
	(0.005)	(0.500)
Specifications:		
Mean of Dep. Var.	0.33	38.9
Kernel Type	Triangular	Triangular
Order Loc. Poly. (p)	1	1
BW Selection	mserd	mserd
BW left	901	1,262
BW right	901	1,262
N (deaths)	65,933	93,666
N (individuals)	192,412	271,133
ER 64		
RD Estimate	0.010	2.473***
	(0.009)	(0.569)
Specifications:		
Mean of Dep. Var.	0.53	51.01
Kernel Type	Triangular	Triangular
Order Loc. Poly. (p)	1	1
BW Selection	mserd	mserd
BW left	1,135	1,371
BW right	266	1,141
N (deaths)	75,756	129,164
N (individuals)	128,086	216,737

Notes: This table shows the overall effect of the policy introduction on mortality (reduced form). The upper half shows the RD estimate for the reform that introduced early retirement at 63, the lower half early retirement at 64. Column (1) shows the effect on the mortality rate of men day of birth relative to men alive in 1990 per day of birth. Column (2) shows the effect on the total number of deaths of men per day of birth. The number of included observations shows the total number of deaths *N(deaths)* and the number of individuals alive in 1990 *N(individuals)* within the data driven bandwidth. Standard errors are included in parentheses. (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$)

FIGURE 3.2: Mortality Effect by Year of Death (ER 63)



Notes: This figure shows when the mortality effects show up for the ER 63 reform. The policy was announced six years before the introduction of the policy. Specifications as indicated in Table 3.3

change. Second, I test for discontinuities at other random cutoffs. To do so, I use the method suggested by Imbens & Lemieux (2008). Thereby, I split the sample at the cutoff value x_0 and run the RDD estimation at the new cutoffs that are themselves the median of the split sample. Table B.1 shows the p-values of these artificial cutoffs. None of the values differ significantly from 0. Then, I split the sample again at the new median cutoffs and create another cutoff point at the median of the new samples. At those newly created four cutoffs, there is no significant discontinuity.

Although manipulation of the birthday is complicated, it could nevertheless be that there are more births at the beginning of the year, which would then also at the margin lead to more deaths after the cutoff. To alleviate those concerns, Figure B.6 checks if the density differs before and after the cutoff (Calonico et al., 2014). For both reform cutoffs, I cannot reject the hypothesis that densities are the same around the cutoff. Thus, there is unlikely to be some sort of manipulation.

In addition, Figure B.5 shows the results for different bandwidths and orders of local polynomials. While the estimate decreases with a higher bandwidth, it is never significantly different from the main specification and almost always significantly higher than zero.

Finally, I check when the mortality effect manifests. If it is indeed the early retirement treatment that triggers the mortality effect, no mortality effect should be visible before the treatment starts or before it was announced. Figure 3.2 measures how the effect varies along years relative to the reform. One can see that the mortality effect does not start before the implementation year of the policy and is significant thereafter.

3.6 Mechanisms

3.6.1 Diseases

Data from the death register allows to analyze which diseases are behind the increase in mortality. For each death, I observe the cause of death and concomitant diseases as ICD-10 codes. I do not distinguish between cause of death and concomitant diseases because this is often arbitrary and causes of death as well as concomitant disease are informative. For example, when categorizing a death related to alcohol diseases, this is true if alcohol related ICD-10 codes occur as a cause of death or concomitant disease. For each of the disease categories, I run a regression discontinuity regression to see whether there is a discontinuity at the date of birth cutoff for this specific disease. As such, individuals can have multiple causes of death and the results do not need to sum up to the net effect of the policy.

Figure 3.3a shows the effect heterogeneity on the highest aggregation level, the ICD-10 main disease groups. The increase is significant and largest for vascular diseases. Cancer and respiratory diseases show a high point estimates too, but the effect is not significantly different from zero. The other groups infectious diseases, mental diseases, and injuries have smaller point estimate and do not differ significantly from zero.

Next, I explore whether diseases influenced by lifestyle behavior increase at the cutoff. Figure 3.3b shows the effect for diabetes 2 (ICD E11.9), diseases attributed to smoking according to Rostron (2012), diseases attributed to alcohol according to Shield et al. (2020), and drug related diseases (ICD F1). Deaths related to alcohol increase strongly at the cutoff. The point estimate for smoking is large as well, but not significantly different from zero. Interestingly, diabetes 2, which is associated with bad nutritional habits, does not increase at the cutoff.

Figure 3.3c digs deeper into alcohol-related causes. There is a discontinuous jump in deaths related to alcohol dependence. Also, alcohol related liver cirrhosis and alcohol liver disease show a positive point estimate but are not significantly different from zero. Figure 3.3d does the same for smoking-related diseases that do not show a significant effect at a more aggregate level in Figure 3.3b. Here, I find a marginally significant increase in COPD, the chronic airway obstruction disease. In terms of effect size, it is only a little less than alcohol dependence. Also, the point estimate for lung cancer is positive, but imprecisely estimated.

Figure 3.3e looks at other frequent diseases or disease groups that are less likely related to lifestyle behavior. Certainly, some of them might still be influenced or correlate with unhealthy behavior such as smoking or drinking, but one would still expect to see a smaller effect—if it is indeed unhealthy behavior that triggers the effect. Indeed, none of the effects are significant and the point estimates are, in general, small. Among psychological diseases not related to lifestyle are, for example, personality disorders that already manifest during childhood.

The results from the analysis of the diseases are in line with the lifestyle hypothesis. Alcohol dependence and COPD, both likely caused by excessive alcohol intake and smoking, show a significant increase at the cutoff. At the same time, diabetes 2 does not increase, which suggests that unhealthy nutrition plays less of a role.

3.6.2 Income

How different is the effect regarding income? Figure 3.4a shows the effect of different income groups. Those are estimates individually calculated, thus the increase in mortality is divided by the probability of retiring two years earlier. The most striking feature is that the effect is precisely zero for the lowest income quintile. For all the higher income quintiles, the point estimate is positive. Although not significantly different from zero, the highest estimates are in quintiles 2, 3, and 5. Thus, there are no empirical grounds to assume that the mortality effect decreases with higher income. Consequently, it is unlikely that lower financial resources mediate the higher mortality.

There are other institutional settings in Switzerland that support the idea that income does not play an important role. Switzerland offers rather generous social welfare. If retirement pension and income do not cover minimum living costs, retirees are entitled to supplementary benefits.

3.6.3 Civil Status

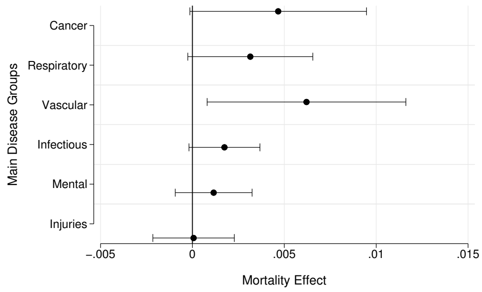
Figure 3.4b shows how the effect differs by civil status. Although a small group, the coefficient for single men is significantly positive. The fact that unmarried men have a higher mortality indicates that the loss of structure plays an important role. One might speculate that whether a marriage is conducive to health depends on the very nature of the marriage. This could explain why the effect is imprecisely estimated.

3.6.4 Geography

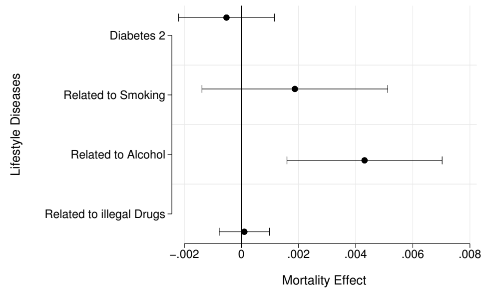
Switzerland inherits the border between two large cultural groups of Europe: the German and the Latin culture. Thus, it provides a convenient setting to test if mortality effects are driven by culture. Figure 3.4c shows the coefficients of a FRD with automatic bandwidth selection for different geographic factors. The first group shows the effects of the language groups. Interestingly, the effect seems to be entirely driven by people

FIGURE 3.3: Mortality Drivers by Diseases

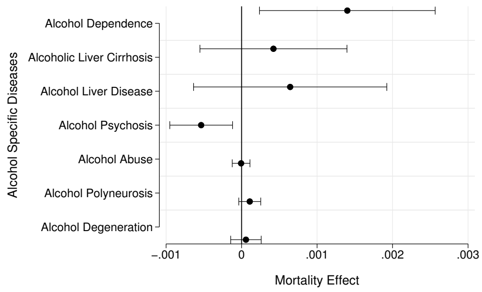
(A) Main Disease Groups



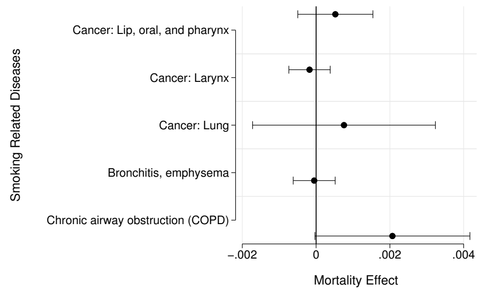
(B) Diseases related to Lifestyle Behavior



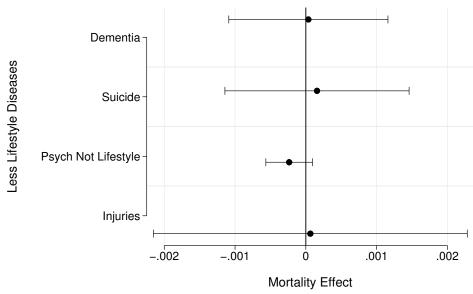
(C) Diseases related to Alcohol



(D) Diseases related to Smoking



(E) Diseases less related to Lifestyle



Notes: This figure shows which diseases drive the mortality effect for the ER 63 reform. Each coefficient and its 90% interval are estimated with local linear approximation with the same bandwidth chosen for the main results (Table 3.3). Panel (a) shows the results for the main ICD disease groups. Panel (b) shows the results for diseases often related to lifestyle behavior. Panel (c) shows diseases that are caused or highly related to alcohol according to Shield et al. (2020). Panel (d) looks at smoking related diseases according to Rostron (2012). Panel (e) looks at diseases *less* related to lifestyle behavior.

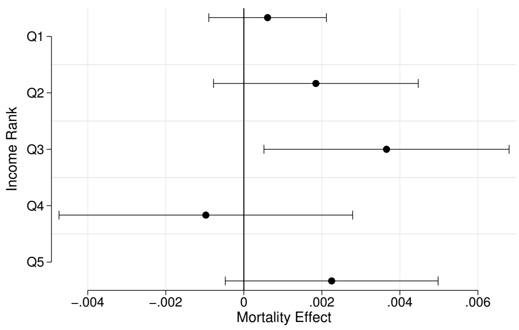
living in German-speaking municipalities. In French- and Italian-speaking municipalities, the point estimate is very close to zero.

Cultural differences have been shown to affect labor market outcomes (Eugster et al., 2017). This indicates that preference and norms towards work are higher in the German-speaking part. Therefore, it is natural to assume that the loss of work due to retirement has stronger consequences for Swiss-Germans. These results are in line with studies from France by Bozio et al. (2020). They do not find a harmful effect of early retirement on mortality. At the same time, Kühntopf & Tivig (2012) show tentative evidence from Germany by documenting an increase in mortality due to early retirement.

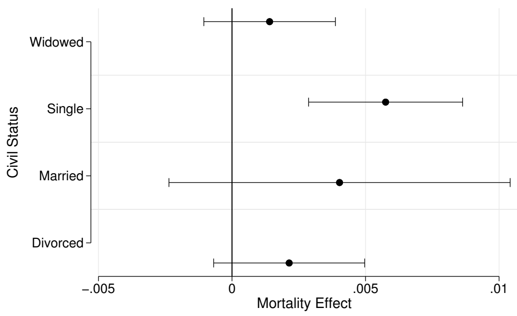
Figure 3.4c also shows how the effect differs between municipality characteristics, such as urbanization (city or no city) and population density. The point estimates of these characteristics are similar. The effect is, however, significant for cities and for regions with high population density. Although this could be driven by chance, one interpretation could be that the loneliness and thus lack of social control is highest in anonymous cities and in regions with high population density.

FIGURE 3.4: Effect Heterogeneity by Personal Characteristics

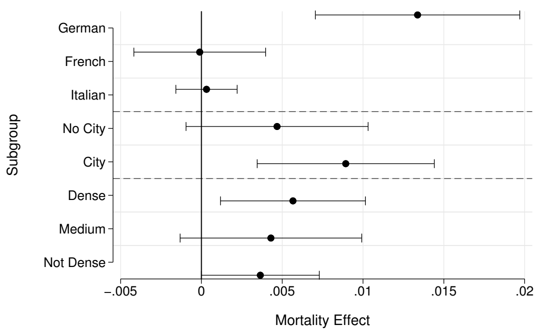
(A) Life Time Income Quintile



(B) Civil Status



(C) Municipality of Residence



Notes: This figure shows how the mortality effect differs by personal characteristics. Each coefficient and its 90% interval are estimated with local linear approximation with the same bandwidth chosen for the main results (Table 3.3). Panel (a) shows the results for five quintiles of the lifetime income before retirement. Panel (b) shows the RD results by civil status. Panel (c) looks at characteristics of the municipality of residence when diseased: the language region (German, French, Italian), whether the municipality is a city, and on the population density (dense, medium dense, and not dense). Life time income quintiles have the following median yearly life time income in 2017 CHF (almost equal to USD): $Q_1 = 16,458; Q_2 = 39,246; Q_3 = 53,172; Q_4 = 68,364; Q_5 = 98,748$

3.7 Summary and Concluding Remarks

This paper studies the effect of early retirement policies on male mortality. Because regressing early retirement on mortality is likely to yield biased coefficients, I exploit exogenous variation provided by two policy reforms. Those policy reforms allowed men born after a certain day of birth to draw public pension one and two years earlier. The treatment assignment around those cutoff dates is therefore as good as random, and I can identify the causal effect of early retirement on mortality in a regression discontinuity design. I draw from two full sample administrative data sets which include precise information on retirement, lifetime income, and mortality causes. Combining a credible identification strategy with precise data and an objective health measure allows to yield trustworthy estimates on the effect of early retirement on mortality. Compared to other studies with a credible identification that uses the default retirement age as an RDD cutoff, my design has the advantage that I can analyze long-term mortality effects.

Giving men the option to retire two years earlier leads to a strong increase in male mortality. On average, men born after the cutoff experience an increase in the mortality rate of 1.1 percentage points. On the other hand, retiring one year earlier does not show a significant increase, but the point estimates are similar. The effect of an increase in mortality for two years of early retirement is robust to several robustness checks. There are no jumps at other non-policy year cutoffs and no jumps at other random cutoffs. Further, there is no discontinuity for women at the same cutoff date. Also, the mortality effect does not materialize before the announcement of the policy. Manipulation or selection around the cutoff can be ruled out by construction because the day of birth is very hard to manipulate and the density of days of birth is smooth around the cutoff.

Analyzing mechanisms and heterogeneity suggests that the increase in mortality by the «two years reform» is driven by an unhealthy lifestyle behavior as a coping mechanism. Deaths related to alcohol dependence and COPD show a strong and significant increase at the cutoff. Other frequent diseases, such as injuries or infectious diseases that are less likely to be influenced by lifestyle behavior, show no significant increase. Also, the effect is significantly higher for unmarried men, giving further strength to the «loss of structure» argument. Interestingly, the effect is almost entirely driven by men living in the German part of Switzerland—which, admittedly, is also the largest language group. The German culture has, in general, a stronger social norm towards work. As pointed out by Eugster et al. (2017), attitudes toward work differ between language groups and it is thus likely that the effect of job loss due to retirement differs as well. I do not find evidence for the competing hypothesis arguing that the effect is driven by a loss of income due to the forgone work income and the actuarial reduction in the annuity. Heterogeneity analysis shows that the effect is highest in the middle of the income distribution. The finding that alcohol plays a major role—or a major mediator—in the increase in mortality is also substantiated by findings of other studies. Several research documents an increase in alcohol consumption during retirement (Zins et al., 2011; Zantinge et al., 2014; Wang et al., 2014; Halonen et al., 2017). This does not mean that men drink themselves to death, but it rather indicates that something in terms of lifestyle is at odds.

How do my results compare to the literature looking at mortality effects of retirement? Despite methodological differences between early retirement and retiring at the default retirement age, it is striking that many results are similar to the ones found by Fitzpatrick & Moore (2018). They look at differences in mortality around the social security cutoff at the age of 62 in the US. Although their focus is on short-term mortality differences around the age cutoff 62, they also find that the increase in mortality is highest for non-married men. Also similar to this study, is that they show that COPD increases. They also note that these causes of death have previously been found to be related to job loss. In a broader view, job loss and early retirement might indeed have similar consequences. Therefore, it is also noteworthy that many of the results show parallels to the *deaths of despair* by Case & Deaton (2015) in the US. In that sense, it is also interesting that the study from Austria by Kuhn et al. (2020) also found an increase in mortality due to early retirement policies.

The primary limitation of this study is its external validity. The very fact that the positive mortality effect is mostly driven by men from the German-speaking part of Switzerland shows that the results cannot be carried over to other countries without limitations. This limitation is not limited to this study, but applicable to all studies in this field. As such, effects might not only differ by country specific institutional settings but also by cultural attributes. By the very nature of the research question, there is always a limitation to certain generations. The generation under study is the World War II generation, also called the «generation silent» which is often characterized as having a high work ethic and being financially prudent (Bialik & Fry, 2019). Absence of work could, therefore, have a more negative health effect. It is, however, possible that younger generations, for example, the «Generation X» (1961 bis 1981) for which work-life balance is in general more important, will also react differently to early retirement.

The results of this study have important policy implications. Although flexible retirement might increase welfare because it allows retiring according to one's own preferences, it can also decrease welfare by reducing life expectancy and increasing health care costs. The argument that early retirement leads to shortened life expectancy and could consequently ease financial pressure of public pension is not only ethically but also economically questionable, mainly because lower health can also lead to higher health care costs. Perhaps a flexible (and earlier) retirement age is in general beneficial and thus one should fight the negative health consequences thereof. For example, with policies that combat the loss of structure due to retirement. Such policies could be to support initiatives that fight loneliness in retirement, such as collaborative forms of living in old age.

Chapter 4

Switzer-Land of Opportunity: Intergenerational Income Mobility in the Land of Vocational Education

joint with Veronica Schmiedgen-Grassi

Abstract This study documents intergenerational income mobility in Switzerland and analyzes the role of educational tracks, local policies, and socio-demographic characteristics. We match the universe of labor incomes over generations and add census and survey data. Using over 900,000 observations from 18 cohorts (1967-1984), we show that income mobility in terms of rank-rank slope (0.14) is higher than in the US and even higher than in Nordic countries. At the same time, educational mobility is low. This shows that low educational mobility does not need to translate into low income mobility. We find high income mobility for individuals with vocational education and training (VET), suggesting that the divergence between educational and income mobility is due to the prominent VET system. Further, children of immigrants show higher mobility rates than children of Swiss born parents. Besides, regions with higher public expenditures, lower tax rates, and higher income inequality exhibit greater income mobility.

Keywords: social mobility, inequality, vocational education and training;
JEL classification: H0, J0, R0

We are grateful for support, advice, and inspiration by Dominik Ullmann, Monika Bütler, Caroline Chuard, Janosch Brenzel-Weiss, Nadia Myohl, Urs Birchler, David Dorn, Reto Fölmi, Jürg Schweri, Guido Neidhöfer, Richard Baldwin, Uschi Backes-Gellner, Immanuel Lampe, Sandy Black, Raphaël Parchet, Laura Casellini-Fontana as well as to the participants of the SOLE 2021 meeting, the ZEW Social Mobility Workshop, the Young Swiss Economist Meeting, Swiss workshop on Local Public Finance and Regional Economics, the PhD Seminar in St. Gallen, the Swiss Leading House of Education, members of the Swiss Parliament. We also thank the Federal Office of Statistics and the Central Compensation Office for generously providing the data. Any errors are our own.

4.1 Introduction

Inequality is one of society's primary concerns. While the desired amount of inequality differs along the political spectrum, the notion that «every child should have the same chance to succeed» is a common denominator among all parties. The American Dream is the moral foundation on which most Western societies are built. Equal opportunity is not only morally desirable, but it also matters for economic growth. Economic growth can suffer when children from poorer parents are impeded from living up to their economic potential—a phenomenon sometimes referred to as «Lost Einsteins» (Bell et al., 2019).

One important facet of equal opportunities is intergenerational income mobility. How much does children's income depend on their parents' income? Despite its importance, only a handful of studies have reliably estimated intergenerational mobility. This is because of the demanding data requirements. To minimize bias, longitudinal income data and information on parent-child relationships are required. In recent years, some notable exceptions succeeded in analyzing high-quality data, for example Chetty et al. (2014a); Heidrich (2017); Bratberg et al. (2017); Acciari et al. (2019); Deutscher & Mazumder (2020); Connolly et al. (2019); Corak (2020a).

Policies that boost upward mobility are urgently needed. Vocational education and training (VET) might be such an option. Rodrik & Stantcheva (2021) declare vocational training as a policy option to intervene at the pre-production stage, targeting bottom incomes. Theoretically, the persistence of income inequality across generations is caused by the socioeconomic gap in human capital investment, that is, poorer parents underinvesting in child education. This underinvestment can be explained by credit constraints or informational frictions on the return to education (Becker & Tomes, 1986a; Cunha & Heckman, 2007; Heckman & Mosso, 2014; Barone et al., 2017; Stuhler et al., 2018; Black et al., 2020). With VET, credit constraints might cause less underinvestment in children's human capital because this kind of training comes at low costs for parents and still gives children ample options for further education after the apprenticeship. However, so far there are no empirical studies analyzing the role of VET in intergenerational mobility.

In this paper, we study intergenerational income and educational mobility in Switzerland. We use administrative high-quality data that cover the universe of labor incomes between 1982 and 2017, combined with administrative linkages between parents and children. We provide national mobility estimates for country benchmarking and we analyze variations across regions. A strong focus lies on the role of Switzerland's VET system. The country is an interesting case since most children opt for VET after compulsory school. In addition, we study how tax rates and public expenditures are related to upward mobility and which personal characteristics best predict upward mobility.

Our study contributes to the literature in several ways. We are first to provide reliable estimates for intergenerational income mobility in a country with a strong VET system. Importantly, our data allows us to link information on education, religion, and other characteristics at an individual level. Second, we add an interesting data point on intergenerational mobility for country comparison. This data point is significant because Switzerland considerably differs from countries for which recent high-quality estimates exist, such as Italy, the US, or Sweden. Third, we analyze how the substantial variation across tax rates and public expenditures is correlated to upward mobility.

We find that intergenerational income mobility is high in Switzerland. A child with parents in the highest income rank 100 can expect to achieve rank 57, whereas a child with parents from the lowest rank 1 can expect to achieve rank 43. This wedge of 14 ranks translates to approximately 11,000 CHF ($\approxeq$ 11,000 USD) in the early thirties[1], which corresponds to roughly two median monthly salaries in Switzerland. This difference in ranks is lower than, for example, in Sweden (18 ranks), Italy (25 ranks), and the US (34 ranks) (Heidrich, 2017; Chetty et al., 2014b; Acciari et al., 2019).

These high income mobility estimates are surprising not only because Switzerland differs from egalitarian welfare states like Sweden, but also, and primarily, because educational mobility is low in the country. The educational track—whether a child opts for VET or high school after compulsory school—highly depends on parental income. Only a little over 10 percent of children with parents below the median income opt for high school. In the top decile, this share amounts to over 40 percent. Also, children's odds of frequenting a high school are five times higher if one of their parents did so. These results are quite fascinating. They show that low educational mobility does not need to translate into low income mobility. However, we also see that regions with high educational mobility have also high income mobility. This suggests that high educational mobility does lead to high income mobility, although in Switzerland to a lesser degree.

To test whether the country's prominent VET system can explain this divergence between educational and income mobility at a national level, we analyze upward mobility for educational tracks that start either with VET or high school. For each track, we calculate the share of children from poorer backgrounds and the share of those who move up the income ladder. This analysis shows that there is a trade-off: tracks with a higher probability of moving up the income ladder provide little access to children from poorer backgrounds. However, there are also tracks with better trade-offs. Those are the ones that start with VET and lead to further education after the apprenticeship. They provide relatively ample access to children from poorer backgrounds and still give them a high probability of climbing up the income ladder.

[1] 12,300 CHF in the early forties.

There are good theoretical reasons that support the thesis that VET can boost upward mobility. Theory informs us that a major factor for the persistence of income inequality across generations is that parents are financially constrained and therefore invest too little in their children's education, which consequently lowers their earnings (Becker & Tomes, 1986a; Solon, 1992, 1999). With VET, this financial constraint might be less important. VET comes at low costs for parents since children even earn a small salary during the apprenticeship. At the same time, VET provides ample options for further education after the apprenticeship. Many of these options take place complementary to the job. This facilitates financing human capital investment.

We further find that mobility differs between regions in Switzerland. Income mobility is more heterogeneous than in Sweden but less heterogeneous than in the US. We do not find any clear and significant trend in income or educational mobility over time, while we see that inequality (GINI index) increased slightly over the child cohorts 1967 to 1984[2].

At a regional level, we see that higher expenditures and lower tax rates are correlated with higher mobility. Relating the income mobility estimates to inequality, we do not find evidence for a «Great Gatsby» curve, which states that higher inequality leads to lower income mobility. In contrast, we see that regions with higher inequality also have higher income mobility. There are, however, two exceptions. When looking at the «cycle of poverty» and «cycle of privileges» measures , we see that inequality is negatively related to mobility.

The strongest individual predictor of upward mobility is gender. Men are almost three times as likely as women to climb the ladder from the bottom to the top quintile. Also, children of immigrants show higher upward mobility than children of parents born in Switzerland. In terms of religion, protestants show the lowest probability to achieve the American Dream—-even though the origins of the American Dream date back to the Protestant Revolution.

We structure this paper as follows. First, we summarize the literature on intergenerational mobility. Then, we outline the different measures to estimate income and educational mobility and describe our sample. Next, we present the mobility estimates at the national level and compare them to other countries. We also show how mobility estimates are correlated to each other. Then, we look at variations across time and space. We study the drivers of mobility: education, inequality, and initial conditions. In the robustness section, we show that our results are robust to several specifications. The last section concludes the discussion.

[2]For family income: From around 30 to 31

4.2 Literature

Since the pioneer contribution of Solon (1992), several scholars have analyzed income transmission across generations (see Solon (1992) and Black & Devereux (2011) for a review). Virtually all of those former studies rely on small-scale survey data and are therefore prone to several biases (e.g. sample selection, attenuation, or life-cycle bias). With the increasing access to large databases in the last decades, research on intergenerational income mobility has experienced a revival.

Our study adds to the current literature on intergenerational income mobility that uses administrative data to analyze income mobility in and within a country. In their innovative study, Chetty et al. (2014a) use tax data and provide a set of measures of relative and absolute income-mobility not just in the United States but also within the United States. They document sizeable geographical variation in mobility across commuting zones. For example, the probability that a child from a family in the bottom quintile reaches the top quintile is 4.5 percent in Atlanta, while it is over two times higher in Washington. Among others, the research stands out by its systematic within-country investigation to shed light on drivers of mobility and the use of directional measures.

Similar in and within-country analyzes have been performed for Sweden, Denmark, Italy, Canada, and Australia. Heidrich (2017) shows that income persistence in Sweden is lower than in the United States and that relative mobility is quite homogeneous across regions, while absolute mobility differs more. Income mobility is also quite evenly distributed across the Danish municipalities (Eriksen, 2018). In contrast, Italian provinces exhibit substantial variation in income mobility (Güell et al., 2018; Acciari et al., 2019). Corak (2020a) and Connolly et al. (2019) analyze mobility in and within Canada and find sizeable variation in mobility across regions. More recently, Deutscher & Mazumder (2020) analyze intergenerational income mobility in and within Australia. They conclude that there is high mobility in Australia, but also substantial dispersion across regions.

For Switzerland, no study has analyzed intergenerational income mobility with administrative data. Bauer (2006) looks at intergenerational income mobility in Switzerland. He estimates an intergenerational income elasticity (IGE) of 0.35, suggesting that an increase of 1 percent in the parent's income is associated with an increase of 0.35 percent in the child's income. The results from this study have to be interpreted with caution, as they are based on predicted incomes from a small-scale survey. Several studies analyze the broader concept of social mobility. Favre et al. (2018) uses historical data from the City of Zurich to examine the extent of occupational persistence during the 1780s and 1870s. Unexpectedly, their results show a decrease in occupational mobility. A more recent study by Häner & Schaltegger (2020) investigates education persistence across 15 generations in the Swiss canton of Basel using a surname-based approach.

We aim to fill the gap in the literature by providing the first estimate of intergenerational mobility for Switzerland, based on administrative data. Thus, we add a further data point to the small set of reliable country estimates for international comparison.

Despite being interesting per se, national and regional measures of income mobility alone are not enough for understanding the process of income transmission. In other words, the ultimate goal is to understand what shapes intergenerational income mobility and to inform policymakers on which policies to implement.

Since the seminal contribution of Becker & Tomes (1986a), the theoretical literature on intergenerational income transmission has long emphasized the idea that institutions significantly affect economic opportunities (see Ichino et al. (2011) for a review). However, empirical evidence is rare.

The lack of past information about local conditions and socio-demographic characteristics probably contributes to the scarcity of descriptive evidence. Indeed, current information about local conditions, such as current local tax rates, does not necessarily reflect the local conditions when children were growing.

Findings of previous literature show that income mobility positively correlates with education, social capital, and economic activity, and negatively with inequality. The role of local tax policies is, however, less clear (Chetty et al., 2014a; Güell et al., 2018; Acciari et al., 2019). Chetty et al. (2014a) find a positive, while not robust, correlation between local taxes and upward mobility. Characterized by high decentralization, Switzerland is an ideal setting for analyzing the role of local policies. Despite our purpose being purely descriptive, we aim to advance the understanding of mobility drivers by exploiting rich socio-demographic information and historical local public finance data. Our data has the advantage that we can directly link individual characteristics, such as education, religion, or family characteristics, which are arguably fixed after a certain age. Most important, municipal-level variables are available since the eighties. To sum up, we add to the literature not only by providing a country estimate but also by describing determinants of mobility.

4.3 Measuring Intergenerational Mobility

4.3.1 Income Mobility

Income mobility aims to describe how a child's income depends on the parent's income. In this section, we describe the measures of income mobility. We largely follow the previous literature, specifically Chetty et al. (2014a) and Corak (2020a). This is to ensure that we can compare our estimates to those of other countries.

It is important to distinguish between two concepts of intergenerational income mobility: *relative mobility* and *absolute mobility*. Relative mobility captures the idea that all children should have equal opportunities to succeed—independent of the economic status of their parents. Absolute mobility measures where children end up in the income distribution when they come from a specific parent rank. Usually, one is interested in the economic outcome of children coming from low-income parents. Absolute mobility captures the idea that parents want their children to do better than themselves—independent of where they rank in the distribution.

Relative mobility has been the focus of most prior work. It aims to answer the following question: «To which extent does my income depend on my parent's position in the income distribution?» In a society with perfect equality of opportunities, the relative ranking of parent's and children's income should be uncorrelated—assuming that genetic dispositions in ability are uncorrelated to a parent's income.

Relative upward mobility occurs when children increase their position in the income distribution compared to their parents. However, if someone moves up in relative terms, someone has to move down. When comparing relative mobility between units, higher relative mobility could also happen if children from rich parents do worse. Similarly, if all children increase their income compared to their parents in such a way that the income ranking stays constant, relative income mobility does not increase. Thus, the impact of changes in relative mobility on welfare is ambiguous.

Absolute mobility might thus be more important from a normative perspective. Relative mobility is not necessarily informative to capture the opportunities of poor children. Relative mobility can also be high when all children have the same low income or if rich children do worse. From a normative perspective, absolute mobility might, therefore, be more meaningful than relative mobility.

4.3.1.1 Logarithmized Income Mobility Measures

Intergenerational Income Elasticity (IGE) has been the most used measure for income mobility, probably because of its intuitive appeal. The IGE is estimated by regressing the logarithm of child income $log(Y_c)$ on the logarithm of parent (usually father or family) income $log(Y_f)$:

$$log(Y_c) = \alpha + \beta log(Y_f) + \epsilon \tag{4.1}$$

The IGE results from Equation 4.1 as the estimated coefficient $\hat{\beta}$:

$$IGE = \hat{\beta} = \rho_{Y_c Y_f} \frac{SD(logY_c)}{SD(logY_f)} \tag{4.2}$$

where $\rho_{Y_c Y_f}$ is the correlation between the logarithm of child income and the logarithm of parent income. SD is the standard deviation.

The IGE measure the differences in income between children from high-income families versus children from low-income families. Thus, it captures the rate of regression to the mean. An IGE of 0.4 means that if parents earn 10 percent more, the income of their children is 4 percent higher.

The intuitive approach of the IGE comes with some drawbacks. The IGE does not only capture the parent child relationship. Equation 4.2 shows that higher inequality in parent's income can lead to higher $SD(logY_f)$ and thus to a lower IGE. The most important drawback is that the relationship between log incomes of parents and log incomes of children is not well approximated by a linear regression. As a result the elasticity might not reflect income mobility at all points of the distribution. A further problem when estimating the IGE is the handling of zeros because the logarithm of zero is not defined. Dropping zeros can lead to overestimated mobility if observations with zeros are more prevalent within children of low-income parents.

4.3.1.2 Rank transformed Income Mobility Measures

Despite the shortcomings of the IGE, a parsimonious statistic facilitates the comparison of intergenerational mobility estimates between units (Black & Devereux, 2011). Another parsimonious statistic is the *rank-rank slope (RRS)*. It gained attention in recent years because it overcomes several drawbacks of the IGE. The rank-rank slope is a positional measure: Income of parents and children are transformed into their percentile ranks. Then, child income rank is regressed on parent income rank. The estimated slope of the linear regression is called the rank-rank slope (RRS). Formally,

$$R_c = \zeta + \omega R_f + \varepsilon \tag{4.3}$$

where R_c is the rank of the child in the child cohort specific income distribution, and R_f is the parent's rank in the child cohort specific parent income distribution. The estimated coefficient $\hat{\omega}$ yields the rank-rank slope (RRS). The estimated constant $\hat{\zeta}$ is then the absolute rank mobility (Corak, 2020a). It indicates which rank children from the lowest parental income rank can expect to achieve.

The rank-rank slope measures the correlation between a child's position and its parent's position in the income distribution. Values close to one indicate a society in which chance of succeeding depends highly on parent's rank and thus with low mobility. Values close to zero denote a society with high mobility. The $RRS \times 100$ equals the child rank difference, also called wedge, between children from the richest and lowest parent income percentile.

Compared to the IGE, the RRS has several advantages. First, zero incomes are preserved. Second, previous studies using rank-rank measures have discovered a strikingly linear functional form (Chetty et al., 2014a; Dahl & DeLeire, 2008; Heidrich, 2017; Acciari et al., 2019; Corak, 2020a; Deutscher & Mazumder, 2020). This makes it a parsimonious statistics across the whole parental income distribution. The IGE in contrast shows non-linearities (compare Figure 4.2). Furthermore, the transformation leads to the same standard deviation for parent and child income (both have a uniform distribution). Thus, the RRS is independent of changes in inequality between parents and children.

Equation 4.3 can also be estimated for subgroups, such as geographic entities. When keeping the national income ranks, the estimates can be interpreted as *absolute mobility* estimates (Chetty et al., 2014a). This is because the ranks on small geographic entities are only weakly related to the ranks on a national level. For example, one might ask: «What is the income that children from poor parents can expect?» This is called *absolute upward mobility (AUM25)*. Following Chetty et al. (2014a), we define AUM25 as the mean adult rank of children whose parents were located at a the 25th percentile in the parent income distribution. Thus, we also refer to it as AUM25.

When looking at large sample, AUM25 can inferred non-parametrically by simply calculating the mean rank of children with parents at rank 25. However, for smaller samples, e.g. at the regional level, noise might distort the measure and this estimate at precisely that point. Therefore, we use a statistical model to increase stability of the estimate. This statistical model is again the linear rank-rank regression stated in Equation 4.3. Instead of using the observed rank at parents rank 25, we use the rank that our linear model predicts at $R_f = 25$. Because the relationship is linear, the mean child outcome at the 25th percentile of parent's income, is the same as the mean outcome for parent's below the median. That is, the AUM25 measures the mean outcome of children born in the poorer half of the society.

4.3.1.3 Directional Mobility

Following Corak (2020a), we use three statistics to indicate directional mobility. Those measures look at specific parts of the quintile transition matrix. The *American Dream (Q1Q5)* measure (sometimes also called the *rags to riches* measure) describes the probability of a child born to parents in the bottom quintile to move up to the top quintile (Corak & Heisz, 1999; Hertz, 2006; Chetty et al., 2014a).

$$Q1Q5 = Pr[R_c > 80 | R_f \leq 20] \tag{4.4}$$

The *cycle of poverty measure (Q1Q1)* indicates the share of children from parents in the bottom quintile that stay in the bottom quintile.

$$Q1Q1 = Pr[R_c > 20 | R_f > 20] \tag{4.5}$$

Likewise, the *cycle of privileges (Q5Q5)* measures the share of children from parents in the top quintile that stay in the top quintile themselves.

$$Q5Q5 = Pr[R_c \geq 80 | R_f \geq 80] \tag{4.6}$$

4.3.1.4 Rate of Absolute Mobility (RAM)

The *rate of absolute mobility (RAM)* measures the fraction of children who earn more than their parents at the same age in real monetary units.

$$RAM = \frac{1}{N} \sum_{i}^{N} 1[Y_{ci} > Y_{fi}] \tag{4.7}$$

N is the number of children in the respective cohort. Incomes of parents Y_{fi} and children Y_{ci} have to be adjusted for inflation. Besides, income is usually measured around mid-life with the intention to minimize life-cycle bias.

4.3.2 Educational Mobility

4.3.2.1 Switzerland's Education System

Switzerland is well known for its vocational education and training (VET) system. In contrast to other countries, where vocational education is seen as the «last resort», the VET system in Switzerland is highly regarded and the «standard» track for the majority of adolescents after the lower-secondary level. Almost 70 percent of the children earn a vocational degree after compulsory school. Only around 20 percent opt for a high school, also called gymnasium or baccalaureate schools.[3] A high school degree allows children to take up studies at the university level in almost all fields. (Educa, 2021)

The education system differs from most foreign systems of vocational and professional education and training. VET is usually based on a dual system: It comprises practical

[3]Smaller tracks include specialized schools. They provide students with preparation for tertiary level professional education in specific occupational fields at colleges of higher education.

training (apprenticeship) on three to four days at a company and is supplemented by formal classes on one to two days at a VET school. Currently, there are around 250 VET programs for different occupations. Therefore, many children obtain professional qualifications already at an upper secondary level, while in other countries the same qualifications are received on a tertiary level. (Educa, 2021)

After lower-secondary level, children can apply for apprenticeship at a training company. The training company can decide which children to employ. Usually, the firm's selection is based on the student's performance in school, the application documents, and on an interview—similar to «adults labor market procedures». Depending on the VET program, the duration ranges from two to four years. (Educa, 2021)

4.3.2.2 Educational Mobility Measures

We use three different measures for educational mobility. The correlation in years of schooling between parents and children, the share of children from the bottom quintile that visit a high school, and the change in likelihood to visit a high school when one of the parents visited a high school.

The *correlation in years of schooling* measures how the years of schooling between parents and children correlate. To do so, we use the highest education from any of the parents and approximate this with years of education.[4] Since vocational education is highly common in Switzerland, this measure might not capture the full persistence of educational inequality over generations. The reason is that a large share of time in the vocational apprenticeship is spent on hands-on training in a firm and not necessarily in a school. Thus, children in VET have far less formal education than children in a high school, even though the difference might not be large in terms of years of schooling if one approximates three years of VET to three years of schooling.

Therefore, we specifically look at the persistence in educational tracks over generations. First, we do so by measuring the share of children in the high school track by parental income. We refer to this measure as *Share Bottom 20 in HS* . We estimate the share of children in high school in the bottom quintile of the parental income distribution. Second, we look at how much more likely children are to visit a high school, if one of their parents visited high school. We call this measure *Child in HS if parent was*. To keep things simple, we run a linear regression model with the binary outcome variable of child high school on parental high school. The slope then indicates how much more likely children from parents with a high school degree are to visit a high school too.

[4]The conversion table from highest completed education into years of schooling can be found in the appendix in Table C.1

4.4 Data and Variable Construction

4.4.1 Data sources

We use several data sources for our analysis. We combine individual-level demographic information and official (mandatory) survey records, both provided by the Federal Statistical Office (FSO), with social security earnings records (SSER) from the Central Compensation Office (CCO).

We derive data on demographic characteristics, family ties, and citizenship from the «Population and Households Statistics» (STATPOP), a collection of several registers.[5] To establish the intergenerational connection, we use information from the INFOSTAR register. This register contains around 85 percent of all parent-child relationships of the Swiss population. Family ties for non-natives are less likely to be identified since births occurring in foreign countries are not recorded. We will take this into account by excluding non-natives.[6]

We match individual information to the longitudinal «social security earnings records» (SSER). The register includes every legal labor income in Switzerland. It provides complete earnings information for employed and self-employed in Switzerland since 1982. Its purpose is to calculate public old-age insurance. Earning records are not top-coded, allowing us to depict the labor income distribution accurately.

We complement the matched STATPOP-SSER with information from the structural survey (SE). This data set is available since 2010 and surveys roughly 200,000 persons per year (2 percent of the population). Participation is mandatory and non-participation is sanctioned. As we have nine years available, we have a sample size of over 1,600,000 unique observations (some individuals are surveyed multiple times). The survey includes, for example, information on education, religion, and occupation. Although this data is only available since 2010, this is not a drawback for us. Most variables we use, such as educational attainment or religion, can be assumed to stay constant after the age of 30.

4.4.2 Sample Selection

The core sample comprises native child cohorts from 1967 to 1984. Conditional on being born in Switzerland, we link 72 percent of children to their father. The share varies

[5]This data is also known as the «New Population Census» or the «Register Survey».

[6]Excluding immigrants might be seen as a limitation. However, regarding comparability, we are in line with other studies, which exclude immigrants from the sample (Chetty et al., 2014a) or (Heidrich, 2017). One could also argue that intergenerational mobility, which is also a measure of opportunities during childhood, should focus only on children that spent their entire childhood in a country. Children of immigrants are included in our sample if they are born in Switzerland.

between 88 percent for the 1984 cohort and 56 percent for the 1967 cohort. As we observe the universe of Swiss residents between the years 2010 and 2017, non-identified intergenerational links are because of death, emigration, or missing register updates in the IT system. Following the previous literature, we measure child income between the age of 30 and 33, while we measure lifetime family income when the child is between 15 and 20 years old. Therefore, our sample includes cohorts aged at least 15 years in 1982 and at least 33 years in 2017. Virtually every individual in our sample has at least one income record. For over 97 percent of children born between 1967 and 1984, we observe at least one non-negative income record between the age of 30 and 33. [7]

The core sample comprises individuals born between 1967 and 1984, for whom we can at least identify the father, whose mean income is non-negative between the age of 30 and 33, and whose mean parent income is non-negative between child age 15 to 20. The justification for requiring the father to be identified lies in the relatively strong gender difference in labor income. Using children for which only the mother needs to be identified, leads to higher income mobility estimates. A concern might then be that if the father passed away before 2010, we cannot accurately capture the persistence in income inequality and introduce attenuation bias. In a robustness sections, we also use different samples. For example, for those for whom we can identify either father or mother or for whom we can identify both father and mother. In a robustness check, we also use different samples.

Although we have an extensive coverage of child-parent relationships, we still check whether it represents Switzerland accurately. Table C.2 shows sample means for the full population (1967 to 1984 cohorts) and the core sample and alternative samples. In the population, we have 1,266,376 individuals born in Switzerland between 1967 to 1984 (Column 1). For almost 90 percent of those individuals, we have been able to identify either the mother or the father (Column 2). Conditioning on observations for whom we can identify at least the father, which is our core sample, the share declines to 72 percent (Column 3). The last column restricts to individuals for which we can identify the father and the mother (Column 4). Panel (A) shows the summary statistics for the income sample. Our core sample (Column 3) is slightly younger than the full sample, as parents of older cohorts are more likely to be dead relative to parents of younger cohorts. The lower Part of the Table, Panel (B), reports descriptive statistics for the education sample. We do not observe the education level for everybody in the core sample since information on education stems from the structural survey (SE). However, the size of the education sample is still large. Overall, sample differences are minimal

[7]Negative income records occur because of accounting techniques. When the income has to be corrected, the correction is recorded with a minus, and the amount has to be subtracted. Less than 0.03 percent of observations have either negative mean parent income or negative mean income.

or explainable and our core sample represents the total population of children born in Switzerland.

4.4.3 Income Definition

Income is the sum of wage earnings (employment and self-employment income), unemployment benefits, military compensation, maternity leave payments, and disability benefits. We deflate all incomes with the consumer price index (Swiss CPI) to 2017 CHF.

Child Income. In the core specification of rank mobility, we measure child income during the ages 30 to 33. The principal reason for this choice is to compare our estimates to other countries, specifically to the US. To smooth out transitory income shocks, we average income over four years. We set income equal to zero if we do not observe any income during those four years. In Section 4.8, we test the robustness of our baseline estimates using alternative age definitions. In Table C.3, we assess the sensitivity to alternative income definitions.

Family Income: In the core specification, family income is the average of father and mother income when the child is between 15 and 20 years old. The reason for this choice is threefold: First, we aim to capture a child's opportunities while it is growing. The age between 15 to 20 is decisive in Switzerland because children decide which educational track they will follow. Second, parents are on average in their mid-forties, making their rank in the income distribution stable and the life-cycle bias negligible. Third, to ensure comparability to the US (Chetty et al., 2014a). In Table C.3, we evaluate the sensitivity of this choice by varying child age at which parent income is measured.

4.4.4 Summary Statistics

Table 4.1 provides summary statistics for the core sample. In the main specification, we use 923,107 observations for the income and 308,622 observations for the education analysis. Panel (B) and Panel (C) also show child and parent income at different points of the distribution. More detailed description of the child and parent income distribution can be found in the Appendix in Table C.5.

We base the geographic assignment of a child on the mother's municipality in 2010. If the mother is missing, we use the father's municipality in 2010. We do not have longitudinal information on geographic location of individuals before 2010. Data allows, however, to figure out since when an individual lives in a municipality. Geographic mobility is low in Switzerland. Over 70 percent of mothers live in the same municipality as they used to live in 1995. In addition, we also know the place of birth of individuals, which is usually in a hospital close to their municipality. Thus, we use the municipality of birth in a robustness check.

TABLE 4.1: Summary Statistics of Core Sample

	(1) mean	(2) sd	(3) min	(4) max
Panel A: General				
Personal Characteristics				
Year of birth	1975.65	5.27	1967	1984
Father age at childbirth	30.26	5.20	13	68
Mother age at childbirth	27.47	4.62	13	57
Female	0.49	0.50	0	1
Married	0.45	0.50	0	1
Non-Native Father[a]	0.11	0.31	0	1
Geography				
Same municipality as in 1995[b]	0.71	0.46	0	1
Lake Geneva Region	0.15	0.36	0	1
Espace Mittelland	0.25	0.43	0	1
Northwestern Switzerland	0.13	0.34	0	1
Zürich	0.18	0.38	0	1
Eastern Switzerland	0.14	0.35	0	1
Central Switzerland	0.11	0.32	0	1
Ticino	0.04	0.18	0	1
Language Region				
German	0.76	0.43	0	1
French	0.20	0.40	0	1
Latin	0.04	0.20	0	1
Education[c]				
High-school	0.21	0.41	0	1
VET	0.66	0.47	0	1
Master	0.16	0.37	0	1
Income				
Child: at least one income record	0.96	0.21	0	1
Family: at least one income record	1.00	0.06	0	1
Child income at age 30-33	60,598.21	39,451.96	0	7,385,721
Child income at age 40-43[d]	75,358.69	80,547.33	0	12,711,202
Family income at child age 15-20	64,214.18	65,426.26	0	13,654,888
Panel B: Child Income at Age 30-33				
Bottom 20	10,970	9,041		
At Rank 25	34,426	3,476		
At Rank 50	61,857	2,690		
Top 20%	109,995	42,932		
Top 10%	126,510	54,118		
Top 1%	186,735	105,673		
Panel C: Family Income at Child Age 15-20				
Bottom 20	26,042	9,282		
At Rank 25	40,901	1,517		
At Rank 50	55,049	2,332		
Top 20%	123,132	122,051		
Top 10%	155,176	162,008		
Top 1%	291,319	345,799		
Obs.	923,107			

[a]Father not born in Switzerland
[b]Mother
[c]Observations: 308,622
[d]Observations: 365,573

Notes: Table 4.1 provide a description of the core sample. Panel A reports general characteristics. Panel B and Panel C show the average income and the standard deviation of children's income between 30 and 33 and family income (measured when the child is between 15 and 20). The number of observations refers to the " Income sample". All amounts are in 2017 CHF.

4.5 National Mobility Estimates

4.5.1 Income

4.5.1.1 Income Mobility Estimates

Figure 4.1 presents the *rank-mobility estimates*. The points show the expected income rank of children for every parent income rank. The relationship is almost linear. This justifies the use of a linear regression to summarize the rank-rank relationship and the rank-rank slope is an insightful and parsimonious statistic across the parental income distribution.

The line in Figure 4.1 is the prediction of a linear regression of child rank on parent rank. A higher slope (RRS) means lower intergenerational mobility. Here, the slope is 0.14. Since there are 100 ranks, the difference between the lowest and the highest income rank is 14 ($0.14\times$ 100), which is sometimes referred to as the wedge between rich and poor children. We can also translate the rank back into monetary units to increase the interpretability. This difference of 14 ranks translates to approximately 11,000 CHF ($\approxeq$ 11,000 USD) in the early thirties, which corresponds to roughly two median monthly salaries in Switzerland.[8] The constant in the regression in Figure 4.1 is 44. This is the rank which a child with parents in the lowest rank can expect to achieve. The R^2 of the regression is only 0.02. While there is clearly a positive relationship between parent and child rank, parental income rank is only a weak predictor of the child's income rank. The *AUM25* can be calculated by the slope and the constant and is 46. Thus children from the bottom half of the income distribution can expect to achieve income rank 46.[9]

The *directional mobility* estimates are shown in the quintile transition matrix in Table 4.2. It describes in which quintile children end up conditional on their parent quintile. If child and parent income were, one would expect to see 20 percent in each cell. This would be the case with perfect mobility.[10]

The *American Dream (Q1Q5)* measure is presented in column 1 and row 5. It reveals the share of children with parents at the bottom quintile that make it to the top quintile. In Switzerland, this share is around 12 percent. The *cycle of poverty measure (Q1Q1)* is shown in row 1 and column 1. It indicates the share of children from the bottom quintile, which stay in the bottom quintile. In Switzerland, this share is around 24 percent. The *cycle of privileges measure (Q5Q5)* is shown in row 5 and column 5. This share is around 30 percent. Thus, around 30 percent of children from the top quintile stick in the top quintile.

[8] Mean income for parent and child rank can be found in Table C.5 in the appendix.

[9] In the national sample, due to its large size, the AUM25 can also be retrieved by simply looking at the mean rank of children with parents at rank 25.

[10] Table C.7 shows the same for the distribution with both parents identified

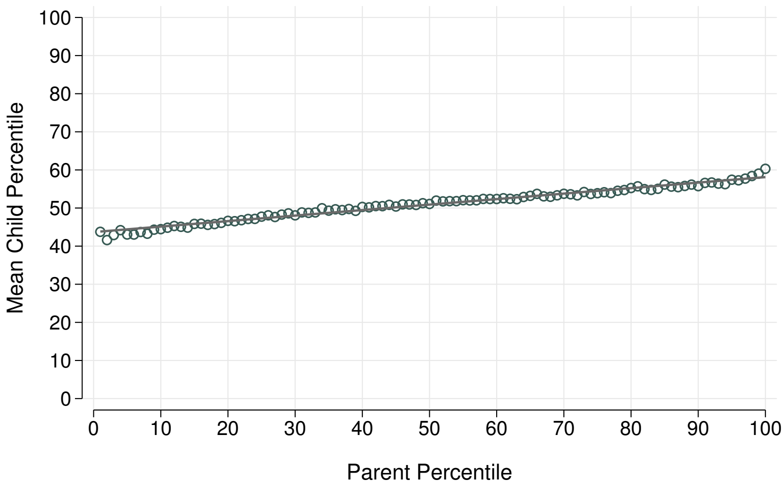


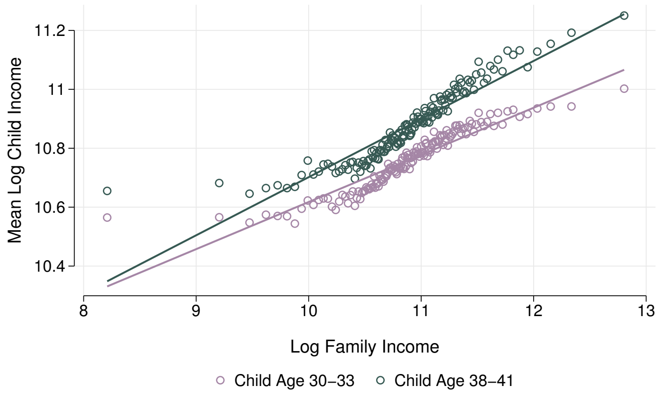
FIGURE 4.1: Child and Family Rank Relationship

Notes: The points show the mean child rank for each parent rank. Ranks are in percentiles. The pink line is the prediction of an OLS regression of child rank on parent rank based on 923,262 observations. The OLS regression yields a constant of 43.7 and a (rank-rank) slope of 0.14 (RRS). The R^2 is 0.02. The estimated rank-rank slope of 0.14 is a measure of relative income mobility. The higher this slope, the more child income depends on parent income, hence the lower income mobility. The rank difference between children from the poorest and the richest parents equals the slope $\times 100$, and is 14 in this case. This is sometimes also called the «wedge» between children from the highest and lowest parent percentile.

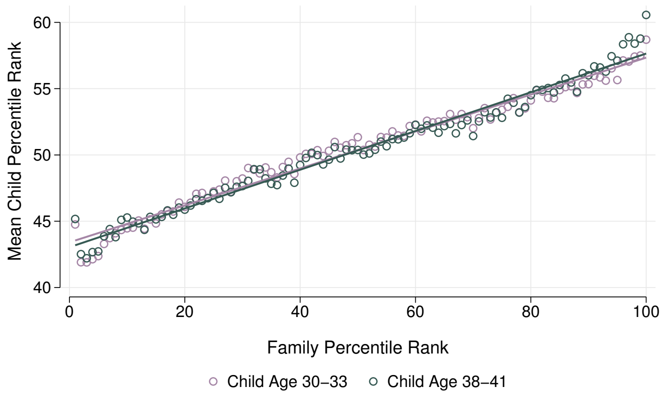
TABLE 4.2: Quintile Transition Matrix

	Family quintile				
Child quintile	1	2	3	4	5
1	23.66	21.55	19.43	17.93	17.43
2	21.46	20.26	19.48	19.26	19.54
3	20.94	21.92	20.37	19.02	17.75
4	17.66	21.07	21.38	20.56	19.33
5	11.87	15.49	19.35	22.98	30.30

Notes: Each cell describes in which quintile (row) children end up conditional on the parent quintile (column). Parent quintile is based on mother and father income conditional that father income is identified. For example, 17.44% of children from parents in the top quintile of the income distribution will end up in the bottom quintile of the income distribution. 11.8% of children from parents of the bottom quintile of the income distribution end up in the top quintile («American Dream measure»). The table includes children born in Switzerland from 1967 to 1984 for which at least the father can be identified. It is based on 923,262 observations.



(A) Intergenerational Income Elasticity (IGE)



(B) Rank-Rank Slope

FIGURE 4.2: Log-Log and Rank-Rank Relationship by Child Age

Notes: This figure shows how the IGE and the rank-rank slope depend on the age at which child income is measured. The estimated IGE in Panel (a), represented by the lines, is higher when measured at later ages. Thus, the IGE is subject to severe life-cycle bias when child income is measured too early. The relationship between log-child and log-parent income is not linear. The rank-rank slope in Panel (b) is almost similar when child income is measured at later ages.

Figure 4.2 compares the parent-child income relationship in logs to the relationship in ranks for different ages at which child income is measured. The slope of the line in Panel (a) is the *intergenerational income elasticity (IGE)*. The figure reveals two insights: First, the log relationship is not linear. Thus, the IGE differs along the parental income distribution. It is lower at the bottom and at the top of the parental income distribution. While those non-linearities are interesting per se, they make it difficult to use the IGE as a parsimonious statistic and compare it to other countries. Second, the slope is higher if child income is measured at later ages (life-cycle bias). Therefore, the IGE should be interpreted with caution.[11]

The *IGE* is 0.16 when child income is measured between 30 and 33, and 0.21 if child income is measured between 38 and 41 (see Table C.8). Nybom & Stuhler (2016) show that life-cycle bias is lowest when income is measured around mid-life. Thus, 0.21 would be our baseline estimate. Furthermore, the IGE varies along the parental income distribution. It is around 0.09 in the lowest parent tertile, 0.42 in the middle tertile, and 0.22 in the highest tertile. Around half of the size of the IGE is driven by changes in inequality over generations. This can be seen in Column (2) and (7) in Table C.8. When log income is standardized, thus divided by the standard deviation, the IGE becomes 0.10 for income at age 30 to 33 and 0.12 for age 38 to 41 (also see Equation 4.2. Thus the IGE is not only smaller, but also less prone to life-cycle bias.

Finally, we move to the *rate of absolute mobility (RAM)*. Table 4.3 shows the share of children earning more than their parents at the same age. Income is averaged between the ages 40 and 45. This reduces the number of included cohorts and the sample size (n=451,491). The results show that 39 percent of children earn more than their father did, at the same time, 83 percent of children earn more than their mother. Due to the gender specific labor market participation over time, comparing sons to fathers might be most sensible. Here, we see that almost 58 percent of sons earn more than their father.

4.5.1.2 International Comparison

How can these national estimates be interpreted? Table 4.4 puts the mobility estimates of Switzerland in context to other countries. As always, one has to be careful when comparing estimates across countries since data processing might differ for example with respect to the analyzed cohorts or the definition of income. To enhance comparability, we have only chosen studies that use large and/or administrative data and were conducted in the last 10 years.

Looking at the rank-based mobility results shows that Switzerland has very high income mobility estimates. One exception is the «American Dream (Q1Q5)» measure: In

[11] Estimates for different samples are shown in Table C.6 in the Appendix

TABLE 4.3: RAM: Rate of Absolute Mobility

Child Sex	Share Child Income > Father Income	Share Child Income > Mother Income	Observations
All	0.386 (0.0007)	0.834 (0.0005)	451,491
Female	0.183 (0.0008)	0.739 (0.0009)	230,931
Male	0.579 (0.0010)	0.925 (0.0005)	220,560

Notes: This figures shows the share of children earning more in real terms than their mother or father at the same age. Income of children and parents is measured at the same ages between 40 to 45 (child cohorts: 1967 to 1977). For example, 18.3% of women earn more between 40 and 45 than their father did between 40 and 45. Income is deflated with the consumer price index 2017. Standard errors are shown in parentheses.

Sweden, for example, children from the bottom quintile are more likely to reach the top quintile than in Switzerland. Also, Switzerland is doing a little worse in terms of IGE. However, as shown above, this measure is highly susceptible to the location in the parental income distribution, to changes in inequality over time, and to the age at which child income is measured. This makes comparison between countries very unreliable. Thus, we argue that in the most comprehensive measure, the rank-rank slope (RRS), Switzerland is doing better than any other country.

4.5.2 Educational Mobility

Despite the high income mobility estimates in the previous section, we find that educational mobility is low in several dimensions. Figure 4.3 shows how the educational track depends on the parental income rank. In the bottom quintile, only around ten percent of children visit a high school. In the top decile, however, more children opt for a high school than VET.

Table 4.5 provides further evidence of low intergenerational educational mobility. Column (1) shows the slope coefficient of a linear regression of years of schooling of the child on years of schooling of the parents. We associate one year of schooling more of the parents with 0.33 years more schooling of the child. This is a relatively high-estimate compared to other countries, and especially considering the high income mobility estimates (Hertz et al., 2008a). Column (2) shows that children with at least one parent with a high school degree are around 5 times more likely to visit a high school themselves. Similarly to Figure 4.3, columns (3) and (4) show how much more likely children are to visit a high school when parents are in the top quintile or the top percentile of the income distribution.

TABLE 4.4: International Comparison

Country	Mobility Measure						Source	Observations
	RRS	IGE	Q1Q5	Q1Q1	Q5Q5	AUM25		
Switzerland	0.14	0.22	11.87	23.67	30.3	46	Chuard-Keller and Grassi (2021)	923,262
US	0.34	0.45	7.50	33.7	36.5	41.4	Chetty et al. (2014a)	40,000,000
Sweden	0.18	0.29	15.7	26.3	34.5	43.6	Heidrich (2017)	927,008
	0.22	0.231	-	-	-	-	Bratberg et al. (2017)	252,745
Italy	0.25	0.25	9.9	28.66	35.6	44	Acciari et al. (2019)	647,662
Canada	0.24	0.20	11.4	30.1	32.3	44.35	Corak (2020a)	3,002,950
Australia	0.21	0.19	12.3	31	30.7	45.1	Deutscher & Mazumder (2020)	1,025,800
Norway	0.22	0.19	-	-	-	-	Bratberg et al. (2017)	324,870
Denmark	0.20	0.17	-	-	-	-	Helsø (2021)	151,360
	-	-	11	31	35	-	Eriksen (2018)	205,625

Notes: This table compares results of recent studies on intergenerational income mobility that use high quality data and are thus likely to provide reliable results. *RRS* stands for rank-rank slope. The higher the RRS, the lower income mobility. *IGE* stands for intergenerational elasticity. For Switzerland, IGE is measured at age 38 to 41. The higher the IGE, the lower mobility. *Q1Q5* is the «American Dream» measure. It reports the share of children from the bottom quintile that make it to the top quintile. The higher this measure, the higher mobility. *Q1Q1* is the «cycle of poverty» measures. It reports the share of children from the bottom quintile that stay in the bottom quintile. The higher this measure, the lower mobility. *Q5Q5* is the «cycle of privileges» measures. It reports the share of children from the top quintile that stay in the bottom quintile. The higher this measure, the lower mobility. *AUM25* stands for absolute upward mobility at percentile 25. It shows where children with parents at the 25th percentile of the income distribution can expect to end up. This follows from the prediction of the rank-rank slope regression. It also shows where children with parents below the median can expect to end up.

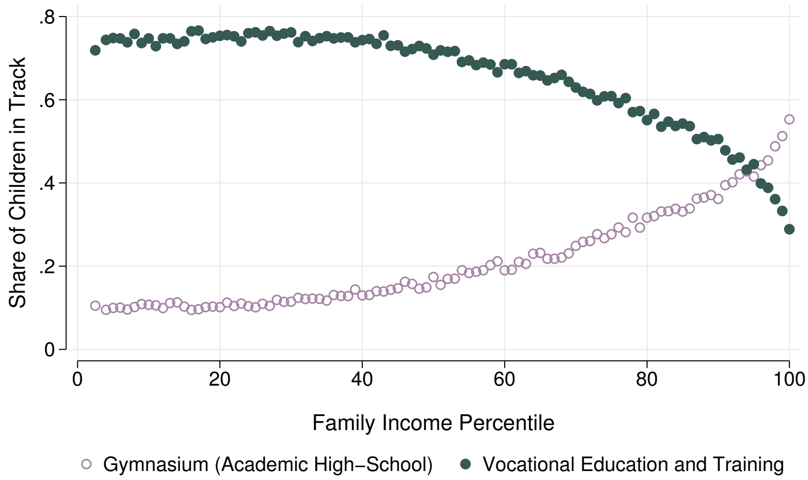


FIGURE 4.3: Educational Track by Parent Income

Notes: This figure shows the share of children by family income rank in the high school (gymnasium) and the vocational education and training track (VET).

The results of low educational mobility are in line with previous research. For example, Bauer & Riphahn (2007) find high intergenerational persistence in terms of educational track. Also, Hertz et al. (2008a) rank Switzerland's educational mobility similar to the ones of the US or Pakistan.

4.5.3 Correlation of Mobility Measures

Figure 4.4 shows how mobility measures correlate with each other on a labor market regions level.[12] When looking at the rank-rank slope (RRS), we find high correlations with most income mobility measures. One exception is the cycle of privileges measure (Q5Q5): The persistence in the top quintile of the income distribution seems to be only weakly correlated to the rank-rank slope. Interestingly, the cycle of privileges measure is strongly correlated with the American Dream measure. This result is somewhat counter-intuitive. It is telling us that regions with high top income persistence over generations also provide good opportunities to children from the bottom quintile to climb to the top quintile.

Moving further to educational mobility, we see that there is a fairly strong correlation of most income mobility measures with educational mobility. For example, a higher rank-rank slope (lower income mobility) is associated with a lower share of children from the bottom quintile in high schools and also with a higher persistence in the educational track («Child in HS if Parent was»). The educational mobility measure of «correlation in years of education» is only weakly related to most other measures. As mentioned

[12]Figure C.9 in the Appendix shows the same on a cantonal level.

TABLE 4.5: Educational Mobility

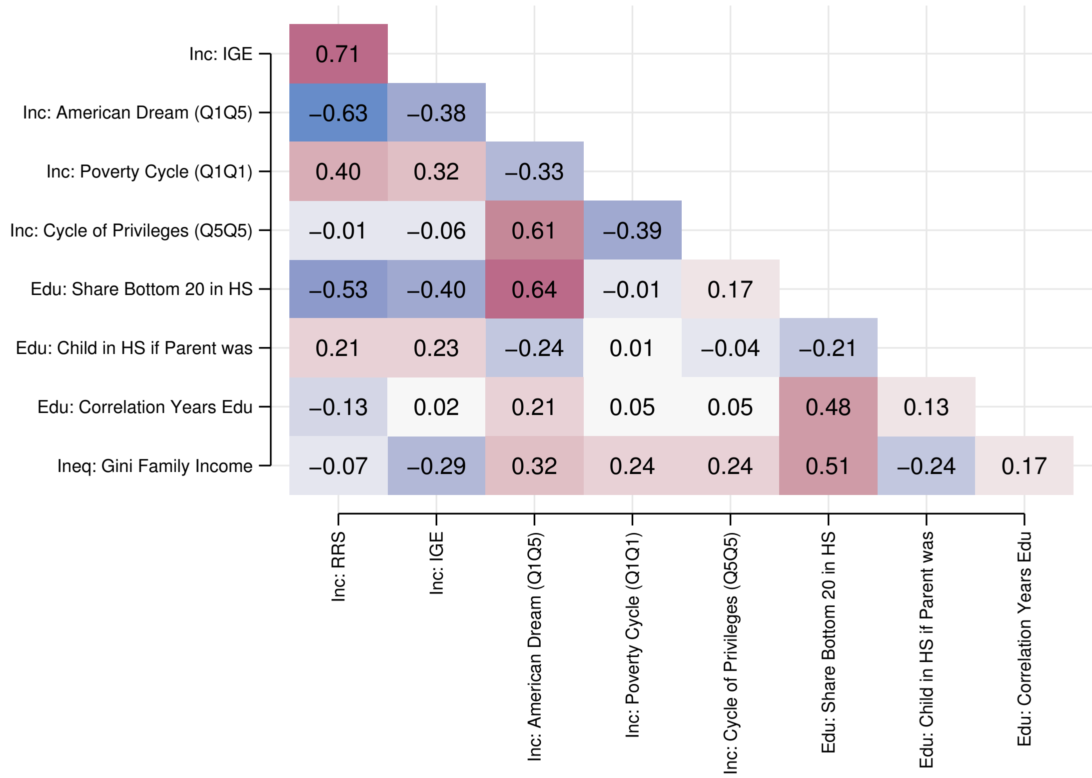
	OLS	Logit		
	(1)	(2)	(3)	(4)
	Child Yrs. Schooling	Child HS	Child HS	Child HS
Yrs. School Parent	0.334***			
	(0.0029)			
Parents HS		5.237***		
		(0.0793)		
Parents > Rank 80			3.436***	
			(0.0337)	
Parents > Rank 99				4.753***
				(0.1707)
Observations	152,334	182,501	308,673	308,673
Mean Dep. Var.	14.233	0.219	0.210	0.210

Exponentiated coefficients in Col (2) to (4); Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table shows different measures of educational mobility. Column (1) shows how years of schooling of parents and children are correlated (yrs of schooling are standardized). Column (2) shows how much more likely children are to visit a high-school if one of their children visited a high-school. Column (3) shows how much more likely children are to visit a high-school if their parents are in the top quintile of the income distribution compared if they are below. Column (4) shows the same as column (3) but does so for the top percentile of the parent income distribution.

FIGURE 4.4: Correlation Mobility Measures (Unit: LM Region)



Notes: This graph shows the correlation of different mobility and inequality measures on a labor market units (n=106). *RRS* is the rank-rank slope, *IGE* is the intergenerational income elasticity, *Share Bottom 20 in HS* measures the share of children from parents in the bottom quintile that go to high school, *Child in HS if Parent was* is the slope coefficient of a linear provability model that regresses high school attendance of the child on high school attendance of the parents, *correlation Years Edu* measures the correlation in years of education between parents and children, *Gini Family Income* measures the GINI index for family income at child age 15 to 20.

before, this measure might not capture the full persistence in educational inequality because of the idiosyncratic education system in Switzerland.

The finding that intergenerational educational mobility is correlated with income mobility is important. It implies that our previous (national) finding of high income mobility but low educational mobility does not mean that education does not matter. Although this is not a causal estimate, it is in line with predictions of theoretical models (e.g. Becker & Tomes (1986a)) which state that educational mobility is the main pathway to income mobility. The crucial point is that in Switzerland, educational inequality translates only weakly into income inequality over generations. One can argue that this is due to its VET system, which provides good wage outcomes with comparably little (formal) education. We further elaborate this point in Section 4.7.1

4.6 Mobility Across Time and Space

4.6.1 Mobility over Time

Figure 4.5 shows how *income mobility* developed. It does so for three measures of relative mobility: The American Dream measure (Q1Q5), the poverty cycle measure (Q1Q1), and the rank-rank slope (RRS). In general, there is no clear and significant trend in any of the income mobility measures. If anything, one can discern a small upward trend in the poverty cycle measure, which would suggest lower mobility. However, the same can be said about the American Dream measure, which would then suggest higher mobility.

The rank-rank slope increases slightly since cohort 1975, but decreased before. One interpretation for the peak around the cohort 1972 might be the boom before the financial crisis in 2008. For Children born in 1972, their income is measured between 2002 and 2005. It is conceivable that children from high income parents profited more of this upswing. This would then mechanically lead to lower mobility. It is also interesting that the increase in the rank-rank slope around this time did not affect children from the bottom quintile—since the Q1Q1 and Q1Q5 measure stay almost constant.

Figure 4.6 shows how educational mobility developed. There is a trend towards higher educational mobility in terms of educational tracks: The share of children from the bottom quintile that visit a high school increases from around 8 percent in the late 60s cohort to around 12 percent for the 80s cohort. Also, whether a child visits a high school depends less on whether parents visited a high school since cohort 1975. At the same time, the correlation in years of schooling between parents and children increases slightly. However, as mentioned before, the years of schooling measure should be interpreted with caution.

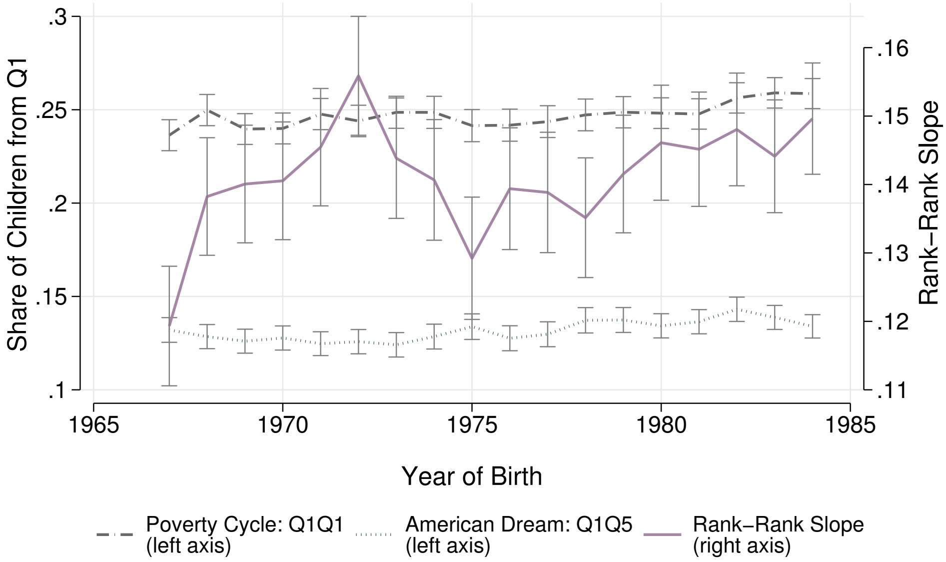


FIGURE 4.5: Income Mobility over Time

Notes: This figure shows how income mobility varies across cohorts. The *poverty cycle (Q1Q1)* measure shows the share of children from the bottom quintile staying in the bottom quintile. An increase means lower mobility. The *American Dream (Q1Q5)* measure shows the share of children from the bottom quintile moving to the top quintile. A decrease shows lower mobility. The *rank-rank slope (RRS)* on the right axis shows the rank-rank slope. An increase shows lower mobility. The bars represent 95% confidence intervals.

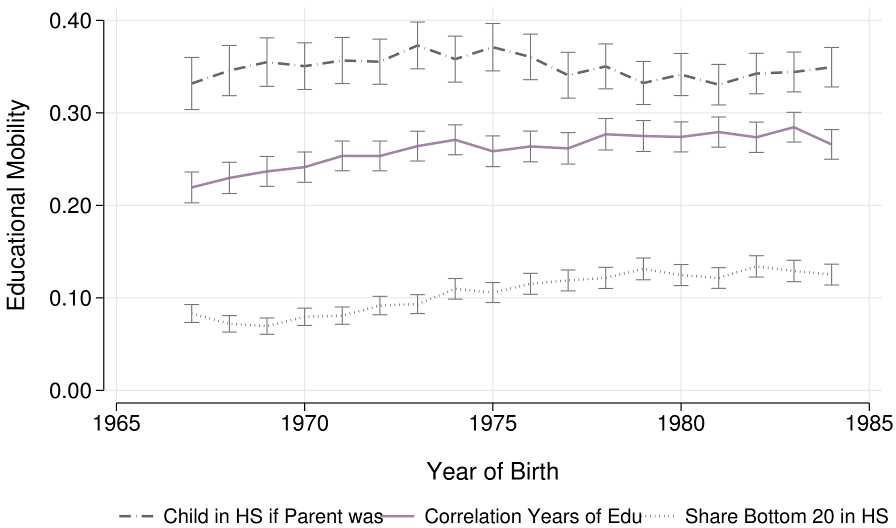


FIGURE 4.6: Educational Mobility over Time

Notes: This figure shows how educational mobility varies across cohorts. «Child in HS if Parent was» shows how much more likely children are to be in high school if one of their parents was in high school. «Correlation Years Edu» shows the correlation between years of education of parents and children. «Share Bottom 20 in HS» shows the share of children from the bottom quintile of the income distribution that visit a high school.

4.6.2 Geographical Variation

In this section, we analyse if Switzerland is a land of opportunity overall, or whether there are some regions with a high intergenerational (income) mobility driving the result. This is to say, that the national mobility estimates in the previous section could mask differences in mobility across regions. Regional analysis can be important to guide further research in finding policies that promote upward mobility. We will look more deeply at regional covariates in Section 4.7.

We analyze mobility on two geographical entities: labor market regions (n=106) and cantons (n=26). Cantons are the main political entities with substantial authority in policy setting. Labor market regions depict commuting patterns and are constructed by the Swiss Federal Office of Statistics, similarly to commuting zones in the US.[13]

Figure 4.7 shows heat-maps of the different income mobility measures on a labor market level. The precise estimates can be found in Table C.9 in the Appendix. Figure C.1 and Table C.10 in the Appendix do the same on a cantonal level. The darker the colors, the higher the mobility estimates. One can see that there is spatial correlation: Regions with higher mobility are more likely close to regions with high mobility. In general, there is a pattern with higher mobility in urban regions and lower mobility in the mountains. Again, the cycle of privileges (Q5Q5) measure seems to be less related to the other measures. In urban regions, we also see that children from rich parents are also more likely to stay rich.

Figure 4.8 shows the same for educational mobility on a labor market level.[14] Here, brighter colors highlight regions with higher educational mobility. The patterns is similar to the on from income mobility before: Urban regions show higher mobility than regions in the mountains.

Table 4.6 summarizes how the mobility estimates vary across labor market regions. Looking at the variation coefficient, we see the highest variation in the educational mobility measure «Share Bottom 20 in HS». This varies between 2.6 percent in Schanfigg and 27.4 percent in Geneva. The cycle of privileges (Q5Q5) and the cycle of poverty (Q1Q1) estimates show a small variation coefficient (CV) and seem to be more homogeneous across Switzerland.

The highest absolute mobility at $p = 25$ (AUM25) can be found in Limmattal (close to the city of Zurich) with 56, the lowest in Kandertal (in the mountains) with 41. This means

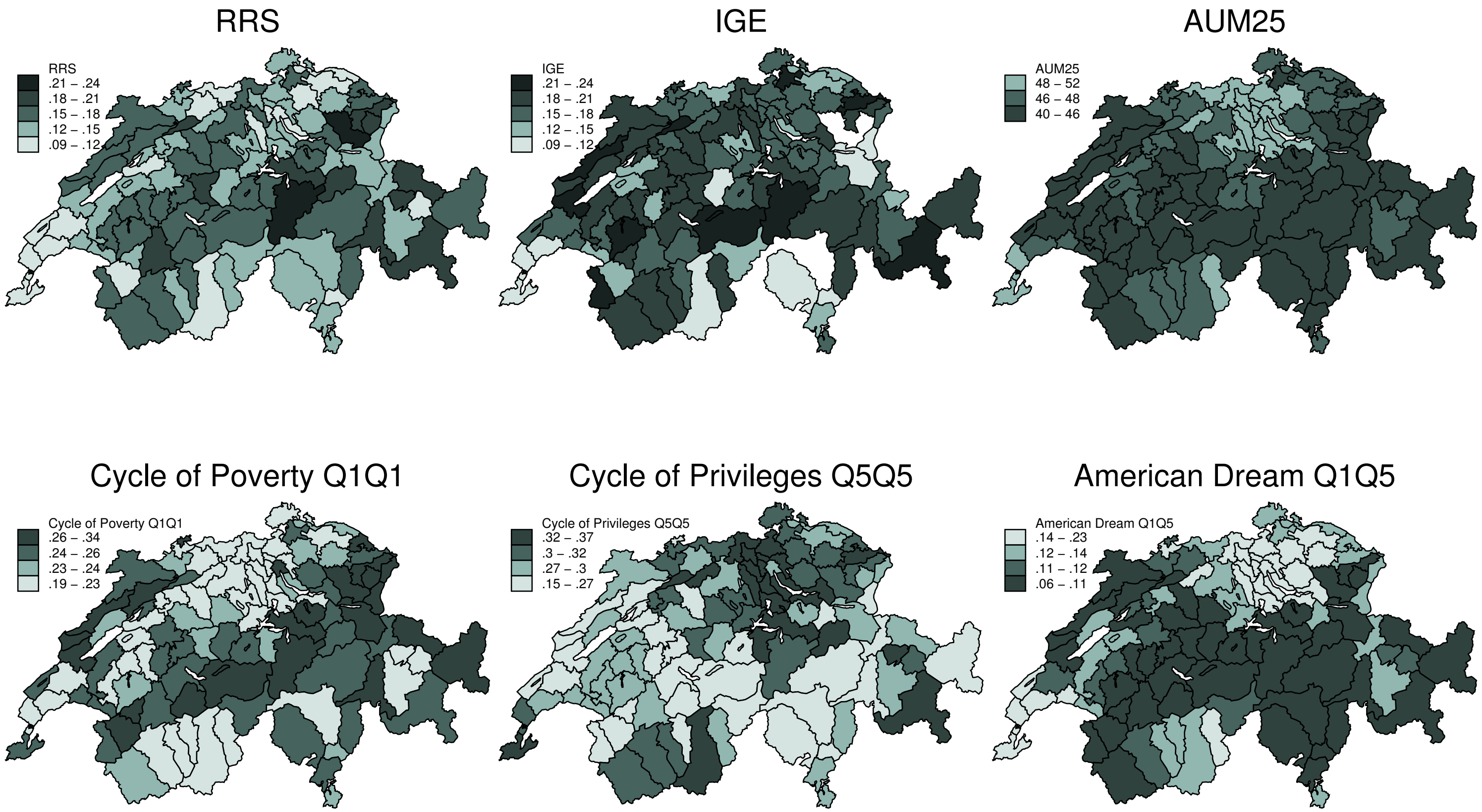
[13] In the main specification, we use the mother's municipality in 2010 to approximate childhood location because we do not have panel information on the exact location until 2010. In the robustness section we show that the results of our maps are robust to several location specifications, such as place of birth or place of residence of the child (see Section 4.8.3).

[14] Again, the precise estimates can be found in the Appendix in Table C.12

children from the bottom half of the income distribution will on average reach rank 56 in the national income distribution in Limmattal, while children from Kandertal reach 41. When we calculate this difference back to income levels, this amounts to around 12,500 Swiss Francs[15], which is around twice the median monthly salary. The standard deviation in absolute mobility is 2.3 (and 1.8 when using cantonal units). This is slightly higher than the standard deviation in Sweden (1.6) (Heidrich, 2017). Also, the range in Sweden is smaller: Absolute mobility at $p = 25$ varies from 41 in Arjäng to 49 in Värnoma. In terms of income, this difference mounts to 90 percent of a monthly salary in Sweden. The variation in Switzerland is however smaller than in the US, where absolute mobility at $p = 25$ varies between 36 ($\approxeq$ \$26,300) in Charlotte and 46 ($\approxeq$ \$37,900) in Salt Lake City (Chetty et al., 2014a).

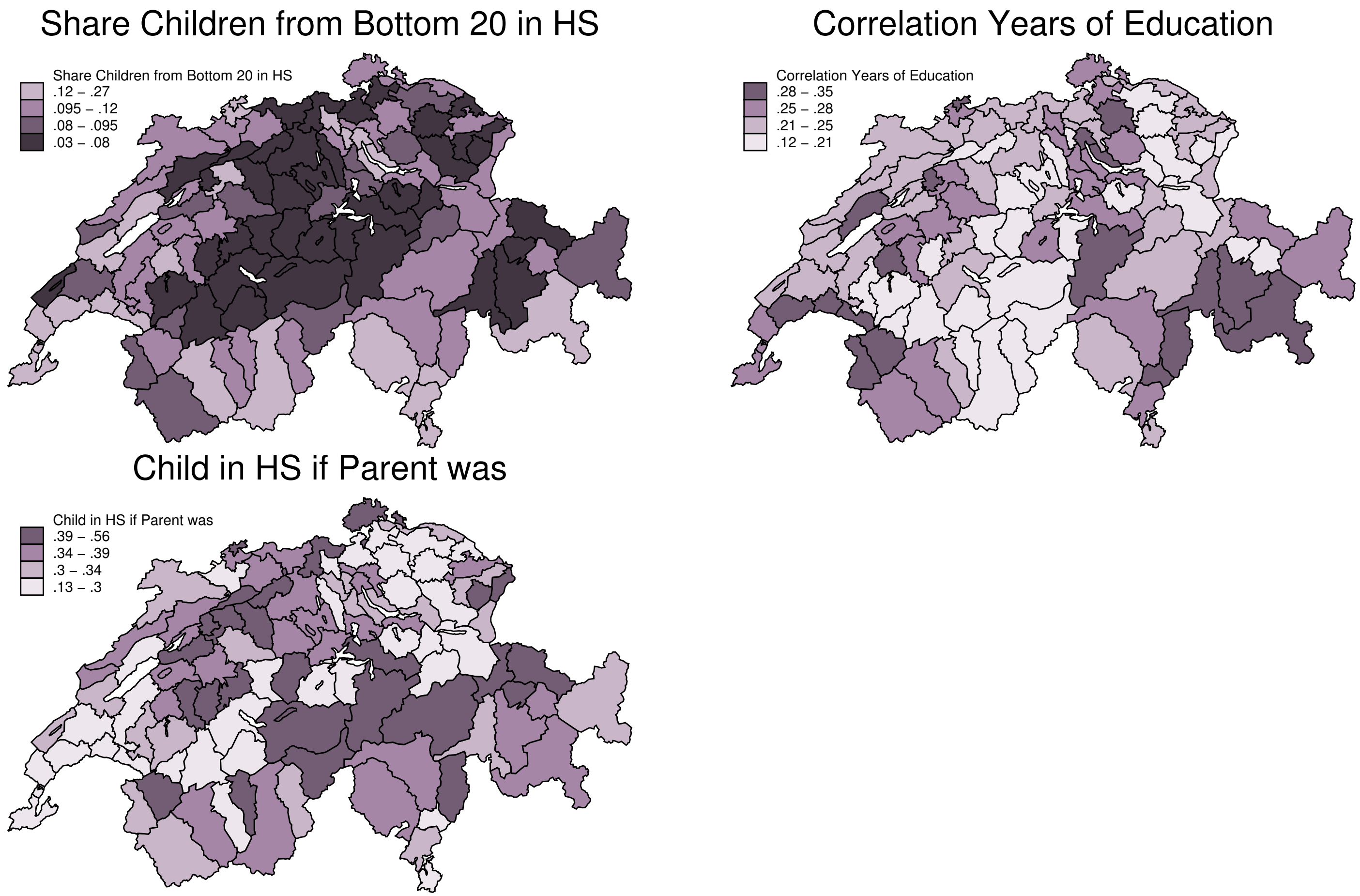
[15] 12,572 Francs at age 30-33 and 14,271 at age 39 to 42

FIGURE 4.7: Income Mobility Estimates Labor Market Regions



Notes: This figures shows how income mobility estimates vary across labor market regions. Brighter colors indicate higher mobility.

FIGURE 4.8: Educational Mobility Estimates Labor Market Regions



Notes: This figures shows how educational mobility varies across labor market regions. Brighter colors indicate higher mobility.

4.7 Drivers of Mobility

4.7.1 Educational Tracks

In the previous paragraphs, we were looking at income and educational mobility separately. Now we are interested in how upward income mobility depends on the educational track. This might help in understanding the puzzle, why income mobility is high, even though educational mobility is low. We take up the approach by Chetty et al. (2020). In their paper, they analyze the background of children and the probability to move up the income ladder for different colleges in the US. Instead of colleges, we are looking at educational tracks.

If there are a lot of children that move up the income ladder in a specific educational track, this might be (a) because of high probability for poor children to move up the income ladder or (b) because there are many children from poor families in this track. The product of (a) and (b) yields the mobility rate—the share of children from poor backgrounds that reach the top. Say differently, we can decompose the mobility rate into the *probability of moving up* and into *access* it provides to children from low-income families.

We define educational tracks as a two element decision to simplify the analysis. The first element is the education after compulsory school: the upper secondary track. This can be: VET, gymnasium (high school), specialised middle school, or none (only mandatory). VET is the most common track, gymnasium—which is like a high school— the second most common.[16] The second element is the highest education on the tertiary or post-secondary level, if children decide to take up such an education. It comprises Bachelor, Master, PhD, HF, HFP, and vocational matura[17]. Bachelor's and Master's degrees can be obtained at a university or at an University of Applied Sciences. PhD can only be obtained at universities. HF and HFP are specialized, vocational specific, further educations that require a VET diploma.

In the first step, we are looking at the American Dream measure. Thus, access is defined as the share of children from the bottom quintile of the income distribution. Upmover rate is defined as the share of children from the bottom quintile that make it to the top quintile. The same with «medium upward mobility», the probability that children from the bottom half make it to the top half, is shown in the Appendix in Figure C.3.

Figure 4.9 shows the up-mover rate on the vertical and the access rate on the horizontal for the specified tracks. It sticks out that most tracks lie on a curve and that there is

[16] Since there are very few children in specialised middle school, we refrain from showing their further paths

[17] Strictly speaking, the vocational matura (baccalaureate) is also part of the upper-secondary level.

TABLE 4.6: Variation in Mobility on Labor Market Level

Variable	Mean	Std.	Min	Max	CV
Income Mobility					
RRS	0.145	0.026	0.099	0.233	0.181
IGE	0.170	0.037	0.081	0.257	0.215
AUM25	46.794	2.279	40.815	51.731	0.049
Q1Q1	0.245	0.028	0.188	0.336	0.114
Q5Q5	0.296	0.038	0.145	0.368	0.127
Q1Q5	0.126	0.032	0.064	0.232	0.255
Educational Mobility					
Years Education	0.243	0.046	0.121	0.350	0.189
Share Bottom 20 in HS	0.104	0.044	0.026	0.274	0.427
Child in HS if Parent was	0.348	0.074	0.130	0.560	0.214

Notes: This table shows summary statistics for the labor market regions (n=106) for different income and educational mobility estimates. CV is the variation coefficient. The mean represents the unweighted mean and therefore differs from the national mean.

a trade-off: Tracks providing a high probability of moving up, give only little access to poor children, for example «Gym+Master». There are only around 10 percent of children with parents from the bottom quintile in this track. However, if a child from the bottom quintile is in this track, it has a high likelihood of 37 percent of moving to the top quintile. This could either be by selection, e.g. more able children from poor families select into this track (selection effect) or because the track really adds some value to those children (causal effect).

The most interesting part is that there are some tracks that lie off this curve and provide a better trade-off. Those are the tracks on the upper-right. They provide relatively high access *and* relatively high chances to move up the ladder. It is striking that all those tracks start with VET and add some higher education. Thus, one can conclude that tracks that start with *VET and add some other higher education* inhabit lot of children that move up, because they provide relatively high access and relatively good wage outcomes. Interestingly, children that have a vocational degree «only» have relatively low mobility rate. Thus, the high mobility in Switzerland is not necessarily because of the VET system, but because there is a high permeability to further education when children start with VET.

The numbers are shown in Panel (a) in Table 4.7. More importantly, it also shows that the mobility rate—the product of the up-mover rate and the access rate—is highest for those tracks that start with VET and add some further education. Panel (b) in Table 4.7 shows the same results but defines access as the share of children from the bottom half of the parent income distribution, and the upmover rate as making it to the top half of the income distribution. Again, tracks starting with VET and add some further education show the highest mobility rate. It is interesting that with this measure also «VET only» shows a relatively high mobility rate. When looking at medium upward mobility, «VET only» might still drive upward mobility. Children are, however, less likely to make it to the top quintile with VET only.

Taken together, the low educational mobility in terms of educational tracks does not matter too much for income and therefore only weakly translates into low income mobility. «Medium upward mobility» is high, even when children «only» conduct vocational education. Even more promising for policy advice are the tracks that start with VET and add some further education. There are many children in those tracks that achieve the American Dream. This is likely because children can opt for this kind of education even if parents are credit constraint. Children of poorer parents can opt for VET, which comes at very little costs for parents and even gives the children a small wage. After the children received their VET diploma they can opt for further education. A large share of this further education can be done parallel to a job, which further facilitates financing this human capital investment. This finding is in line with recent and seminal evidence on the importance of credit constraints in human capital accumulation (e.g. Black et al. (2020); Card & Solis (2020); Bettinger et al. (2019); Chu & Cuffe (2020); Denning & Jones (2019); Brown et al. (2012); Carneiro & Heckman (2002)).

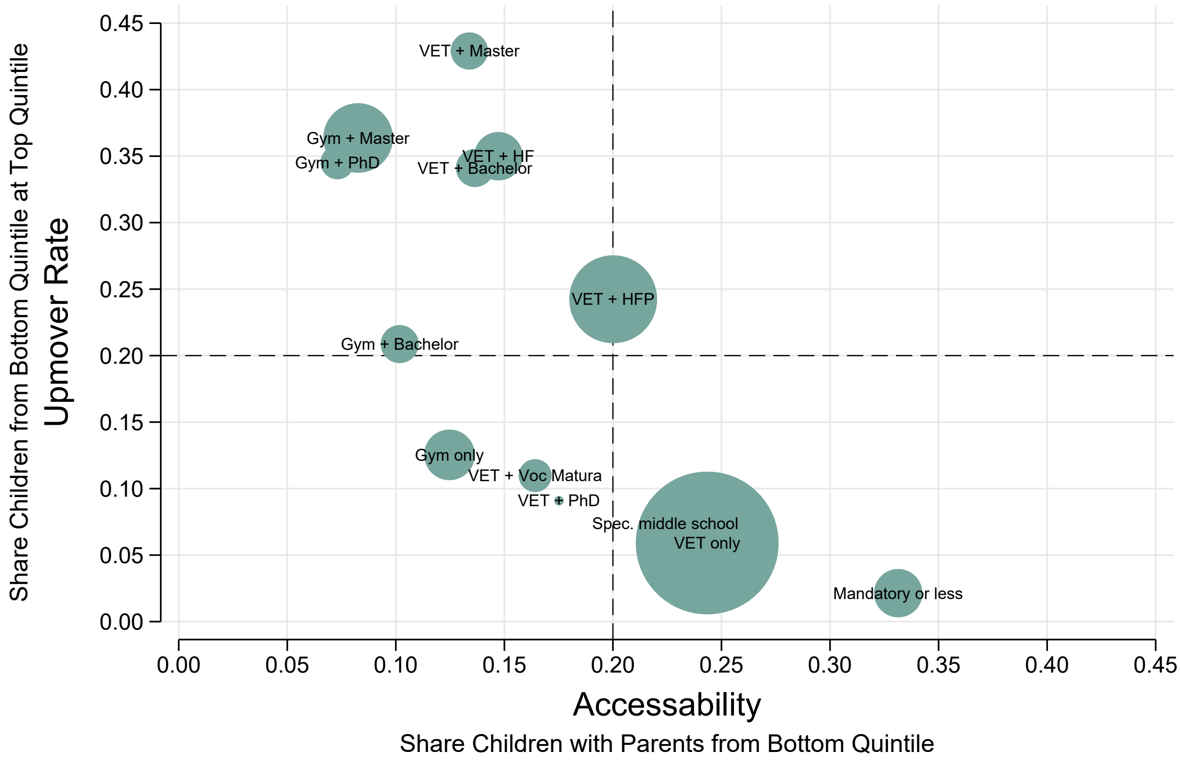


FIGURE 4.9: Access and Upward Mobility Rate by Educational Track (Q1Q5)

Notes: This graph shows how upward mobility (upmover rate) and access differ between educational tracks. Upmover rate is defined as the share of children with parents in the bottom quintile that move to the top quintile (Q1Q5). Accessibility is defined as the share of children from the bottom quintile relative to the total number of children in the educational track. The size of the points is proportional to the number of children in that track. The syntax of the educational tracks is defined as follows: [Upper Secondary Education] + [Highest Post Secondary Education]. *Gym* refers to gymnasium (academic high school), *VET* to vocational education and training. *HF* and *HFP* are occupation specific higher educations. *Voc Matura* refers to "vocational matura", which is VET with more formal education, *spec. middle school* refers to specialized middle schools, which is like a professional high school and not as selective as the gymnasium.
The upmover rate multiplied with the access rate equals the mobility rate, which is the share of children that climb from the bottom to the top quintile *relative to all children in that track*. In contrast, the upmover rate is the share of children that climb to the top quintile *relative to children from parents in the bottom quintile.*

TABLE 4.7: Access and Upmover Rate by Educational Track

Educational Track	Upmover Rate		Access Rate		Mobility Rate		N
	$P[R_C > 80 \| R_P \leq 20]$		$P[R_P \leq 20]$		$P[R_C > 80 \wedge R_P \leq 20]$		
VET only	0.059	(0.001)	0.243	(0.001)	0.014	(0.000)	31,936
Gym only	0.125	(0.008)	0.125	(0.003)	0.016	(0.001)	1,922
Gym + Bachelor	0.209	(0.014)	0.102	(0.003)	0.021	(0.002)	853
Gym + Master	0.364	(0.010)	0.083	(0.002)	0.030	(0.001)	2,472
Gym + PhD	0.345	(0.022)	0.073	(0.003)	0.025	(0.002)	475
VET + Bachelor	0.341	(0.014)	0.136	(0.004)	0.047	(0.002)	1,155
VET + HF	0.350	(0.010)	0.147	(0.003)	0.052	(0.002)	2,066
VET + HFP	0.242	(0.004)	0.200	(0.002)	0.049	(0.001)	9,723
VET + PhD	0.091	(0.039)	0.175	(0.021)	0.016	(0.007)	55
VET + Voc Matura	0.110	(0.010)	0.164	(0.005)	0.018	(0.002)	1,011
Mandatory or less	0.021	(0.002)	0.331	(0.004)	0.007	(0.001)	4,627
Spec. middle school	0.074	(0.006)	0.224	(0.004)	0.017	(0.001)	2,208

Notes: This table shows how the upward mobility (upmover rate) and access differs between educational tracks as shown in Figure 4.9. The upmover rate is defined as the share of children with parents in the bottom quintile that move to the top quintile (Q1Q5). Accessibility is defined as the share of children from the bottom quintile relative to the total number of children in the educational track. The mobility rate shows the share of children who move from the bottom to the top quintile relative to all children. In Table C.4 in the Appendix, the upmover rate is defined as the share of children with parents in the bottom half that move to the top half.
Standard errors of the mean (se) are shown in parentheses. N refers to the observations in either the bottom quintile for Panel (a) or the bottom half Panel(b).

4.7.2 Regional characteristics

4.7.2.1 Public Goods and Fiscal Policies

For policymakers, it is important to understand how public policy can increase upward mobility. To shed first light on a broad level, we analyse how income mobility is related to tax rates and public expenditures. In Switzerland, there is considerable variation in tax rates and public expenditures since it is organized federally.

To capture tax policies, we rely on local personal income tax rates from Parchet (2019). The author computes the consolidated (cantonal, municipal and, church) tax rates for all municipalities in Switzerland between 1983 and 2012. We define the local tax rate for four income brackets in each labor market (LM) region as the (unweighted) average consolidated tax rate across the LM region's municipalities. To proxy local public goods provision, we use data on local public finances from Fontana-Casellini (2020). The author has collected information on local (municipal and cantonal) expenditures (by functional category) since 1950. The coverage rate ranges between 2 percent in 1982 to 74 percent in 2004, with an average of 50 percent. We define local government spending as the (unweighted) mean of (per capita) municipal and cantonal spending in the municipalities in each LM region. Ideally, we aspire to capture local conditions when children grow up. Therefore, we take the mean local tax rate and government spending between 1982 and 2004, when our children are between 15 and 20 years old.

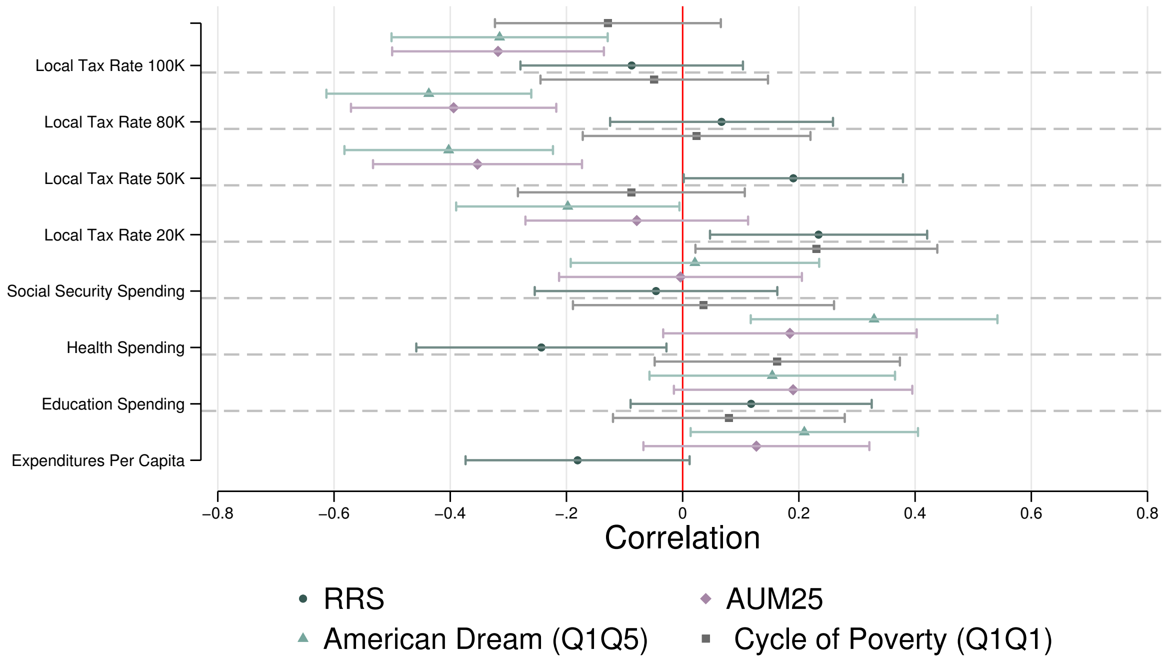
Figure 4.10 shows how tax rates at different income levels and different expenditures are related to income mobility. Symbols display the value of the correlation of our mobility measures with regional characteristics, while the lines show the 95% confidence interval based on standard errors clustered at the LM region level.

The figure reveals some interesting patterns. First, there is a negative correlation between tax rates and the share of children achieving the American Dream (Q1Q5) and between absolute upward mobility (AUM25). Thus, the lower tax rates, the higher income mobility. Interestingly, in terms of Q1Q5 and AUM25, mobility is also higher for total per capita expenditures, health and education spending. It is also interesting that the higher social security spending, the higher the number of children trapped in the poverty cycle. Of course, causality could go in both directions. Overall, we find that regions that invest more in health and education exhibit higher levels of upward mobility. In contrast, regions with higher tax rates tend to have lower levels of upward mobility.

4.7.2.2 Income Inequality («Great Gatbsy Curves»)

There is currently concern that increasing income inequality could also lead to lower income mobility. The relationship between inequality and income mobility has been

FIGURE 4.10: Public Expenditures and Taxes



Notes: This figure displays how income mobility estimates are related to public expenditures and tax rates. On the y-axis, we list the local Labor Market (LM) characteristics. Each symbol represents a different intergenerational income mobility measure (RRS, AUM25, Q1Q5, Q1Q1) and plots the unweighted correlation of intergenerational income mobility with local conditions across LM regions. The lines represent 95% confidence intervals, calculated using standard errors clustered at LM region level. We evaluate the tax rate of four income levels: 20,000 CHF, 50,000 CHF, between 80,000 CHF and 100,000 CHF. Regarding local spending, we consider per capita expenditures and three (per capita) spending categories: Education, health, and social security.

named the «Great-Gatsby Curve» (Corak, 2013). It is based on the empirical finding that countries with higher inequality also show lower mobility. Here, we test whether we can also observe this relationship on a within-country level.

We find mixed evidence for the existence of a Great Gatsby curve in Switzerland. Figure 4.11 plots income mobility against income inequality on a cantonal level. The linearly fitted line shows the relationship weighted by the population size of the canton. Income inequality is measured on family income level with the Gini Index. This index ranges from 0 to 1. The higher the index, the higher inequality. Figure C.4 in the Appendix shows the results are similar when using labor market regions instead of cantons.

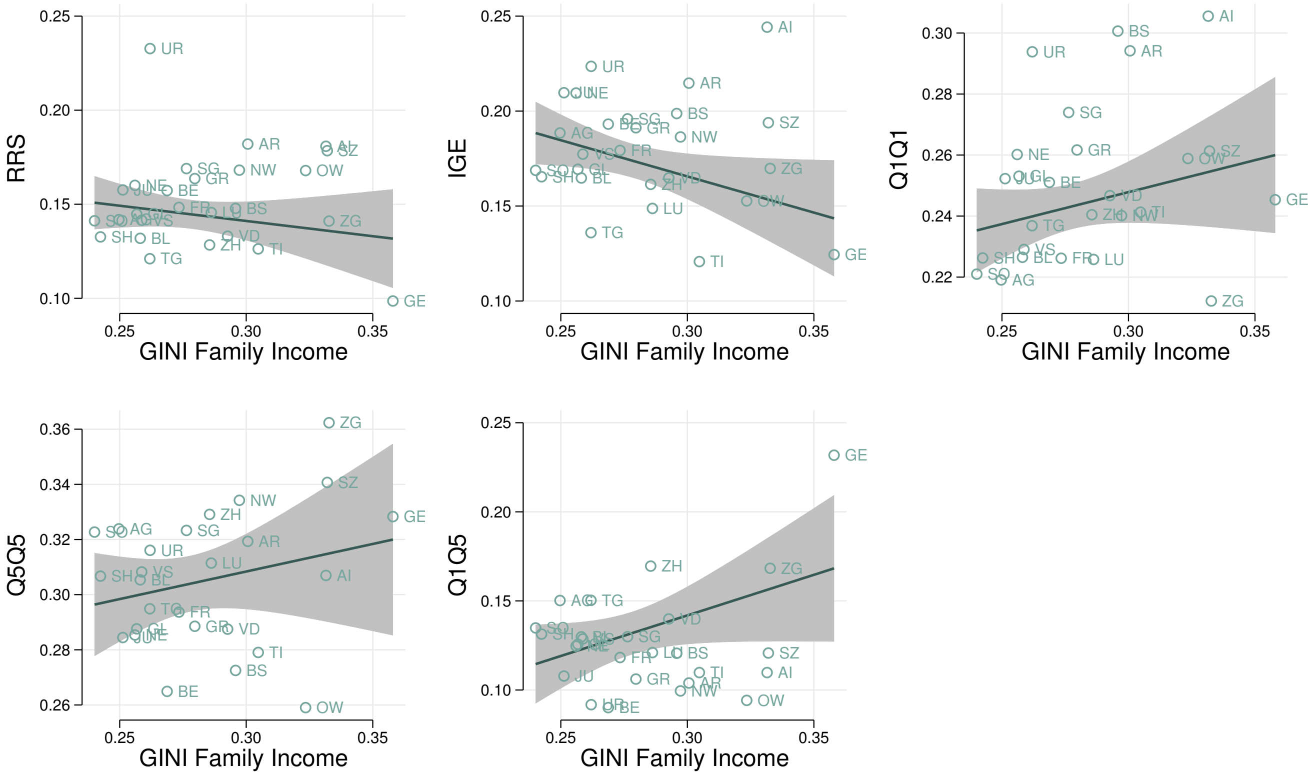
The rank-rank slope (RRS) and the IGE show a negative correlation: The higher inequality, the higher mobility. The results of the American Dream (Q1Q5) go in the same direction. If inequality is higher, children from the bottom quintile are also more likely to reach the top quintile. Thus, when looking at the above mentioned measure, there is no support for the Great Gatsby curve. There is, however, evidence for a Great-Gatsby curve when looking the directional mobility measures Q5Q5 and Q1Q1. In cantons with higher inequality, there seems to be higher persistence at the top and at the bottom of the parental income distribution. How can this counter-intuitive finding on Q1Q5 and Q1Q1 be interpreted? Maybe more inequality leads to more polarized outcomes. It could inspire children to a «all or nothing» mentality. Some children are really incentivized to make it to the top, while others do not even try to do so.

4.7.3 Individual Socio Demographic Characteristics

For policymakers, it is also important to know which individual characteristics are associated with low income mobility, for example, to create programs targeting specific groups. Column (1) in Table 4.8 shows how the American Dream (Q1Q5) measure varies for different personal characteristics. Column (2) shows how large the share is of children in the bottom quintile for this characteristic. Column (3), the mobility rate, shows the share of children achieving the American dream with a certain characteristic.

The most striking difference occurs between women and men. While almost 19 percent of men from the bottom quintile end up in the top quintile of the income distribution— in the distribution with men and women— this share is only around 6 percent for women. Men are, thus, around three times more likely to achieve the American Dream. Of course, this number reflects individual labor income and not household income and is subject to within-household labor division within a household. Thus, gender specific consumption inequality is most likely lower.

FIGURE 4.11: Great Gatsby Curves (cantonal level)



Notes: This graph shows how income mobility is relates to income inequality on a cantonal level. It does so for different income mobility measures. Income inequality is measured on a family level when children are between 15 and 20 years old. The grey line shows the fitted slope between the values weighted by the size of observations in each canton.

Looking at different religions reveals other interesting insights: Jewish children (although small in sample size), show the highest likelihood to climb up the ladder. Also, the point estimate for Muslim children is high. Protestant children have to lowest likelihood to achieve the American Dream—which is noteworthy since the origins of the American Dream trace back to the Protestant Reformation in Europe.

The next interesting finding is that children with parents born abroad have a higher likelihood to achieve the American Dream. This is not only true if parents were born in high-income countries, like the Germany, the UK, or France, but also for low or middle-income countries, such as Turkey, Poland, or Bosnia. Thus, Switzerland provides good opportunities for second-generation immigrants.[18]

Many variables in Table 4.8 are correlated. Therefore, it is hard to figure out which variable indeed predict upward mobility. To further understand which personal variables drive upward mobility, we conduct a LASSO regression (Tibshirani, 1996).

Table 4.9 shows the «post-selection OLS coefficients». Also in this multivariate regression for which the variables were selected by LASSO, gender is the strongest predictor of the American Dream measure. Similarly, children of immigrated parents show a higher upward probability and Protestants the lower mobility. Also, regions and language regions are predictive of upward mobility.

4.8 Robustness

4.8.1 Attenuation Bias

Attenuation bias arises when transitory income shocks are not filtered out. This will attenuate the correlation between child and parents' earnings, leading to upward biased estimates of mobility. It is easy to see when using a single point in time. If transitory fluctuations are not serially correlated, averaging income across more years of observations eases the attenuation bias (Solon, 1992; Mazumder, 2005).

To understand whether our estimates suffer from such attenuation bias, we vary the number of years to calculate the average parental income. Figure C.5 Panel (a) shows the results from an OLS regression of child rank on parent rank varying the number of years over which we aggregate parent mean income. We start with one year, the year when the child is 15, up to fourteen years, the years when the child is between 15 and 28. In our baseline estimates, we measure parent mean income when the child is between 15 and 20. Thus income is averaged across six years. In the graph, the baseline estimate

[18]Switzerland experienced strong immigration during the Yugoslav Wars 1991 to 2001. Since we are looking at cohorts born until 1984 in Switzerland, those children are not yet in the sample.

TABLE 4.8: American Dream by Personal Characteristics

		(1) AD (Q1Q5)		(2) Share Bottom 20		(3) Mobility Rate		(4) N (Q1)	(5) N
Sex	Male	0.188	(0.001)	0.200	(0.001)	0.038	(0.000)	94,273	472,072
	Female	0.057	(0.001)	0.200	(0.001)	0.011	(0.000)	90,368	451,190
Parents Divorced	no	0.127	(0.001)	0.208	(0.000)	0.026	(0.000)	158,891	765,513
	yes	0.108	(0.002)	0.163	(0.001)	0.018	(0.000)	25,750	157,749
Religion	Catholic	0.134	(0.002)	0.232	(0.001)	0.031	(0.001)	22,523	97,047
	Protestant	0.117	(0.003)	0.183	(0.001)	0.021	(0.001)	14,857	81,402
	Other Christian	0.132	(0.007)	0.241	(0.004)	0.032	(0.002)	2,370	9,854
	Jewish	0.333	(0.066)	0.106	(0.014)	0.035	(0.008)	51	481
	Islamic	0.150	(0.020)	0.354	(0.016)	0.053	(0.008)	307	867
	Other	0.133	(0.027)	0.166	(0.012)	0.022	(0.005)	158	951
	No confession	0.160	(0.004)	0.148	(0.001)	0.024	(0.001)	9,378	63,257
Language Region	German	0.124	(0.001)	0.196	(0.000)	0.024	(0.000)	137,178	699,709
	French	0.133	(0.002)	0.197	(0.001)	0.026	(0.000)	36,643	186,126
	Italian	0.097	(0.003)	0.282	(0.002)	0.027	(0.001)	9,549	33,903
	Romanian	0.062	(0.007)	0.355	(0.009)	0.022	(0.003)	1,103	3,111
Any Parent born abroad	no	0.119	(0.001)	0.212	(0.000)	0.025	(0.000)	148,169	698,614
	yes	0.145	(0.002)	0.162	(0.001)	0.023	(0.000)	35,345	218,416
Boths Parents born abroad	no	0.121	(0.001)	0.200	(0.000)	0.024	(0.000)	175,425	877,655
	yes	0.181	(0.004)	0.205	(0.002)	0.037	(0.001)	8,089	39,375
Country of Birth Father	Switzerland	0.120	(0.001)	0.201	(0.000)	0.024	(0.000)	164,750	820,572
	Italy	0.141	(0.004)	0.248	(0.002)	0.035	(0.001)	8,865	35,798
	Germany	0.177	(0.009)	0.118	(0.002)	0.021	(0.001)	2,002	16,939
	France	0.162	(0.010)	0.164	(0.004)	0.027	(0.002)	1,330	8,133
	Austria	0.164	(0.013)	0.140	(0.004)	0.023	(0.002)	837	5,995
	Spain	0.191	(0.015)	0.176	(0.006)	0.034	(0.003)	686	3,904
	Turkey	0.147	(0.010)	0.391	(0.009)	0.058	(0.004)	1,201	3,068
	Cechia	0.149	(0.032)	0.081	(0.007)	0.012	(0.003)	121	1,489
	UK	0.178	(0.028)	0.132	(0.009)	0.024	(0.004)	185	1,404
	Netherlands	0.134	(0.029)	0.125	(0.010)	0.017	(0.004)	142	1,136
	Croatia	0.197	(0.035)	0.119	(0.010)	0.024	(0.005)	127	1,063
	Poland	0.159	(0.034)	0.120	(0.011)	0.019	(0.004)	113	943
	Greece	0.196	(0.033)	0.158	(0.012)	0.031	(0.006)	148	934
	Algeria	0.162	(0.024)	0.272	(0.015)	0.044	(0.007)	228	838
	Serbia	0.138	(0.027)	0.199	(0.014)	0.027	(0.006)	160	803
	Portugal	0.194	(0.034)	0.179	(0.014)	0.035	(0.007)	139	777
	Bosnia	0.208	(0.034)	0.190	(0.014)	0.040	(0.007)	144	759
Country of Birth Mother	Switzerland	0.121	(0.001)	0.210	(0.000)	0.025	(0.000)	159,385	758,597
	Italy	0.145	(0.005)	0.252	(0.003)	0.036	(0.001)	4,524	17,953
	Germany	0.157	(0.007)	0.129	(0.002)	0.020	(0.001)	2,673	20,722
	France	0.140	(0.008)	0.161	(0.004)	0.023	(0.001)	1,746	10,838
	Austria	0.153	(0.009)	0.177	(0.004)	0.027	(0.002)	1,445	8,179
	Spain	0.208	(0.016)	0.165	(0.006)	0.034	(0.003)	665	4,035
	Turkey	0.167	(0.012)	0.404	(0.010)	0.067	(0.005)	1,014	2,512
	Cechia	0.203	(0.034)	0.099	(0.008)	0.020	(0.004)	138	1,401
	UK	0.230	(0.023)	0.127	(0.007)	0.029	(0.003)	322	2,539
	Netherlands	0.119	(0.018)	0.121	(0.006)	0.014	(0.002)	312	2,572
	Croatia	0.179	(0.029)	0.125	(0.009)	0.022	(0.004)	173	1,388
	Poland	0.190	(0.031)	0.138	(0.010)	0.026	(0.005)	163	1,183
	Greece	0.117	(0.031)	0.152	(0.013)	0.018	(0.005)	111	730
	Algeria	0.171	(0.037)	0.222	(0.019)	0.038	(0.009)	105	474
	Serbia	0.119	(0.026)	0.181	(0.013)	0.022	(0.005)	151	835
	Portugal	0.149	(0.044)	0.135	(0.015)	0.020	(0.006)	67	498
	Bosnia	0.172	(0.030)	0.203	(0.014)	0.035	(0.006)	163	804

Notes: This table shows how upward mobility varies for different personal characteristics. *AD (Q1Q5)* shows the share of children from the bottom quintile that move to the top quintile for a given characteristics, *Share Bottom 20* indicates the share of children in the bottom quintile in this group, *Mobility Rate* shows the share of children achieving the American Dream overall in this group, *N (Q1)* shows the number of observations in the bottom quintile, *N* shows the number of observations for all parent income groups. Standard errors of the mean are shown in parentheses.

TABLE 4.9: Post LASSO Regression Results

Dep.Var.:		American Dream (Q1Q5) (Post LASSO Coefficient)
Sex	Male	0.149
Any Parent born abroad	No	-0.019
Both Parents born abroad	No	-0.033
Religion	Protestant	-0.020
NUTS-2 Region	Lake Geneva	0.035
	Mittelland	-0.013
	NW	0.036
	Zurich	0.098
Language Region	Italian	-0.031
Intercept		0.102

Notes: This table shows the results of an OLS regression where the coefficients are select using LASSO. The set of potential (factor) variables includes sex, religion of child parents immigration, country of father, country of mother, language region, NUTS-2 region, maritial status of mother, maritial status of father. Year of birth fixed effects included everywhere. Total number of covariates: 85. Number of observations included: 46,158. Implemented in Stata with *rlasso* by Ahrens et al. (2019). λ is determined by the with heteroskedastic plugin method.

corresponds to the vertical line. The rank-rank slope based on one year of income data is 0.132, which is lower than the rank-rank slope based on six years (0.141).

This attenuation bias is much smaller than the one encountered by Solon (1992). His IGE estimates were 0.3 for a single year and 0.4 when using a five-year average. Mazumder (2005) reports that even five-year averages suffer from attenuation bias. However, we find that the rank-rank slope is virtually unaffected by adding more years of observations beyond six years: The rank-rank slope is 0.144 when we use 12 years of observations and 0.144 when we use 16 years. The quality of our data and the rank-rank specification lead to stable estimates. The magnitude of the attenuation bias is comparable to the one found by Chetty et al. (2014a). They noticed an increase of 6.6 percent in the rank-rank slope, when five years of observations were used instead of a single year and nearly no changes in estimates when adding more years beyond five years.

Panel (b) tests how robust our estimates are to the number of years used to average child's income. The first point uses only the year when the child is 30 years old. This yields a rank-rank slope of 0.125. The vertical line corresponds to the baseline specification with a rank-rank slope of 0.141. Beyond this point, the number of cohorts is decreasing in the number of years. This is because in the core sample we can observe income for every cohort up to the age of 33. The rank-rank slope in the last point is

0.151, the sample includes only the 1967 and 1970 cohorts, and uses mean child income between the age of 30 and 47 (17 year average). The rank-rank slope increases when we aggregate child income over a larger time span. Thus, this bias is of small magnitude. Moreover, the bias includes also part of the life-cycle bias, as children are on average older. Even with this «upper-bound» estimate, Switzerland would rank among the countries with the highest relative mobility in terms of rank-rank slope.

4.8.2 Life-Cycle Bias

Life-cycle bias arises when income measured at the life-cycle stage systematically deviates from lifetime income. This might be the case when child income is measured earlier than their parent's income or when only a short snapshot of lifetime income is used. Life-cycle bias imposes a danger to understate income for those with steeper income profiles, like the more educated children. This can lead to an overestimation of mobility.

Figure C.6 evaluates the sensitivity of our baseline estimates to changes in age at which child income is measured. We plot the coefficients of separate rank-rank slopes by varying the age at which a child incomes are measured for three samples. Parents' income is measured when the child is between the age of 15 and 20. Parents are ranked relative to other parents of children in the same birth cohorts. Child incomes are averaged across four years, at different ages up to the age 47. In the first point, the mean income is averaged over the age of 21 and 24. The straight line plots the coefficient of the core sample, the vertical line shows the baseline estimates. As before, beyond that point the number of cohorts decreases in child age. Around the age of 33—which is defined as the mean of age 30 to 33—, the rank-rank slope is reasonably stable. Life-cycle bias should not be an issue for our rank-rank estimates estimates.

When varying the age at which a child incomes are measured, we implicitly vary the number of cohorts and the calendar years at which child income is measured. However, we get similar results if we keep calendar year 2017 fixed and vary the cohorts, and if we restrict the sample to the 1967 to 1970 cohorts. The dashed line shows the RRS for the 1967 and 1970 cohorts, for which we observe income up to age of 47. The dotted line reports the coefficients when keeping calendar year fixed from 2014 to 2017 and varying the cohorts.

There is, however, substantial life-cycle bias when looking at the IGE. We have shown this in Panel (b) of Figure 4.2 and in Table C.8. Focusing on the rank-rank slope therefore allows us to look at more cohorts since we can measure child income at earlier ages.

A similar bias emerges if parents' income is measured too early or too late. In Table C.3, we evaluate the robustness of our estimates to the age at which parents' income is measured. To simplify the analysis, we focus on father's age. We also report the coefficient of the rank-rank slope when parents' income is measured at father's age 45. For a subset of cohorts, our data allows us to measure parents' income when the child is very young. We also want to test whether financial resources during early childhood matter more for child outcomes than resources at later ages of the childhood. Therefore, we restrict our sample to the cohorts from 1979 to 1981. Then, we measure parents' income when the child is between three and eight, and between nine and fourteen years old. The estimates reveal virtually no variation with father's age between 30 and 50 years old.

4.8.3 Location Choice

In the main specification, we use the mother's municipality in 2010 to approximate childhood location. This is because we do not have panel information on the exact location until 2010. However, we know when a person arrived in a specific municipality in 2010 and in which municipality a person was born. The municipality of birth is usually the location of the hospital in which the mother gave birth.

Table C.13 shows how robust the regional mobility estimates are for different location assignment rules. We use three alternative specifications and compare its correlation with our base specification. «Mother Location 16» restricts the sample to children for which we know for sure that the mother lived in this place when the child was 16 75 percent of mother's still living in the same municipality where they lived when their child was 16. «Child Place of Birth» is the place where the child was born. 77 percent of mothers still live in the same canton where their child was born. Finally, we use child location in 2010 used the location where the child lives when adult. This assignment rule should be taken with caution, as children are more mobile and this location does not present the place where they grew up.

There is a very high correlation between the main specification and the alternative mobility measures. Interestingly, the IGE is also the less robust to income mobility estimate when looking at geographic assignment rules. The correlation is lowest when looking at the location of the child in 2010. However, this is most likely not a good approximation for the place where the child grew up.[19]

[19]We provide the precise income and educational mobility estimates, including standard errors for all geographic assignment rules here: https://www.dropbox.com/s/489coqmuue2tam5/mobility$_geo.7z?dl = 0Zip$

4.8.4 Regional Deflator

Our regional mobility estimates could be affected by differences in purchasing power. Purchasing power is likely to vary between regions in Switzerland. Regional deflation might therefore affect the ranks of the parents and children in the national income distribution. In mountainous regions, prices might be lower and a nominal income might be valued higher than in urban areas.

Although no general regional price indices are available, we can draw from price indices for housing. Table C.14 shows how the rank-rank slope varies when using different regional price indices. Column(1) reports our baseline; we adjust income using the national consumer price index. Column (2) uses the *Residential Property Privately Owned Apartments Price Index*, column(3) the *Residential Property Regional Housing Price Index* and column(4) the *Rented properties, rental housing units price index*.

The rank-rank slope decreases when we account for regional real estate price differences, which means a lower correlation between a child's position and family position in the income distribution. The drop is in line with previous studies (Acciari et al., 2019; Chetty et al., 2014a) and does not substantially affect the estimates. Indeed, one can expect regional price differences to have only a minor effect on intergenerational correlation when most children live close to their parent's place, and regional differences do not significantly change over time.[20]

4.8.5 Capital Income

How susceptible are our estimates to the definition of income? Since our data only used labor income, a natural concern is that our estimates would be different when including capital income. We argue that this is unlikely the case when using rank transformed income measures—which are our main estimates of interest (RRS, Q1Q5, Q1Q1, Q5Q5, AUM25). Accordingly, our rank based mobility measures are well suited to be compared to other countries. We base our argument on two pillars: empirical findings of previous studies and and theoretical arguments.

For Australia, Deutscher & Mazumder (2020) show that the RRS changes only sightly when adding capital to labor income. They find the RRS of wages to be 0.19 and the rank-rank slope and only slightly smaller than the rank-rank slope based on total income 0.22. In contrast, the IGE is much more susceptible to changes in income definition. Based on wages, the IGE is 0.11, while based on total income, the IGE is 0.19. Assuming that capital income in Australia is similarly distributed as in Switzerland, this relative or absolute increase in the RRS would still leave the RRS small compared

[20]E.g. 50 percent of children live closer than 10 miles from their mother's place in 2010

to other countries. We can also compare our RRS to the study by Heidrich (2017) for Sweden, which also only uses labor income. Here, we see that Switzerland (0.14) still has a lower RRS than the Sweden (0.18).

We can also draw from other empirical studies and analyzing the joint distribution of capital and labor income and combine them with theoretical arguments to infer how our results would change with capital income. For Switzerland, Martínez (2020) analyzes the joints distribution for capital and labor income using administrative data. Since we (and other studies) measure at relatively early ages (around 30 and 45), capital income is unlikely to play a major role since it is most prevalent in old ages. It is therefore unlikely that there are major changes when assigning children to their income rank. The same holds true for parents, since most parents are measured around father's age of 45. Rank based measures could also be biased if parents with high capital income have low labor income. We would wrongly classify rich parents as poor, which would then attenuate the RRS. Non-working capital income rich individual is a myth in Switzerland. Capital income is highly correlated with labor income and much more right-skewed than labor income. Therefore, such wrong rank assignments should be rare.

4.8.6 Comparisons with US-Distribution

When comparing the rank-rank slopes between different countries, one concern is that the distributions can differ. For example, inequality in the US is considerably larger than in Switzerland. Therefore, we convert the Swiss income into PPP adjusted US dollars and assign the ranks according to the US distribution according to Chetty et al. (2014a).[21]

Figure C.8 shows rank mobility between the US and Switzerland which permits analysing absolute and relative mobility. For better comparison, we converted the Swiss incomes into the US income distribution. For Switzerland, the constant is higher and the rank-rank slope is considerably lower. A higher rank implies that absolute mobility is higher. Children from similarly poor parents can expect to have much higher wage outcomes in Switzerland than in the US. Only for the top income percentiles, children in the US have higher wage outcomes. The relatively good wage outcomes for Switzerland are at least partly because of the high valuation of the Swiss Franc since the financial crisis.

4.9 Discussion and Conclusion

In this paper, we use administrative income, census, and survey data to document intergenerational income and educational mobility in Switzerland. We analyze how upward

[21]PPP data is retrieved from https://stats.oecd.org/Index.aspx?DataSetCode=CPL. Data for the US from https://opportunityinsights.org/data/.

mobility varies across regions and which personal and regional characteristics correlate with upward mobility. Most importantly, we analyze how income mobility varies between educational tracks.

We find that intergenerational mobility is high in Switzerland. Income mobility in terms of rank-rank slope (RRS) is 0.14 and higher than in all other countries for which high-quality estimates exist. We also see that, compared to other countries, children from the bottom quintile of the parental income distribution are less likely to stay in the bottom quintile themselves («cycle of poverty»). Also, children from the top quintile are less likely to stay in the top quintile themselves («cycle of privileges»). In terms of the «American Dream» measure, which indicates the share of children from the bottom quintile that makes it to the top quintile, only Sweden has a higher share than Switzerland. Taken together, almost all mobility estimates are higher in Switzerland than in the US, Italy, Canada, Denmark, Australia, or Sweden.

Despite the high income mobility estimates, we find that educational mobility is low. Children's educational track and years of education depend considerably on parental income and education. The socioeconomic gap is especially strong when looking at whether children frequent a high school or a VET program (educational track). Children from the top decile of the parental income distribution are almost five times more likely to frequent a high school than those below the top quintile.

We investigate the reasons behind this divergence of educational and income mobility. First, we find that in regions with high educational mobility, there is—in general—also high income mobility. This suggests that educational mobility is still related to income mobility, as seminal theoretical papers suggest (Becker & Tomes, 1986a). However, in Switzerland, low educational mobility translates only weakly into low levels of income mobility. Second, the reason for this weak link might be the permeability of the VET system, not the system per se. We find that educational tracks that start with VET and add some further education account for a large share of upward mobility. Conceptually, this makes sense. VET comes at almost no costs for parents and, therefore, credit constraints are less binding for children's human capital accumulation—if there is ample scope for further education.

Intergenerational mobility varies across regions in Switzerland. This variation is slightly higher than in Sweden but lower than in the US or Italy (Chetty et al., 2014a; Heidrich, 2017). Looking at regional correlates, we find higher income mobility in regions with higher public spending but lower mobility in regions with higher tax rates. Looking at the relationship between mobility and inequality («Great Gatsby Curve»), we find a weak positive relationship with most measures but a negative one when using the cycle of poverty or the cycle of privileges measures.

Our results have potentially important policy implications. First, they show that low educational mobility does not necessarily translate into low income mobility. This is an important and maybe even comforting finding for countries that—by lack of administrative income data—try to infer income mobility from educational mobility. Second, although we currently lack causal evidence, there are good reasons to think of VET as a driver of upward mobility. This system could thus be an interesting policy option for countries with low intergenerational income mobility.

Chapter 5

Multigenerational Mobility in Earnings and Education: Evidence from Administrative Data

Abstract This study measures the persistence in income and education over three generations in Switzerland. I use administrative data covering the universe of labor income since 1982 and family linkages over three generations. Most studies rely on mobility estimates from two generations to predict the long-term dynamics of inequality. However, recent studies show that extrapolating estimates from two generations to more than two generations underestimate the persistence in inequality. My results show that the traditional two-generation paradigm is well suited to predict persistence in earnings, but strongly underestimates persistence in education. The intergenerational income elasticity (IGE) is around 0.22 between two and 0.05 between three generations. The elasticity for education is 0.31 between two and 0.08 between three generations. I test two theories that could explain this excess persistence: the «Clark hypothesis» and the «direct grandparental effects» model. I can reject Clark's hypothesis of a latent persistence as large as 0.75, but cannot reject the direct grandparental effects model.

Keywords: social mobility, intergenerational mobility, multigenerational mobility;
JEL classification: H0, J0, R0

I thank the Federal Office of Statistics and the Central Compensation Office for generously providing the data. Any errors are my own.

5.1 Introduction

The transmission of socio-economic status over multiple generations matters for the long-term dynamics of inequality. Not only does low mobility violate the moral norm of equal opportunities, it also decreases efficiency, since child potentials from lower socio-economic family can lie idle.

To measure mobility over generations, researchers mostly rely on the parent-child paradigm: The measurement of socio-economic status between two adjacent generations. Because of data restrictions, there is little empirical evidence on the degree of socio-economic persistence between multiple generations, the long-term mobility. To still make predictions on the long-term mobility in a society, researchers mostly rely on extrapolating parent-child correlations. Thereby, one assumes that the intergenerational transmission follows a first order auto-regressive process.

In this paper, I measure the correlation in income and education over multiple generations in Switzerland and test several theories on long-term persistence in inequality over generations. To do so, I use administrative data on earnings and education, as well as family linkages, which allows to identify multiple generations. This grants to test whether the two-generation paradigm is well suited to predict transmission of inequality over more than two generations.

The standard workhorse model in intergenerational income mobility research, the Becker-Tomes model, relates parental resources to the outcomes of the child generation (Becker & Tomes, 1986b). Transmission of economic status between grandparents and children only works via parents. Measures on intergenerational mobility then focus on the elasticity of child and parent income. Since this is a stationary first order auto-regressive process, economic status decays at a geometric rate. If, for example, the intergenerational income elasticity equals 0.5 between parents and children, the elasticity would be 0.25 after two generations, and 0.125 after three, and 0.0625 after four generations[1]. Privileges would literally disappear within a century. Therefore, the saying «from shirtsleeve to shirtsleeve after three generations» (Becker & Tomes, 1986b). Parents then have little impact on the earnings of their grandchildren and later descendants.

However, this geometric decay of economic status has been challenged by recent findings. The rare set of studies measuring persistence in economic status over multiple generations find that the decay is lower than the iteration from the two-generation model suggests (Braun & Stuhler, 2018; Solon, 2018; Lindahl et al., 2015; Zeng & Xie, 2014). The

[1]Elasticity and correlation are used interchangeably. The elasticity equals the correlation if the variation (inequality) in the two generations stayed the same

process therefore generates *excess persistence*—the difference between the predicted correlation from a two-generation model and the actual correlation between three or more generations.

Mainly two theories exist to rationalize this excess persistence in economic status. In his controversial book «The Son Also Rises», Clark (2014) claims that the transmission of economic status is determined by an unobserved latent factor which is considerably larger than the usual correlations observed between parents and children. The observed two-generation correlation suffers from an «errors-in-variable» problem and underestimates economic persistence. Based on rare surname correlations, he claims that this latent factor is around 0.75 and universal across time and space. This is a large coefficient, since elasticities in income and education are estimated around 0.5 in countries with low-mobility such as the US. With such a large coefficient, economic status of families would persist over centuries.

Other scholars, such as Mare (2011), purport the idea that grandparents could directly influence child outcomes. In the standard Becker-Tomes model, this transmission over three generations only works indirectly via the parents. However, it is conceivable that grandparents could also directly influence their grandchildren. For example, with financial support, (educational) time spent with their grandchildren, or even via genes that are sometimes only expressed after leaping a generation. Only measuring parent-child correlation would underestimate the persistence of inequality over generations and could, therefore, explain excess persistence. However, as Braun & Stuhler (2018) point out, any process that creates excess persistence beyond the two generation paradigm will also result in a positive grandparent coefficient when regressing child outcome on parent and grandparent outcomes. Therefore, a positive grandparent coefficient is no prove a direct effect of grandparents.

In this paper, I proceed as follows. First, I test whether there is excess persistence beyond the parent-child correlation in Switzerland. To do so, I measure the correlation in outcomes between children and grandparents. Second, using direct empirical multigenerational estimates allows me to identify Clark's underlying latent factor. Thus, I test whether this latent factor is larger than the two-generation reduced form correlation suggests, and if it is indeed as large as 0.75 as claimed by Clark. Second, I assess the direct grandparental effect model. To test whether grandparents could directly influence their grandchildren, I exploit the multi-linguistic nature of Switzerland. If grandparents indeed have a direct influence on children, by spending more time with them, I argue the coefficient would be smaller for grandparents living in a different language region.

A handful of studies measure the persistence of inequality over generations. Lindahl et al. (2015) analyze 900 families and their long-term persistence of human capital and earnings in Malmö, Sweden. They find estimates obtained from data on two generations underestimate long-run intergenerational persistence in labor earnings and education. Thus, long-run social mobility is much lower than previously thought. Braun & Stuhler (2018) analyze the persistence of occupational and educational attainment in Germany. They also conclude that persistence over multiple generations is higher than the two generations paradigm suggests. However, they also find that the persistence is not as high as Clark's hypothesis suggests. Also, their results do not support direct grandparental effects. Long & Ferrie (2018) study occupational data from the US and Britain and find that the two-generation estimates overstate the correct amount of social mobility. Colagrossi et al. (2020) analyze educational outcomes and occupational status in a retrospective survey from 28 European countries. They find that by estimating an errors-in-variables model, that the persistence is indeed as large as Clark suggests. Furthermore, they cannot reject the hypothesis of a direct grandparental effect for some countries. Importantly, they show that there is no single data-generating process to describe multigenerational persistence which would speak against the idea of an «universal law». Further, Zeng & Xie (2014) find support of direct grandparent effects by showing that the educational correlation is higher between grandparents co-residing with grandchildren, while the results of Ferguson & Ready (2011) speak against such a direct effect.

I contribute to the literature in several ways. I add to the rare set of empirical studies that measure multigenerational income mobility with administrative data. With the notable exemption of Lindahl et al. (2015)[2], most studies focus on the educational or occupational domain when measuring social status across multiple generations. However, the (long-term) transmission of income might differ from the transmission of education and occupational choice, since it is easier for parents to influence the educational or occupational choice than the actual earnings (the return to education). Second, I can directly test the Clark hypothesis of a large universal latent factor for the income and education domain. The multi-cultural nature of Switzerland further allows to assess whether the underlying factor is indeed universal, or whether it differs between cultural backgrounds. Third, I can suggestively assess the direct-grandparental effects model by testing, whether the correlation between grandparents and grandchildren is smaller for those pairs that live in the same language region.

The main results show that there is little excess persistence over three generations when looking at earnings. However, there is substantial excess persistence when looking at education. Thus, with Switzerland, the Becker-Tomes model is well suited to explain

[2]In principle, Adermon et al. (2019) also analyze income over multiple generations. However, they have a broader focus and take the entire family dynasty into account.

the multigenerational persistence in earnings, but performs poorly when looking at education. Consequently, the estimates for Clark's latent factor differ between education and income. While it is around 0.17 for the income domain, it is around 0.48 for years of schooling. Also, I cannot reject the direct grandparental effects hypothesis since the grandparent coefficient is higher, although not significantly, for children speaking the same language as their grandfather and living closer to their grandfather.

The discrepancy between income and education can be explained because it is harder for parents to influence earnings of the child than education. For example, Lindahl et al. (2015) also finds lower excess persistence in income than in education, although the difference is smaller than in Switzerland. Another potential explanation for the strong discrepancy in Switzerland might be the vocational education system (VET). Over 70% of children opt for an apprenticeship after mandatory school, which comes at almost no cost for parents and still gives room to many sorts of further education. Also, an apprenticeship usually promotes entry into the labor market and yields good earnings prospects. Thus, years of schooling is might be less linked to earnings compared to other countries.

The finding that persistence over multiple generations can be domain specific has important consequences for interpreting results in inter- and multi-generational mobility. Many studies measuring persistence over generations in education find that excess persistence is high. However, this does not imply that there is also higher persistence in income.

I structure the rest of this paper as follows. First, I explain the theoretical background behind the iterated regression procedure, the latent factor model, and the direct grandparent effect model. Then, I explain the data and show the estimated parameters. The last section concludes.

5.2 Theoretical Background

In this section, I lay out the theoretical models that will be tested in the empirical section. I start with the Becker-Tomes model, which predicts that persistence over generations follows a Markovian process. Then I move to Clark's latent factor hypothesis, which states that persistence is much larger because of a measurement issue in the two-generation case. It also states that the persistence is universal across time and space. The last model that could explain excess persistence is the direct grandparental effects model.

5.2.1 Iterated Regression (Becker-Tomes)

The idea that the inheritance of endowments follows a Markovian process of order 1 grounds on the seminal theoretic papers on intergenerational mobility, such as Becker & Tomes (1979, 1986b) and Loury (1981). For example, Becker & Tomes (1986b) state that social (income) mobility is governed by a simple law:

$$log(y_t) = \alpha + \beta log(y_{t-1}) + \epsilon \tag{5.1}$$

where y_t is child income, y_{t-1} is parents' income, and β is the parameter of interest, the famous «intergenerational income elasticity» (IGE). The higher β, the lower mobility between generations.

The important implication of this AR(1) model is that when looking after two generations, the IGE elasticity will be β^2 and, consequently, $(\beta)^m$ after m generations. Since β is usually estimated to be between 0.2 and 0.5, privileges will disappear quickly. Even for countries with high intergenerational elasticities such as the US for which Hertz et al. (2008b) estimate an elasticity of 0.47, this elasticity would reduce to $0.47^2 = 0.22$ after two generations and to $0.47^3 = 0.10$ after three generations. As noted by Stuhler (2012), this «extrapolation by exponentiation» is crucial for interpreting the intergenerational evidence. Moreover, this «regression-to-the-mean» process is often seen as consolidating for countries experiencing low mobility between two generations as even relatively large elasticities fade out relatively quickly.

As described previously, several studies that actually measure persistence in economic status find that the iterated regression procedure underestimates the persistence in a society. That is, status decays at a slower rate than inferred from the two-generation estimates.

More formally, let's define excess persistence as the difference between the actual persistence and the empirical persistence after m generations. Excess persistence exists if

$$(\beta_1)^m < (\beta_m) \tag{5.2}$$

where β_m with $m > 1$ is the actual persistence and $(\beta_1)^m$ the iterated persistence from the two-generation estimate β_1.

5.2.2 Latent Factor Model

Clark (2014) and Clark & Cummins (2015) claim that the usual β estimate as shown in Equation 5.1 is severely biased downward. The true persistence in social status might be much higher, around 0.75. Furthermore, the value is uniform across all countries and time. Their estimates are based on rare sure-name correlations in wealth and spans over multiple centuries.

Formally, the concept behind the underestimation of social status in Equation 5.1 is based on an errors-in-variable problem:

$$y_t = \rho e_t + u_t \tag{5.3}$$

$$e_t = \lambda e_{t-1} + v_t \tag{5.4}$$

where y_t is the outcome, such as education or income. e_t is the unobservable endowment that is inherited from parent to child according to the heritability coefficient λ. The endowment that children receive from their parents is then translated into the outcome according to the transferability coefficient ρ. Since this transmission is not perfect, reduced form estimates will underestimate the true persistence.[3]

This model compares to the two-generation reduced form model as described in Equation (5.1) like an errors-in-variable model. Thus, if $\rho < 1$, then the reduced form estimate β_1 will underestimate the true persistence and, consequently, $\beta_m \geq (\beta_1)^m$.

In this study, I can directly identify the inheritability coefficient λ. As shown by Braun & Stuhler (2018), the heritability coefficient λ can be identified with multigenerational data. To see this, I can rewrite the slope parameter as follows:

$$\beta_{t-1} = \frac{Cov(y_t, y_{t-1})}{Var(y_{t-1})} = \rho^2 \lambda \tag{5.5}$$

Similarly, the empirical three-generation coefficient (child-grandparent) is:

$$\beta_{t-2} = \frac{Cov(y_t, y_{t-1})}{Var(y_{t-2})} = \rho^2 \lambda^2 \tag{5.6}$$

Consequently,

$$\frac{\beta_{t-2}}{\beta_{t-1}} = \frac{\frac{Cov(y_t, y_{t-2})}{Var(y_{t-2})}}{\frac{Cov(y_t, y_{t-1})}{Var(y_{t-1})}} = \lambda \tag{5.7}$$

identifies the heritability coefficient λ.

[3]The variances of y_t and e_t are normalized to one to allow that the slopes can be interpreted as correlations.

Similarly, the transferability coefficient can be identified as follows:

$$\sqrt{\frac{\beta_{t-1}^2}{\beta_{t-2}}} = \rho \tag{5.8}$$

5.2.3 Direct Grandparent Effects

Scholars, such as Mare (2011) interpret the empirical excess persistence differently. He argues grandparents might affect grandchildren's outcomes. For example, richer or more educated grandparents might increase grandchildren's learning or give them financial support, which would then increase persistence.

Formally, the transmission process would then resemble an AR(2) process, with both the grandparent and the parent coefficient included in the model:

$$y_t = \tau_{t-1} y_{t-1} + \tau_{t-2} y_{t-2} + \epsilon \tag{5.9}$$

where τ_{t-1} captures the impact of parents and τ_{t-2} the impact of grandparents, conditional on parent outcomes.

However, as pointed out by Braun & Stuhler (2018), a positive grandparent coefficient does not prove a direct effect of grandparents. Any process that creates excess persistence beyond the parent-child correlation creates a positive grandparent coefficient τ_{t-2} in Equation 5.9.

Therefore, studies are testing whether proximity of grandparents is associated with a higher grandparent coefficient. Certainly, those tests rely on strong assumptions, for example, that the distance itself is not related to intergenerational persistence.

5.3 Data

5.3.1 Sample

This study combines several data sources based on administrative income, survey, and census data. Information on *income* is based on individual labor income data from the «social security earnings record» (SSER). The SSER is used is to calculate public old age insurance. It contains longitudinal income data on all individuals who were every employed or self-employed in Switzerland. The available record period starts in 1982 and ends in 2017. The SSER is kindly provided by the Public Compensation Office.

The income data is matched to the register-based population census (STATPOP). This data includes all people living in Switzerland from 2010 to 2018 and provides information on demographics, such as the municipality of residence or date of birth. Most importantly, it also contains IDs for other family members, such as children, parents, grandparents, or household members, which are crucial to measure the intergenerational dependency of income. In addition, I merge data from the Structural Survey (SE) which includes information on education and occupation. This survey does not cover the entire population, but a random sample, which includes roughly one third of people living in Switzerland. Thus, when analyzing information on education, the sample size is smaller than when analyzing income.

As shown in Chuard & Grassi (2020), the linked parent-child data provides good and representative coverage. In the most recent cohort 1987, I can link roughly 37 percent to any grandfather, while in cohort 1967, only 0.3 percent of children can be linked to any grandfather. To increase sample size, I use child cohorts starting in 1950, even though only very few children can be matched to their grandparents.

Intergenerational linkages can be missing if family members died or emigrated before the year 2010, if fathers are unknown, or due to lack of updating the civil register. The younger the cohorts, the larger the share of children that can be matched to parents. Obviously, linking children to grandparents is not complete. I can only link children to grandparents if they are still alive in 2010 and if the grandparents' children (the child's parent) still are alive and live in Switzerland.

The sample for grandparents is nevertheless representative. Table D.1 shows means of variables conditional on the sample. The first column shows the full sample. It includes all children born in Switzerland between 1950 and 1987 (that are alive in 2010). Column (2) and (3) show the sample of children for which I can identify the mother and the father. Column (4) shows the sample for which I can identify any grandfather. Column (5) shows the same for the sample for which I can identify the maternal grandfather of the child, column (6) the same for the paternal grandfather. When analyzing the grandparental generation, I focus on grandfathers instead of grandmothers, simply because labor market participation of women in these generations is very small and the income therefore not informative. In all columns, the means stay remarkably similar to the full sample means in column (1)— except for the year of birth. As explained above, the younger the child cohort, the more likely to match them to their ancestors. Although the sample is smaller for the grandfather-generation, it is still very large compared to other studies in multigenerational research.

The data does not cover information on parent-child relationships when children were not born in Switzerland. Thus, the analysis excludes immigrants, but includes Swiss-born children of immigrants.[4]

5.3.2 Measuring Life Time Income

The longitudinal labor income history is available for 35 years. While this is a long time, it is still not enough to capture three generations at the same age. I therefore residualize the log income similar to Lindahl et al. (2015) to construct life-time income. Also, I want to average over as many years as possible to get a stable approximation of lifetime income. I therefore run the following regression on the entire set of individual-year (it) observations ($n = 251,868,060$):

$$log(income)_{it} = \alpha + \phi_1 age_{it} + \phi_2 age_{it}^2 + \phi_3 age_{it}^3 + \gamma_1 c_i + \gamma_2 c_i^2 + \gamma_3 c_i^3 + year_t + \epsilon_{it} \quad (5.10)$$

I include a polynomial for age (age) and cohort (c) effects up to order three to account for age and cohort effects. I also add year fixed effects to account for business cycle fluctuations. Then, I receive the stable part of the income by averaging the residuals ϵ_{it} over all available years, at which individuals are at least 28 and not older than 65 years. This should then be a good approximation for life-time income.

When analyzing mobility, I distinguish between the elasticity and the correlation. The elasticity refers to the β coefficient in Equation 5.1. In the income domain, the elasticity is called «intergenerational income elasticity» (IGE). The elasticity increases if the standard deviation in the child generation increases. Thus, if inequality increases over time, persistence will increase, even though equality of opportunity in relative terms did not necessarily change. Therefore, I also show the correlation which abstracts from changes in inequality over time. The elasticity is directly related to the correlation. If the variance in child and parent income is standardized to one, the elasticity equals the correlation coefficient. Since both measures can be of interest, I will provide results for the correlation and the elasticity.

5.3.3 Measuring Years of Education

Educational attainment is measured as years of schooling. Data provides information on highest completed education, which I translate into years of schooling according to the following scheme in Table 5.1.

[4]In studies on intergenerational mobility, such as Chetty et al. (2014a) or Heidrich (2017) this is usually the case. Not only because of data restrictions but also because including children born in another country would dilute the measurement of intergenerational mobility of a country when not the full childhood can be attributed to a specific country.

TABLE 5.1: Highest Education and Years of Schooling

Highest Education	Years of Schooling
No education	0
Max 7 years mandatory school	7
Mandatory school only	9
Vocational Training and Education	12
High school (gymnasium)	13
Higher professional degree	14
Bachelor degree	16
Master degree	18
PhD, habilitation	21

TABLE 5.2: Correlation Coefficients

	Paternal Line		Maternal Line		Maximum	
	(1)	(2)	(3)	(4)	(5)	(6)
	Inc	Edu	Inc	Edu	Inc	Edu
G1-G2	0.13	0.33	0.13	0.30	0.13	0.31
	(0.010)	(0.037)	(0.0047)	(0.016)	(0.0094)	(0.0039)
G2-G3	0.28	0.32	0.21	0.27	0.24	0.29
	(0.011)	(0.023)	(0.021)	(0.034)	(0.018)	(0.0063)
G1-G3: Actual	0.036	0.16	0.042	0.15	0.040	0.15
	(0.0015)	(0.016)	(0.0011)	(0.0076)	(0.00097)	(0.0026)
G1-G3: Predicted	0.037	0.11	0.028	0.081	0.031	0.090
	(0.00044)	(0.020)	(0.0051)	(0.015)	(0.0085)	(0.00064)
Δ Actual-Predicted	-0.0018	0.050	0.014	0.067	0.0070	0.059
t-test	(-1.13)	(1.95)	(2.64)	(4.01)	(0.82)	(22.0)
Obs.	94,229	6,934	143,650	11,252	212,227	17,599
mean(G1)	0.076	13.9	0.066	13.9	0.066	13.9
sd(G1)	0.85	2.58	0.84	2.59	0.85	2.59
mean(G3)	0.56	11.3	0.54	11.3	0.54	11.3
sd(G3)	0.69	3.35	0.67	3.29	0.67	3.30

Notes: This table shows the correlation coefficients between adjacent and skipping generations. Row *G1-G2* indicates the correlation between child and parents for income (log-residualized) and education (years of schooling) and the corresponding standard error in parentheses. It does so for different family lineages: Columns (1) and (2) use the paternal grandfather's income and education. Columns (3) and (4) look at the maternal grandfather and columns (5) and (6) use the maximum value of any grandparent (including grandmothers).
Row *G2-G3* looks at the correlation between parents and grandfather. Row *G1-G3:Actual* shows the correlation between child and grandparents. The next row, *G1-G3:Predicted* shows the predicted value (the multiplication of row *G1-G2* and *G2-G3*.
Row Δ *Actual-Predicted* shows the absolute difference between the predicted and the actual G1-G3 correlation and a corresponding t-test.

TABLE 5.3: Slope Coefficients (Elasticity)

	Paternal Line		Maternal Line		Maximum	
	(1)	(2)	(3)	(4)	(5)	(6)
	Inc	Edu	Inc	Edu	Inc	Edu
G1-G2	0.22	0.31	0.21	0.27	0.21	0.28
	(0.0060)	(0.020)	(0.0050)	(0.018)	(0.0041)	(0.013)
G2-G3	0.21	0.27	0.17	0.23	0.20	0.25
	(0.0042)	(0.015)	(0.0035)	(0.013)	(0.0029)	(0.010)
G1-G3: Actual	0.044	0.08	0.052	0.12	0.050	0.12
	(0.0040)	(0.0091)	(0.0033)	(0.0074)	(0.0027)	(0.0059)
G1-G3: Predicted	0.046	0.081	0.035	0.061	0.041	0.070
	(0.0016)	(0.0069)	(0.0011)	(0.0052)	(0.0010)	(0.0043)
Δ Actual-Predicted	-0.0023	0.039	0.017	0.055	0.0089	0.048
t	(-0.54)	(3.45)	(5.01)	6.07	(3.05)	(6.57)
Obs.	94, 229	6, 934	143, 650	11, 252	212, 227	17, 599
mean(G1)	0.076	13.9	0.066	13.9	0.066	13.9
sd(G1)	0.85	2.58	0.84	2.59	0.85	2.59
mean(G3)	0.56	11.3	0.54	11.3	0.54	11.3
sd(G3)	0.69	3.35	0.67	3.29	0.67	3.30

Notes: This table shows the slope coefficients between adjacent and skipping generations. In contrast to the correlation coefficient, the slope coefficient also incorporates changes in the standard deviation between generations. Higher inequality over time therefore results in a higher slope estimate. Row *G1-G2* indicates the slope coefficient between child and parents for income (log-residualized) and education (years of schooling) and the corresponding standard error in parentheses. It does so for different family lineages: Columns (1) and (2) uses the paternal grandfather's income and education, Columns (3) and (4) look at the maternal grandfather and columns (5) and (6) use the maximum value of any grandparent (including grandmothers).
Row *G2-G3* looks at the slope coefficient between parents and grandfather. Row *G1-G3:Actual* shows the slope coefficient between child and grandparents. The next row, *G1-G3:Predicted* shows the predicted value (the multiplication of row *G1-G2* and *G2-G3*.
Row Δ *Actual-Predicted* shows the absolute difference between the predicted and the actual G1-G3 slope coefficient and a corresponding t-test.

5.4 Results

5.4.1 Empirical Excess Persistence in Education and Income

First, I test whether the empirical persistence in educational attainment and income is higher over three generations than the iterated estimates of the two generation correlation suggests. Table 5.2 shows the results for the correlation coefficients. The first two rows, G1-G2 and G2-G3, show the intergenerational correlation for income (inc) and education (edu). I use different measures for grandparental outcomes. The *maximum* values use the highest income or years of schooling for any grandparent. The *paternal line* values use the outcomes of the father's father and *maternal line* use the outcomes of the mother's father.

Thereby, several things are interesting. First, the persistence in education is considerably larger than the persistence in income. Second, while the persistence in education remains similar over time (G1-G2 vs G2-G3), the persistence in income decreases and is remarkably low in the youngest two generations (G1-G2).

Let us now move to the key point of interest, the correlation between three generations. Here, row «G1-G3: Actual» shows the correlation between grandparents and child. For all family lineages and domains (education and income), the correlation is significantly larger than zero. Now, we can compare the actual correlation to the predicted correlation. The predicted correlation is shown in row «G1-G3:Predicted» and is calculated by multiplying G1-G2 and G2-G3.[5] The difference between the actual and the predicted correlation is shown in the next row «Δ Actual-Predicted». When looking at the paternal line, which is standard in most studies, it is striking that there is literally no excess persistence in life-time income. When looking at educational attainment, there is substantial excess persistence. This leads to the conclusion that the iterated regression procedure can adequately capture the persistence between multiple generations when looking at income, but not when looking at education. More precisely, the Becker-Tomes model can not be rejected when looking at income, but it can be rejected when looking at education.

When looking at the maternal line, however, there seems to be excess persistence in income as well. The reason for this is likely that the ancestors in the two generation coefficient are the sum between mother and father. Since the correlation between mothers and children is low, missing paternal income can result in low income between parent and children. Therefore, the correlation in the two-generation case is underestimated,

[5] An other possibility would be to multiply the correlation of G1-G2 with itself. However, since the two-generation correlation can change over time, it would not be clear whether the difference in actual and predicted correlation is because the iterated model is wrong or because the correlation in G2-G3 differs from G1-G2.

which will lead to excess persistence when looking at the three-generation case, since there, I only use male grandparents. Therefore, the paternal line might be the most appropriate to test the Becker-Tomes model.

Table 5.3 shows the result of the same exercise for the slope parameter instead of the correlation. Thus, the coefficients here also measure changes in inequality over time. As inequality increased over time, the persistence in terms of the slope parameter is higher than the correlation shown in Table 5.2.

5.4.2 Testing the Latent Factor Model

Table 5.4 shows parameter estimates of the latent factor model. Columns 3 presents the average correlation between two adjacent generations: β_1. Specifically, the correlation between parents and between parents and grandfathers. λ shows the heritability coefficient. Following Braun & Stuhler (2018), I also calculate λ_{alt}, which does not average over two generations, but only takes the β_1 correlation of the child and parent outcome. Column 4 shows the correlation between the child and the grandfather β_2. As shown in previous studies, the latent factor λ is determined by dividing β_2 by β_1. The result is shown in column 5. Thereby, the upper half of the table is looking at correlation in income, whereas the bottom half is looking at correlation in education. Also, the estimates are shown for different grandparent samples. The father of the father (Pat.Gf.), the father of the mother (Mat.Gf), and the maximum of both grandfathers (Max.Gf).

The underlying latent factor λ is always significantly larger than zero. The most striking point is the divergence of the λ between the two domains: income and education. While λ is estimated to be around 0.15-0.20 for the income domain, it is more than twice as large for the education domain.

How do those results for the latent factor λ compare to other studies? The first important thing to notice is that even when looking at the large estimate in the education domain, the estimates are still much lower than the one proposed by the Clark Hypothesis of 0.75.

Compared to the study in Germany by Braun & Stuhler (2018), the education estimates for λ are fairly similar. In their study, the estimates range from 0.49 to 0.61. Compared to Sweden, the schooling estimates are slightly smaller in Switzerland. Lindahl et al. (2015) estimate schooling correlations of around 0.61. In terms of earnings, the pattern is similar in Malmö (Sweden) in the sense that the λ for earnings is smaller than the one for education. However, the estimate is still fairly larger in Sweden with 0.49.

Clark also claims that the underlying factor would not vary over places. To test this hypothesis, I calculate the latent factor (income) by different cantons of births of the

TABLE 5.4: Latent Factor Model

Sample	Outcome	Avg. 2-Gen Correlation	G1-G3 Correlation	λ	λ_{alt}	ρ
Pat.Gf	Inc	0.206	0.036	0.173	0.128	1.092
	(se)	(0.004)	(0.003)	(0.015)	(0.011)	(0.047)
Mat.Gf	Inc	0.171	0.041	0.241	0.198	0.844
	(se)	(0.002)	(0.002)	(0.012)	(0.009)	(0.019)
Max.Gf	Inc	0.189	0.039	0.208	0.162	0.953
	(se)	(0.001)	(0.002)	(0.012)	(0.010)	(0.030)
Pat.Gf	Edu	0.327	0.157	0.480	0.486	0.826
	(se)	(0.010)	(0.013)	(0.043)	(0.047)	(0.043)
Mat.Gf	Edu	0.281	0.152	0.542	0.573	0.719
	(se)	(0.013)	(0.005)	(0.030)	(0.049)	(0.034)
Max.Gf	Edu	0.301	0.153	0.510	0.526	0.768
	(se)	(0.008)	(0.003)	(0.015)	(0.022)	(0.019)

Notes: This table shows the estimated parameters for the latent factor model. Column *sample* indicates the used sample to measure the three generation correlations (paternal grandfather, maternal grandfather and the maximum value of either the maternal or paternal grandfather). Column (3) shows the average of the correlation between generation G1-G2 and G2-G3. Column (4) shows the empirical correlation between child and grandparent (G1-G3). The latent factor λ is then calculated by dividing column (4) by column (3) and shown in column (5). $\lambda(alt)$ is the latent factor calculated not by averaging over the two two-generation coefficients G1-G2 and G2-G3, but by using only the coefficient of G1-G2. ρ shows the transferability coefficient. If ρ is smaller than 1, the two-generation coefficient will underestimate the persistence over generations. Standard errors are shown in parentheses and are calculated by a bootstrap procedure using 1,000 replications.

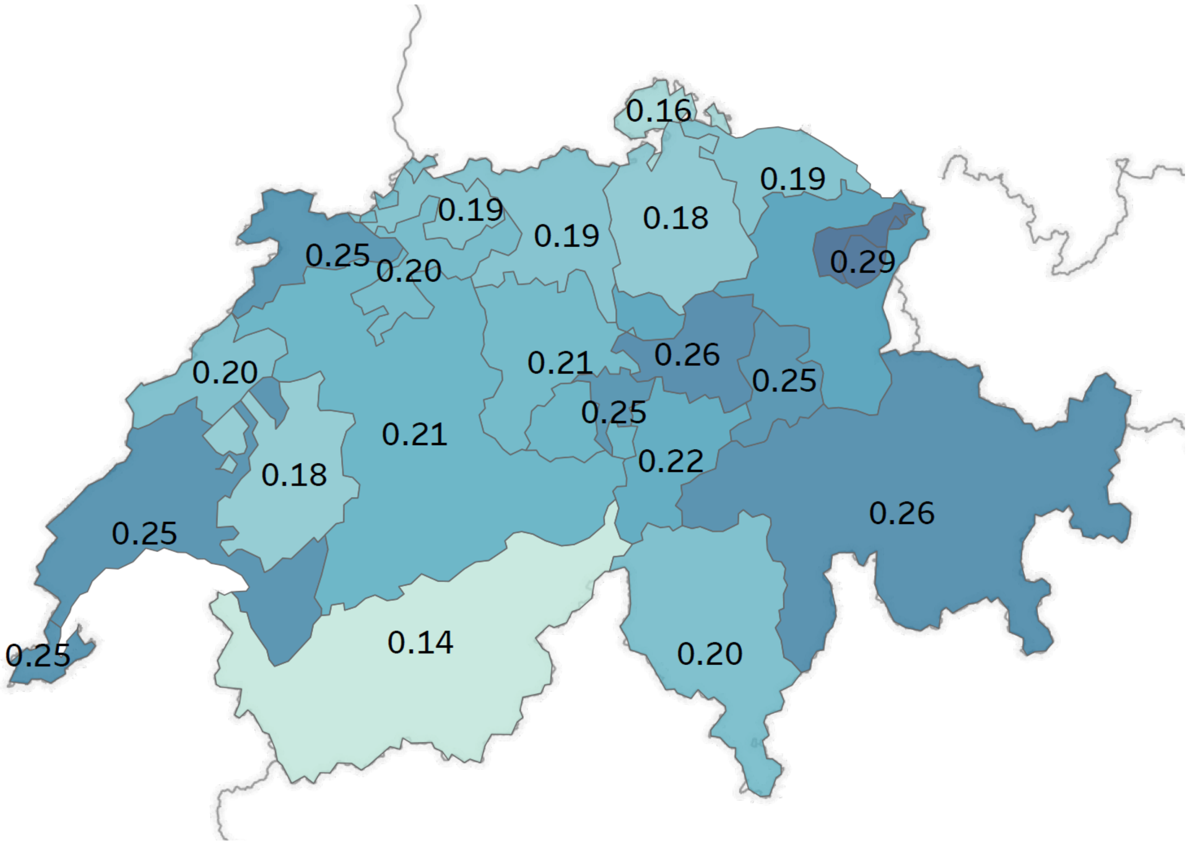
paternal grandfather. The results are illustrated in Figure 5.1 and Table D.3 in the appendix. Clearly, the underlying factor differs by cantons—-even though the differences are not large. In general, more rural cantons have a higher inheritability factor. I do not show the coefficients for education, since the sample size is too small for such an analysis. Thus, I can reject the hypothesis that the latent factor is uniform across space.

Another hypothesis of Clark is that the underlying factor λ stays constant over time. To test this hypothesis, Figure 5.2 shows the latent factor λ for different cohorts. The dashed lines show a 95%-confidence interval. For the early cohorts, the point estimates are slightly smaller than for younger cohorts. However, the differences are not significant. Since the early seventies, the coefficients are remarkably stable. Therefore, I cannot reject Clark's hypothesis of a time constant underlying factor.

5.4.3 Evidence on direct grandparental effects

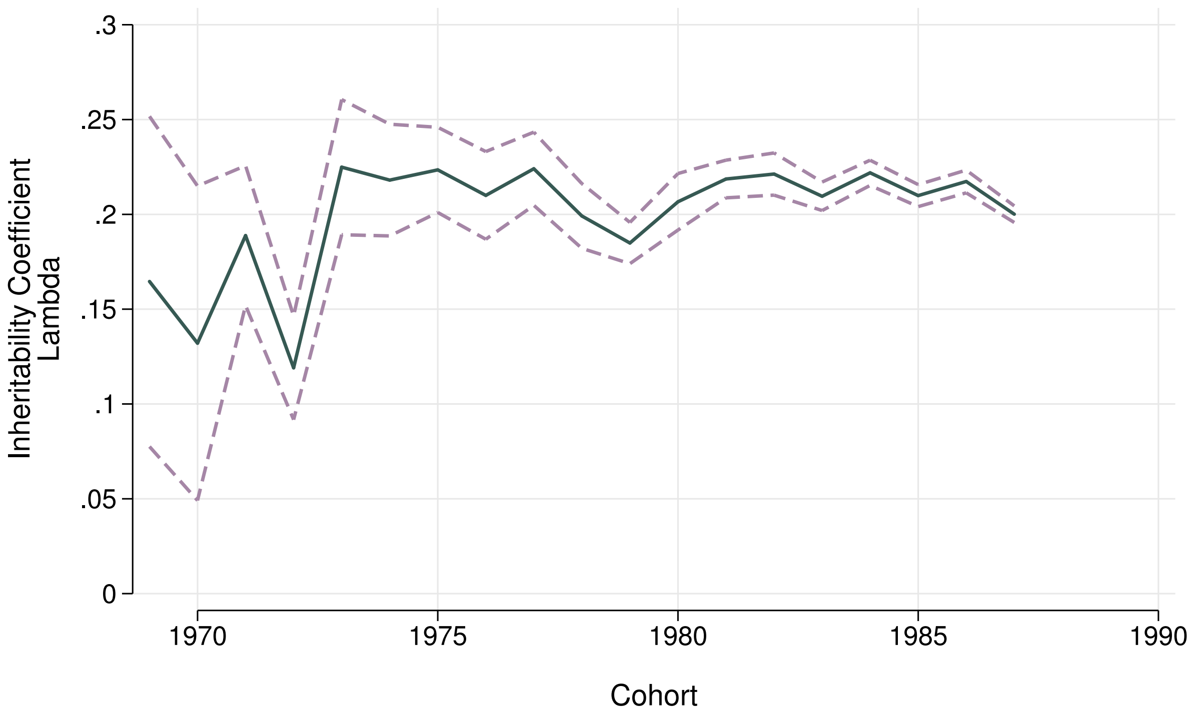
Table 5.5 presents the results of Equation 5.9 for the income and education domain. The grandparent coefficient τ_{t-2} from the AR(2) is showed as «Education Gf.» and «Income Gf.». Column (1) and Column (4) show the equation for all child-father-grandfather

FIGURE 5.1: Latent Factor (λ) by Cantons



This figure shows the regional distribution of the latent factor (λ) based on the place of birth of the paternal grandfather.

FIGURE 5.2: Latent Factor (λ) over Time



This figure shows how the latent factor (λ) for income changes over time. The dashed lines represent a 95% confidence interval.

observations. Column (2) and Column (5) show the results only for grandfathers that live in the same language region as their grandchild. Column (3) and Column (6) show only the coefficients for grandfathers that live in different language regions than their grandchild.

Looking at the results for the full sample (1) and (4) shows that there is a positive grandparent coefficient for the education domain, while for the income domain there is no significantly positive grandparent coefficient—even though the sample is larger when looking at income. This can be explained by the previous finding that there is more excess persistence in education than in income.

If we now compare the grandparent coefficient for those grandparents in the same and in different language regions, we see that the point estimate is larger for those grandparents that live in the same language region—for the education and for the income domain. This would in principle be in line with the direct grand-parental effect hypothesis. However, the difference is far from significant—therefore it is hard to draw final conclusions.

Another test for the direct grandparents effect theory is shown in Figure 5.3. Here, I run Equation 5.9 on different quintiles of distance from grandfather to grandchild. If there

TABLE 5.5: Direct Grandparental Effects by Language Region

	Education			Income		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Same Lang.	Diff. Lang.	All	Same Lang.	Diff. Lang.
Father	0.289***	0.269***	0.444***	0.217***	0.219***	0.200***
	(0.0196)	(0.0206)	(0.0553)	(0.0060)	(0.0064)	(0.0181)
Grandfather	0.032*	0.038*	-0.010	-0.001	0.001	-0.020
	(0.0159)	(0.0168)	(0.0470)	(0.0042)	(0.0045)	(0.0123)
Observations	2,432	2,185	247	91,303	82,167	9,136
R-Squared	.11	.099	.2	.018	.019	.016

This table shows the coefficients of a regression of child outcomes (education and income) on father and grandfather outcomes. The regression thus resembles an AR(2) process. Columns (2) and (5) analyze the subgroup of grandfather-grandchildren pairs that live in the same language region. Columns (3) and (6) analyze the subgroup of grandfather-grandchildren pairs that live in different language regions.
*p<.10; **p<.05; ***p<.01

are indeed some direct effects, by spending time or by educating grandchildren directly, one would expect the effect to be stronger for children living closer to the grandparent.[6] Panel (a) looks at years of education and Panel (b) at lifetime income. When looking at education in Panel (a), the grandparent coefficient is indeed larger for grandparents living close to their grandchild—except for the highest quintile. Also, for income in Panel (b), the coefficient is indeed largest for short distances, although it is harder to pin down a conclusive pattern.

Taken together, the tests on direct grandparent effects by distance and language region cannot reject the hypothesis of no direct effect of grandparent on grandchildren.

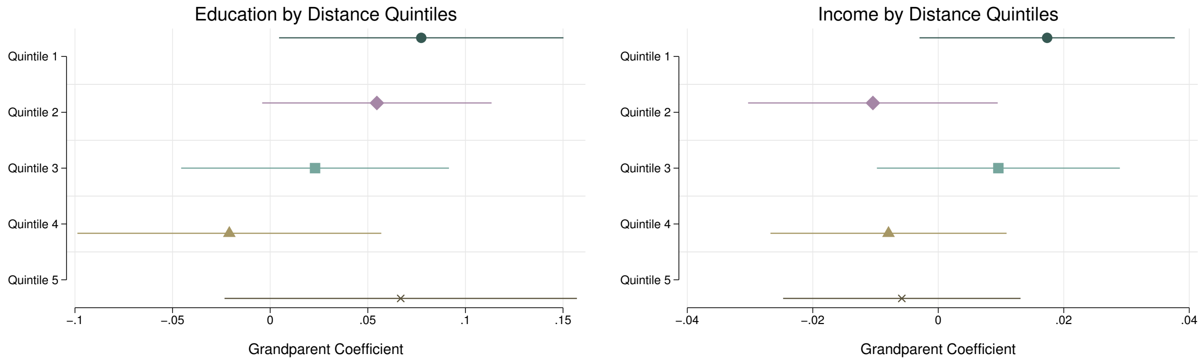
5.5 Discussion

This study measures the persistence of inequality in income and education over three generations in Switzerland. I find that the two-generation paradigm is well suited to predict persistence in income over three generations, but overestimates educational mobility. The key take-away thus is that long-term persistence in economic status might depend on the domain that is under study.

[6]Distance might itself be related to other characteristics that influence intergenerational mobility. Thus, this is certainly not a powerful test to reject the direct grandparental effect hypothesis.

FIGURE 5.3: Direct Grandparental Effect by Distance

(A) Years of Education (B) Lifetime Income



This figure shows the grandparent coefficients of an AR(2) regression while controlling for parental outcome. Panel (a) shows the grandparent coefficient for years of education. Panel(b) shows the grandparent coefficient for lifetime income. The coefficients are shown for different quintiles of distances between grandchild's place of birth and grandfather's residence in 2010 (average of maternal and paternal grandfather distance). Quintile 1 refers to the quintile with the lowest distance between parent and child. The average travel-time in the Quintile 1 is 11 minutes and 106 minutes in Quintile 5.

The results on the excess persistence are similar to previous studies: The two-generation paradigm overestimates educational mobility (Braun & Stuhler, 2018; Lindahl et al., 2015; Zeng & Xie, 2014). The fact that inequality in income decays faster than inequality in education can be rationalized because education has to be «transferred» again into income, which adds a layer of uncertainty and thus reduces the correlation. Intuitively, it is easier for parents to influence education of their children than their income — at least in a society in which nepotism plays a minor role.

The domain specific difference in long-term inequality over generations might be important for countries that lack administrative income data and infer social mobility from educational data. Low educational mobility, and even excess-persistence in educational inequality, does not need to translate into low income mobility.

Future research could analyze the origins of this discrepancy between income and education, and whether this discrepancy can also be observed in other countries. One hypothesis is that this difference is especially large in countries with a strong vocational education and training system (VET). Since in such a system, fewer years of (formal) schooling might not necessarily lead to a strong reduction in income because VET provides usually good labor market prospects with fewer years of schooling.

Chapter 6

The Effect of Immigration on Intergenerational Income Mobility of Natives

Abstract This study analyzes the effect of immigration on intergenerational income mobility of natives. I exploit a large quasi-natural policy experiment that led to variation across time and space in exposure to cross-border immigrants: The incremental removal of restrictions for cross-border immigrants in Switzerland in regions close to the border. The study draws from several administrative data covering the universe of labor income, family linkages, and census data. Results show that children from low-income parents experience a stronger decrease in labor income than children from high-income backgrounds. While children from the bottom quintile experience a 3 to 6 percent decrease in labor income, the effect for children from the top quintile is close to zero or slightly positive. Subsequently, intergenerational income mobility, measured as the rank difference between children from the lowest and the highest parent percentile, decreases by 1.5 percentile ranks or 12 percent in relative terms. The income of children from low-income backgrounds decreases more because they are more likely to learn occupations and to choose educational tracks which are more negatively affected by the immigrant influx. Also, there is no adaption in educational or occupational choices that could mitigate the negative labor income effect. The decrease in intergenerational mobility correlates with sentiments against immigration: Cantons experiencing a more negative effect are more likely to vote in favor of anti-immigration policies.

Keywords: immigration, intergenerational income mobility, inequality;
JEL classification: E24, F22, J26

I thank the Federal Office of Statistics and the Central Compensation Office for generously providing the data. Any errors are my own.

6.1 Introduction

Immigration and inequality are two burning issues. The past three decades have seen a substantial rise in income inequality within many Western countries. Inequality is perceived as particularly unfair if it results from unequal opportunities—if factors outside of an individual's control determine the well-being. If equality of opportunity is low, senses of disempowerment and feelings of social exclusion can grow and increase sentiments against established parties. At the same time, rising immigration provides a breeding ground for populist parties. Rhetoric of such anti-immigration parties fuels the idea that immigration is at least partially responsible for rising inequality and lower opportunities.

This paper studies if immigration affects intergenerational income mobility of natives. Can immigration affect this «unfair» inequality and decrease opportunities for children from low-income parents? Specifically, it asks whether children from poor parental backgrounds are affected differently by the influx of immigrants. And if this—in turn—is related to sentiments against immigration.

To estimate the effect of immigration on intergenerational mobility, I take advantage of a large quasi-natural experiment in Switzerland which gave rise to variation across space and time in exposure to cross-border immigrants. Specifically, I exploit the incremental removal of restrictions for cross-border workers in administratively defined border-regions, which led to a strong increase in the inflow of foreign workers (Beerli et al., 2020; Parenti & Tealdi, 2019). The study draws from several administrative data, including the universe of labor income over three decades, as well as administrative family linkages, census, and survey data.

While a bunch of studies have analyzed how immigration effects labor market outcomes, there has been little discussion about how immigration can affect intergenerational income mobility of natives. Standard economic theory predicts that immigration reduces the labor opportunities of workers competing with immigrants and increases the opportunities for complementary workers. Thus, depending on whether children from relatively poorer parents are more or less likely to compete with immigrants, intergenerational mobility might be affected positively or negatively. However, there have been no empirical studies linking immigration to intergenerational income mobility.

This paper contributes to the literature in several ways. The principal contribution is to analyze the effect of immigration on intergenerational income mobility of natives. While several studies analyze intergenerational mobility of immigrants, to the best of my knowledge, no previous paper focuses on natives. Another important contribution is to link the effects of intergenerational mobility on political outcomes. Swiss citizens have regular ballots in which they vote on a myriad of topics, including topics on immigration. Those ballots can be compared to measures of intergenerational mobility. Considering growing resentments against immigrants and populism gaining momentum and the fact that individuals consider decreases in income mobility as unfair, I argue that this might be an overlooked aspect in understanding resentments against established political parties.

The results show that labor income of natives from low-income parental backgrounds is reduced more severely than labor income of children from high-income parental backgrounds. Children with parents from the bottom quintile of the income distribution experience a loss of total labor income of around 3 to 6%, while there is little or only a small positive effect for children with parents from the top quintile. The increased wedge between incomes of children from poorer and richer backgrounds lowers intergenerational mobility. In terms of rank-rank slope—a broadly accepted measure for the dependence of child and parental income—the increase in cross-border immigrants increases the rank-rank slope by 1.6 percentile ranks. In relative terms, this corresponds to a decrease in intergenerational mobility of 12%.

Does the decrease in intergenerational income mobility due to cross-border immigration lead to resentments against immigration? Although it is hard to pin down a causal pathway, several empirical findings are in line with such an interpretation. First, cantons whose children's labor income from the bottom quintile are more negatively affected by cross-border immigration show a higher tendency to vote in favor of anti-immigration policies. Second, municipalities with a high share of children from poorer backgrounds also vote more often in favor of anti-immigration policies. Third, parties favoring more restrictive policies increase their voting share by around two percentage points in the post period in border regions.

Why do children from poorer parents experience a stronger drop in labor income? Occupations whose labor incomes are most negatively affected by the inflow of immigrant workers are also those with the highest share of children from low-income parents. Further, I observe no reactions that might mitigate this effect. For example, children do not alter their educational track or their learned occupation. One might expect that children exposed to immigrants opt for higher education or different occupations, considering increased competition by immigrants. One might also assume that children from higher income parents might be more informed on the consequences of immigration and therefore more likely increase their educational level. The empirical results do not support those hypotheses: Educational pathways do not change at all, neither for children from poorer backgrounds nor for such of richer backgrounds. Also, they do not move to the non-border region. That low-income children are more often in occupations more negatively affected by the policy, therefore leads to lower intergenerational income mobility—and the stickiness of the educational and occupational pathway to parental background does not mitigate this effect.

The results are robust to several specifications. Restricting the treatment group to municipalities closer to the border increases the coefficients. This is in line with the fact that higher exposure of cross-border immigrants increases the effect. Placebo tests show insignificant differences between border and non-border regions. Also, the large pre-period time frame of the data allows to assure that trends in labor income of treatment and control group were parallel before the policy introduction. The results do not change when including higher order region specific time trends. Further, the results are robust to specifications concerning the specified variables, such as the age at which I measure child income or ages at which the parental income rank is measured. Also, geographic mobility analysis—comparing place of birth with place of residence—shows that mobility patterns do not change around the onset of the policy.

These results have important policy implications. They show that the economic regulatory environment can influence income mobility. Even if welfare overall increases because of an opening of the border, children from more disadvantaged backgrounds might lose. Thus, when explaining resentments against immigration, not only might diffuse feelings against the foreign play a role, but also rational economic thinking. Apart from compensatory measures, education might once again be the key to solving this issue. Children from poor parents do not experience a decrease in income when they completed an academic education.

This paper adds to two strands of literature. First, it adds to the literature on the effects of immigration on labor market outcomes of natives and, more closely, its distributional effect. Second, it adds to the growing literature on determinants of intergenerational mobility.

Studies looking at the distributional effects of immigration yield mixed results. For example, Dustmann et al. (2013) find that immigration depresses wages below the 20th percentile of the wage distribution but leads to slight wage increases in the upper part of the wage distribution. Similarly, Borjas et al. (1997) find that immigration negatively affects the wage of individuals at the bottom quintile. Card (2009) concludes that immigration had little effect on native wage inequality in the US. However, because immigrants cluster at the top and the bottom of the income distribution, inequality is still higher than it would be without immigration.

Other studies find that immigration has a positive effect on wages of poor individuals. For example, Foged & Peri (2016) find that refugee immigrants pushed less educated native workers to pursue less manual-intensive occupations. Thus, immigration had a positive effect on the income of unskilled native workers. In a similar fashion, Ottaviano & Peri (2012) find that immigration increased wages of natives without a high-school degree slightly. In terms of policy change and identification strategy, the study most closely related to this paper is the one by Beerli et al. (2020). They thoroughly analyze the labor market effects of the abolition of immigration restrictions for cross-border workers in Switzerland. Their results show that the policy change increased foreign employment substantially. Even though many cross-border workers are highly educated, wages for highly educated natives increased as well, pointing to complementarity of immigrants and natives. The effect on less educated workers shows a negative coefficient, that is, however, not significantly different from zero.

The literature on intergenerational mobility focusing on reliable country estimates of equality of opportunity recently gained a revival due to newly gained access to large administrative data, such as Chetty et al. (2014a); Heidrich (2017); Corak (2020b) or Chuard & Grassi (2020). In contrast, studies on the determinants of intergenerational mobility focus mostly on the causal transmission between children and parents (see Black & Devereux (2010) for a summary).

At the intersection between labor market effects, immigration, and intergenerational mobility are studies looking at intergenerational mobility of immigrants (Borjas, 1993; Abramitzky et al., 2019). There are, however, no studies looking at intergenerational income mobility of natives in light of immigration. This study aim to fill this gap in the literature.

The structure of this paper is as follows. First, I explain the background of the policy. Then I lay out the empirical strategy and explain the data. The results section shows the effects on wages conditional on parental background and what this implies on a customary measure of intergenerational income mobility. The section on mechanisms analyzes the origins of the labor income effect by education, occupation, and parental background. Then I check how the effect is correlated to political outcomes. The last section concludes.

6.2 Institutional Background

Although at the heart of Europe, Switzerland is not part of the European Union (EU). To facilitate trade and movement of persons, the EU and Switzerland signed several bilateral agreements. One such agreement is the «Agreement on the free movement of persons» (AFMP). It enables citizens of Switzerland and of member states of the European Union (EU) and members of the European Free Trade Agreement (EFTA) to choose their place of employment and residence. Individuals need a valid employment contract or be self-employed.

The AFMP was signed on June 1999 and came into force on June 1st, 2002. Unsurprisingly, the contract was highly controversial. Opponents feared that the opening of the border to immigrants could harm Swiss workers and increase the financial burden on the social security system. To allay fears of rapid changes, the Swiss government implemented changes gradually until the full implementation in 2007.

Importantly, the implementation process was different for resident immigrants and cross-border workers. Cross-border workers (also called frontier workers) represent a special immigrant group. They work in Switzerland but have their main residence abroad, where they must return regularly. During the transition to the full implementation of the AFMP, they were subject to several regulatory changes.

Before 1999, Swiss companies were only allowed to employ cross-border workers if they could find an equally qualified resident worker for a particular job. This is called the «priority requirement». Further, cross-border workers were only allowed to work in the border regions of Switzerland, which is an administratively defined region around 30 minutes from the border. The number of resident immigrants was subject to national quota and had to satisfy the priority requirement as well.

Between 1999 and 2004, Switzerland gradually removed impediments on cross-border workers. For example, cross-border workers were allowed to return home only weekly instead of daily, the permit was valid up to 5 years, and cross-border workers did not have to live in the contiguous border region for six months. In this analysis, I follow Beerli et al. (2020) and refer to this phase as «transition period». On 1 June 2004, accompanying measures on the labor market came into force to better protect workers against the risk of wage and social undercutting. Thus, work inspectors can carry out checks on compliance with minimum or customary working and pay conditions in the workplace.

The next period starts in 2004. Labor markets in the border region became fully open to cross-border workers. Importantly, cross-border workers were still only allowed to work in border regions. This is convenient for our analysis, as it allows to compare developments in border and non-border regions.

On June 1, 2007, all areas in Switzerland underwent full liberalization of cross-border commuters and for resident immigrants from the EU/ETFTA and citizens. At the end of 2008, Switzerland joined the Schengen area. Thereby, border controls were removed, which further facilitated commuting due to abolished border checks. Thus, even after the full implementation period in 2007, incentives for cross-border workers to work in a region close to the border increase.

The government defines which municipalities belong to the border region. In general, they are within a 30 minutes car drive from the next border-crossing. Figure 6.1 depicts the border and non-border regions. The greenish municipalities belong to the border region, while the purple ones belong to the non-border region. Importantly, border regions are defined on a municipality and not on a cantonal level. Thus, cantons—which represent the main political body in Switzerland—can have border and non-border municipalities and thus an overlap of treatment and control groups.

Cross-border commuters make up a considerable share of the working population in Switzerland. They depict a share of 6.8% of the total working population in 2017, while it is naturally larger in cantons close to the border. For example, in the canton Ticino, which is adjoined to Italy and has the same language, cross-border workers constitute 27.3% of the working population (FOS, 2019). Formally, cross-border commuters must be a citizen of an EU/EFTA member country and return to their place of residence, that is outside of Switzerland, at least once per week. Further, if they prove they are employed or self-employed, they receive a specific working permit. If unemployed, cross-border commuters receive their unemployment benefit in their home country.

6.3 Data

6.3.1 Data Sources

This study uses several administrative data sources: individual level income, census, and survey data for the main analysis. In addition, for the analysis of political outcomes, it uses publicly available data provided by the Federal Office of Statistics based on municipality level.[1]

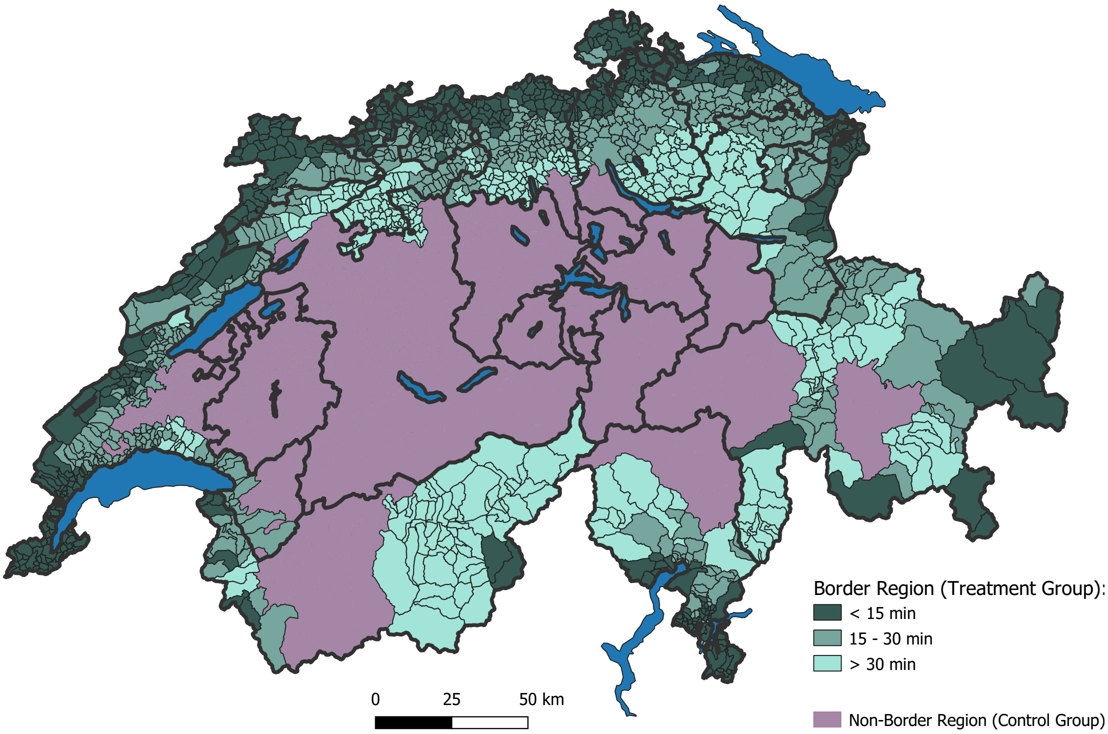
Information on income is based on individual labor income data from the «social security earnings record» (SSER). The SSER is kindly provided by the Public Compensation Office. The purpose of the SSER is to calculate public old age insurance. It covers the full population and provides full, longitudinal earnings information for employed and self-employed adults. The available record period starts in 1982 and ends in 2017.

I then match the SSER data with the register-based population census (STATPOP). The STATPOP is provided by the Swiss Federal Office of Statistics for the years 2010 to 2018. STATPOP includes all people living in Switzerland and provides information on demographics, such as the municipality of residence, date of birth. Importantly, it also contains IDs for other family members, such as children, parents, or household members, which are crucial to measure the dependency of parent and child income.

As shown in Chuard & Grassi (2020), this data provides good and representative coverage of parent-child relationships. Roughly 91% of all children alive in 2010 and born between 1967 and 1984 can be matched to either the mother, the father, or both parents.

[1] This data can be freely accessed on https://www.bfs.admin.ch/bfs/de/home/statistiken/politik/

FIGURE 6.1: Treatment and Control Region



Notes: This figure shows the control and treatment regions. The greenish municipalities belong to the treatment region (border region), while the purple municipalities belong to the control region (non-border region). The border region is administratively defined as only 30 minutes away from the closest border crossing. The black lines indicate borders of cantons (federal staates). The figures shows that cantons have borders and non-border municipalities. Some border regions, that are adjecent to the border, are non-border region because mountains forbid crossing the border by car within 30 minutes.

Linkages between parents and children can be missing if parents died or emigrated before the year 2010, if fathers are unknown, or due to lack of updating the civil register. The data does not contain—or only incompletely—information on parent-child relationships when children were not born in Switzerland. Thus, our analysis excludes immigrants, but includes Swiss-born children of immigrants. In studies on intergenerational mobility, such as Chetty et al. (2014a) or Heidrich (2017) this is usually the case. Not only because of data restrictions but also because including children born in another country would dilute the measurement of intergenerational mobility of a country when not the full childhood can be attributed to a specific country. In this study, this is even less of a problem because it intentionally wants to measure the labor effects on natives.

The last data set is the structural survey (SE). This data provides valuable information on education and learned occupations. It is available for the years 2010 to 2018. Further, it covers a little over 200,000 persons per year, which corresponds to roughly 2% of the population. As I have nine years available and only need one measurement per person, I have a sample size of roughly 1,8 Mio. unique observations, which then corresponds to around 20% of the population.

The sample size differs depending on the analysis. This is because data on education and occupation stems from the SE, which does not cover the full population. For our main analysis, we do not need information on education and occupation, thus we can draw from the (almost) full sample.

Table 6.1 describes the sample and yields information on variables conditional on being in the treatment or control region. The overall sample comprises roughly 1,5 Million children matched to their parents with year of birth between 1960 and 1984. Child income is measured between the ages 30 to 33, parent income when children are between 12 and 22. When analyzing information on education or occupation, the sample size reduces to roughly half a million. The treatment group (border region) is considerably larger than the control region in terms of population, which reflects the relatively small surface of Switzerland. Also, mean child income and parent quintile is slightly higher in the treatment group.

6.3.2 Variable Construction

Child Income: In the baseline specification, child income is measured as the mean income between the ages 30 and 33. Several studies on intergenerational mobility show that relationship between child and parent income stabilizes if child income is measured in the early thirties (Chetty et al., 2014a; Chuard & Grassi, 2020). Measuring income too early could bias the results because children from rich parents usually have a steeper earnings path because of longer education. As the focus of this paper is to estimate the causal effect of a policy, this would only be relevant if treatment and control group would experience different trends. Nevertheless, I test the sensitivity of this specification with measuring child income at different ages.

Parent Rank: Parent income is measured as the sum of father and mother income averaged over the child ages 12 to 22 years. If the link of one parent is missing, the income is set to zero, assuming that this parent is dead or otherwise not available to support the child. If we do not measure parent income in one year, I set this income-year to missing and do not use it to measure the average income. Finally, the goal is to determine the parent rank in terms of quintiles or percentiles. Therefore, I rank the parents by child cohort. As the parent rank is always relative to the birth cohort, this should ease concerns on the selection of different age groups. As the principal focus is to classify parents in quintiles, small measurement errors are unlikely to affect the classification.

TABLE 6.1: Summary Statistics

	(1) All				(2) Treatment (Border Region)				(3) Control (Non-Border Region)			
	Mean	Min	Max	Obs	Mean	Min	Max	Obs	Mean	Min	Max	Obs
Year of Birth	1,971.64	1,960.00	1984.00	1,480,449	1,971.67	1,960.00	1984.00	993,009	1,971.58	1,960.00	1984.00	487,440
Share Female	0.47	0.00	1.00	1,480,449	0.47	0.00	1.00	993,009	0.47	0.00	1.00	487,440
Married	0.50	0.00	1.00	1,480,449	0.49	0.00	1.00	993,009	0.52	0.00	1.00	487,440
Swiss Citizenship	0.99	0.00	1.00	1,480,449	0.99	0.00	1.00	993,009	1.00	0.00	1.00	487,440
Mean Income 30 to 33	62,915.03	3.96	20,638,628	1,480,449	63,682.56	4.95	20,638,628	993,009	61,351.42	3.96	7,393,222	487,440
Mean Log Income 30 to 33	10.79	1.38	16.84	1,480,449	10.80	1.60	16.84	993,009	10.77	1.38	15.82	487,440
Parent Quintile	3.02	1.00	5.00	1,480,449	3.10	1.00	5.00	993,009	2.86	1.00	5.00	487,440
Parent Income 12 to 22	107627	0.00	1.88e+08	1,480,449	112254	0.00	1.88e+08	993,009	98,200	0.00	14,095,484	487,440
Tertiary Education	0.42	0.00	1.00	500,990	0.42	0.00	1.00	336,247	0.40	0.00	1.00	164,743
Gymnasium	0.19	0.00	1.00	500,990	0.20	0.00	1.00	336,247	0.17	0.00	1.00	164,743
Vocational Degree	0.66	0.00	1.00	500,990	0.64	0.00	1.00	336,247	0.70	0.00	1.00	164,743
Senior officials and managers	0.03	0.00	1.00	465,127	0.03	0.00	1.00	311,151	0.03	0.00	1.00	153,976
Professionals	0.32	0.00	1.00	465,127	0.33	0.00	1.00	311,151	0.30	0.00	1.00	153,976
Technicians and associate	0.14	0.00	1.00	465,127	0.14	0.00	1.00	311,151	0.14	0.00	1.00	153,976
Clerks	0.16	0.00	1.00	465,127	0.17	0.00	1.00	311,151	0.15	0.00	1.00	153,976
Service/shop sales woker	0.12	0.00	1.00	465,127	0.12	0.00	1.00	311,151	0.14	0.00	1.00	153,976
Skilled agricultural / fisher	0.03	0.00	1.00	465,127	0.03	0.00	1.00	311,151	0.04	0.00	1.00	153,976
Craft and related workers	0.18	0.00	1.00	465,127	0.17	0.00	1.00	311,151	0.19	0.00	1.00	153,976
Plant operators/assemblers	0.01	0.00	1.00	465,127	0.01	0.00	1.00	311,151	0.01	0.00	1.00	153,976
Elementary occupations	0.00	0.00	1.00	465,127	0.00	0.00	1.00	311,151	0.00	0.00	1.00	153,976
N	1,480,449				993,009				487,440			

Notes: This table summarizes several variables for treatment (border region) and control group (non-border region). The main data including cohorts from 1967 to 1984 consists of 1,48 million observations. This data is used to estimate the effect of immigration on different parts of the parental income distribution. Variables on education and occupation are used for the analysis of mechanism. Here, the data relies on a representative survey which does not cover the full population. Thus, the sample size is reduced to roughly 0,51 million observations.

6.4 Empirical Strategy

6.4.1 Estimation

The goal of this study is to estimate the effect of exposure to crossborder-workers on labor income conditional on parental background and infer the effect on intergenerational income mobility of natives. As exposure to cross-border immigrants is likely endogenous to labor market outcomes of natives, naïve OLS would yield biased results. To circumvent endogeneity, I analyse a large quasi-random experiment in Switzerland. Thereby, restrictions for cross-border workers were incrementally removed in certain municipalities close to the border. This leads to variation over time and space, which I exploit in a difference-in-difference setting.[2]

I divide municipalities in treatment and control groups. In the baseline specification, municipalities within a 30 minutes car drive from the border passing are ascribed to the treatment group.[3] In those municipalities, restrictions for cross-border worker were removed earlier. As Figure 6.1 shows, there are also few municipalities further away than 30 minutes that belong to the border region according to the government. However, factual exposure to cross-border workers is low in those municipalities, therefore they are counted to the control group in the baseline specification.[4] Municipalities further away are defined as control group (non-border region). In a robustness check, I also distinguish between high-exposure treatment group (less than 15 minutes away) and low-exposure treatment group (between 15 and 30 minutes away).

I ascribe children to the municipality in which their mother lives in 2010. If there is no information about the mother, I take the municipality of the father in 2010. Data does not provide panel data information on the municipality of residence until 2010. However, the residence of the children is known as of 2010, and in 2010 there is also an indicator telling since when a person lives in that municipality. In section 6.8.3, I test if the results are robust to other definitions of the child's municipality, such as place of birth or residence in 2010. In addition, I restrict the sample to children for whom parents still living in the same municipality as in 2010 when the child was 16.

Figure 6.1 shows a map of the treatment and control municipalities. Importantly, the treatment and control groups can be within one canton. This is important because Switzerland is organized federally and cantons have most jurisdiction. The overlap between treatment and control group within cantons alleviates concerns that other policy changes interfere with the abolition of the cross-border restrictions.

[2]This identification strategy has been used in other studies before, most notably in Beerli et al. (2020)
[3]Section 6.8.1 tests the robustness of the results with different definitions of the treatment groups.
[4]Section 6.8.1 shows that results do not change when they are ascribed to the treatment group.

In terms of variation across time, I follow Beerli et al. (2020) and divide the years 1991 to 2017 into three periods. I define the years before 1999 as pre-treatment period. The years between 1999 and 2004 are defined as transition period. Years after 2004 are defined as post-treatment period.

The resulting variation in time and space is analyzed in two (nested) models. A difference-in-difference model and—as a generalized version of it—an event study model showing year lags and pre-treatment effects.

The difference-in-difference specification is defined as follows:

$$\begin{aligned} y_{irnt} &= \alpha + \gamma BorderRegion_r + \lambda_{trans} Trans_t + \delta_{trans}(BorderRegion_r \times Trans_t) \\ &\quad + \delta_{post} Post_t + \delta_{post}(BorderRegion_r \times Post_t) + X'_{irnt}\beta + \lambda_{1n} t + \lambda_{2n} t^2 + \epsilon_{irnt} \end{aligned} \tag{6.1}$$

where i stands for individuals, r for regions (Border and Non-Border Region), n for NUTS-2 regions, and t for years. $BorderRegion$ equals one if an individual is in a border region, $Trans_t$ is one if the year belongs to the transition period, $Post_t$ equals one if the year belongs to the post-treatment period. X' is a vector of control variables, including gender and region fixed effects, t and t^2 are linear and quadratic trends. The outcome variable y stands either for real income in logs between the ages 30 to 33 or for the rank of the mean income between 30 to 33 within a child's cohort. λ_{trans} captures treatment effects in the transition period, δ_{post} measures treatment effects in the post-treatment period. Importantly, the equation is estimated for different child subgroups defined by the parents' rank in the income distribution.

Besides the standard difference-in-difference model above, I also show estimates of the generalized event study model. Thereby, I estimate the treatment effects for every year—except for the year before the policy implementation ($t = 1999$) which serves as a reference year. Treatment effects before introducing the policy serve as a placebo test. This has the advantage that one can analyze how the effects change after implementing the policy and whether effects are zero before the introduction. Formally, this yields the following equation:

$$\begin{aligned} y_{irnt} &= \alpha + \gamma BorderRegion_r + \sum_{\substack{t=1991 \\ t \neq 1999}}^{2017} \delta_t \mathbb{1}(year = t) \times BorderRegion_r \\ &\quad + X'_{irnt}\beta + \lambda_{1n} t + \lambda_{2n} t^2 + \epsilon_{irnt} \end{aligned} \tag{6.2}$$

where δ_t show the yearly treatment effects. Importantly, one would expect δ_t to be zero before implementing the policy (or at least the announcement of the policy). Otherwise, it is doubtful whether the parallel trends assumption actually holds. Throughout the study, I cluster robust standard errors on NUTS-2 regions. NUTS-2 regions consist of seven different regions.[5]

6.4.2 Identification

The crucial identifying assumption in a difference-in-difference setting is that treatment and control group would have followed the same trends in absence of the policy. While this assumption is not testable, I take advantage of the long period of the data set, allowing to show that the trends were parallel before introducing the treatment. In addition, I control for region specific fixed effects (NUTS-2) and test whether the estimates are robust to including region specific linear and quadratic time trends.

Figure 6.2 shows the development of yearly real labor income by parental background and treatment group assignment status. The graph shows the trends for women and men separately to abstract from gender-specific labor supply changes. Specifically, it shows the log wage for children from parents in the top quintile in the border and non-border region, and the log wage for children from parents of the bottom quintile. The first vertical line in 1999 shows the onset of the transition period, the second vertical line in 2004 shows the commencement of the free movement phase in the border region.

When looking at men in Panel (a) of Figure 6.2, this first descriptive evidence already shows a relatively severe drop in real income in the year 2004. Also in 1999, when first restrictions were removed, real income declines in border regions for children from the bottom quintile—whereas real income of children from the bottom quintile living in the non-border region did not experience such a decline. The graph also shows that there is no growing difference in border vs. non-border regions for children from the top quintile of the parental income distribution. For women in Panel (b) of Figure 6.2, there does not seem to be a strong difference in the bottom quintile between border and non-border region at the times of the policy change. Nevertheless, real income growth decelerates more for women in the border region compared to women in the non-border region.

This different development in treatment and control group cannot yet be interpreted as a causal effect, because it might still be confounded by some time-varying variables. However, the crucial point of this graph is that trends before the policy changes in 1999 and 2004 are remarkably similar. This suggests that the parallel trends assumption is

[5] Those seven regions are: Lake Geneva region, Espace Mittelland, Northwestern Switzerland, Zurich, Eastern Switzerland, Central Switzerland, and Ticino.

plausible because—at least in the pre-treatment period—-there is no difference in the development over time between groups.

A less discussed, but important, assumption in a difference-in-difference setting is the stable treatment unit value assumption (SUTVA). This assumption requires the treatment groups to be well defined and rules out spillovers. In this study, people are, in principle, free to move to another municipality or work in another region. In Section 6.8.2, I analyze whether there is a geographic mobility reaction around the policy implementation date. I do not find any evidence that there is a reaction in mobility around the cut-off date. I also test if the effects change when the municipality of birth is used to assign children to treatment and control group.

6.5 Results

6.5.1 First Stage

The abolition of restrictions for cross-border workers led to a strong increase in the number of cross-border workers in the exposed region. This has already been shown by Beerli et al. (2020) and Parenti & Tealdi (2019). Figure B.4 replicates those results by showing the share of cross-border workers for the treatment and control region relative to the number of employed people in 1995. As of the transition period (1999), there is a strong increase in the number of cross-border workers.

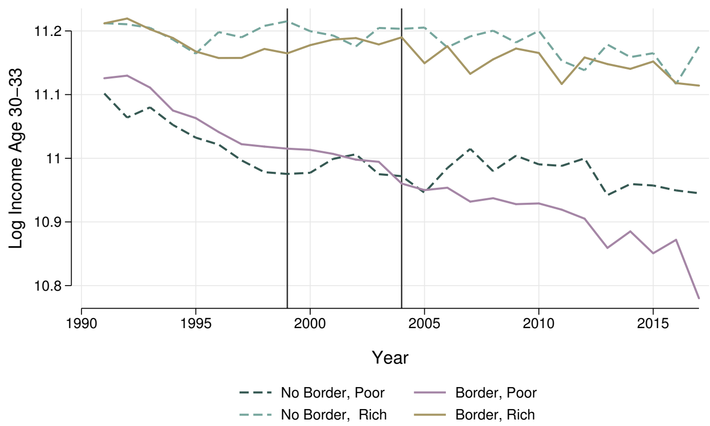
6.5.2 Effect on Child Income by Parental Background

Figure 6.4 shows the event study coefficients δ_t as described in Equation 6.2 for children with parents at the bottom quintile (Parent-Q1) and for children with parents at the top (Parent-Q5) quintile of the income distribution. The year 1999 is the omitted; thus the coefficients are relative to this «pre-introductory year».

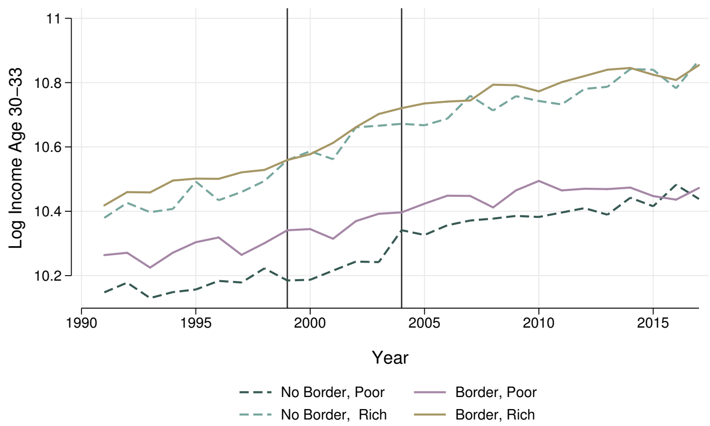
Looking at the coefficients of children with parents from the bottom quintile Parent-Q1, we can see that the coefficients are always insignificantly different from zero in the pre-treatment and also mostly during the transition period. This is in line with the parallel trends assumption required to identify a causal effect: Before the full introduction of the policy in 2004, treatment and control group do not diverge. As the post period starts, so does the decrease of the coefficients: The treatment effects become negative. Figure 6.4 also shows that the negative effects for children from the bottom quintile become more pronounced over time. The increase in the treatment effect over time for poor background children coincides with the intensity of the treatment shown in Figure B.4.

FIGURE 6.2: Real Income by Region, Parental Background and Gender

(A) Men

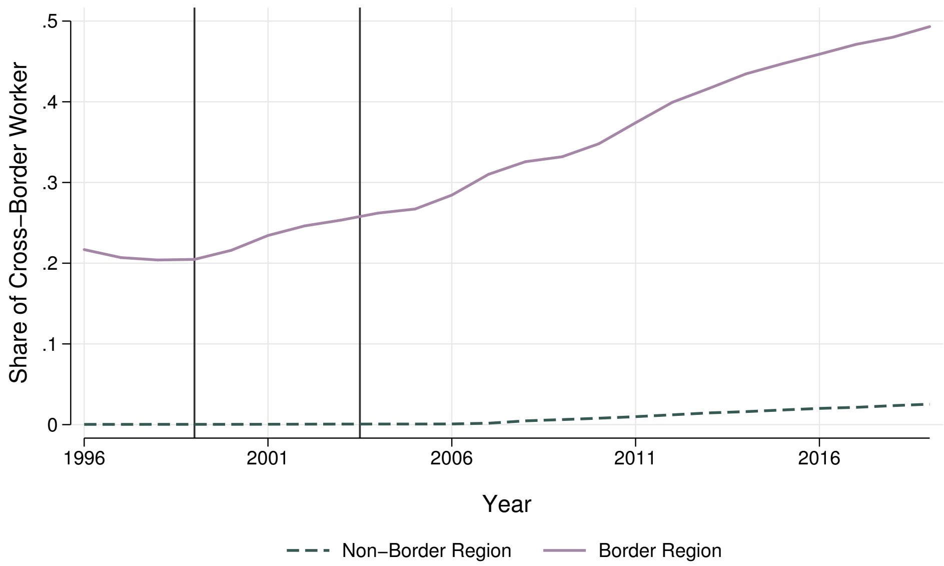


(B) Women



Notes: This figure shows how real log income of children from different parental backgrounds develops in the border and non-border region. Child income is defined as real log income averaged over the ages 30 to 33. I split the figure between women and men to abstract from changes in gender-specific labor supply. Especially for men in the bottom quintile of the income distribution, the trends in treatment and control group start to diverge in 1999, when cross-border immigration restrictions were removed in cross-border regions. Most importantly, the trends of treatment and control group shows a parallel pattern before the onset of the policy.

FIGURE 6.3: First Stage



Notes: This figure shows the share of cross-border workers in border and non-border regions by municipality. Until 2007, cross-border workers were not allowed to work in non-border regions. However, even after the complete liberalization in 2007, the number of cross-border workers in non-border regions remains tiny.

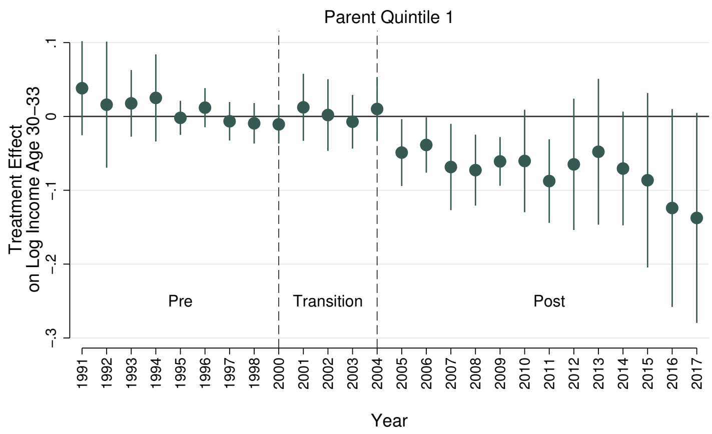
When looking at children from the top quintile (Parent-Q5), the overall effect is less pronounced but is pointing in the opposite direction. Labor earnings increased mildly for those children in the post-treatment period. Again, the coefficients are mostly insignificant in the pre-treatment period, which would be in line with the parallel trends assumption.

Table 6.2 summarizes the event study estimates and provides difference-in-difference coefficients for the post and transition period as specified in Equation 6.1. The columns (1) to (10) are grouped by parental quintile, where the odd column numbers show the estimates of a specification without region specific trends and the even numbered columns show estimates controlling for region specific trends.

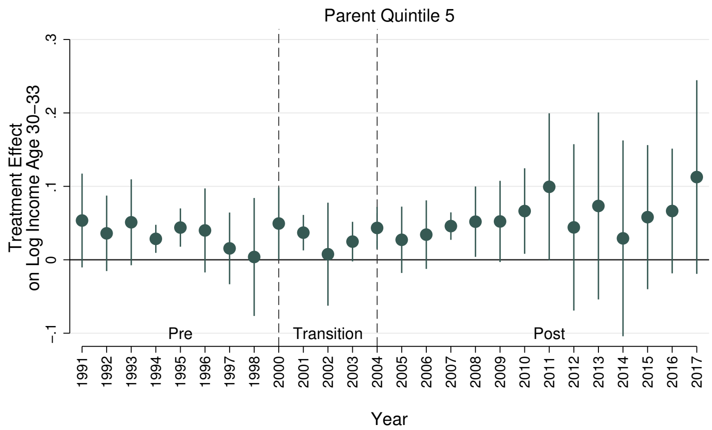
For the first and second parent quintile depicted in columns (1) to (4), the coefficients of the post treatment period are significantly negative and remarkably similar to each other. They suggest that labor income is reduced by around 3.2 to 6.8 percents. This is similar to the event studies estimates shown in Figure 6.4. Children from the third parent quintile shown in columns (5) and (6) experience a slighter loss in labor income. The coefficients are roughly half as large as in the two lower quintiles mentioned before. In the fourth parent quintile, the coefficients are quantitatively similar to the third quintile. When moving up to the top quintile shown in columns (9) and (10) the effect is not significantly different from zero anymore.

FIGURE 6.4: Event Study Estimates on Log Income of the Child

(A) Parent-Q1



(B) Parent-Q5



Notes: This figure shows the event study difference-in-difference coefficients δ_t at every year t as described in Equation 6.2 for children with parents from the bottom quintile (Parent-Q1) and from the top quintile (Parent-Q5) of the parental income distribution. The corresponding lines shows a 90% confidence interval. Standard errors are clustered on regional NUTS-II level. Year dummy 1999 is omitted. The specification includes region fixed effects (NUTS-II), region specific trends, and sex. The income of the children is measured as log of the mean real income over the ages 30 to 33. Parental rank is measured as child cohort specific rank of the mean income when the child is between 12 and 22. Figure E.1 in the Appendix shows the event study estimates for all parent quintiles and for different region-specific trend specifications.

TABLE 6.2: Treatment Effects on Log Child Income by Parent Quintile

	Dependent Variable: Child Log Income									
Subgroup:	Parent-Q1		Parent-Q2		Parent-Q3		Parent-Q4		Parent-Q5	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment Post =1	-0.068***	-0.032**	-0.067***	-0.029*	-0.037**	-0.008	-0.031**	-0.007	-0.017	0.004
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Treatment Trans = 1	0.011	0.020	-0.034**	-0.015	-0.009	0.004	-0.017*	-0.007	-0.021*	-0.012**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional Trends	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	290,907	290,907	291,629	291,629	294,715	294,715	299,218	299,218	303,980	303,980
R2	0.126	0.127	0.144	0.144	0.153	0.153	0.135	0.136	0.104	0.104

Notes: This table shows the treatment effects of the regression on Equation 6.1 for different subgroups according to parental income. Thereby, the outcome variable is the log of the child's income distribution around the age 30 to 33. Parent-Q1 refers to parent quintile 1 etc.. The coefficient «Treatment Post» shows the difference-in-difference coefficient for the post-treatment period, the coefficient «Treatment Trans» shows the difference-in-difference coefficient for the transition period. Regional Trends include region specific quadratic trends. Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

Table 6.3 shows the same estimates for child rank instead of child log income. The picture is essentially the same. The effect is stronger for children from poorer parental backgrounds. Interestingly, the effect here is strongest for the second quintile. Figure E.2 in the Appendix also shows the event study estimates with child rank as an outcome variable.

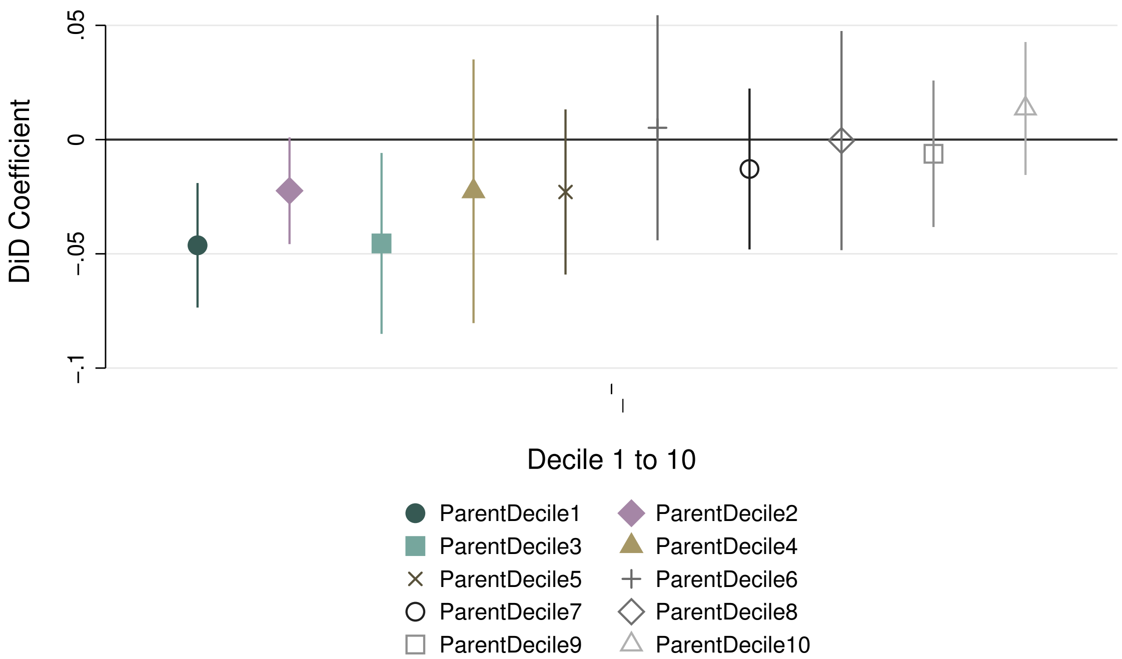
TABLE 6.3: Treatment Effects on Rank Child Income by Parent Quintile

	Dependent Variable: Child Income Rank				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post = 1	-1.431*	-1.862**	-0.970*	-0.358	0.190
	(0.63)	(0.54)	(0.44)	(0.48)	(0.44)
Treatment Trans= 1	0.772	-0.874	-0.232	-0.006	-0.190
	(0.47)	(0.55)	(0.35)	(0.57)	(0.69)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
Regional Trends	Yes	Yes	Yes	Yes	Yes
N	278,873	280,317	283,695	288,645	293,909
R2	0.138	0.153	0.17	0.153	0.115

Notes: This table shows the treatment coefficients as specified in Equation 6.1 for different quintiles of the parental income distribution. The outcome variable here is the percentile rank of children in the income distribution.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

Figure 6.5 shows the estimated difference-in-difference coefficient again, but now for deciles of the parental income distribution instead of quintiles. The points indicate the estimates of the difference-in-difference estimate of the post-period for each decile of the parental income distribution. Again, the coefficients show a similar picture. The poorer the parental background, the larger the negative effect of the policy. Children from the highest decile even show a positive point estimate although not significantly different from zero. Remember that the sub-samples here is two times smaller, which decreases statistical power. The numbers can also be found in Table E.1.

FIGURE 6.5: Difference-in-Difference Treatment Coefficient on Log Child Income by Parent Decile



Notes: These points show the difference-in-difference coefficients by parental income deciles and its corresponding 90% confidence interval. The specification includes regional fixed effects, sex fixed effects, and regional specific trends.

To summarize, total labor income of children decreases more, the lower parental income is. Mechanically, this leads to lower intergenerational mobility because it increases the already positive correlation between child and parent income. In the next subsection, I will quantify how strong the decrease in intergenerational income mobility is.

6.5.3 Effect on Intergenerational Income Mobility

The previous section has shown that labor income of children from poorer parental background decreases more in light of immigrant influx. What do those parental background specific treatment effects mean for intergenerational income mobility?

Several measures exist to quantify intergenerational (income) mobility. Traditionally, the «intergenerational elasticity» has been used widely (IGE). It measures how a percentage increase in parental income is associated with a percentage increase in child income. The IGE is estimated as the slope coefficient of a univariate linear regression of logarithmized child income on logarithmized parent income. However, recently this measure has been replaced by the rank-rank slope (RRS), where child and parent income are not logarithmized, but transformed into percentile ranks. In contrast to the IGE, the RRS has the advantage that the rank-rank relationship is almost linear and the slope coefficient can therefore be used as a meaningful, parsimonious statistic. Further, the RRS takes zero incomes into account (Black & Devereux, 2010; Lee & Solon, 2009). Therefore, I will estimate the effect of cross-border immigration on the rank-rank slope (RRS) among natives.

To estimate the effect of the immigrant influx on the RRS of natives, I proceed as follows. First, I estimate the effect of the policy for each percentile of the parental income distribution (instead of quintiles used before). Thereby, the outcome variable is now income in percentile ranks instead of log income.[6]

I depict the estimated post treatment coefficients of those 100 regression in Figure 6.6. The figure shows the effect for each parent percentile rank. Confidence intervals are not shown because due to the 100 times reduced sample size, they mostly overlap the zero. Nevertheless, one can perceive a pattern that is consistent with the coarser results based on quintiles from the previous sections: The effect is most negative for children with low-income parents.

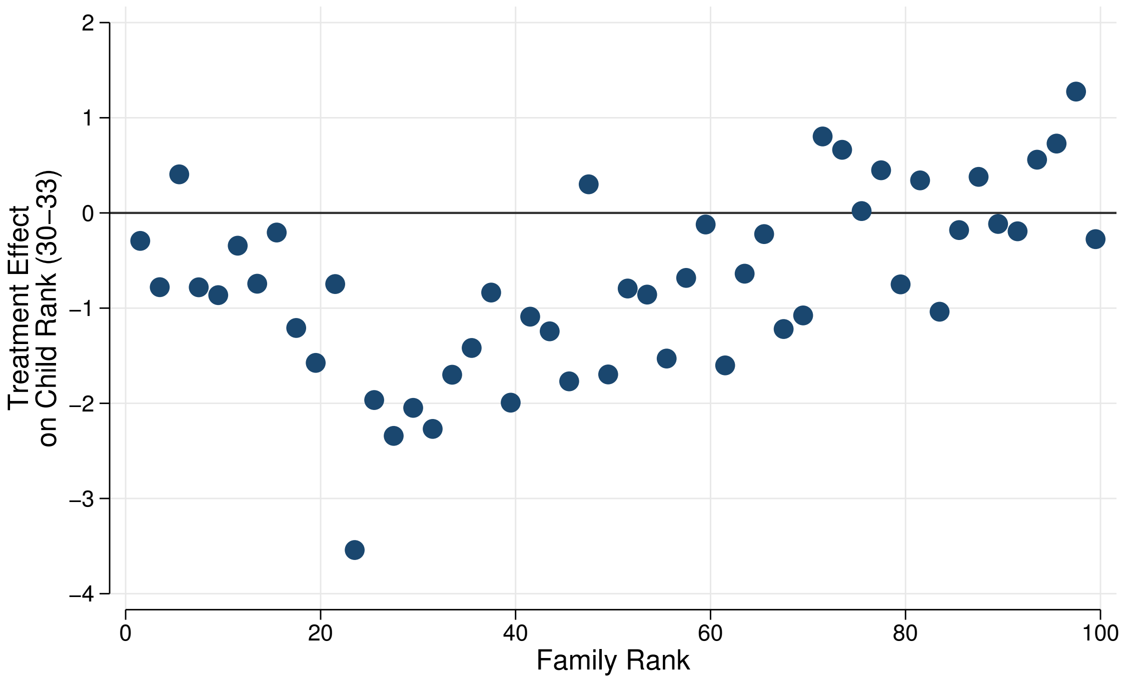
Next, I will subtract those estimated coefficients from the observed mean child rank in the post-treatment period. This generates a counter-factual «child rank» for every parent percentile that would have occurred in the absence of the policy. Then, I estimate the rank-rank slope for this counter-factual child-parent distribution and compare the resulting rank-rank correlation to the observed child-parent distribution. Then, I can compare the observed and the counter-factual RRS and infer the effect on intergenerational mobility.

The results are shown in Table 6.4. Column (1) shows the results of the observed rank-rank regression in the post-treatment period in the border-region. Column (2) shows the counterfactual slope. The variable *Family Rank* refers to the rank-rank-slope estimates (RRS). The RRS of the observed incomes is 0.134: An increase in parent income ranks is associated with an increase in 0.134 child ranks (Column 1). In the absence of the policy, the estimates slope would be 0.118. Thus, the policy led to a reduced intergenerational mobility by 1.6 percentile ranks—or in relative terms to a decrease in 12%.[7] Here it is important to mention that Switzerland has a high intergenerational income mobility compared to other countries (Chuard & Grassi, 2020). Thus, the relative increase is naturally large.

[6]Child rank is the rank of the mean child income between 30 to 33 sorted by child year of birth. As the parallel trends assumption is sensitive to different transformation, parallel trends in log does not imply parallel trends in ranks. However, Figure E.2 in the Appendix shows that trends in the pre-treatment period are also similar for a rank transformation.

[7]Formally, the slope increases by 0.016. However, the RRS is often multiplied by 100 because this depicts the «wedge» between children from rank 1 and children from rank 100.

FIGURE 6.6: Effect of Policy on Child Rank by Parent Percentile



Notes: This graph shows the treatment effects by percentile of the parent income distribution. Dependent variable is the child cohort rank in the income distribution at age 30 to 33.

6.6 Mechanism

What could be the reason behind this decrease in intergenerational mobility of natives? The canonical partial equilibrium model aiming to explain the impact of migration, as for example in Altonji & Card (1991), makes straightforward and intuitive predictions in the short run: An increase in immigrant labor supply lowers the labor opportunities for workers competing with immigrants, while it increases the opportunities for workers complementing immigrant workers. Consequently, if children from poor parental backgrounds are more often in segments competing with immigrants, intergenerational mobility would fall. Over time, one could expect that the potential negative income effects would be attenuated because natives move to segments with less competition by immigrants. To do so, children might opt for different educational or occupational tracks.

To see if this is the case, I run the difference-in-difference regression based on Equation 6.1 for different learned occupations (ISCO-08) and different educational tracks. Then, I compare the resulting treatment effects in those specific segments with the share of children from poorer parental backgrounds in this segment. If more negative treatment effects are correlated with more children from poorer parents, this would explain the source of the decrease in intergenerational mobility.

TABLE 6.4: Counterfactual Rank-Rank Slope

	(1) Observed Child Income Rank	(2) Counterfactual Child Income Rank
Family Rank	0.134***	0.118***
	(0.002)	(0.009)
Constant	44.075***	45.320***
	(0.128)	(0.500)
Observations	100	100

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This graph shows the results of regressions of child rank on parent rank. «Family Rank» indicates the slope parameter (RRS). The first column shows the observed rank-rank slope (RRS). The second column shows the counterfactual RRS—the rank-rank slope that would result in the absence of the immigration influx.

6.6.1 Educational Track

Does the educational track impact the size of the effect? And if so, does the educational track depend on parental income? In Switzerland, around 70% of children opt for vocational education and training (VET) after compulsory school. Only a small share of children opts for the academic track which starts with the gymnasium (similar to an academic high school) and which can traditionally lead to university or less often to a University of Applied Sciences (UAS), where the first degree would be a bachelor therefore similar to a college degree in the US. The answer to the second question is already clear: In Switzerland, education depends highly on parental income. This is true, even though intergenerational income mobility is among the highest (Chuard & Grassi, 2020).

Table 6.5 shows the treatment effects for children opting for the academic track and for children opting for the non-academic track (VET). To exemplify how this effect depends on parental income, columns (1) and (2) show the effects for children from the bottom quintile and columns (3) and (4) for children with parents from the top quintile.

The first result that sticks out is that children with an academic education are less affected by the negative income effect by immigrant influx. Column (1) and column (3) show that the point estimate is above zero, suggesting a positive income effect independent of parental background. Column (1) shows that children from Parent-Q1, that belong to the small share of poorer children with an academic education, do not suffer from a decrease in labor income. At the same time, Column (4) shows that children with parents from the top quintile of the income distribution also experience a negative effect, if they are not opting for an academic education—although slightly smaller than the non-academics from parents in the first quintile shown in column (2).

The row «Share children in track» notes how education correlates with parent background. Specifically, it shows the share of children in either the academic or the non-academic track. The differences are stark: While 22.5% of children from parents in the top quintile end up with a university master's degree, this share is only 6.6% for children with parents from the bottom quintile of the income distribution.

This has direct effects on intergenerational mobility. Immigration decreases labor income for non-academic individuals who are more prevalent among poorer parents. To summarize the role of education: Highly educated individuals do not seem to be negatively affected by the immigrant influx. Thus, intergenerational mobility is reduced because children from the bottom quintile are almost four times less likely to complete an academic education.

Do children opt for other educational tracks in light of the policy to ease the negative wage effect? Figure 6.7 shows the development for several tracks. Trends between border and non-border regions do not systematically diverge around the policy cut-off or in the aftermath. Thus, there is no evidence that children adopt their educational track because they fear a negative labor income effect because of immigrants.

6.6.2 Learned Occupation

To analyze how learned occupations are affected and correlated with parental income, I look at different levels of granularity of the ISCO-08 occupations classification (International Standard Classification of Occupations). First, the ten major groups and then, the sub-major groups which include 43 different occupational descriptions.

Figure 6.8 depicts the treatment effects (in the post period) for the ISCO-08 major groups and the share of children with parents below the median of the income distribution. The pink line is a linear regression line of a regression weighted by the number of children in this regression. The size of the point is a proxy for the relative size of this segment in the sample. Likely because of the reduced sample size, the treatment effects are almost never significantly different from zero. Nevertheless, one can spot a pattern: The regression line shows that there is a negative relationship between the segment specific treatment effect and share of children from below the median in that segment. To see whether this pattern also holds for more fine-grained occupational segments, Figure 6.9 shows the same for sub-major occupational groups. Again, there is a negative relationship between share of children from parents below the median and the treatment effect.

TABLE 6.5: Treatment Effects on Log Child Income by Education and Parental Background

	Dependent Variable: Child Log Income			
Parental Background:	Parent-Q1		Parent Q-5	
Educational Track	Academic	Non-Academic	Academic	Non-Academic
	(1)	(2)	(3)	(4)
Treatment Post = 1	0.012	-0.069***	0.046	-0.056***
	(0.04)	(0.01)	(0.04)	(0.01)
Treatment Trans = 1	0.064	-0.000	0.001	-0.041**
	(0.08)	(0.01)	(0.02)	(0.02)
Region FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes
Share children in track	6.6%	93.4%	22.5%	77.5%
N	6,341	89,606	23,856	82,308
R2	0.0496	0.15	0.0639	0.125

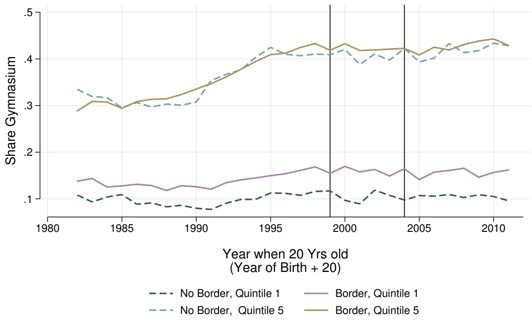
Standard Errors clustered on Regional (NUTS-2) Level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

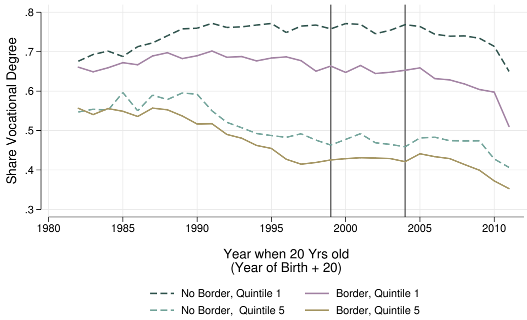
Notes: This table shows the treatment effects for four subgrous along two criteria: Parental background and educational track. E.g. Column (1) shows the treatment effects for children in the academic track (gymnasium and university) and for children with parents from the bottom quintile of the income distribution. In addition, the row «share children in track» shows how many children of the parent quintile are in a specific school track. E.g. in column (1) 6.6% of children from parents in the bottom quinitle of the income distribution attend an academic education track. The results indicate that the treatment effect is negative for children from the bottom and the top quintile if they do not attend an academi track. Thus, the overall negative effect for children with parents from the bottom quintile results from the low share of children in the academic track.

FIGURE 6.7: Educational Track over Time, Treatment Region, and Parental Background

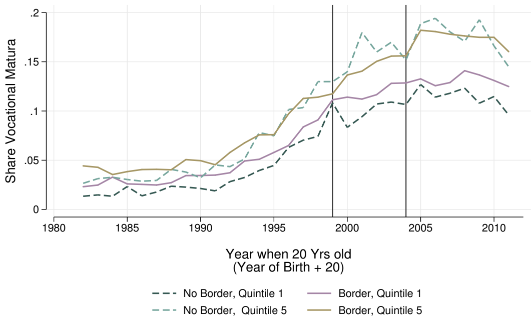
(A) Gymnasium (Academic High School)



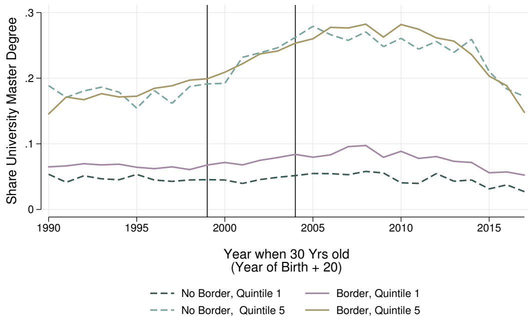
(B) Vocational Education and Training (VET)



(C) Vocational Matura (VET with formal education)



(D) Gymnasium and Master Degree (likely University)



Notes: This figure shows the share of children in several educational tracks by cohort year, treatment group, and parent quintile. Treatment and control group (dashed vs. solid lines) in the two parent income groups do not seem to diverge around the policy cut-off. Here: the policy cut-off is the year when we would expect the education to be completed.

The negative relationship of the share of low-income children and the learned occupation explains why intergenerational income mobility declines due to the policy: Children from poorer backgrounds disproportionately learned occupations that are or will become negatively affected by the policy.

If certain occupations are more negatively affected due to the immigrant influx, this raises the question whether children react by opting for other occupations? To test this, I plot how several learned occupations develop by cohort and by border and non-border region. However, children do not react by choosing different occupations. Figure 6.10 depicts how the learned occupations for the ISCO-08 major groups develop over cohort, by treatment assignment and by parental background. The year indicates when children are supposed to have completed the learned education. However, I do not know exactly when children completed educations, thus this year is simply an approximation. It could also be some years earlier or later. Nevertheless, if there are no strong differences in trends between treatment and control group, there is unlikely a reaction. Indeed, this is what the graphs tell us. There is no stark difference between treatment and control group around the policy cut-off, and the trends also do not strongly diverge in the aftermath.

6.7 Political Consequences

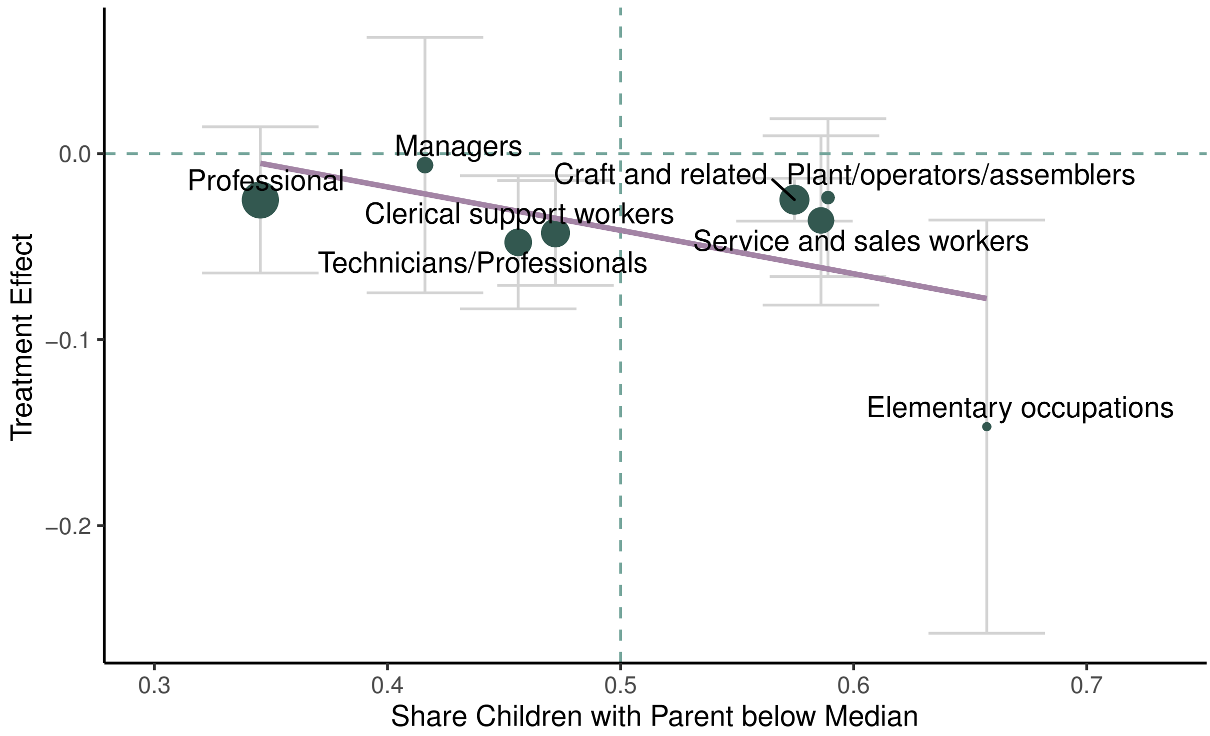
Opponents against immigration often rely on economic arguments saying that the increased competition could harm workers. The previous chapters have shown that immigration can decrease intergenerational income mobility and subsequently increase inequality over generations. This sort of inequality is often perceived as unfair and raises the question whether it could lead to sentiments against immigration.

This setting provides a way to test if economic drawbacks by immigrants also correlate with sentiments against immigration. Switzerland provides a convenient setting to measure such sentiments because it regularly holds ballots in which citizens vote on several aspects, also on such concerning immigration. In 2014, there has been an initiative against «mass-immigration» and in 2020 an initiative for «modest immigration». Besides data from plebiscites, I also use data on municipality specific support for immigration opposing parties in the quadrennial national elections.

To test this hypothesis whether a decrease in intergenerational mobility could be responsible for sentiments against immigration, I run several tests.

First, I check if there is larger support for immigration restrictions in municipalities with many children from poor-income backgrounds. As children from poorer backgrounds experience more negative wage effects, one would expect to observe more

FIGURE 6.8: Treatment Effect on Log Child Income and Parental Background by ISCO-08 Major Groups

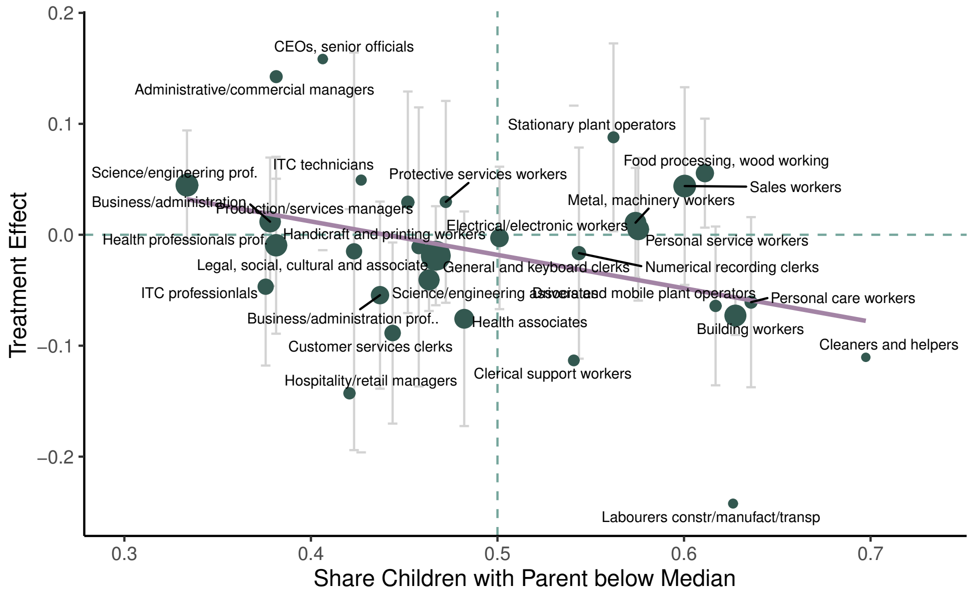


Notes: This figures shows how treatment effects vary by learned occupation and how large the share of children with parents from below the median income is. The vertical lines show how large a 95% confidence interval. The size of the point indicates the number of children in this group. The purple line is a weighted regression line and shows that a more negative treatment effect correlates with a larger share of parents below the median in those occupations. Learned occupations are based on the ISCO-08 major groups (without army and agriculture).

votes in favor of immigration restrictions in municipalities with more children from poorer backgrounds—given that the loss in intergenerational mobility actually leads to anti-immigration sentiments. Figure 6.11 shows that the higher the share of children from poor parental backgrounds in a municipality, the higher the votes in favor of anti-immigration policies.

Second, I test whether the support for anti-immigration votes is also highest in cantons experiencing the highest decline in intergenerational mobility due to immigration. Figure 6.12 shows that this is in general true. Cantons with more negative effects for children from the bottom quintile also show a higher voting share for anti-migration policies.

FIGURE 6.9: Treatment Effect on Log Child Income and Parental Background by ISCO-08 Submajor Groups



Notes: This figures shows how treatment effects vary by learned occupation and how large the share of children with parents from below the median income is. The vertical lines show how large a 95% confidence interval. The purple line is a weighted regression line and shows that a more negative treatment effect correlates with a larger share of parents below the median in those occupations. The size of the point indicates the number of children in this group. Learned occupations are based on the ISCO-08 sub-major groups (without army and agriculture)

Third, I check if parties in favor for more restrictive gain due to the increase in cross-border worker. Thus, I run the difference-in-difference regression as specified in Equation 6.1 with party votes on a municipality level for every four years when the voting takes place. Table 6.6 shows the result of this regression. Parties on the right favoring anti-migration policies increased their voter share by 2.4 percentage points due to exposure to immigration.

Those tests cannot fully prove that it is indeed the decline in intergenerational mobility that led to an increase in sentiments against immigrants. However, the arguments do not contradict such a hypothesis.

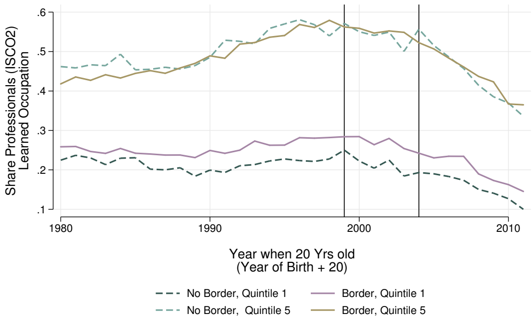
6.8 Robustness

6.8.1 Alternative Control Groups

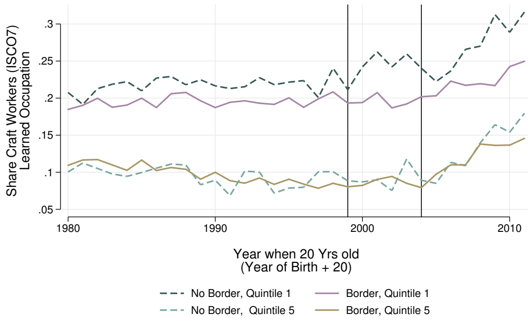
In the main specification, treatment groups are defined as municipalities within a 30 minutes car drive to the next border passing. The results are, however, robust to several other treatment and control group specifications.

FIGURE 6.10: Learned Occupations over Time, Treatment Status, and Year

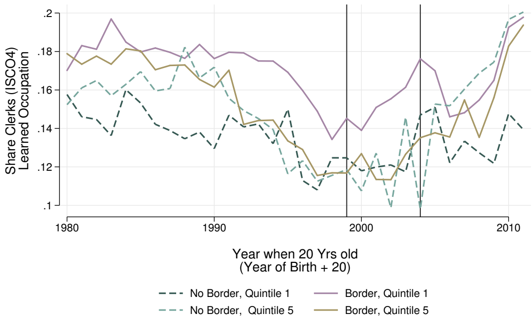
(A) Professionals (ISCO-08 = 2)



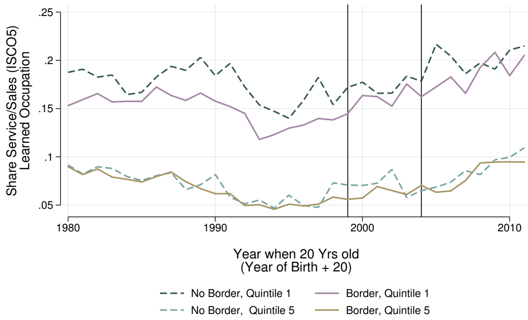
(B) Craft Workers (ISCO-08 = 7)



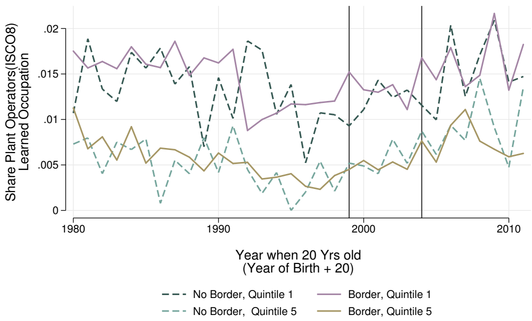
(C) Clerks (ISCO-08 = 4)



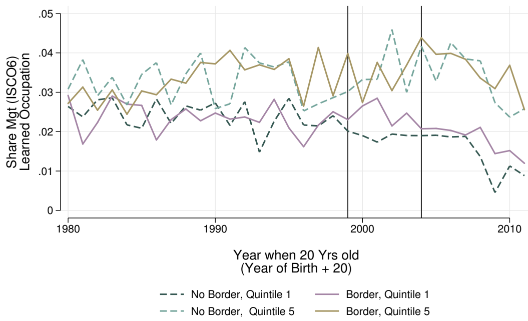
(D) Service (ISCO-08 = 5)



(E) Plant Operators (ISCO-08 = 8)



(F) Management (ISCO-08 = 1)



Notes: This figure shows the share of children in several learned occupations (ISCO-08 major groups) by cohort year, treatment group, and parent quintile. Treatment and control group (dashed vs. solid lines) in the two parent income groups do not seem to diverge around the policy cut-off. Here: the policy cut-off is the year when we would expect the learned education to be completed.

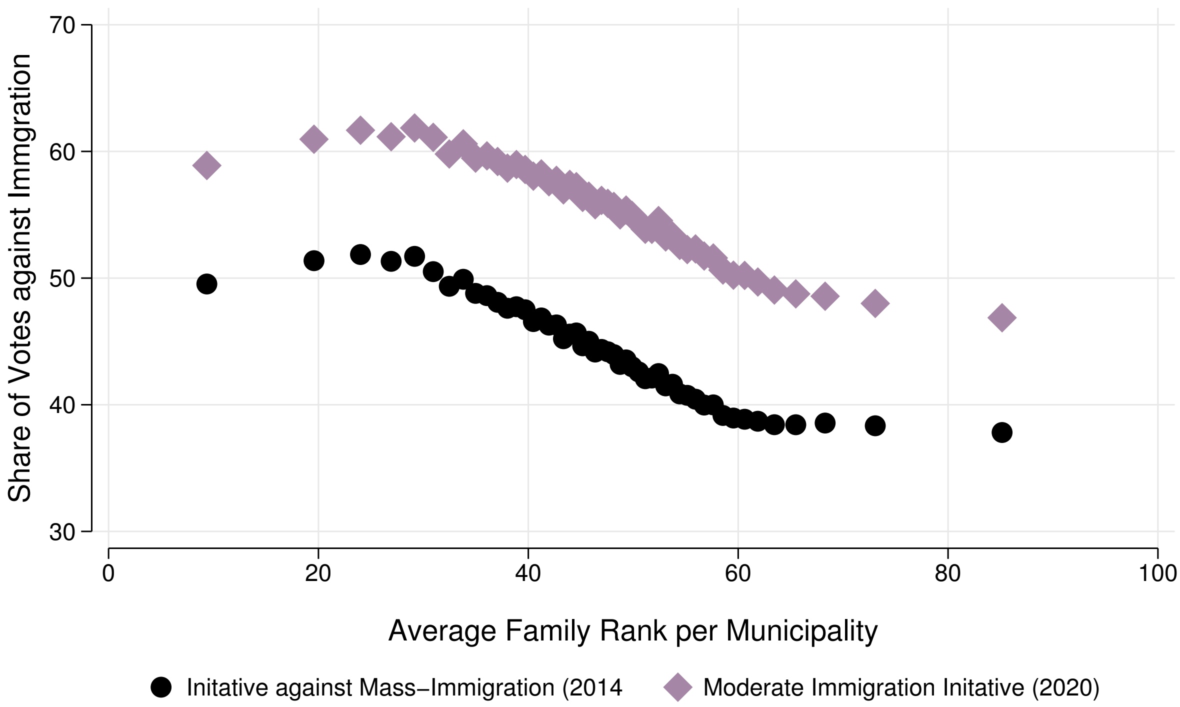
TABLE 6.6: Difference-in-Difference Estimates on Party Votes

	(1)	(2)	(3)
Dependent Variable:	Share Populist Right	Share Center	Share Left
Treatment Post=1	0.024*	-0.024***	-0.001
	(0.01)	(0.01)	(0.01)
Treatment Transition=1	0.037***	-0.013***	-0.001
	(0.01)	(0.00)	(0.01)
Region FE	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes
N (Years x Individuals)	1.45e+07	1.45e+07	1.45e+07
R2	0.367	0.364	0.217

Notes: This table shows the difference-in-difference treatment effects according to Equation6.1, but using support for political parties in the quadrennial national elections as an outcome variable. Column (1) shows that the support for populist right parties increases by 2.4 pp—likely at the costs of the center parties shown in column (2). Column (3) shows that support for left parties does not change in light of cross-border immigration influx.
Robust standard errors clustered on regional (NUTS-2) level

(*p<0.10; **p<0.05; ***p<0.01)

FIGURE 6.11: Anti Immigration Votes and Parental Background

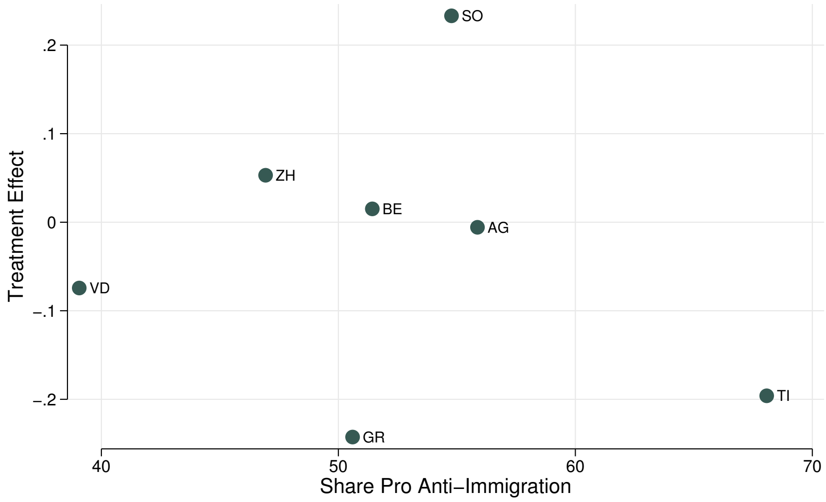


Notes: This figure shows how parental background and share of votes in favor of anti-migration policies correlate. The black dots depict votes for the initiative against mass immigration in 2014. The purple dots show the votes in favor of the moderate immigration initiative in 2020. In general: The lower the average family rank in the municipality, the higher the support for anti-migration policies.

Table E.2 shows the results of the high-intensity treatment. Here, municipalities closer than 15 minutes to the next border passing are included in the treatment groups. This group shows the largest increase in cross-border workers. This is also reflected in the estimates, as the coefficients are even more negative. The treatment effect, however, shows the same pattern—with more negative labor income effects for children from poorer parents.

Table E.3 shows the results with municipalities closer than 15 minutes in the treatment group, but the control group are municipalities that are also administratively in the border region, but further away than 30 minutes (colored brightest green in Figure 6.1). While this control group could still be affected by the treatment—because legally cross-border workers were also allowed to work in those municipalities—, the proximity to the treatment region could increase comparability between those regions in potentially unobserved characteristics. Also, even though they are legally allowed to work in those municipalities, there are in fact little cross-border workers working there. The municipalities between 15 and 30 minutes is excluded from the analysis and serves as spillover

FIGURE 6.12: Effect Heterogeneity and Voting in Favor of Immigration Restriction



Notes: This figure shows how the treatment effect varies with the share of the votes in favor to restrict immigration («mass immigration initiative in 2014»). The unit of measurement here are cantons for which there is common support of control and treatment groups.

region. The results show that the point estimates are qualitatively similar to the main specification. However, the results are not significantly different from zero. This is, however, likely because of the lost power as the sample size is less than half as large as in the main specification.

In Table E.4 the treatment group includes all municipalities legally in the border-region. The control group only comprises municipality in the non-border region (in Figure 6.1 this represents green vs. purple). Again, the results show the same pattern as in the main specification.

6.8.2 Geographic Mobility

A main concern in the when estimating the effect of an immigration influx to natives labor market outcomes is that natives might respond by moving into a region less affected by the immigration. Thus, in this case this might be true if children in the treatment region moved to the other region. This could then dilute the effect. In terms of parental income distribution, if children from richer parents show a higher geographic mobility, this could explain why their wages are less affected by the immigrant influx. Therefore, I check how mobility differs by parental background and regions and, especially, whether there are some changes around the onset of the policy.

Figure E.5 shows the fraction of children born in the border or non-border region that live in the region in the year 2010 for both regions and the top and bottom quintile of the parental income distribution. Children born in the border-region are less likely to move to the non-border region than children born in the non-border region. For both regions, children from a richer parental background show a higher mobility.

The graph does, however, not support the hypothesis that children from the border region move to the non-border region to eschew the increase in labor supply. Rather, it is the opposite way. Also, there are no abrupt changes in mobility around the implementation of the policy for the border group. There is, however, a slight drop in mobility for children from poorer parental backgrounds during the transition phase—which could potentially be a response to the immigration influx. This could potentially attenuate the estimated wage effect slightly. The estimated treatment effect might therefore constitute a lower bound in absolute effect size.

6.8.3 Definition of Child's Location

In the baseline specification, child location is defined as the municipality of the mother in 2010. This is most likely the municipality where the child grew up. Unfortunately, data does not provide yearly panel information on place of residence. The data provides the place of birth and, as of 2010, the municipality of residence, and the arrival date in this municipality.

Table E.5 shows the wage effect of the policy when using place of birth. The coefficients point qualitatively in the same direction, but are a little more pronounced. This could point to the fact that there was some geographic mobility response in light of the policy, and thus, that the effect of the main specification indeed consists a lower bound.

Table E.6 uses the municipality a person is living in 2010. Here, the results are very similar to the main specification. This could be because roughly two-thirds of the children live within a 15 minutes car drive of their mother.

6.8.4 Age of Child's Income

I measure the income of children between the age 30 to 33 for our baseline estimates. This is essentially a trade-off between sample size and a sufficient approximation of life time labor income.

To test whether later income measurements yield different results. Table E.7 and Table E.8 show this for ages 35 to 40 and for ages 30 to 38. The estimates become a little less pronounced, but show the same pattern.

6.8.5 Placebo Tests

I conduct two placebo test to make sure that the effects are not solely driven by the treatment design. In the first placebo test, I use only observations in the pre-treatment period. Thus I shift the policy cut-off to the year 1995. Table E.10 shows that there is no significant effect in this placebo specification. Next, I use a false treatment group. Therefore, I randomly assign half of the children in the non-border region to a placebo treatment group, and the other half to a placebo control group. Again, as Table E.11 there are no meaningful significant effects in this placebo specification.

6.9 Summary and Concluding Remarks

This study analyzes the effect of cross-border immigration on intergenerational income mobility on natives. It exploits the incremental removal of restrictions for cross-border workers in regions close to the border of Switzerland. The study draws from several administrative data sets.

The results show that the increase in cross-border workers decreases intergenerational income mobility for natives. Labor income of children from poorer parents is more negatively affected than income of children of parents from the top of the income distribution. The reason for this decrease is that children from parents with less income learn occupations and choose educational tracks that are more negatively affected by the immigrant influx. Further, they do not adapt their educational or learned occupation to avert the increased competition.

One limitation of this study is that it only analyzes the intergenerational income mobility of natives. Thus, I do not consider the potential gains in income and intergenerational mobility that accrue for the cross-border workers. The study also only uses total labor income. Thus, I cannot how much of the decrease in labor income is because of fewer hours worked and how much because of lower or stagnating real wages.

The results line up with other studies finding negative effects on labor outcomes of natives considering immigrant influx, such as Dustmann et al. (2017) or Borjas et al. (1997). At first sight, the results might be at odds with the study by Beerli et al. (2020) who also analyze the same policy change in Switzerland—-albeit with different data. They find no significantly negative effect on wages of lower educated and a positive effect on wages of highly educated individuals. A closer look, however, reveals that the coefficient of wages of less educated is also negative, although insignificantly different from zero. Furthermore, I do also find a marginally positive effect for children from the highest parental background—which are also likely to be highly educated because education and parental incomes are highly correlated in Switzerland (Chuard & Grassi, 2020). [8]

[8] The remaining differences might be explained by different data sources and data selection. I use the social security earnings register in order to link parents and children and my data selection is slightly different. My sample does not include foreign-born immigrants already in Switzerland. Furthermore, I use total labor income and not wage and employment information.

These results have important policy implications. It shows that the labor income of children from poorer parental background is most susceptible to a labor supply shock by immigrants. Policy makers should take those negative effects into account and potentially take measures against this. Such measures could range from compensatory measures to nudging into occupational choices that are supposedly less affected by the increase in labor supply. Indicative evidence shows that lower equality of opportunity could indeed lead to sentiments against immigrants and thereby decrease political support for «open border policies» whose net effect on the economy is likely positive.

Bibliography

(2019). *Alkoholkonsum 2017*. 10887924. Neuchâtel: Bundesamt für Statistik (BFS).

Abramitzky, R., Boustan, L. P., Jácome, E., & Pérez, S. (2019). Intergenerational mobility of immigrants in the us over two centuries.

Acciari, P., Polo, A., & Violante, G. L. (2019). "and yet it moves": Intergenerational mobility in italy. Working Paper 25732, National Bureau of Economic Research.

Adermon, A., Lindahl, M., & Palme, M. (2019). Dynastic human capital, inequality and intergenerational mobility.

Ahrens, A., Hansen, C. B., & Schaffer, M. (2019). Lassopack: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation.

Aizer, A. (2011). Poverty, violence, and health: The impact of domestic violence during pregnancy on newborn health. *Journal of Human Resources*, *46*(3), 518–538.

Almond, D., Currie, J., & Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, *56*(4), 1360–1446.

Almond, D., Hoynes, H. W., & Schanzenbach, D. W. (2011). Inside the war on poverty: The impact of food stamps on birth outcomes. *The Review of Economics and Statistics*, *93*(2), 387–403.

Altonji, J. G., & Card, D. (1991). *The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives*. University of Chicago Press.

Antecol, H., & Bedard, K. (2006). Unhealthy assimilation: Why do immigrants converge to American health status levels? *Demography*, *43*(2), 337–360.

Athey, S., & Imbens, G. W. (2018). Design-based analysis in difference-in-differences settings with staggered adoption. Tech. rep., National Bureau of Economic Research.

Barone, C., Schizzerotto, A., Abbiati, G., & Argentin, G. (2017). Information barriers, social inequality, and plans for higher education: evidence from a field experiment. *European Sociological Review*, *33*(1), 84–96.

Bauer, P. (2006). The intergenerational transmission of income in switzerland a comparison between natives and immigrants. Wwz discussion paper 06/01.

Bauer, P., & Riphahn, R. T. (2007). Heterogeneity in the intergenerational transmission of educational attainment: evidence from switzerland on natives and second-generation immigrants. *Journal of Population Economics*, *20*(1), 121–148.

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, *81*(2, Part 2), S279–S288.

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of political Economy*, *87*(6), 1153–1189.

Becker, G. S., & Tomes, N. (1986a). Human capital and the rise and fall of families. *Journal of Labor Economics*, *4*(3), S1–S39.

Becker, G. S., & Tomes, N. (1986b). Human capital and the rise and fall of families. *Journal of Labor Economics*, *4*(3), S1–S39.
URL http://www.jstor.org/stable/2534952

Beerli, A., Ruffner, J., Siegenthaler, M., & Peri, G. (2020). The abolition of immigration restrictions and the performance of firms and workers: Evidence from switzerland. (25302).

Behncke, S. (2012). Does retirement trigger ill health? *Health economics*, *21*(3), 282–300.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2019). Who becomes an inventor in America? The importance of exposure to innovation. *Quarterly Journal of Economics*, *134*(2), 647–713.

Bettinger, E., Gurantz, O., Kawano, L., Sacerdote, B., & Stevens, M. (2019). The long-run impacts of financial aid: Evidence from california's cal grant. *American Economic Journal: Economic Policy*, *11*(1), 64–94.

Bialik, K., & Fry, R. (2019). Millennial life: How young adulthood today compares with prior generations. *Pew Research Center*, *14*.

Biddle, N., Kennedy, S., & McDonald, J. T. (2007). Health assimilation patterns amongst Australian immigrants. *Economic Record*, *83*(260), 16–30.

Black, S. E., Denning, J. T., Dettling, L. J., Goodman, S., & Turner, L. J. (2020). Taking it to the limit: Effects of increased student loan availability on attainment, earnings, and financial well-being. Tech. rep., National Bureau of Economic Research.

Black, S. E., & Devereux, P. J. (2010). Recent developments in intergenerational mobility. Tech. rep., National Bureau of Economic Research.

Black, S. E., & Devereux, P. J. (2011). Recent developments in intergenerational mobility. vol. 4 of *Handbook of Labor Economics*, chap. 16, (pp. 1487–1541). Elsevier.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2016). Does grief transfer across generations? Bereavements during pregnancy and child outcomes. *American Economic Journal: Applied Economics*, *8*(1), 193–223.

Blake, H., & Garrouste, C. (2013). Killing me softly: Work and mortality among french seniors, health, econometrics and data group (hedg) working papers 13/25, university of york.

Bloemen, H., Hochguertel, S., & Zweerink, J. (2017). The causal effect of retirement on mortality: Evidence from targeted incentives to retire early. *Health economics*, *26*(12), e204–e218.

Borjas, G. J. (1993). The intergenerational mobility of immigrants. *Journal of Labor Economics*, *11*(1, Part 1), 113–135.

Borjas, G. J., Freeman, R. B., Katz, L. F., DiNardo, J., & Abowd, J. M. (1997). How much do immigration and trade affect labor market outcomes? *Brookings papers on economic activity*, *1997*(1), 1–90.

Borra, C., González, L., & Sevilla, A. (2019). The impact of scheduling birth early on infant health. *Journal of the European Economic Association*, *17*(1), 30–78.

Bozio, A., Garrouste, C., & Perdrix, E. (2020). Impact of later retirement on mortality: Evidence from france.

Bratberg, E., Davis, J., Mazumder, B., Nybom, M., Schnitzlein, D. D., & Vaage, K. (2017). A comparison of intergenerational mobility curves in germany, norway, sweden, and the us. *The Scandinavian Journal of Economics*, *119*(1), 72–101.

Braun, S. T., & Stuhler, J. (2018). The transmission of inequality across multiple generations: testing recent theories with evidence from germany. *The Economic Journal*, *128*(609), 576–611.

Brown, M., Karl Scholz, J., & Seshadri, A. (2012). A new test of borrowing constraints for education. *The Review of Economic Studies*, *79*(2), 511–538.

Brunner, B., & Kuhn, A. (2014). Announcement effects of health policy reforms: Evidence from the abolition of Austria's baby bonus. *European Journal of Health Economics*, *15*(4), 373–388.

Bundesamt für Sozialversicherungen (1969–1992). Arten und Ansätze der Familienzulagen. Zeitschrift für die Ausgleichskassen (ZAK).

Bundesamt für Sozialversicherungen (1993–2004). Fz: Arten und Ansätze der Familienzulagen. AHI Praxis.

Bundesamt für Sozialversicherungen (2005–2020). Arten und Ansätze der Familienzulagen. https://www.bsv.admin.ch/bsv/de/home/sozialversicherungen/l.

Callaway, B., & Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326.

Camacho, A. (2008). Stress and birth weight: Evidence from terrorist attacks. *American Economic Review*, *98*(2), 511–15.

Card, D. (2009). Immigration and inequality. *American Economic Review*, *99*(2), 1–21.

Card, D., & Solis, A. (2020). Measuring the effect of student loans on college persistence. Tech. rep., National Bureau of Economic Research.

Carneiro, P., & Heckman, J. J. (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, *112*(482), 705–734.

Carneiro, P., Løken, K. V., & Salvanes, K. G. (2015). A flying start? Maternity leave benefits and long-run outcomes of children. *Journal of Political Economy*, *123*(2), 365–412.

Case, A., & Deaton, A. (2015). Rising morbidity and mortality in midlife among white non-hispanic americans in the 21st century. *Proceedings of the National Academy of Sciences*, *112*(49), 15078–15083.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the united states. *The Quarterly Journal of Economics*, *135*(3), 1567–1633.

Chetty, R., Hendre, N., Kline, P., & Saez, E. (2014a). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, *129*(4), 1553–1623.

Chetty, R., Hendren, N., Kline, P., Saez, E., & Turner, N. (2014b). Is the united states still a land of opportunity? recent trends in intergenerational mobility. *The American Economic Review*, *104*(5), 141–147.

Chiswick, B. R., Lee, Y. L., & Miller, P. W. (2008). Immigrant selection systems and immigrant health. *Contemporary Economic Policy*, *26*(4), 555–578.

Chu, Y.-W. L., & Cuffe, H. E. (2020). Do struggling students benefit from continued student loan access? evidence from university and beyond. *Evidence From University and Beyond (March 31, 2020)*.

Chuard, P., & Grassi, V. (2020). Switzer-land of opportunity: Intergenerational income mobility in the land of vocational education. *Available at SSRN 3662560*.

Clark, A. E. (2014). Son of my father? the life-cycle analysis of well-being: Introduction. *The Economic Journal*, *124*(580), F684–F687.

Clark, G., & Cummins, N. (2015). Intergenerational wealth mobility in england, 1858–2012: surnames and social mobility. *The Economic Journal*, *125*(582), 61–85.

Cohen, A., Dehejia, R., & Romanov, D. (2013). Financial incentives and fertility. *Review of Economics and Statistics*, *95*(1), 1–20.

Colagrossi, M., d'Hombres, B., & Schnepf, S. V. (2020). Like (grand) parent, like child? multigenerational mobility across the eu. *European Economic Review*, *130*, 103600.

Connolly, M., Corak, M., & Haeck, C. (2019). Intergenerational mobility between and within canada and the united states. *Journal of Labor Economics*, *37*(S2), S595–S641.

Constant, A. F., García-Muñoz, T., Neuman, S., & Neuman, T. (2018). A "healthy immigrant effect" or a "sick immigrant effect"? Selection and policies matter. *The European Journal of Health Economics*, *19*(1), 103–121.

Coomarasamy, A., Knox, E. M., Gee, H., Song, F., & Khan, K. S. (2003). Effectiveness of nifedipine versus atosiban for tocolysis in preterm labour: A meta-analysis with an indirect comparison of randomised trials. *BJOG: An International Journal of Obstetrics & Gynaecology*, *110*(12), 1045–1049.

Corak, M. (2013). Inequality from generation to generation: The united states in comparison. *The economics of inequality, poverty, and discrimination in the 21st century*, *1*, 107–126.

Corak, M. (2020a). The canadian geography of intergenerational income mobility. *The Economic Journal*, *130*(631), 2134–2174.

Corak, M. (2020b). The canadian geography of intergenerational income mobility. *The Economic Journal*, *130*(631), 2134–2174.

Corak, M., & Heisz, A. (1999). The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data. *Journal of Human Resources*, (pp. 504–533).

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.

Currie, J., & Rossin-Slater, M. (2013). Weathering the storm: Hurricanes and birth outcomes. *Journal of Health Economics*, *32*(3), 487–503.

Dahl, M. W., & DeLeire, T. (2008). The association between children's earnings and fathers' lifetime earnings: Estimates using administrative data.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–96.

Denning, J. T., & Jones, T. R. (2019). Maxed out? the effect of larger student loan limits on borrowing and education outcomes. *Journal of Human Resources*, (pp. 0419–10167R1).

Deutscher, N., & Mazumder, B. (2020). Intergenerational mobility across australia and the stability of regional estimates. *Labour Economics*, *66*, 101861.

Dickert-Conlin, S., & Chandra, A. (1999). Taxes and the timing of births. *Journal of Political Economy*, *107*(1), 161–177.

Dustmann, C., Frattini, T., & Preston, I. P. (2013). The effect of immigration along the distribution of wages. *Review of Economic Studies*, *80*(1), 145–173.

Dustmann, C., Schönberg, U., & Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, *132*(1), 435–483.

Educa (2021). Upper secondary level. *Educa*.
URL `swisseducation.educa.ch`

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

Eriksen, J. (2018). Finding the land of opportunity intergenerational mobility in denmark. Tech. rep., Mimeo.

Eugster, B., Lalive, R., Steinhauer, A., & Zweimüller, J. (2011). The demand for social insurance: does culture matter? *The Economic Journal*, *121*(556), F413–F448.

Eugster, B., Lalive, R., Steinhauer, A., & Zweimüller, J. (2017). Culture, work attitudes, and job search: Evidence from the swiss language border. *Journal of the European Economic Association*, *15*(5), 1056–1100.

Favre, G., Floris, J., & Woitek, U. (2018). Intergenerational Mobility in the 19th Century: Micro-Level Evidence from the City of Zurich Intergenerational Mobility in the 19 th Century Micro-Level Evidence from the City of Zurich.

Fé, E., & Hollingsworth, B. P. (2016). Short-and long-run estimates of the local effects of retirement on health. *Journal of the Royal Statistical Society: Series A Statistics in Society*, *179*(4), 1051–1067.

Ferguson, J. L., & Ready, D. D. (2011). Expanding notions of social reproduction: Grandparents' educational attainment and grandchildren's cognitive skills. *Early Childhood Research Quarterly*, *26*(2), 216–226.

Fitzpatrick, M. D., & Moore, T. J. (2018). The mortality effects of retirement: Evidence from Social Security eligibility at age 62. *Journal of Public Economics*, *157*, 121–137.

Foged, M., & Peri, G. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, *8*(2), 1–34.

Fontana-Casellini, L. (2020). The effect of municipal budget balanace rules: the swiss case.

FOS (2019). Grenzgängerstatistik. *Bundesamt für Statistik*.
URL `https://www.bfs.admin.ch/bfs/de/home/statistiken/`

Gans, J. S., & Leigh, A. (2009). Born on the first of July: An (un)natural experiment in birth timing. *Journal of Public Economics*, *93*(1-2), 246–263.

Güell, M., Pellizzari, M., Pica, G., & Rodríguez Mora, J. V. (2018). Correlating social mobility and economic outcomes. *The Economic Journal*, *128*(612), F353–F403.

Gelman, A., & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, *37*(3), 447–456.

González, L. (2013). The effect of a universal child benefit on conceptions, abortions, and early maternal labor supply. *American Economic Journal: Economic Policy*, *5*(3), 160–88.

González, L., & Trommlerová, S. K. (2021). Cash transfers and fertility: How the introduction and cancellation of a child benefit affected births and abortions. *Journal of Human Resources*, (pp. 0220–10725R2).

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*. Forthcoming.

Hagen, J. (2018). The effects of increasing the normal retirement age on health care utilization and mortality. *Journal of Population Economics*, *31*(1), 193–234.

Hallberg, D., Johansson, P., & Josephson, M. (2015). Is an early retirement offer good for your health? Quasi-experimental evidence from the army. *Journal of Health Economics*, *44*, 274–285.

Halonen, J. I., Stenholm, S., Pulakka, A., Kawachi, I., Aalto, V., Pentti, J., Lallukka, T., Virtanen, M., Vahtera, J., & Kivimäki, M. (2017). Trajectories of risky drinking around the time of statutory retirement: a longitudinal latent class analysis. *Addiction*, *112*(7), 1163–1170.

Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility. *Annu. Rev. Econ.*, *6*(1), 689–733.

Heidrich, S. (2017). Intergenerational mobility in sweden: a regional perspective. *Journal of Population Economics*, *30*(4), 1241–1280.

Heller-Sahlgren, G. (2017). Retirement blues. *Journal of Health Economics*, *54*, 66–78.

Helsø, A.-L. (2021). Intergenerational income mobility in denmark and the united states*. *The Scandinavian Journal of Economics*, *123*(2), 508–531.

Hernaes, E., Markussen, S., Piggott, J., & Vestad, O. L. (2013). Does retirement age impact mortality? *Journal of Health Economics*, *32*(3), 586–598.

Hertz, T. (2006). Understanding mobility in America. *Center for American Progress Discussion Paper*.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., & Verashchagina, A. (2008a). The inheritance of educational inequality: International comparisons and fifty-year trends. *The BE Journal of Economic Analysis & Policy*, *7*(2).

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., & Verashchagina, A. (2008b). The inheritance of educational inequality: International comparisons and fifty-year trends. *The BE Journal of Economic Analysis & Policy*, *7*(2).

Häner, M., & Schaltegger, C. A. (2020). The name says it all. multigenerational social mobility in switzerland, 1550-2019. *IFF-HSG Working Papers*.

Hoynes, H., Miller, D., & Simon, D. (2015). Income, the earned income tax credit, and infant health. *American Economic Journal: Economic Policy*, *7*(1), 172–211.

Ichino, A., Karabarbounis, L., & Moretti, E. (2011). The political economy of intergenerational income mobility. *Economic Inquiry*, *49*(1), 47–69.

Imbens, G., & Lemieux, T. (2008). The regression discontinuity design—theory and applications. *Journal of Econometrics*.

Insler, M. (2014). The health consequences of retirement. *Journal of Human Resources*, *49*(1), 195–233.

Johnston, D. W., & Lee, W.-S. (2009). Retiring to the good life? the short-term effects of retirement on health. *Economics Letters*, *103*(1), 8–11.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kearney, M. S. (2004). Is there an effect of incremental welfare benefits on fertility behavior? A look at the family cap. *The Journal of Human Resources*, *39*(2), 295.

Kuhn, A. (2018). The complex effects of retirement on health. *IZA World of Labor*.

Kuhn, A., Staubli, S., Wuellrich, J.-P., & Zweimüller, J. (2020). Fatal attraction? Extended unemployment benefits, labor force exits, and mortality. *Journal of Public Economics*, (p. 104087).

Kühntopf, S., & Tivig, T. (2012). Early retirement and mortality in germany. *European journal of epidemiology*, *27*(2), 85–89.

Lalive, R., & Staubli, S. (2015). How does raising women's full retirement age affect labor supply, income, and mortality. *NBER working paper*, *18660*.

LaLumia, S., Sallee, J. M., & Turner, N. (2015). New evidence on taxes and the timing of birth. *American Economic Journal: Economic Policy*, *7*(2), 258–293.

Laroque, G., & Salanié, B. (2014). Identifiyng the response fertility to financial incentives. *Journal of Applied Econometrics*, *29*(2), 314–332.

Lee, C. (2014). Intergenerational health consequences of in utero exposure to maternal stress: Evidence from the 1980 Kwangju uprising. *Social Science & Medicine*, *119*, 284–291.

Lee, C.-I., & Solon, G. (2009). Trends in intergenerational income mobility. *The Review of Economics and Statistics*, *91*(4), 766–772.

Lima, S. A. M., El Dib, R. P., Rodrigues, M. R. K., Ferraz, G. A. R., Molina, A. C., Neto, C. A. P., De Lima, M. A. F., & Rudge, M. V. C. (2018). Is the risk of low birth weight or preterm labor greater when maternal stress is experienced during pregnancy? A systematic review and meta-analysis of cohort studies. *PloS one*, *13*(7), e0200594.

Lindahl, M., Palme, M., Massih, S. S., & Sjögren, A. (2015). Long-term intergenerational persistence of human capital an empirical analysis of four generations. *Journal of Human Resources*, *50*(1), 1–33.

Long, J., & Ferrie, J. (2018). Grandfathers matter (ed): occupational mobility across three generations in the us and britain, 1850–1911. *The Economic Journal*, *128*(612), F422–F445.

Loury, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica: Journal of the Econometric Society*, (pp. 843–867).

Mare, R. D. (2011). A multigenerational view of inequality. *Demography*, *48*(1), 1–23.

Martínez, I. Z. (2020). *Evidence from Unique Swiss Tax Data on the Composition and Joint Distribution of Income and Wealth*. University of Chicago Press.

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the united states using social security earnings data. *87(2)*, 235–255.

Mazzonna, F., & Peracchi, F. (2017). Unhealthy retirement? *Journal of Human Resources*, *52*(1), 128–151.

McDonald, J. T., & Kennedy, S. (2005). Is migration to Canada associated with unhealthy weight gain? Overweight and obesity among Canada's immigrants. *Social science & medicine*, *61*(12), 2469–2481.

Milligan, K. (2005). Subsidizing the stork: New evidence on tax incentives and fertility. *Review of Economics and Statistics*, *87*(3), 539–555.

Müller, T., & Shaikh, M. (2018). Your retirement and my health behavior: Evidence on retirement externalities from a fuzzy regression discontinuity design. *Journal of health economics*, *57*, 45–59.

Neugart, M., & Ohlsson, H. (2013). Economic incentives and the timing of births: Evidence from the German parental benefit reform of 2007. *Journal of Population Economics*, *26*(1), 87–108.

Nielsen, N. F. (2019). Sick of retirement? *Journal of Health Economics*, *65*, 133–152.

Nybom, M., & Stuhler, J. (2016). Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation. *Journal of Human Resources*, *51*(1), 239–268.

Ottaviano, G. I., & Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, *10*(1), 152–197.

Parchet, R. (2019). Are local tax rates strategic complements or strategic substitutes? *American Economic Journal: Economic Policy*, *11*(2), 189–224.

Parenti, A., & Tealdi, C. (2019). Does the implementation of the schengen agreement boost cross-border commuting? evidence from switzerland.

Quintana-Domeque, C., & Ródenas-Serrano, P. (2017). The hidden costs of terrorism: The effects on health at birth. *Journal of Health Economics*, *56*, 47–60.

Richard, A., Rohrmann, S., Vandeleur, C. L., Schmid, M., Barth, J., & Eichholzer, M. (2017). Loneliness is adversely associated with physical and mental health and lifestyle factors: Results from a swiss national survey. *PLOS ONE*, *12*(7), 1–18.

Rodrik, D., & Stantcheva, S. (2021). A policy matrix for inclusive prosperity. Tech. rep., National Bureau of Economic Research.

Rose, L. (2020). Retirement and health: Evidence from england. *Journal of Health Economics*, *73*, 102352.

Rostron, B. (2012). Smoking-attributable mortality by cause in the united states: revising the cdc's data and estimates. *Nicotine & Tobacco Research*, *15*(1), 238–246.

Sanders, N. J., & Stoecker, C. (2015). Where have all the young men gone? Using sex ratios to measure fetal death rates. *Journal of Health Economics*, *41*, 30–45.

Scalone, F., & Rettaroli, R. (2015). Exploring the variations of the sex ratio at birth from an historical perspective. *Statistica*, *75*(2), 213–226.

Schulkind, L., & Shapiro, T. M. (2014). What a difference a day makes: Quantifying the effects of birth timing manipulation on infant health. *Journal of Health Economics*, *33*(1), 139–158.

Schwandt, H. (2018). The lasting legacy of seasonal influenza: In-utero exposure and labor market outcomes. *CEPR Discussion Paper*.

Shapiro, G. D., Fraser, W. D., Frasch, M. G., & Séguin, J. R. (2013). Psychosocial stress in pregnancy and preterm birth: Associations and mechanisms. *Journal of Perinatal Medicine*, *41*(6), 631–645.

Shield, K., Manthey, J., Rylett, M., Probst, C., Wettlaufer, A., Parry, C. D., & Rehm, J. (2020). National, regional, and global burdens of disease from 2000 to 2016 attributable to alcohol use: a comparative risk assessment study. *The Lancet Public Health*, *5*(1), e51–e61.

Solon, G. (1992). Intergenerational income mobility in the united states. *The American Economic Review*, *82*(3), 393–408.

Solon, G. (1999). Intergenerational mobility in the labor market. In *Handbook of labor economics*, vol. 3, (pp. 1761–1800). Elsevier.

Solon, G. (2018). What do we know so far about multigenerational mobility? *The Economic Journal*, *128*(612), F340–F352.

Stuhler, J. (2012). Mobility across multiple generations: The iterated regression fallacy.

Stuhler, J., et al. (2018). A review of intergenerational mobility and its drivers. *JRC Working Papers*, (JRC112247).

Tamm, M. (2013). The impact of a large parental leave benefit reform on the timing of birth around the day of implementation. *Oxford Bulletin of Economics and Statistics*, *75*(4), 585–601.

The World Bank (2020). Per Capita GDP (current USD).
URL `https://data.worldbank.org/indicator/NY.GDP.PCAP.CD`

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Wang, X., Steier, J. B., & Gallo, W. T. (2014). The effect of retirement on alcohol consumption: results from the us health and retirement study. *The European Journal of Public Health*, *24*(3), 485–489.

WorldBank (2021). Gender statistics.
URL `https://databank.worldbank.org/`

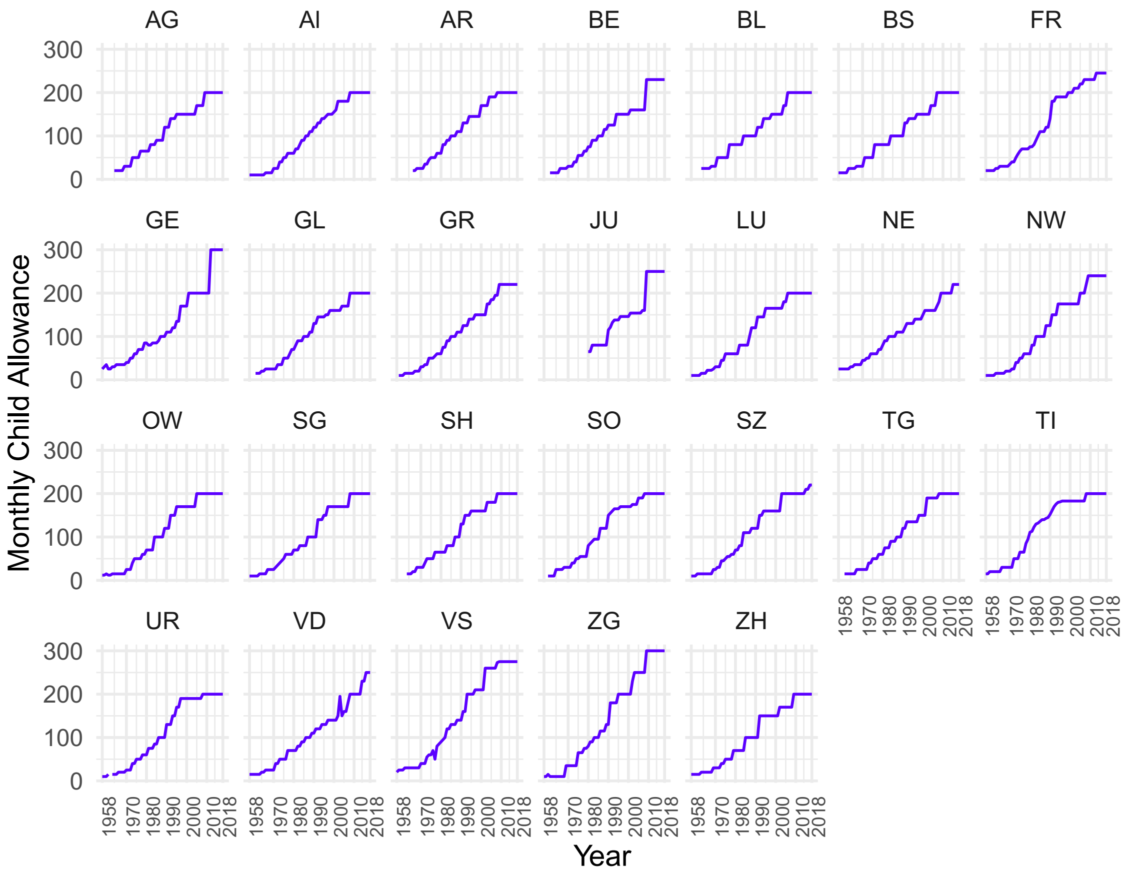
Zantinge, E. M., van den Berg, M., Smit, H. A., & Picavet, H. S. J. (2014). Retirement and a healthy lifestyle: opportunity or pitfall? a narrative review of the literature. *The European Journal of Public Health*, *24*(3), 433–439.

Zeng, Z., & Xie, Y. (2014). The effects of grandparents on children's schooling: Evidence from rural china. *Demography*, *51*(2), 599–617.

Zins, M., Guéguen, A., Kivimaki, M., Singh-Manoux, A., Leclerc, A., Vahtera, J., Westerlund, H., Ferrie, J. E., & Goldberg, M. (2011). Effect of retirement on alcohol consumption: longitudinal evidence from the french gazel cohort study. *PLoS One*, *6*(10).

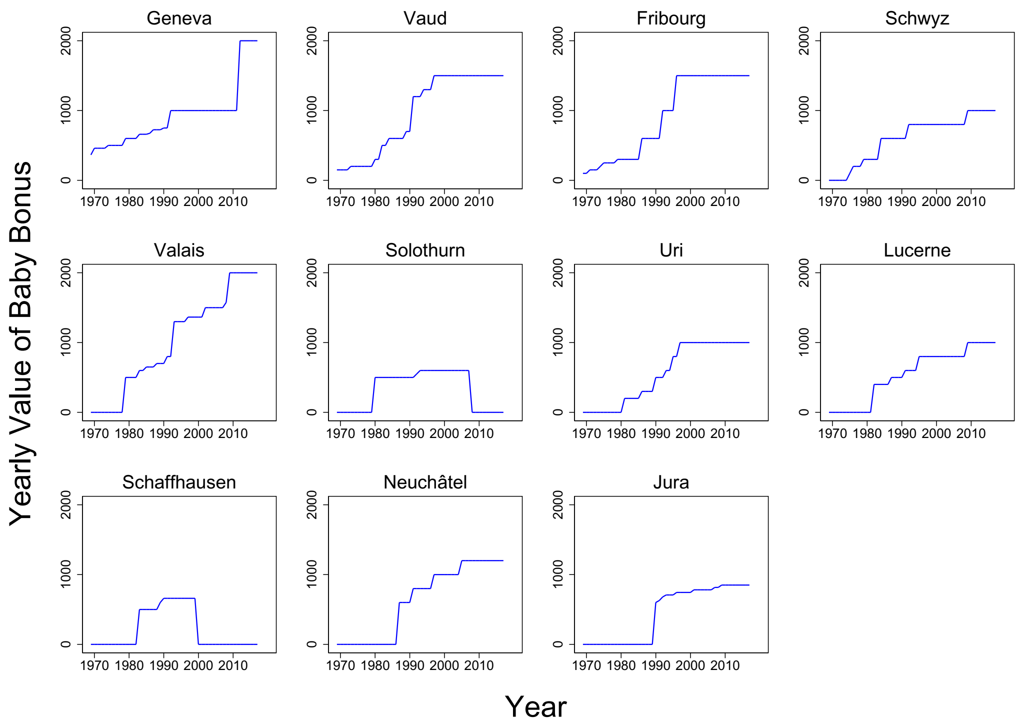
Appendices

FIGURE A.1: Monthly Child Allowances per Canton



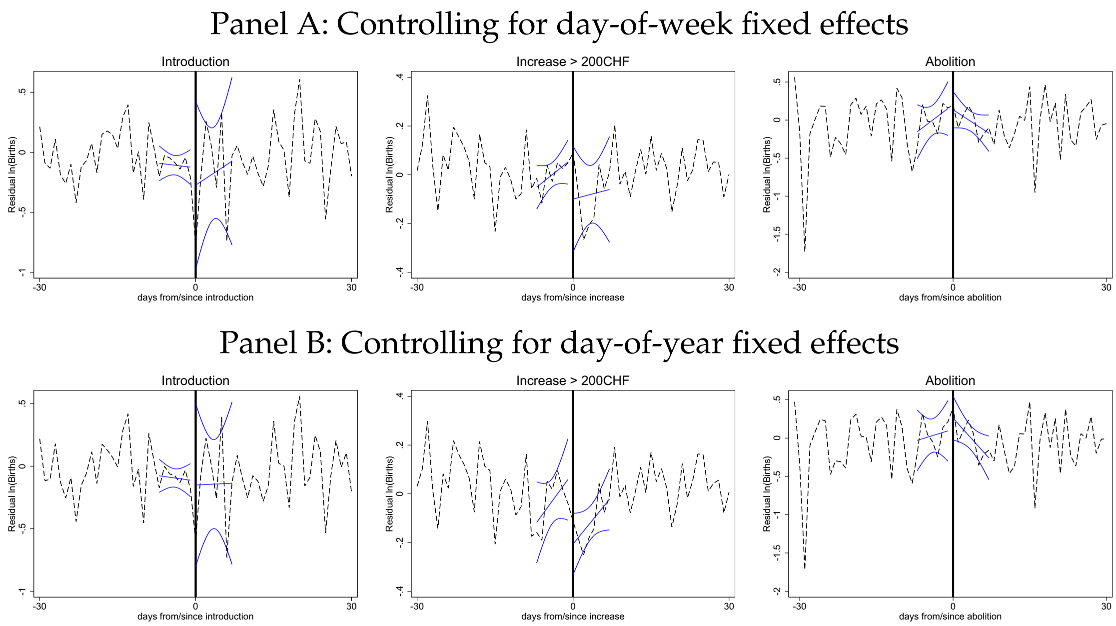
Notes: This figure shows the amount of child allowances provided per child per month per canton in current year values.

FIGURE A.2: Time Variation of Birth Allowances by Treated Cantons



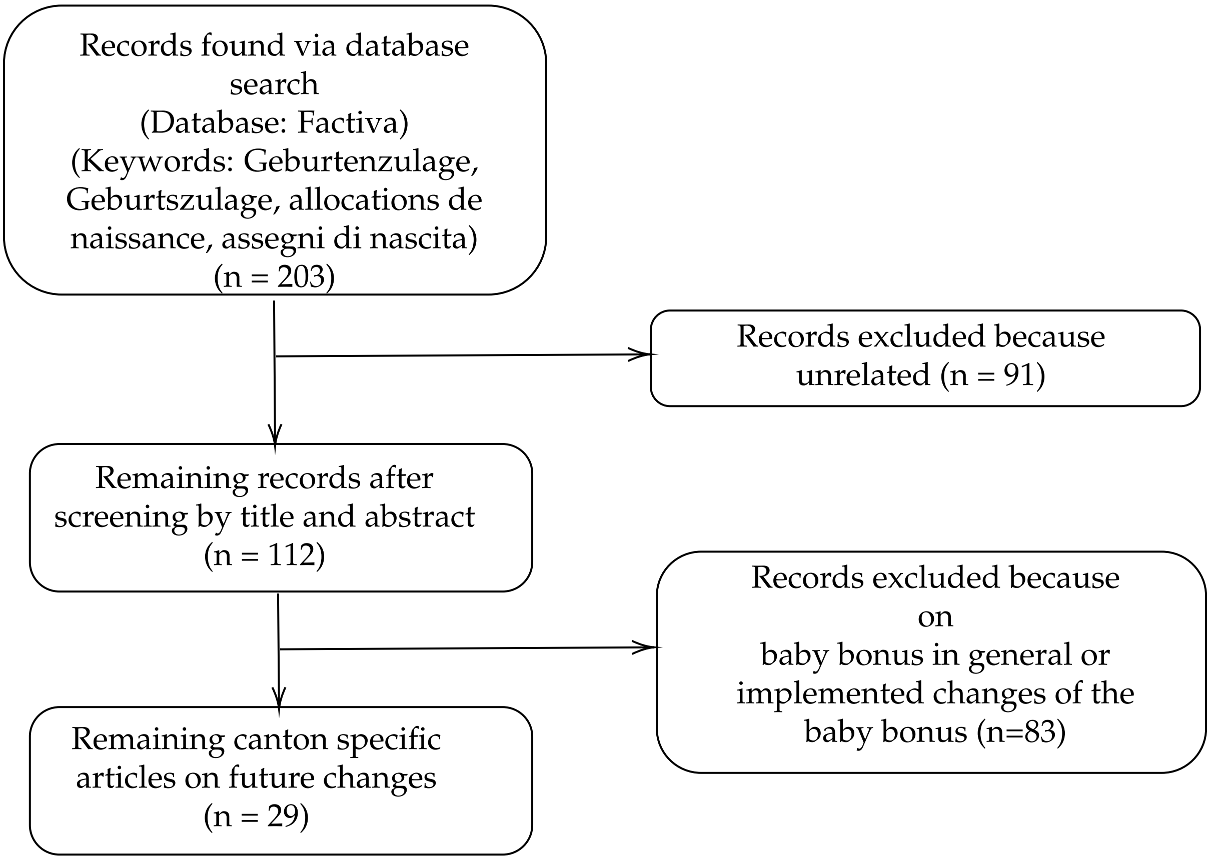
Notes: This figure shows the amount of birth allowances provided per child per canton in current year values. It only shows the movement over time for those cantons which ever introduced a baby bonus at one point in time. The ordering of the cantons is according to their introduction year of the birth allowance.

FIGURE A.3: Birth-Scheduling Log Specification Event Study



Notes: This figure shows the residuals (in dashed black line) from a linear prediction of estimating Equation (2.2) of the log of total births per day and a linear fit including a 95% confidence interval (in blue) for the week before and after the policy change. Panel A reports the residual when ζ_i controls for day-of-week fixed effects and Panel B when ζ_i controls for day-of-year fixed effects. The three event studies combine either all introductions, increases above CHF 200, or abolition across cantons and time.

FIGURE A.4: Media Search



Notes: This figure shows the results of the media search related to the keywords "Geburtenzulage", "Geburtszulage", "allocations de naissance", "assegni di nascita" on Factiva.

TABLE A.1: Overview Policy Changes in Detail

Canton	Announcement Date	Implementation Date
Geneva	14.03.1969	01.05.1969
Fribourg	15.12.1970	01.01.1971
Vaud	27.11.1972	01.01.1973
Geneva	12.06.1973	01.07.1973
Fribourg	24.09.1973	01.01.1974
Schwyz	09.05.1974	01.07.1974
Fribourg	29.10.1974	01.01.1975
Schwyz	05.12.1975	01.01.1976
Schywz	25.09.1977	01.01.1978
Valais	01.12.1977	01.01.1978
Solothurn	12.06.1978	01.01.1979
Geneva	12.10.1978	01.01.1979
Fribourg	10.10.1978	01.07.1979
Vaud	18.09.1979	01.01.1980
Uri	28.09.1980	01.01.1981
Lucerne	10.03.1980	01.07.1981
Vaud	13.11.1981	01.01.1982
Schaffhausen	24.06.1982	01.07.1982
Geneva	07.03.1982	01.07.1982
Valais	12.11.1982	01.01.1983
Schwyz	20.10.1983	01.01.1984
Vaud	12.12.1983	01.01.1984
Valais	16.11.1984	01.01.1985
Geneva	15.02.1985	01.04.1985
Fribourg	25.09.1985	01.01.1986
Uri	08.10.1985	01.01.1986
Geneva	25.06.1986	01.01.1987
Neuchatel	20.10.1986	01.01.1987
Lucerne	14.11.1986	01.01.1987
Valais	13.11.1987	01.07.1988
Schaffhausen	06.06.1988	01.07.1988
Vaud	09.11.1988	01.01.1989
Jura	24.02.1989	01.07.1989
Uri	08.06.1989	01.01.1990
Geneva	27.09.1989	01.01.1990
Schaffhausen	06.11.1989	01.01.1990
Valais	28.09.1990	01.01.1991
Vaud	30.11.1990	01.01.1991
Neuchatel	03.12.1990	01.01.1991
Jura	04.12.1990	01.01.1991
Geneva	12.12.1990	01.01.1991
Lucerne	18.12.1990	01.01.1991
Fribourg	18.02.1991	01.03.1991
Jura	16.04.1991	01.10.1991
Solothurn	15.10.1991	01.01.1992
Schwyz	08.12.1991	01.01.1992
Valais	06.04.1992	01.01.1993
Jura	20.09.1992	01.01.1993
Solothurn	12.11.1992	01.01.1993
Uri	08.12.1992	01.01.1993
Vaud	26.11.1993	01.01.1994
Lucerne	13.09.1994	01.01.1995
Uri	28.09.1994	01.01.1995
Fribourg	13.11.1995	01.01.1996
Jura	21.11.1995	01.01.1996
Valais	11.09.1996	01.01.1997
Vaud	24.09.1996	01.01.1997
Uri	13.11.1996	01.01.1997
Neuchatel	27.11.1996	01.01.1997
Schaffhausen	05.09.1999	01.01.2000
Jura	31.10.2000	01.01.2001
Valais	23.09.2001	01.01.2002
Neuchatel	01.12.2004	01.01.2005
Jura	26.11.2006	01.01.2007
Valais	31.10.2007	01.01.2008
Solothurn	16.11.2007	01.01.2008
Valais	11.09.2008	01.01.2009
Schwyz	28.09.2008	01.01.2009
Jura	25.11.2008	01.01.2009
Lucerne	28.11.2008	01.01.2009
Geneva	23.06.2011	01.01.2012

TABLE A.2: Event Study Estimates without Abolishing Cantons

	(1) Fertility Rate	(2) Crude Birth Rate	(3) Birth Interval	(4) Age Mother	(5) Birth Weight	(6) Share Low Birth Weight	(7) Death Rate	(8) Stillbirth Rate	(9) Sex Ratio
Event Time Estimate: δ_τ									
-5	-0.027	-0.265	-1.184	-0.085	46.348**	-0.008***	0.596	0.081	-0.047*
	(0.039)	(0.207)	(1.632)	(0.053)	(13.619)	(0.002)	(1.422)	(0.953)	(0.019)
-4	-0.014	-0.162	0.094	0.002	5.410	-0.004	1.013	-0.963*	-0.029
	(0.010)	(0.114)	(0.369)	(0.057)	(12.342)	(0.003)	(1.154)	(0.369)	(0.015)
-3	0.031	0.147	-0.371	-0.023	22.718**	-0.008***	-1.828	-0.452	-0.037*
	(0.045)	(0.276)	(0.192)	(0.038)	(7.095)	(0.002)	(1.141)	(1.201)	(0.015)
-2	0.029*	0.158	0.015	-0.022	7.594	-0.002	-0.345	0.309	-0.025
	(0.012)	(0.115)	(0.257)	(0.063)	(29.031)	(0.002)	(1.094)	(1.058)	(0.020)
0	0.064	0.237	-0.475	-0.474***	92.283**	-0.014**	2.052	1.344**	0.044
	(0.053)	(0.299)	(0.848)	(0.110)	(24.301)	(0.004)	(1.916)	(0.460)	(0.023)
1	0.099*	0.554*	-0.332	-0.415**	69.616***	-0.012***	2.873***	1.837	0.027
	(0.046)	(0.259)	(0.916)	(0.112)	(16.197)	(0.003)	(0.539)	(1.263)	(0.019)
2	0.080	0.446	-0.048	-0.343**	81.385***	-0.013**	2.081*	1.820	-0.017
	(0.048)	(0.280)	(0.581)	(0.106)	(20.790)	(0.004)	(0.960)	(1.072)	(0.016)
3	0.111*	0.719*	-0.207	-0.317**	79.729***	-0.021***	2.444	1.419*	0.029
	(0.052)	(0.322)	(0.757)	(0.099)	(20.313)	(0.002)	(1.192)	(0.669)	(0.024)
4	0.017	0.066	-0.479	-0.307**	69.246***	-0.014**	0.901	-0.071	-0.020
	(0.052)	(0.334)	(0.461)	(0.098)	(17.258)	(0.004)	(0.859)	(0.981)	(0.015)
5	0.060	0.386	0.024	-0.233*	46.752***	-0.010***	2.399***	1.230	-0.003
	(0.042)	(0.274)	(0.810)	(0.092)	(7.463)	(0.002)	(0.619)	(0.688)	(0.011)
6	0.038	0.235	-0.421	-0.217*	38.340**	-0.007***	2.105*	0.506	0.021
	(0.050)	(0.338)	(0.490)	(0.081)	(10.394)	(0.002)	(0.945)	(0.741)	(0.011)
7	-0.010	-0.097	-0.848	-0.153	29.290*	-0.003	0.433	0.317	0.014
	(0.032)	(0.211)	(0.706)	(0.080)	(12.355)	(0.002)	(0.833)	(0.596)	(0.020)
8	-0.010	-0.080	-0.428	-0.038	16.158	-0.005	1.743*	0.467	0.013
	(0.032)	(0.220)	(0.409)	(0.045)	(9.792)	(0.003)	(0.812)	(0.833)	(0.021)
Canton FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LinTrends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N (Canton x Years)	617	617	655	830	655	655	830	830	830
R2	0.952	0.94	0.857	0.993	0.715	0.692	0.827	0.673	0.158
Mean Dependent	1.63	11.87	37.15	28.86	3,342	0.053	7.84	5.13	0.94

Notes: This table shows coefficients for δ_τ from Equation (1) for $\tau \in [-5, 8]$ for each specific outcome variable. The event year represents the year relative to the introduction of the baby bonus. The omitted category is event time $\tau = -1$. Estimates in the canton-year cell are weighted corresponding to the following structure: fertility rate with number of fertile women; crude birth rate with total population; birth interval, age mother, birth weight, share low birth weight, death rate, stillbirth rate, and sex ratio with number of births. Robust standard errors (shown in parentheses) are clustered at the cantonal level and significance levels are indicated by * 0.05 ** 0.01 *** 0.001.

TABLE A.3: Event Study Estimates Controlling for Child Allowances

	(1) Fertility Rate	(2) Crude Birth Rate	(3) Birth Interval	(4) Age Mother	(5) Birth Weight	(6) Share Low Birth Weight	(7) Death Rate	(8) Stillbirth Rate	(9) Sex Ratio
Event Time Estimate: δ_τ									
-5	-0.041	0.285	-1.075	-0.083	49.447**	-0.008***	0.203	0.126	-0.033
	(0.043)	(0.169)	(1.702)	(0.051)	(13.889)	(0.002)	(1.115)	(0.886)	(0.018)
-4	-0.029	0.271	0.205	0.003	9.434	-0.005	1.128	-1.077	-0.018
	(0.017)	(0.156)	(0.326)	(0.056)	(13.317)	(0.003)	(0.998)	(0.650)	(0.015)
-3	0.019	0.172	-0.271	-0.022	17.379	-0.007*	-0.859	0.015	-0.026
	(0.048)	(0.167)	(0.227)	(0.038)	(18.247)	(0.003)	(1.304)	(0.920)	(0.015)
-2	0.018	-0.113	0.019	-0.022	14.857	-0.001	-0.019	0.028	-0.017
	(0.018)	(0.132)	(0.258)	(0.064)	(26.996)	(0.002)	(0.848)	(1.117)	(0.018)
0	0.050	-0.495	-0.522	-0.470***	95.224***	-0.016**	1.554	0.296	0.051**
	(0.051)	(0.352)	(0.845)	(0.110)	(24.262)	(0.005)	(1.126)	(0.790)	(0.014)
1	0.081	-0.494	-0.349	-0.409**	77.337***	-0.012**	2.275***	0.968	0.030
	(0.044)	(0.351)	(0.915)	(0.114)	(19.459)	(0.003)	(0.574)	(1.118)	(0.016)
2	0.066	-0.314	-0.042	-0.337**	86.816***	-0.014**	1.300	1.653	-0.002
	(0.048)	(0.298)	(0.578)	(0.108)	(22.398)	(0.004)	(0.734)	(0.854)	(0.015)
3	0.094	-0.162	-0.170	-0.308**	83.785***	-0.021***	1.425	0.547	0.030
	(0.050)	(0.348)	(0.766)	(0.102)	(20.578)	(0.003)	(0.909)	(0.816)	(0.022)
4	0.023	-0.497	-0.452	-0.302**	68.656**	-0.012*	0.763	-0.278	-0.018
	(0.042)	(0.290)	(0.497)	(0.097)	(19.091)	(0.005)	(0.883)	(0.815)	(0.012)
5	0.054	-0.095	0.024	-0.231*	51.267***	-0.010***	1.664**	1.090	-0.005
	(0.035)	(0.258)	(0.827)	(0.095)	(10.451)	(0.002)	(0.570)	(0.549)	(0.011)
6	0.023	-0.200	-0.447	-0.219*	34.373**	-0.006*	2.066*	0.236	0.013
	(0.043)	(0.291)	(0.486)	(0.080)	(11.189)	(0.002)	(0.894)	(0.675)	(0.011)
7	-0.025	-0.442*	-0.877	-0.155	24.394	-0.002	0.271	-0.138	0.012
	(0.033)	(0.177)	(0.699)	(0.079)	(12.685)	(0.002)	(0.708)	(0.594)	(0.017)
8	-0.025	-0.312	-0.443	-0.039	19.020*	-0.004	1.683*	-0.231	0.001
	(0.028)	(0.165)	(0.406)	(0.044)	(9.134)	(0.003)	(0.649)	(0.862)	(0.022)
Child Allowance	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Canton FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LinTrends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N (Canton x Years)	636	636	678	862	678	678	862	862	862
R2	0.952	0.94	0.857	0.993	0.719	0.974	0.827	0.672	0.159
Mean Dependent	1.63	11.87	37.15	28.86	3,342	0.053	7.84	5.13	0.94

Notes: This table shows coefficients for δ_τ from Equation (1) for $\tau \in [-5, 8]$ for each specific outcome variable. The event year represents the year relative to the introduction of the baby bonus. The omitted category is event time $\tau = -1$. Estimates in the canton-year cell are weighted corresponding to the following structure: fertility rate with number of fertile women; crude birth rate with total population; birth interval, age mother, birth weight, share low birth weight, death rate, stillbirth rate, and sex ratio with number of births. Robust standard errors (shown in parentheses) are clustered at the cantonal level and significance levels are indicated by * 0.05 ** 0.01 *** 0.001.

TABLE A.4: Fertility and Birth Weight by Citizenship of Mother

	Fertility Rate		Birth Weight	
	High-Income	LMIC	High-Income	LMIC
	(1)	(2)	(3)	(4)
-5	-0.049	0.023*	47.808*	-29.463
	(0.05)	(0.01)	(14.71)	(141.66)
-4	-0.036	0.017*	9.581	-248.572
	(0.02)	(0.01)	(12.11)	(343.67)
-3	0.016	0.010	15.998	-240.484
	(0.05)	(0.01)	(15.64)	(294.62)
-2	0.019	0.001	14.035	13.205
	(0.02)	(0.01)	(26.19)	(110.70)
0	-0.063	0.152***	94.278***	165.105
	(0.05)	(0.01)	(24.53)	(98.26)
1	-0.033	0.143***	75.224***	170.447
	(0.05)	(0.01)	(19.76)	(92.46)
2	-0.049	0.137***	84.411**	187.007*
	(0.05)	(0.01)	(23.86)	(76.86)
3	-0.008	0.117***	79.178**	110.081
	(0.05)	(0.01)	(22.93)	(143.68)
4	-0.067	0.103***	64.218**	260.095***
	(0.04)	(0.01)	(19.72)	(63.98)
5	-0.013	0.081***	48.037***	119.716
	(0.03)	(0.01)	(11.76)	(68.73)
6	-0.025	0.063***	31.955*	150.339*
	(0.04)	(0.01)	(12.76)	(59.94)
7	-0.048	0.036***	22.330	115.592*
	(0.03)	(0.00)	(13.59)	(55.20)
8	-0.041	0.022***	19.217	-29.219
	(0.03)	(0.00)	(9.50)	(94.70)
Canton FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
LinTrends	Yes	Yes	Yes	Yes
N (Canton x Years)	636	636	678	678
R2	0.974	0.982	0.717	0.33

Notes: This table shows coefficients for δ_τ from Equation (1) for $\tau \in [-5, 8]$ on fertility and birth weight by citizenship of the mother. The event year represents the year relative to the introduction of the baby bonus. The omitted category is event time $\tau = -1$. Estimates in the canton-year cell are weighted with the number of fertile women for the fertility rate and with the number of births for birth weight. Robust standard errors (shown in parentheses) are clustered at the cantonal level and significance levels are indicated by * 0.05 ** 0.01 *** 0.001.

TABLE A.5: Maternal Age at Birth by Citizenship of Mother

	Maternal Age at Birth	
	(1)	(2)
	High-Income	LMIC
-5	0.023	0.277
	(0.05)	(0.36)
-4	0.081	1.130
	(0.05)	(0.56)
-3	0.029	-0.377
	(0.04)	(0.79)
-2	0.003	-1.035
	(0.06)	(1.14)
0	-0.167	-0.107
	(0.09)	(0.78)
1	-0.139	0.172
	(0.09)	(0.71)
2	-0.079	0.272
	(0.09)	(0.85)
3	-0.085	0.320
	(0.09)	(0.89)
4	-0.101	-0.207
	(0.09)	(0.63)
5	-0.084	-0.128
	(0.09)	(0.46)
6	-0.094	0.053
	(0.08)	(0.33)
7	-0.087	0.601
	(0.08)	(0.92)
8	0.003	-0.290
	(0.04)	(0.36)
Canton FE	Yes	Yes
Year FE	Yes	Yes
LinTrends	Yes	Yes
N (Canton x Years)	862	862
R2	0.995	0.816

Notes: This table shows coefficients for δ_τ from Equation (1) for $\tau \in [-5, 8]$ on maternal age by citizenship of mother. The event year represents the year relative to the introduction of the baby bonus. The omitted category is event time $\tau = -1$. Estimates in the canton-year cell are weighted with the number of births. Robust standard errors (shown in parentheses) are clustered at the cantonal level and significance levels are indicated by * 0.05 ** 0.01 *** 0.001.

TABLE A.6: Fertility by Child Rank

	Fertility Rate			
	(1)	(2)	(3)	(4)
	Rank 1	Rank 2	Rank 3	All
-5	-0.011	-0.025*	-0.007	-0.026
	(0.05)	(0.01)	(0.01)	(0.04)
-4	-0.022	-0.008	0.003	-0.018
	(0.01)	(0.01)	(0.01)	(0.02)
-3	0.014	-0.013	0.003	0.026
	(0.03)	(0.01)	(0.01)	(0.05)
-2	-0.011	0.005	-0.011	0.020
	(0.01)	(0.00)	(0.01)	(0.02)
0	0.055	0.038*	-0.011	0.089
	(0.03)	(0.02)	(0.01)	(0.05)
1	0.068	0.021	-0.000	0.110*
	(0.03)	(0.01)	(0.01)	(0.04)
2	0.072*	0.027	-0.017	0.089
	(0.03)	(0.01)	(0.01)	(0.05)
3	0.087*	0.024	-0.012	0.109*
	(0.04)	(0.02)	(0.01)	(0.05)
4	0.049	0.008	-0.024*	0.036
	(0.03)	(0.01)	(0.01)	(0.04)
5	0.049	0.030*	-0.008	0.068
	(0.03)	(0.01)	(0.01)	(0.04)
6	0.048	0.002	-0.011	0.038
	(0.03)	(0.02)	(0.01)	(0.04)
7	0.005	0.008	-0.021*	-0.013
	(0.03)	(0.02)	(0.01)	(0.03)
8	-0.018	0.003	-0.005	-0.019
	(0.03)	(0.01)	(0.00)	(0.03)
Canton FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
LinTrends	Yes	Yes	Yes	Yes
N (Canton x Years)	636	636	636	636
R2	0.913	0.939	0.952	0.952

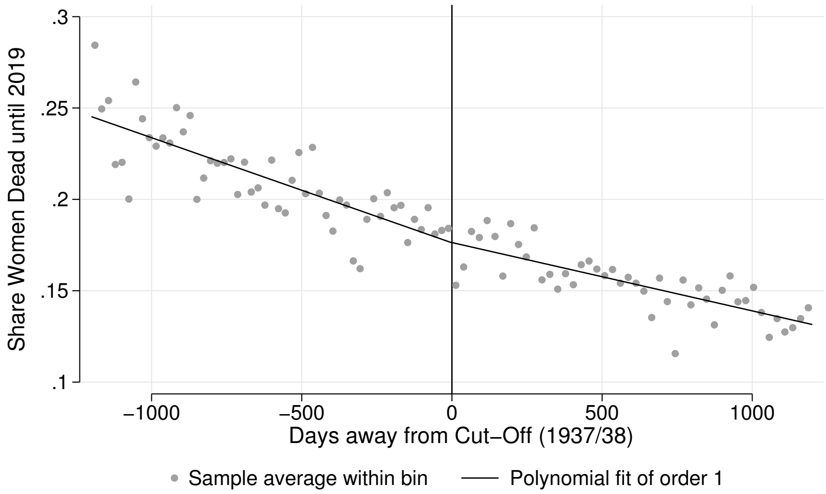
Notes: This table shows coefficients for δ_τ from Equation (1) for $\tau \in [-5, 8]$ on fertility by child rank. The event year represents the year relative to the introduction of the baby bonus. The omitted category is event time $\tau = -1$. Estimates in the canton-year cell are weighted with the number of fertile women. Robust standard errors (shown in parentheses) are clustered at the cantonal level and significance levels are indicated by * 0.05 ** 0.01 *** 0.001.

TABLE A.7: Event Study Estimates with Wild Cluster Bootstrapped Standard Errors

	(1) Fertility Rate	(2) Crude Birth Rate	(3) Birth Interval	(4) Age Mother	(5) Birth Weight	(6) Share Low Birth Weight	(7) Death Rate	(8) Stillbirth Rate	(9) Sex Ratio
Event Time Estimate: δ_τ									
-5	-0.026	-0.350	-1.184	-0.085	47.483	-0.007	0.037	0.140	-0.034
	(0.6346)	(0.5145)	(0.5606)	(0.6877)	(0.2142)	(0.3143)	(0.9730)	(0.8458)	(0.2302)
-4	-0.018	-0.254	0.094	0.002	7.167	-0.004	0.971	-1.071	-0.018
	(0.1471)	(0.1321)	(0.9039)	(0.3243)	(0.7968)	(0.4074)	(0.2743)	(0.2032)	(0.4535)
-3	0.026	0.064	-0.371	-0.023	15.172	-0.006	-0.949	0.021	-0.026
	(0.9009)	(0.8899)	(0.1091)	(0.7878)	(0.3343)	(0.4004)	(0.5145)	(0.9830)	(0.1632)
-2	0.020	0.079	0.015	-0.022	14.659	-0.001	-0.073	0.031	-0.017
	(0.7207)	(0.7928)	(0.9389)	(0.2733)	(0.6066)	(0.5255)	(0.9129)	(0.9630)	(0.3794)
0	0.089	0.189	-0.475	-0.474***	93.340**	-0.015	1.151	0.331	0.049
	(0.4545)	(0.6076)	(0.6396)	(0.2332)	(0.0300)	(0.1001)	(0.4484)	(0.6907)	(0.2663)
1	0.110*	0.423	-0.332	-0.415*	74.973**	-0.011**	2.099	1.027	0.027
	(0.0982)	(0.2492)	(0.7738)	(0.082)	(0.0250)	(0.0431)	(0.1341)	(0.4545)	(0.2252)
2	0.089	0.344	-0.048	-0.343*	83.869**	-0.013**	1.180	1.708	-0.006
	(0.3483)	(0.3914)	(0.9510)	(0.0551)	(0.0310)	(0.0422)	(0.3904)	(0.4194)	(0.7347)
3	0.109*	0.541	-0.207	-0.317*	79.970*	-0.020***	1.558	0.628	0.025
	(0.074)	(0.2793)	(0.7968)	(0.0721)	(0.0501)	(0.0060)	(0.2012)	(0.5285)	(0.4374)
4	0.036	0.061	-0.479	-0.307	65.582**	-0.011**	0.739	-0.232	-0.020
	(0.7457)	(0.8539)	(0.3984)	(0.1461)	(0.0160)	(0.0487)	(0.4034)	(0.7638)	(0.2222)
5	0.068	0.324	0.024	-0.233*	49.411***	-0.009*	1.549	1.112	-0.006
	(0.2683)	(0.2943)	(0.9770)	(0.0861)	(0.0070)	(0.0791)	(0.1902)	(0.1532)	(0.5976)
6	0.038	0.151	-0.421	-0.217*	33.970**	-0.005*	1.780	0.222	0.014
	(0.7337)	(0.7738)	(0.4064)	(0.0630)	(0.0310)	(0.0652)	(0.2002)	(0.7658)	(0.2342)
7	-0.013	-0.174	-0.848	-0.153	24.522	-0.002	-0.031	-0.159	0.014
	(0.5986)	(0.4705)	(0.2913)	(0.1982)	(0.1441)	(0.5345)	(0.9640)	(0.8128)	(0.3674)
8	-0.019	-0.175	-0.428	-0.038	19.108*	-0.004	1.536*	-0.242	0.001
	(0.4545)	(0.238)	(0.3704)	(0.5325)	(0.0881)	(0.3183)	(0.0671)	(0.7938)	(0.9610)
Canton FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LinTrends	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
N (Canton x Years)	636	636	678	862	678	678	862	862	862
R2	0.952	0.937	0.857	0.993	0.719	0.694	0.79	0.672	0.157
Mean Dependent	1.63	11.87	37.15	28.86	3,342	0.053	7.84	5.13	0.94

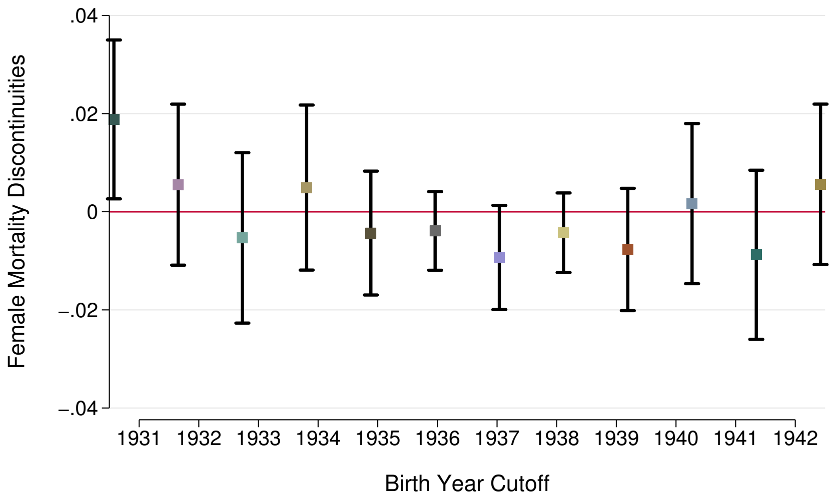
Notes: This table shows coefficients for δ_τ from Equation (1) for $\tau \in [-5, 8]$ for each specific outcome variable. The event year represents the year relative to the introduction of the baby bonus. The omitted category is event time $\tau = -1$. Estimates in the canton-year cell are weighted corresponding to the following structure: fertility rate with number of fertile women; crude birth rate with total population; birth interval, age mother, birth weight, share low birth weight, death rate, stillbirth rate, and sex ratio with number of births. Values in parenthesis indicate the wild-cluster bootstrapped p-values. Stars indicate significance levels of * 0.10 ** 0.05 *** 0.01.

FIGURE B.1: Placebo: RDD at (male) ER 63 Cutoff for Women



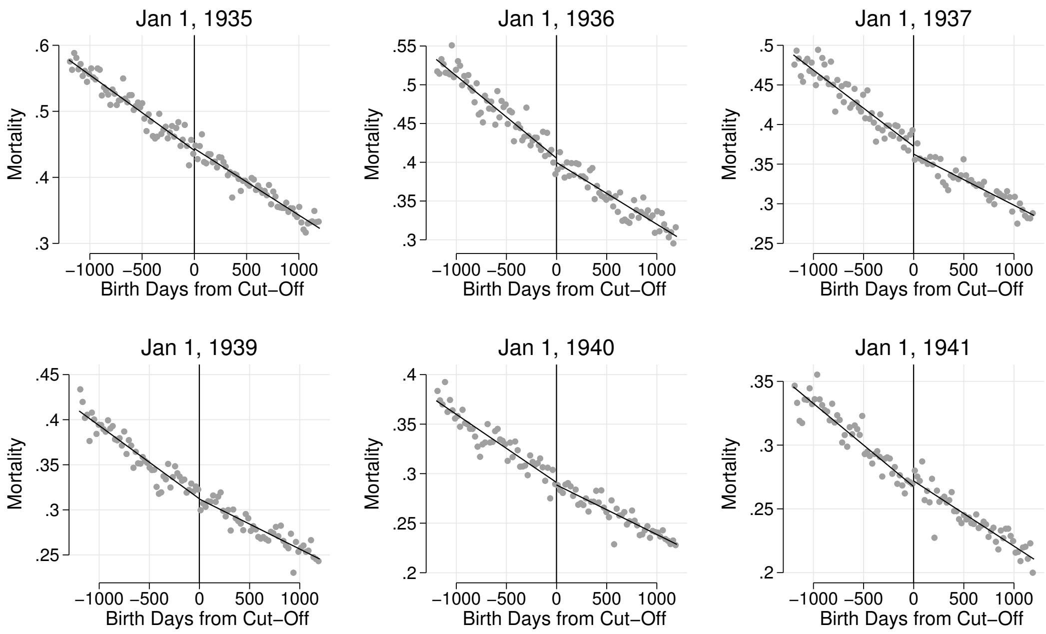
Notes: This graph shows the reduced form estimate at cutoff year 1937/38 for women only. Women are not targeted by the policy and should consequently not show a discontinuity.

FIGURE B.2: Discontinuities for Women



Notes: This graph estimates discontinuities for women at other end-of birth year cutoffs. RD specifications as specified in Table 3.3. The bar shows a 90% confidence interval.

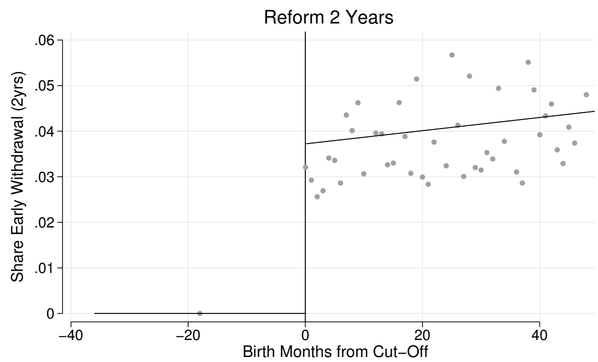
FIGURE B.3: Placebo: Other Cutoff Years for Men



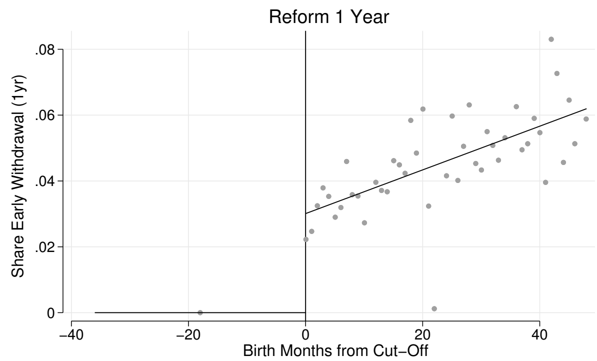
Notes: This figure shows reduced form RDD graphs around several end-of-year cutoffs that are not policy reform cutoffs.

FIGURE B.4: Share of Men Drawing Early Retirement Benefits

(A) 2 years earlier



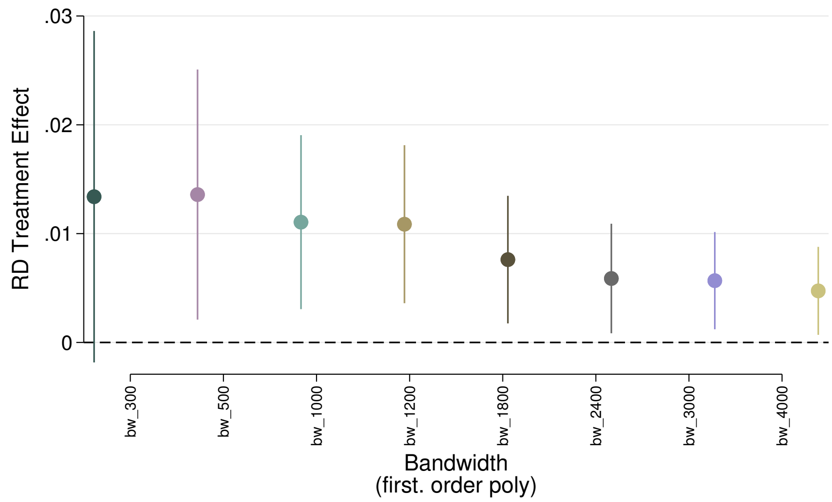
(B) 1 year earlier



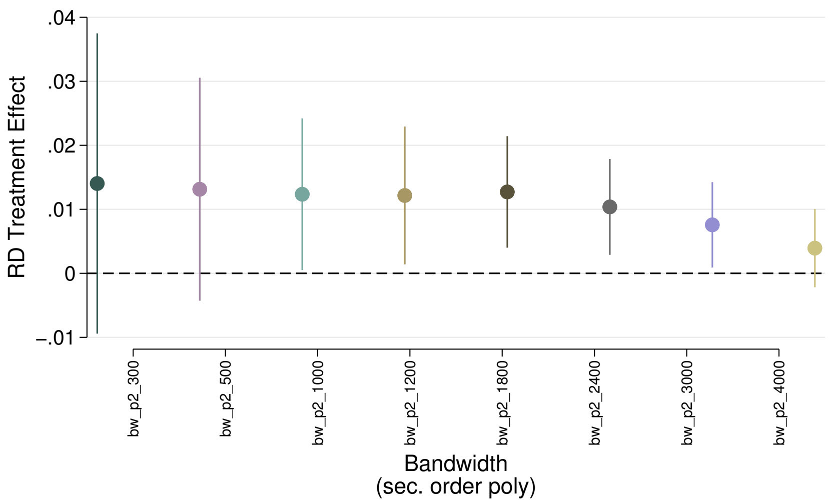
Notes: This figure shows the first stage estimates for the two policy reforms: retiring at age 63 (a) and retiring at age 64 (b).

FIGURE B.5: Robustness: Bandwidths and Order of Local Polynomial

(A) Order Local Polynomial = 1



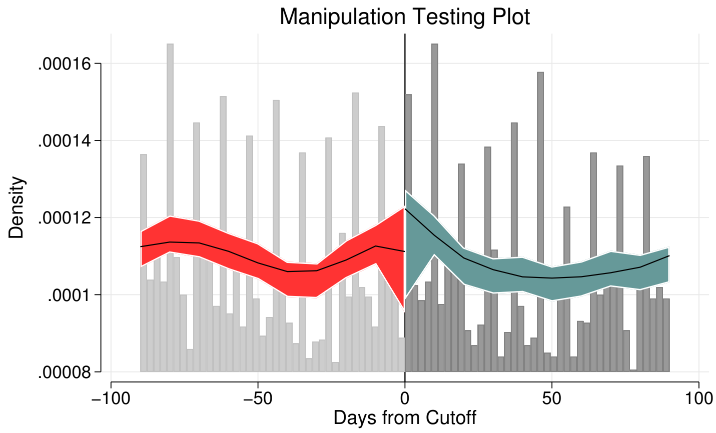
(B) Order Local Polynomial = 2



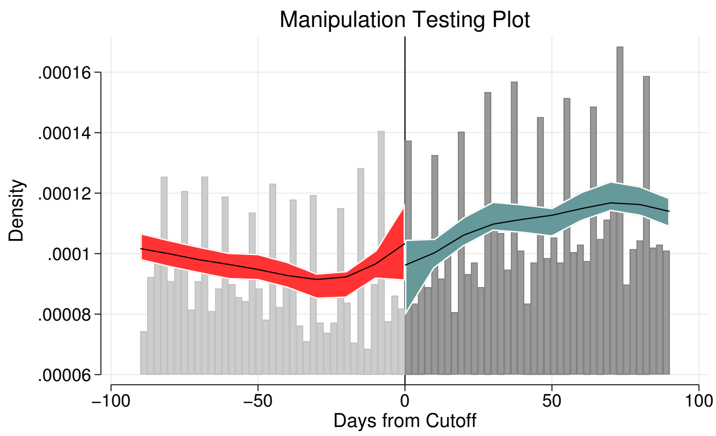
Notes: This figure shows the treatment effect for the 63 reform cutoff for different bandwidths and polynomials. The bar indicates a 90% confidence interval.

FIGURE B.6: Manipulation around Cutoff

(A) ER 63 (Cutoff = 1938)

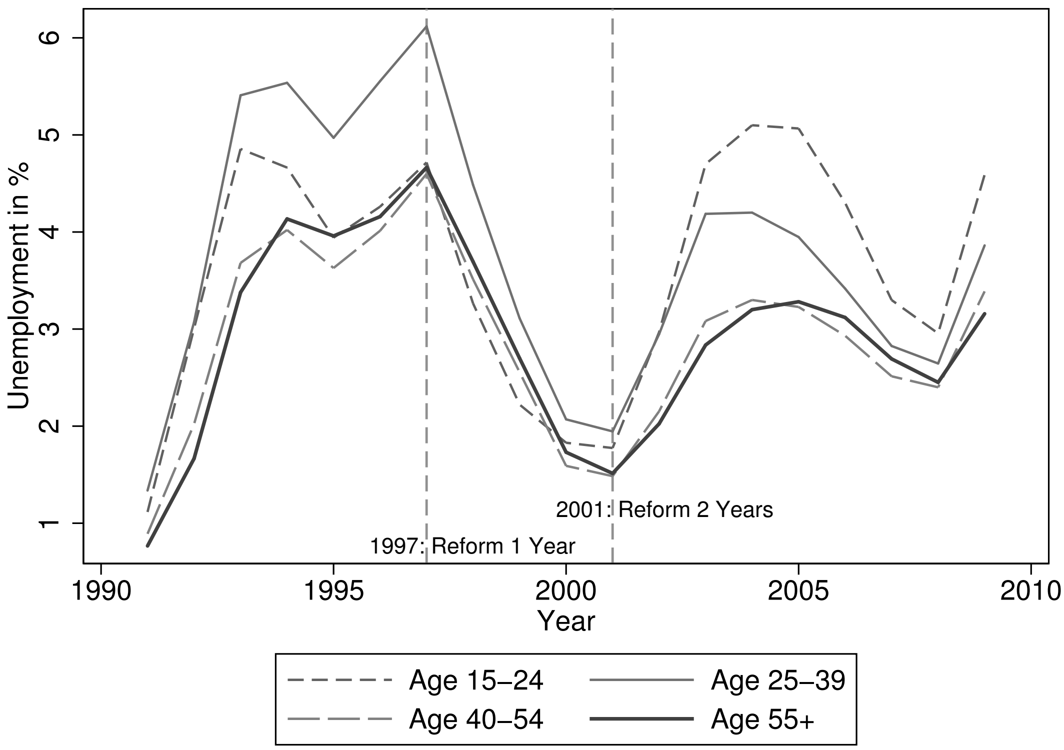


(B) ER 64 (Cutoff = 1933)



*Notes:*This figure shows the density of men alive in 1990 around the two cutoffs in order to perform a manipulation check. In Panel (a), the hypothesis that density changes around the cutoff is rejected with $p = 0.9618$ (bandwidth = 30, polynomial = quadratic). In Panel (b), the hypothesis that density changes around the cutoff is rejected with p = 0.5156 (bandwidth = 30, polynomial = quadratic). The colored area depicts a 90% confidence interval. Implemented with *rddensity* by Calonico et al. (2014).

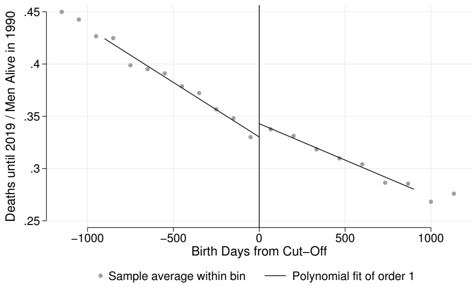
FIGURE B.7: Unemployment by Age Group



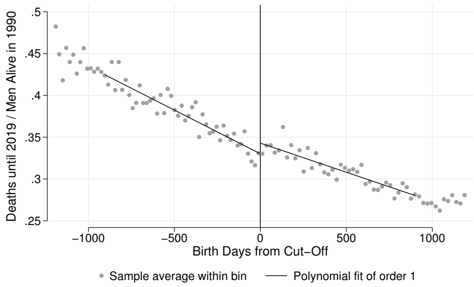
Notes: This figure shows the unemployment rate by age group and year. The age groups include men and women. Source: Federal Office of Statistics

FIGURE B.8: Alternative Graphical RD Plots (ER 63)

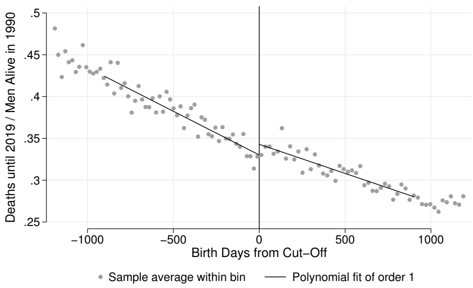
(A) Bin selection = es



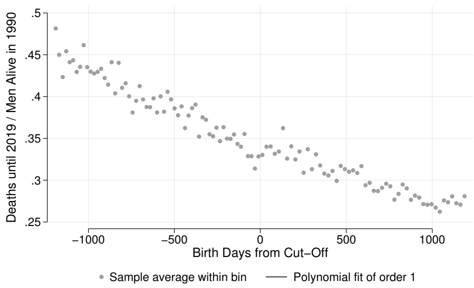
(B) Bin selection = esmvpr



(C) Bin selection = qsmvpr



(D) Without Lines



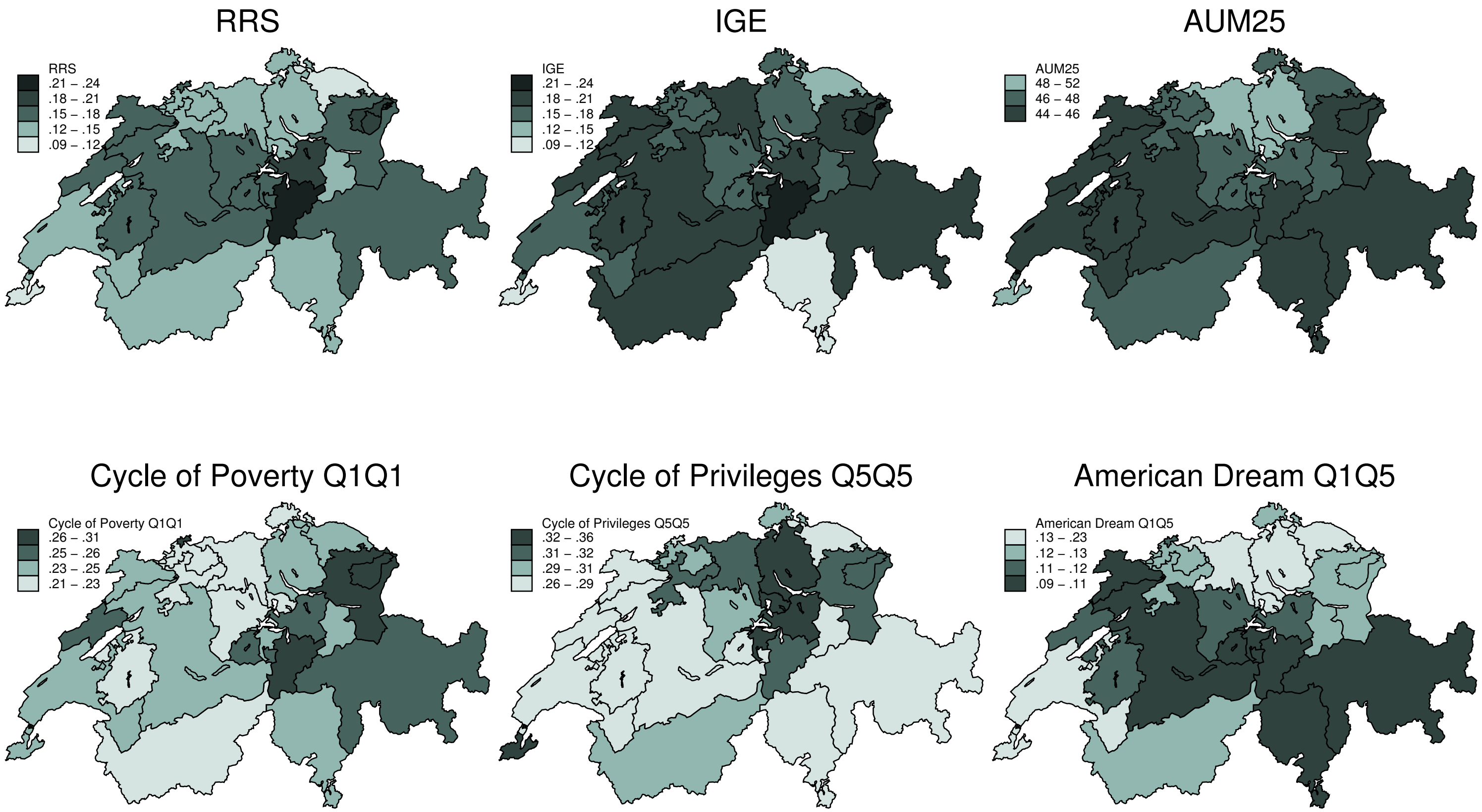
Notes: This figure shows alternative RD plots for the ER 63 reform. The first three graphs use different bin sizes. Panel (B.8a) uses the integrated mean squared error (IMSE)-optimal evenly spaced method using spacing estimators, Panel (B.8b) the mimicking-variance evenly spaced method using polynomial regression and Panel (B.8c) the mimicking-variance quantile-spaced method using polynomial regression. Panel (B.8c) does not show any lines which is a test to see if the discontinuity can be spotted without other visual aid.

TABLE B.1: Testing for jumps at non-discontinuity points

	Cutoff	Estimate (Reduced Form)	p-Value
(1)	+2374	-.00062	0.907
(2)	-1463	0.01347	0.229
(3)	+3561	-0.00121	0.837
(4)	+1187	-0.01257	0.174
(5)	-732	0.01586	0.264
(6)	-2195	-0.01517	0.236

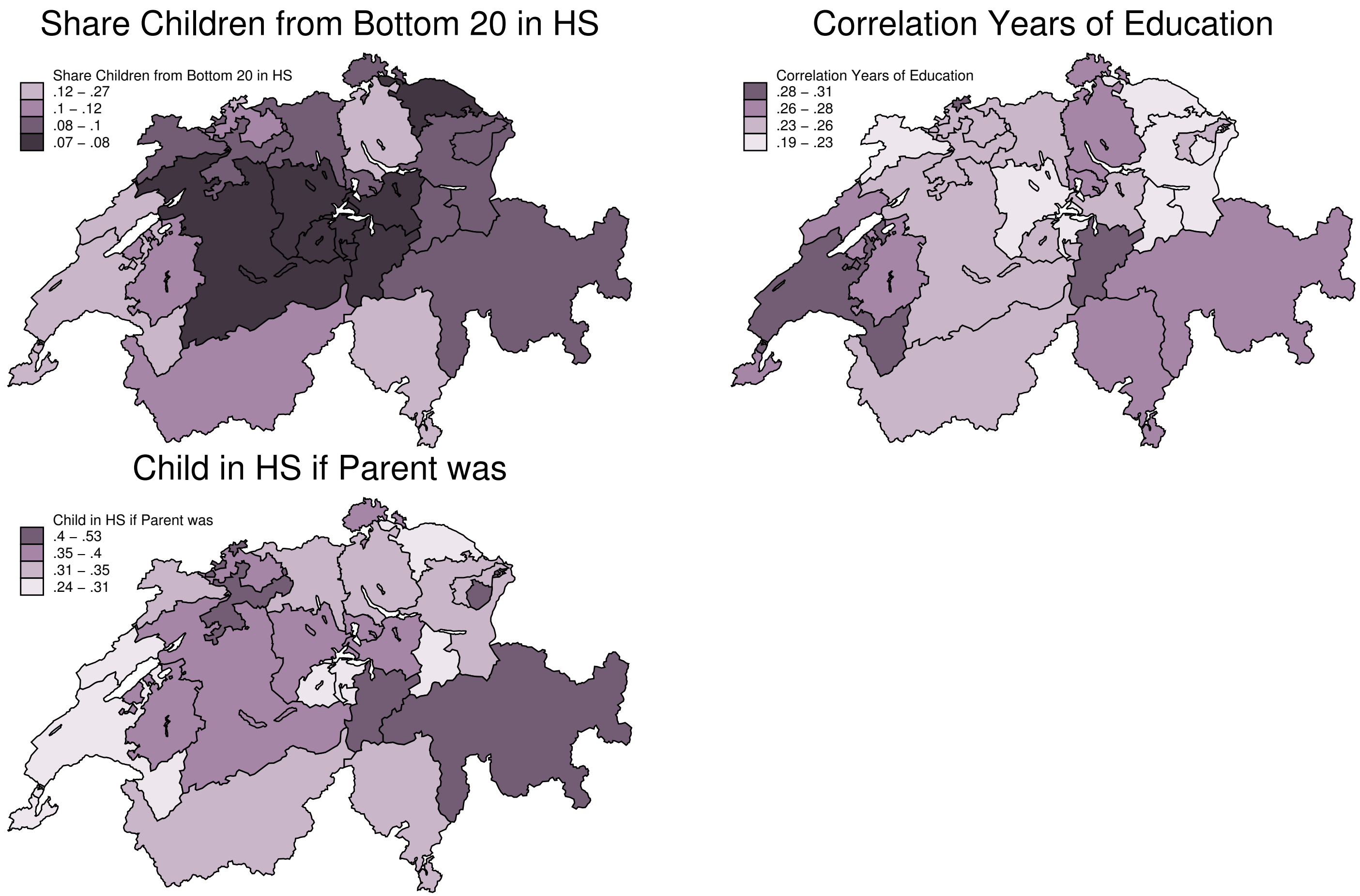
Notes: This table tests if there are no jumps at non-discontinuity points (Imbens & Lemieux, 2008). Row (1) assigns as a placebo cutoff the median above cutoff x_0 and only the sample above x_0 is used. Row (2) estimates the cutoff at the median of the sample below cutoff x_0. Row (3) to (6) split the two samples again in half and test for a discontinuity at the quartile points. All regression discontinuities estimate the reduced form and use quadratic polynomial and automatic bandwidth detection. A p-Value higher than 0.10 indicates that the hypothesis of no discontinuity at the cutoff cannot be rejected.

FIGURE C.1: Income Mobility Estimates by Cantons



Notes: This figure shows how educational mobility varies across cantons. Brighter colors indicate higher mobility.

FIGURE C.2: Educational Mobility Estimates by Cantons



Notes: This figure shows how educational mobility varies across cantons. Brighter colors indicate lower mobility.

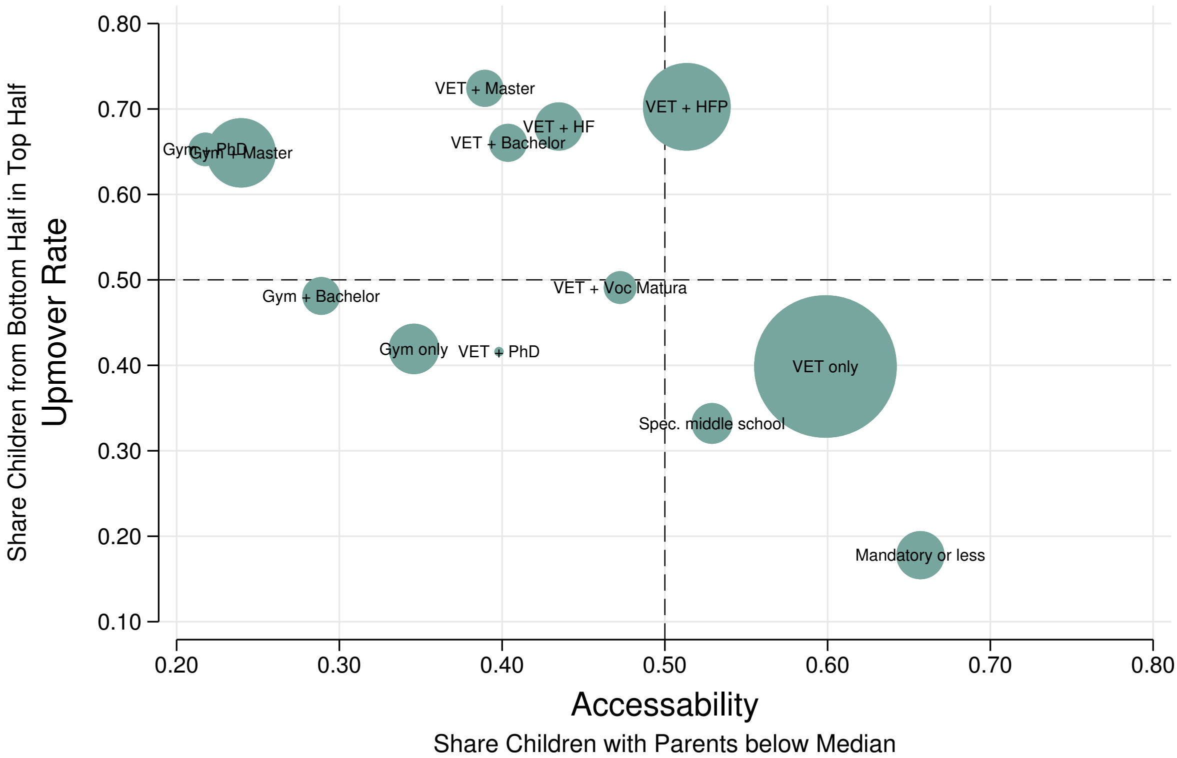
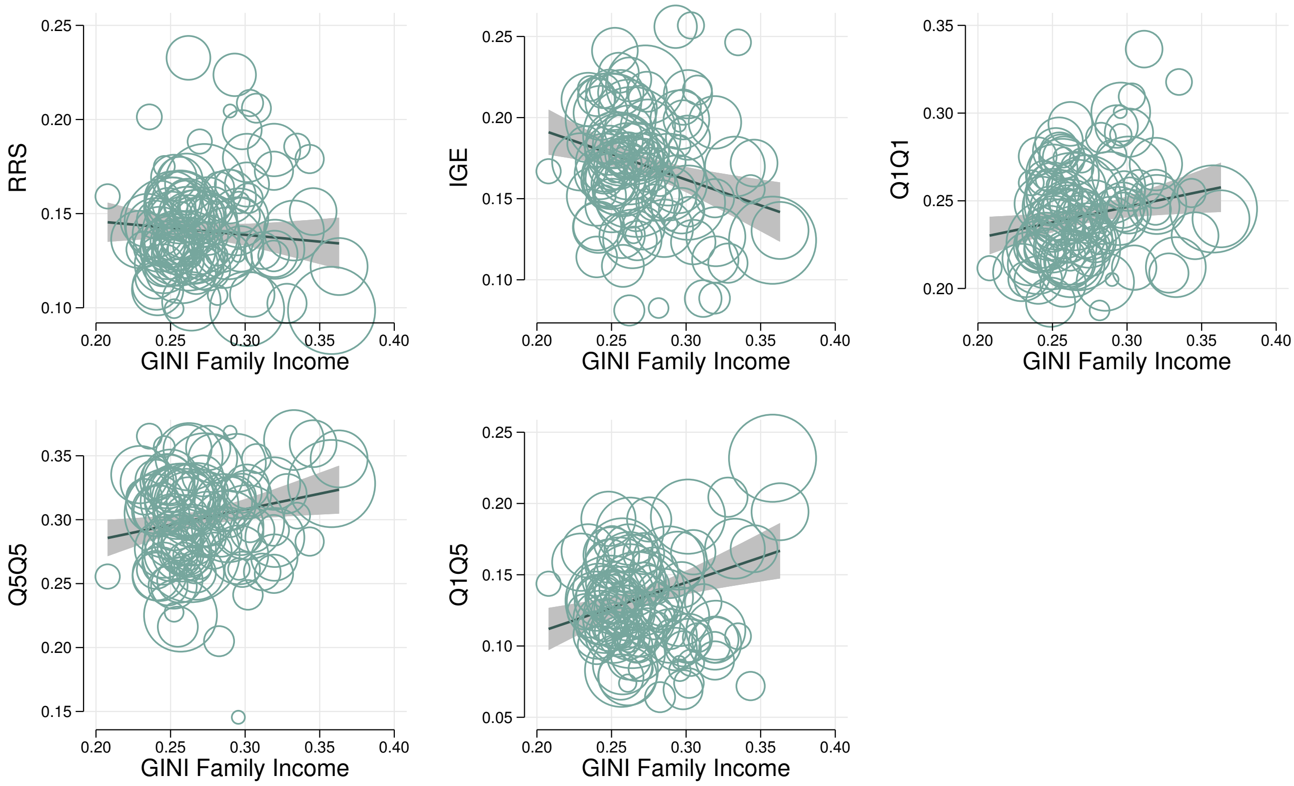


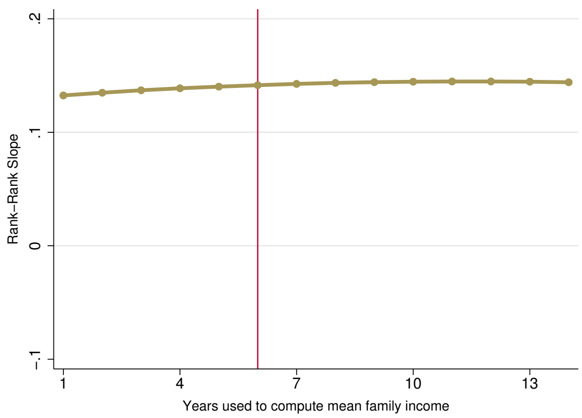
FIGURE C.3: Access and Upward Mobility Rate by Educational Track (Medium Upward Mobility)

Notes: This graph shows how upward mobility (upmover rate) and access differ between educational tracks. Upmover rate is defined as the share of children with parents below the median moving above the median. Accessibility is defined as the share of children from the bottom half relative to the total number of children in the educational track. The size of the points is proportional to the number of children in that track. The syntax of the educational tracks is defined as follows: [Upper Secondary Education] + [Highest Post Secondary Education]. *Gym* refers to gymnasium (academic high school), *VET* to vocational education and training. *HF* and *HFP* are occupation specific higher educations. *Voc Matura* refers to "vocational matura", which is VET with more formal education, *spec. middle school* refers to specialized middle schools, which is like a professional high school and not as selective as the gymnasium.
The upmover rate multiplied with the access rate equals the mobility rate, which is the share of children that climb from the bottom to the top half *relative to all children in that track*. In contrast, the upmover rate is the share of children that climb to the top half *relative to children from parents in the bottom half*.

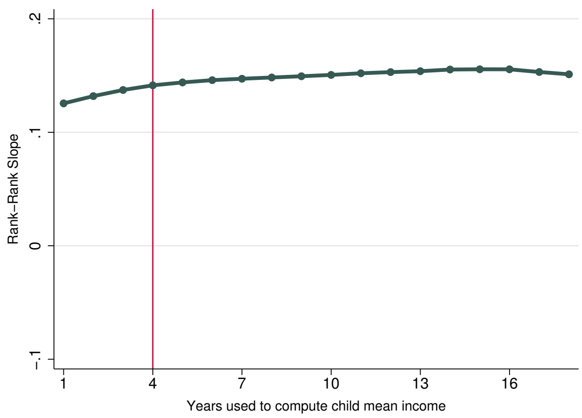
FIGURE C.4: Great Gatsby Curves (Labor Market Level)



Notes: This graph shows how income mobility is relates to income inequality on a labor market regions level. It does so for different income mobility measures. Income inequality is measured on a family level when children are between 15 and 20 years old. The grey line shows the fitted slope between the values weighted by the size of observations in each canton.



(A) RRS by Number of Years of Father Income



(B) RRS by Number of Years of Child Income

FIGURE C.5: Robustness of Rank-Rank Slope Estimates: Attenuation Bias

Notes: In this figure we asses the robustness of the rank-rank slope to changes in the number of years used to measure father income (Panel (a)) and child income (Panel (b)). Fathers are ranked relative to other fathers of children in the same birth cohort. Children are ranked relative to other children in the same birth cohort.

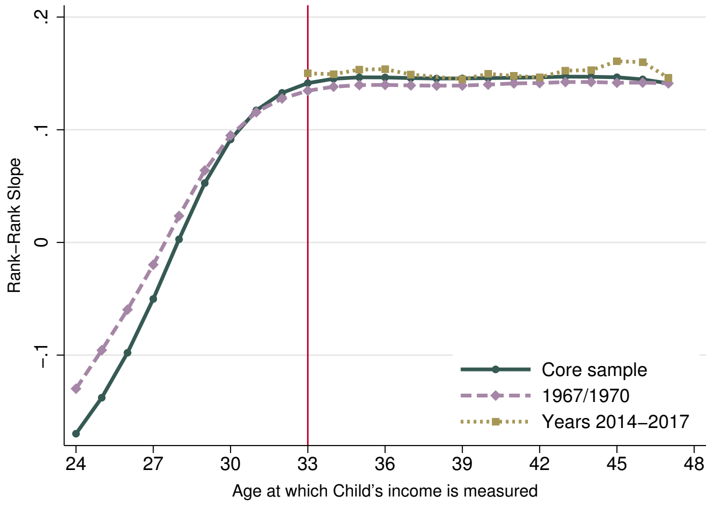
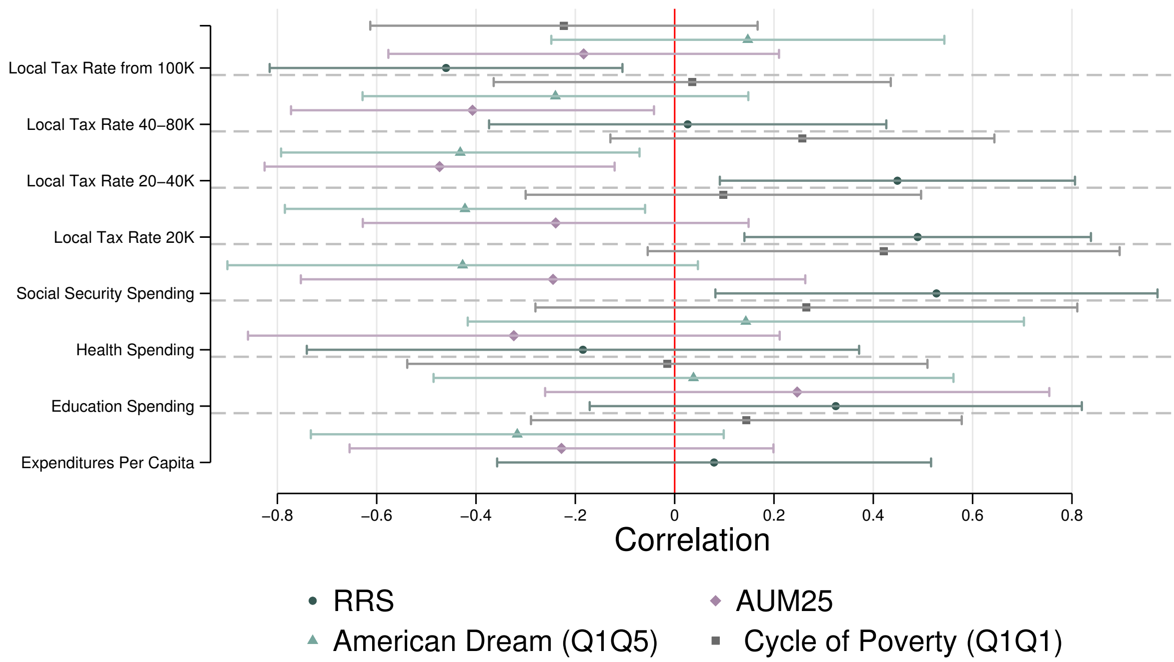


FIGURE C.6: Life-Cycle Bias

Notes: This figure assesses the robustness of the rank-rank slopes estimates. For the baseline estimates, father income is averaged over the years when the child is between 15 and 20 years old. Father rank is defined relative to other fathers of children born between 1969 and 1984. Child mean income is the average income when the child is between the age 30 and 33. Child rank is defined relative to children in the same birth cohort. This corresponds to the point at age 33. The first point corresponds to the rank-rank slope when child mean income is averaged over the ages of 21 and 24. The last point uses average mean income between age 44 and 47 and is only observable for the 1969 and 1970 cohorts. Mean father rank is defined according to father income of children born in those cohorts. The dashed line plots the rank-rank slope coefficients by varying the age at which child income is measured only for the 1969 and 1970 cohorts. The dotted line plots the rank-rank slope coefficients when income is measured in the year 2014 to 2017.

FIGURE C.7: Spatial Correlates: Canton



Notes: This graph shows how public expenditures and taxes are correlated on a cantonal level.

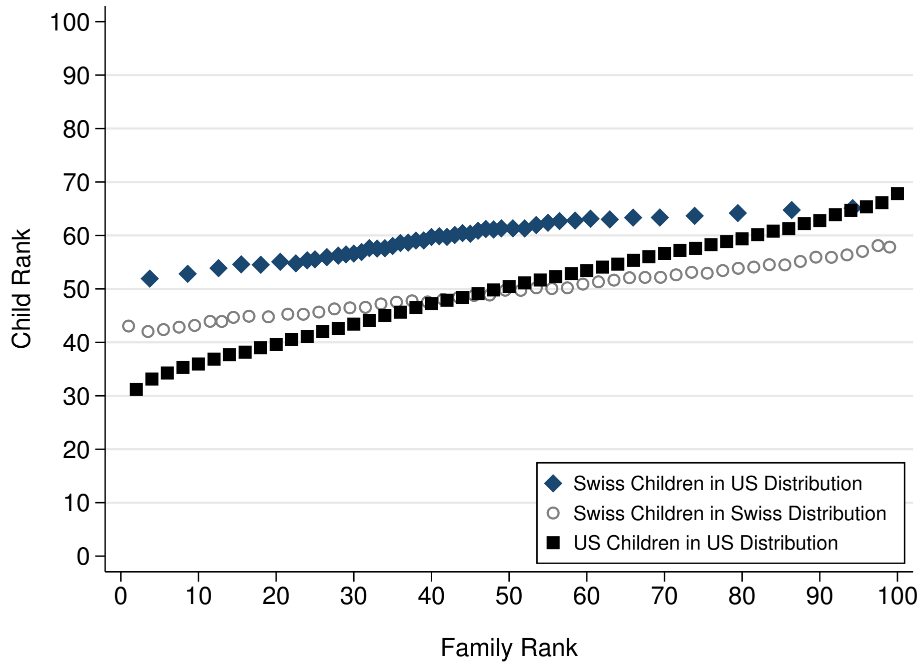
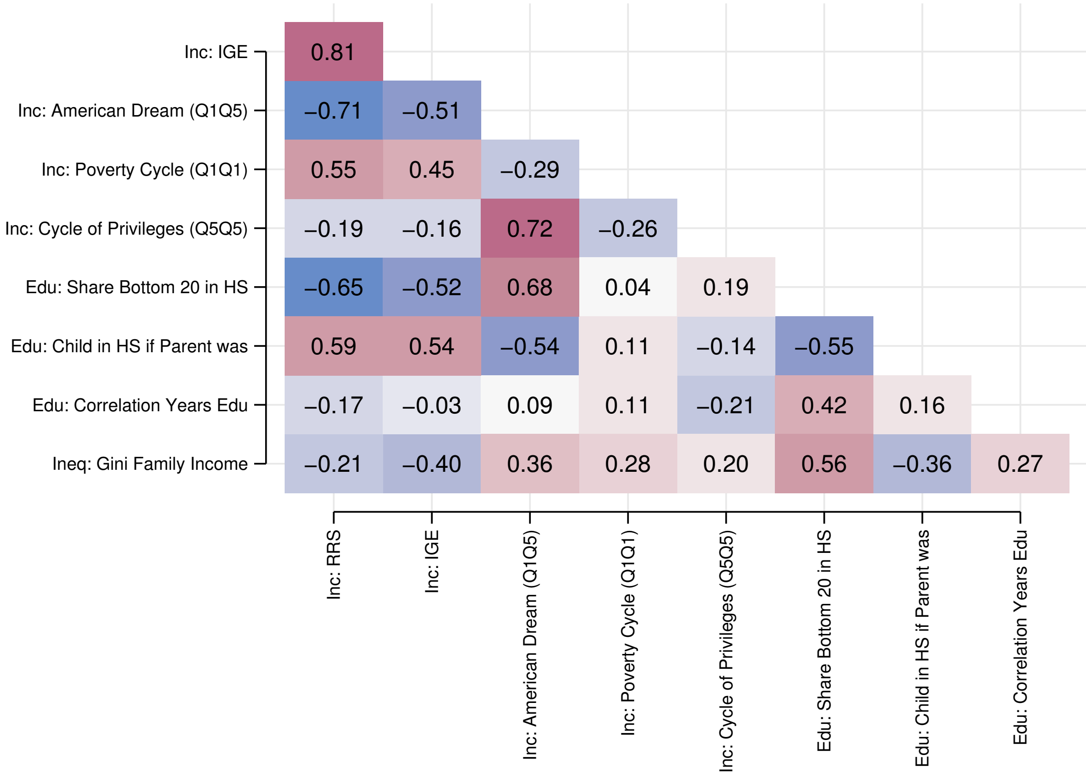


FIGURE C.8: Comparison Switzerland - US

Notes: This graph shows the rank mobility for Switzerland and the US. The black squares correspond to the estimates of Chetty et al. (2014a). The circles correspond to the estimates in our study, where the ranks are assigned according to the Swiss distribution. For the blue diamonds, the income of parents and children is converted into US dollar PPP equivalents and then assigned to ranks according to the US distribution.

FIGURE C.9: Correlation Mobility Measures (Units: Cantons)



Notes: This graph shows the correlation of different mobility and inequality measures on a cantonal unit (n=26). *RRS* is the rank-rank slope, *IGE* is the intergenerational income elasticity, *Share Bottom 20 in HS* measures the share of children from parents in the bottom quintile that go to high school, *Child in HS if Parent was* is the slope coefficient of a linear probability model that regresses high school attendance of the child on high school attendance of the parents, *correlation Years Edu* measures the correlation between years of education of parents and children, *Gini Family Income* shows the Gini index for family income at child age 15 to 20.

TABLE C.1: Highest Education and Years of Schooling

Highest Education	Years of Schooling
No education	0
Max 7 years mandatory school	7
Mandatory school only	9
Vocational Training and Education	12
High school (gymnasium)	13
Higher professional degree	14
Bachelor degree	16
Master degree	18
PhD, habilitation	21

Notes: This table shows how educational attainment is translated into years of schooling.

TABLE C.2: Sample Selection

	Sample			
	(1)	(2)	(3)	(4)
	Full mean	Any Parent mean	Father (Core) mean	Both mean
Panel A: Income Sample				
Year of Birth	1975.10	1975.24	1975.65	1975.79
Female (%)	48.67	48.91	48.87	48.87
Married (%)	46.23	45.77	45.14	44.87
Lake Geneva Region (%)	15.80	15.12	14.84	14.65
Espace Mittelland (%)	24.23	25.07	25.25	25.31
Northwestern Switzerland (%)	13.33	13.05	13.12	13.09
Zürich (%)	17.83	17.57	17.75	17.78
Eastern Switzerland (%)	14.09	14.29	14.23	14.27
Central Switzerland (%)	10.72	11.26	11.28	11.38
Ticino (%)	4.00	3.64	3.53	3.52
German (%)	74.31	75.36	75.79	76.03
French (%)	21.18	20.47	20.16	19.92
Latin (%)	4.50	4.17	4.05	4.05
Child Income at Age 30-33				
Average Income	59,103	59,734	60,598	60,896
Top 10%	124,760	125,350	126,510	126,712
Top 5%	142,275	142,865	144,233	144,387
Top 1%	183,863	184,486	186,735	186,695
Obs.	1,266,376	1,114,543	923,107	859,286
Panel B: Education Sample				
High-School (%)	19.96	20.07	20.99	21.24
VET(%)	66.05	66.09	65.69	65.61
Master (%)	15.70	15.77	16.43	16.62
Obs.	380,018	371,269	308,622	287,848

Notes: The table provides summary statistics by sample. Panel (A) shows the summary statistics for the income sample. Column (1) reports the mean of all individuals born in Switzerland between 1967 and 1984. Column (2) to column (4) are sub-samples of column (1). Column (2) restricts to individuals that can be linked to the mother or the father. Column (3) is our core sample: individuals that can be linked to the father. Column (4) is the most conservative sample, it restricts to individuals matched to both parents. Panel (B) shows the summary statistics for the education sample. All amounts are in 2017 CHF.

TABLE C.3: Relative Mobility Estimates for Different Samples

		Intergenrational correlation			
		Rank-Rank-Slope		IGE	
		(1)		(2)	
A. Varying Child Age					
Core Sample: 1967-1984	Child age 30-33	0.141	(0.0010)	0.166	(0.0017)
Birth cohorts 1967-1981	Child age 33-36	0.146	(0.0011)	0.191	(0.0021)
Birth cohorts 1967-1978	Child age 36-39	0.145	(0.0013)	0.205	(0.0025)
Birth cohorts 1967-1974	Child age 40-43	0.147	(0.0016)	0.219	(0.0030)
Birth cohorts 1967-1971	Child age 43-46	0.145	(0.0020)	0.212	(0.0037)
B. Varying Family Age					
Birth cohorts 1979-1984	Child age 3-8	0.136	(0.0017)	0.177	(0.0033)
Birth cohorts 1973-1984	Child age 9-14	0.138	(0.0013)	0.172	(0.0022)
C. Alternative Income Definitions					
Excl. Missing Incomes		0.096	(0.0174)	0.081	(0.0284)
Recoding non labor income to 0		0.137	(0.0010)	0.165	(0.0017)

Notes: This table reports the baseline estimates and the results of OLS regressions of a measure of child income on a measure of parents' income for several samples. Column (1) reports the coefficient of the rank-rank slope and standard errors in parentheses, column (2) reports the IGE coefficient and the standard error in parentheses. Panel (A) shows the value of the RRS and IGE when child income is measured later in life compared to the baseline specification. Family income is measured when the child is between 15 and 20. Panel (B) shows the value of the RRS and IGE when family income is measured earlier relative to the baseline specification, when the children are between 3 and 8 years old, or between 9 and 14. Panel (C) shows the value of the RRS and the IGE when we only include observations for which we observe every income record between the ages of 30 and 33 for children and between the ages of 15 and 20 for families.

TABLE C.4: Access and Upmover Rate by Educational Track (Medium Upward Mobility)

Educational Track	Upmover Rate		Access Rate		Mobility Rate		N
	$P[R_C > 50 \mid R_P \leq 50]$		$P[R_P \leq 50]$		$P[R_C > 50 \wedge R_P \leq 50]$		
VET only	0.399	(0.002)	0.599	(0.001)	0.239	(0.001)	78,550
Gym only	0.419	(0.007)	0.346	(0.004)	0.145	(0.003)	5,329
Gym + Bachelor	0.481	(0.010)	0.289	(0.005)	0.139	(0.004)	2,421
Gym + Master	0.649	(0.006)	0.240	(0.002)	0.155	(0.002)	7,170
Gym + PhD	0.653	(0.013)	0.218	(0.005)	0.142	(0.004)	1,414
VET + Bachelor	0.660	(0.008)	0.404	(0.005)	0.267	(0.005)	3,416
VET + HF	0.680	(0.006)	0.435	(0.004)	0.295	(0.004)	6,100
VET + HFP	0.703	(0.003)	0.514	(0.002)	0.361	(0.002)	24,949
VET + PhD	0.416	(0.044)	0.398	(0.028)	0.166	(0.021)	125
VET + Voc Matura	0.491	(0.009)	0.473	(0.006)	0.232	(0.005)	2,910
Mandatory or less	0.178	(0.004)	0.657	(0.004)	0.117	(0.003)	9,175
Spec. middle school	0.332	(0.007)	0.529	(0.005)	0.176	(0.004)	5,210

Notes: This table shows th upmover, access, and mobility rate for different educational tracks. Upward mobility is defined as moving from the bottom half to the top half of the income distribution.

TABLE C.5: Mean Child and Family Income by Rank

Rank	Child Income	Family Income	Rank	Child Income	Family Income	Rank	Child Income	Family Income	Rank	Child Income	Family Income
1	0	3,719	26	35,929	41,520	51	62,639	55,631	76	81,896	75,247
2	126	10,554	27	37,367	42,129	52	63,407	56,228	77	82,854	76,394
3	858	14,228	28	38,757	42,730	53	64,156	56,838	78	83,861	77,608
4	2,106	16,830	29	40,107	43,318	54	64,906	57,457	79	84,919	78,869
5	3,372	19,000	30	41,427	43,890	55	65,656	58,079	80	86,004	80,204
6	3,133	20,908	31	42,704	44,458	56	66,383	58,709	81	87,144	81,592
7	4,477	22,644	32	43,946	45,023	57	67,096	59,354	82	88,315	83,071
8	5,772	24,297	33	45,153	45,581	58	67,813	60,013	83	89,566	84,671
9	7,263	25,862	34	46,353	46,141	59	68,522	60,679	84	90,862	86,388
10	8,813	27,322	35	47,497	46,693	60	69,232	61,367	85	92,205	88,193
11	10,454	28,715	36	48,609	47,246	61	69,930	62,078	86	93,649	90,163
12	12,209	30,018	37	49,685	47,798	62	70,630	62,800	87	95,173	92,363
13	14,008	31,228	38	50,754	48,348	63	71,343	63,546	88	96,815	94,760
14	15,829	32,342	39	51,790	48,897	64	72,061	64,305	89	98,585	97,449
15	17,637	33,369	40	52,811	49,450	65	72,789	65,084	90	100,492	100,524
16	19,456	34,304	41	53,811	50,003	66	73,530	65,880	91	102,606	104,020
17	21,271	35,185	42	54,806	50,550	67	74,279	66,689	92	104,933	108,078
18	23,031	36,011	43	55,764	51,104	68	75,048	67,507	93	107,569	112,908
19	24,764	36,791	44	56,705	51,652	69	75,831	68,360	94	110,609	118,867
20	26,471	37,530	45	57,618	52,210	70	76,642	69,264	95	114,237	126,510
21	28,128	38,250	46	58,510	52,760	71	77,462	70,181	96	118,826	136,106
22	29,746	38,935	47	59,373	53,323	72	78,290	71,121	97	124,954	148,861
23	31,355	39,611	48	60,224	53,891	73	79,154	72,096	98	133,953	168,553
24	32,909	40,265	49	61,055	54,462	74	80,040	73,106	99	149,561	203,690
25	34,426	40,901	50	61,857	55,049	75	80,955	74,160	100	223,951	379,043

Notes: This table shows the mean real income in 2017 Swiss Francs for fathers and children.

TABLE C.6: National Estimates

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Sample							
Child's outcome	Parent's inc def	Core sample	1979-1981 Cohorts	Male children	Female Children	Foreign Father	Swiss Father	Fixed age at child birth	Married
IGE:									
Log individual income excluding zeros	Log father income	0.166	0.160	0.111	0.235	0.122	0.168	0.171	0.167
		(0.0017)	(0.0038)	(0.0017)	(0.0028)	(0.0086)	(0.0017)	(0.0028)	(0.0019)
recoding 0 to 1	Log father income	0.188	0.176	0.083	0.299	0.086	0.198	0.213	0.202
		(0.0028)	(0.0059)	(0.0021)	(0.0050)	(0.0092)	(0.0029)	(0.0049)	(0.0031)
recoding 0 to 1000	Log father income	0.192	0.177	0.103	0.284	0.117	0.196	0.199	0.198
		(0.0019)	(0.0042)	(0.0018)	(0.0032)	(0.0086)	(0.0020)	(0.0032)	(0.0021)
RRS:									
Individual income rank	Family income rank	0.141	0.142	0.122	0.162	0.114	0.142	0.143	0.147
		(0.0010)	(0.0025)	(0.0013)	(0.0014)	(0.0057)	(0.0011)	(0.0017)	(0.0011)
Individual income rank	Father income rank	0.153	0.154	0.152	0.152	0.132	0.154	0.154	0.152
		(0.0010)	(0.0025)	(0.0013)	(0.0014)	(0.0059)	(0.0011)	(0.0017)	(0.0011)
Individual income rank	Mother income rank	0.026	0.025	0.004	0.053	0.020	0.027	0.030	0.036
		(0.0009)	(0.0028)	(0.0012)	(0.0012)	(0.0048)	(0.0009)	(0.0015)	(0.0010)
Observations		923,107	155,003	471,989	451,118	35,059	888,048	333,097	775,377

Notes: Each cell reports the coefficient from an OLS regression of a child's outcome on a measure of its family income. Column (1) uses the core sample, which includes all children (i) born in birth cohorts 1969 to 1989 (ii) for whom we have been able to link both parents (iii) whose mean income at age 30 to 33 is positive and (iv) whose mean parent's income when child is between 15-20 is non-negative. Column (2) reports the estimates for birth cohorts 1979 to 1981. Columns (3) and (4) restrict the sample to male and female. Columns(5) and (6) limit the sample to children whose father is either foreign or Swiss. Column (8) estimates income mobility among children whose mother was between 13 and 19 years old at child birth. Column (8) limits the core sample to children whose father fall into a 5-year window of median father age at time of child birth. Column (9) restricts the sample to children whose parents are still married in 2012 and live in the same household in 2012. Child income is the mean of the individual income between age 30 to 33, while parent family income is the mean income when the child is between age 15 and 20. Individual earnings include wage earnings, self-employment earnings, unemployment insurance and disability benefits. Income percentile ranks are constructed by ranking all children relative to other children in the same birth cohort, and ranking parents relative to other parents in the core sample. Ranks are not redefined within sub-samples except in column (2). The number of observations correspond to the specification in row 4.

TABLE C.7: National Quintile Transition Matrix (sample with both parents)

	Family quintile				
Child quintile	1	2	3	4	5
1	23.75	21.72	19.59	17.88	17.06
2	21.57	20.31	19.53	19.11	19.48
3	20.98	22.07	20.52	18.90	17.54
4	17.57	21.00	21.38	20.83	19.22
5	11.74	15.22	19.00	23.02	31.02

Notes: This table shows in which quintile children end up in the income distribution for every quintile of the parent's income distribution. Each cell describes in which quintile (row) children end up conditional on the parent quintile (column). Parent quintile is based on father and mother income. For example, 16.3% of children from parents in the top quintile of the income distribution will end up in the bottom quintile of the income distribution. 11.74% of children from parents of the bottom quintile of the income distribution end up in the top quintile («American Dream measure»). Income ranks are measured relative to child cohort. The table includes children born in Switzerland from 1967 to 1984 and comprises 849,849 observations (child-father pairs).

TABLE C.8: IGE by Age and Parent Tertile

	Child Income Age 30-33					Child Income Age 38-41				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	Std	Fam T1	Fam T2	Fam T3	All	Std	Fam T1	Fam T2	Fam T3
IGE	0.166***		0.100***	0.414***	0.119***	0.215***		0.089***	0.423***	0.222***
	(0.0017)		(0.0037)	(0.0153)	(0.0047)	(0.0027)		(0.0059)	(0.0241)	(0.0078)
IGE (std)		0.103***					0.116***			
		(0.0011)					(0.0015)			
Observations	894, 399	894, 399	295, 535	298, 666	300, 198	476, 098	476, 098	157, 459	159, 087	159, 552
R-Squared	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.01

Notes: This table shows the IGE for different ages at which child income is measured and for different parts of the parental income distribution. T1 refers to the lowest parent income tertile, T3 refers to the highest parental income tertile. Columns (2) and (7) show the standardized IGE coefficient. This means, the log-incomes are divided by the standard deviation in order to abstract from changes in inequality over time.

TABLE C.9: Income Mobility by Labor Market Region

	RRS		Q1Q5		Q1Q1		Q5Q5		IGE		AM P=25	.
LaborMarket												
Aarau	0.145	(0.006)	0.132	(0.005)	0.225	(0.006)	0.305	(0.007)	0.231	(0.022)	48.047	(0.234)
Aaretal	0.148	(0.010)	0.094	(0.006)	0.258	(0.008)	0.259	(0.013)	0.315	(0.030)	45.108	(0.325)
Aigle	0.122	(0.017)	0.100	(0.011)	0.269	(0.016)	0.263	(0.019)	0.291	(0.054)	44.538	(0.599)
Appenzell A.Rh.	0.180	(0.011)	0.106	(0.007)	0.289	(0.010)	0.320	(0.014)	0.251	(0.031)	45.269	(0.388)
Appenzell I.Rh.	0.186	(0.022)	0.107	(0.010)	0.318	(0.014)	0.303	(0.034)	0.192	(0.050)	44.952	(0.647)
Baden	0.130	(0.012)	0.189	(0.014)	0.232	(0.015)	0.330	(0.009)	0.221	(0.039)	50.816	(0.499)
Basel-Stadt	0.148	(0.010)	0.121	(0.008)	0.301	(0.011)	0.273	(0.008)	0.316	(0.031)	44.527	(0.404)
Bellinzona	0.118	(0.015)	0.102	(0.008)	0.242	(0.011)	0.258	(0.019)	0.267	(0.043)	46.802	(0.453)
Bern	0.138	(0.006)	0.119	(0.005)	0.250	(0.007)	0.274	(0.005)	0.270	(0.018)	46.272	(0.241)
Biel/Bienne	0.139	(0.011)	0.120	(0.008)	0.248	(0.011)	0.258	(0.011)	0.283	(0.036)	45.345	(0.392)
Brig	0.135	(0.018)	0.171	(0.012)	0.231	(0.013)	0.304	(0.023)	0.203	(0.055)	50.194	(0.521)
Brugg-Zurzach	0.146	(0.011)	0.167	(0.011)	0.217	(0.012)	0.329	(0.011)	0.247	(0.040)	49.385	(0.429)
Burgdorf	0.166	(0.009)	0.079	(0.005)	0.242	(0.007)	0.278	(0.013)	0.210	(0.029)	45.018	(0.291)
Chur	0.136	(0.011)	0.134	(0.009)	0.257	(0.011)	0.273	(0.012)	0.204	(0.039)	46.570	(0.402)
Davos	0.107	(0.029)	0.117	(0.021)	0.188	(0.025)	0.310	(0.031)	0.195	(0.104)	47.648	(1.068)
Einsiedeln	0.166	(0.019)	0.108	(0.010)	0.276	(0.014)	0.298	(0.027)	0.207	(0.054)	46.087	(0.590)
Engiadina Bassa	0.156	(0.026)	0.083	(0.012)	0.288	(0.020)	0.261	(0.036)	0.174	(0.069)	43.466	(0.776)
Entlebuch	0.128	(0.021)	0.099	(0.007)	0.245	(0.010)	0.292	(0.042)	0.140	(0.049)	45.915	(0.489)
Erlach-Seeland	0.112	(0.012)	0.108	(0.008)	0.205	(0.010)	0.256	(0.015)	0.254	(0.040)	47.155	(0.386)
Freiamt	0.124	(0.012)	0.158	(0.009)	0.192	(0.010)	0.332	(0.014)	0.194	(0.043)	50.466	(0.420)
Fricktal	0.110	(0.012)	0.140	(0.009)	0.224	(0.011)	0.307	(0.013)	0.251	(0.041)	49.140	(0.422)
Genève	0.099	(0.006)	0.232	(0.007)	0.245	(0.007)	0.328	(0.005)	0.137	(0.017)	50.610	(0.304)
Glarner Hinterland	0.175	(0.028)	0.111	(0.015)	0.262	(0.022)	0.357	(0.043)	0.108	(0.078)	46.314	(0.833)
Glarner Unterland	0.131	(0.018)	0.134	(0.013)	0.248	(0.016)	0.269	(0.020)	0.128	(0.049)	47.794	(0.614)
Glattal-Furttal	0.134	(0.009)	0.181	(0.011)	0.214	(0.011)	0.349	(0.007)	0.235	(0.029)	50.877	(0.422)
Glâne-Veveyse	0.155	(0.016)	0.117	(0.008)	0.225	(0.010)	0.276	(0.026)	0.198	(0.042)	46.647	(0.453)
Goms	0.126	(0.038)	0.117	(0.017)	0.255	(0.022)	0.259	(0.058)	0.314	(0.109)	46.098	(0.958)
Grenchen	0.131	(0.017)	0.124	(0.013)	0.205	(0.016)	0.318	(0.020)	0.190	(0.047)	48.329	(0.587)
Gros-de-Vaud	0.123	(0.013)	0.118	(0.010)	0.222	(0.012)	0.266	(0.014)	0.227	(0.050)	45.639	(0.495)
Innerschwyz	0.184	(0.010)	0.102	(0.005)	0.271	(0.008)	0.328	(0.014)	0.229	(0.029)	46.481	(0.329)
Jura	0.158	(0.010)	0.108	(0.005)	0.252	(0.008)	0.284	(0.015)	0.206	(0.030)	44.737	(0.305)
Jura bernois	0.147	(0.015)	0.101	(0.009)	0.275	(0.013)	0.261	(0.020)	0.385	(0.044)	44.410	(0.477)
Kandertal	0.164	(0.020)	0.064	(0.006)	0.286	(0.012)	0.205	(0.032)	0.231	(0.055)	40.815	(0.489)
Knonaueramt	0.126	(0.014)	0.165	(0.015)	0.229	(0.016)	0.334	(0.011)	0.246	(0.046)	50.392	(0.647)
La Broye	0.135	(0.012)	0.125	(0.007)	0.247	(0.009)	0.275	(0.015)	0.156	(0.038)	46.012	(0.375)
La Chaux-de-Fonds	0.169	(0.012)	0.114	(0.008)	0.278	(0.011)	0.248	(0.016)	0.229	(0.037)	44.218	(0.402)
La Gruyère	0.149	(0.014)	0.112	(0.008)	0.237	(0.011)	0.286	(0.020)	0.261	(0.043)	46.302	(0.445)
La Sarine	0.155	(0.010)	0.119	(0.007)	0.227	(0.009)	0.301	(0.011)	0.264	(0.030)	46.994	(0.365)
La Vallée	0.100	(0.031)	0.150	(0.023)	0.259	(0.028)	0.228	(0.035)	0.172	(0.096)	46.711	(1.060)
Laufental	0.135	(0.013)	0.124	(0.010)	0.240	(0.012)	0.323	(0.015)	0.266	(0.043)	48.359	(0.472)
Lausanne	0.134	(0.008)	0.159	(0.008)	0.257	(0.009)	0.295	(0.006)	0.251	(0.024)	47.054	(0.347)
Leuk	0.159	(0.028)	0.144	(0.016)	0.212	(0.018)	0.256	(0.046)	0.213	(0.079)	48.784	(0.712)
Limmattal	0.121	(0.014)	0.189	(0.016)	0.229	(0.018)	0.356	(0.011)	0.194	(0.045)	51.731	(0.631)
Linthgebiet	0.147	(0.012)	0.152	(0.008)	0.248	(0.010)	0.348	(0.013)	0.231	(0.034)	48.952	(0.423)
Locarno	0.129	(0.011)	0.104	(0.006)	0.247	(0.009)	0.270	(0.014)	0.197	(0.032)	44.570	(0.378)
Lugano	0.131	(0.009)	0.111	(0.006)	0.248	(0.008)	0.286	(0.010)	0.192	(0.025)	45.204	(0.327)
Luzern	0.140	(0.007)	0.136	(0.005)	0.223	(0.006)	0.311	(0.007)	0.219	(0.020)	48.538	(0.244)
March	0.151	(0.012)	0.168	(0.010)	0.230	(0.011)	0.359	(0.012)	0.247	(0.035)	50.078	(0.476)
Martigny	0.146	(0.012)	0.109	(0.007)	0.244	(0.009)	0.321	(0.017)	0.264	(0.036)	44.448	(0.372)
Mendrisio	0.130	(0.014)	0.126	(0.009)	0.237	(0.011)	0.287	(0.017)	0.162	(0.039)	47.794	(0.449)
Mesolcina	0.154	(0.033)	0.074	(0.015)	0.257	(0.025)	0.293	(0.051)	0.319	(0.126)	45.098	(0.932)
Mittelbünden	0.128	(0.026)	0.137	(0.016)	0.226	(0.020)	0.302	(0.034)	0.239	(0.070)	47.733	(0.822)
Monthey	0.153	(0.015)	0.110	(0.009)	0.247	(0.013)	0.248	(0.018)	0.265	(0.047)	44.896	(0.490)
Morges	0.107	(0.012)	0.166	(0.012)	0.212	(0.013)	0.292	(0.010)	0.218	(0.038)	47.730	(0.541)
Murten/Morat	0.126	(0.012)	0.140	(0.008)	0.203	(0.010)	0.286	(0.016)	0.163	(0.038)	47.096	(0.410)
Mutschellen	0.156	(0.012)	0.169	(0.013)	0.215	(0.014)	0.355	(0.011)	0.301	(0.041)	50.015	(0.525)
Neuchâtel	0.154	(0.010)	0.132	(0.008)	0.238	(0.010)	0.298	(0.010)	0.237	(0.033)	45.655	(0.391)
Nidwalden	0.165	(0.012)	0.104	(0.007)	0.244	(0.010)	0.326	(0.016)	0.234	(0.035)	47.613	(0.406)
Nyon	0.102	(0.014)	0.204	(0.015)	0.209	(0.016)	0.320	(0.010)	0.116	(0.035)	49.351	(0.682)
Oberaargau	0.157	(0.009)	0.087	(0.005)	0.234	(0.007)	0.275	(0.013)	0.262	(0.030)	45.932	(0.280)
Oberengadin	0.206	(0.019)	0.106	(0.012)	0.259	(0.017)	0.348	(0.022)	0.291	(0.062)	45.292	(0.679)
Oberes Baselbiet	0.123	(0.010)	0.115	(0.007)	0.227	(0.010)	0.277	(0.010)	0.234	(0.033)	47.347	(0.370)
Oberes Emmental	0.195	(0.015)	0.069	(0.005)	0.244	(0.008)	0.321	(0.030)	0.202	(0.039)	43.659	(0.358)
Oberland-Ost	0.168	(0.013)	0.078	(0.006)	0.275	(0.010)	0.253	(0.018)	0.269	(0.039)	43.604	(0.368)
Oberthurgau	0.149	(0.012)	0.133	(0.008)	0.268	(0.010)	0.322	(0.016)	0.239	(0.041)	46.025	(0.403)
Olten	0.136	(0.011)	0.159	(0.009)	0.209	(0.010)	0.335	(0.012)	0.205	(0.034)	49.365	(0.384)
Pays d'Enhaut	0.128	(0.041)	0.088	(0.019)	0.301	(0.031)	0.145	(0.048)	0.232	(0.130)	41.738	(1.205)
Pfannenstiel	0.122	(0.011)	0.194	(0.013)	0.240	(0.014)	0.348	(0.007)	0.179	(0.028)	50.145	(0.547)
Prättigau	0.209	(0.021)	0.088	(0.010)	0.309	(0.016)	0.293	(0.029)	0.233	(0.071)	42.315	(0.639)
Rheintal	0.130	(0.013)	0.125	(0.008)	0.284	(0.011)	0.285	(0.016)	0.158	(0.041)	46.125	(0.404)
Saanen-Obersimmental	0.179	(0.019)	0.072	(0.007)	0.254	(0.012)	0.283	(0.032)	0.295	(0.055)	42.965	(0.520)
Sarganserland	0.137	(0.015)	0.112	(0.008)	0.275	(0.011)	0.281	(0.021)	0.187	(0.050)	44.087	(0.436)
Sarneraatal	0.173	(0.014)	0.090	(0.006)	0.257	(0.010)	0.258	(0.021)	0.208	(0.041)	47.214	(0.412)
Schaffhausen	0.133	(0.012)	0.131	(0.010)	0.226	(0.012)	0.307	(0.012)	0.284	(0.040)	47.443	(0.452)
Schanfigg	0.205	(0.043)	0.122	(0.026)	0.205	(0.032)	0.368	(0.056)	0.281	(0.126)	45.979	(1.382)
Schwarzwasser	0.158	(0.018)	0.074	(0.007)	0.255	(0.012)	0.241	(0.030)	0.233	(0.051)	44.431	(0.501)
Sense	0.156	(0.013)	0.106	(0.007)	0.227	(0.009)	0.302	(0.020)	0.160	(0.038)	47.006	(0.366)
Sierre	0.139	(0.014)	0.143	(0.011)	0.218	(0.012)	0.322	(0.017)	0.260	(0.041)	47.499	(0.483)
Sion	0.150	(0.011)	0.123	(0.007)	0.215	(0.009)	0.318	(0.013)	0.192	(0.030)	47.295	(0.352)
Solothurn	0.143	(0.010)	0.131	(0.008)	0.214	(0.010)	0.312	(0.011)	0.269	(0.032)	47.992	(0.372)
St.Gallen	0.166	(0.007)	0.139	(0.006)	0.253	(0.007)	0.337	(0.008)	0.269	(0.023)	47.364	(0.272)
Sursee-Seetal	0.126	(0.009)	0.134	(0.006)	0.204	(0.007)	0.313	(0.013)	0.185	(0.029)	49.284	(0.301)
Surselva	0.149	(0.016)	0.098	(0.008)	0.257	(0.011)	0.261	(0.021)	0.209	(0.047)	45.319	(0.472)
Thal	0.201	(0.024)	0.115	(0.013)	0.260	(0.018)	0.365	(0.039)	0.384	(0.082)	47.032	(0.703)
Thun	0.154	(0.008)	0.083	(0.004)	0.265	(0.007)	0.225	(0.009)	0.279	(0.024)	44.216	(0.244)
Thurtal	0.103	(0.010)	0.158	(0.007)	0.216	(0.008)	0.287	(0.011)	0.245	(0.034)	48.875	(0.340)
Toggenburg	0.224	(0.013)	0.098	(0.006)	0.292	(0.009)	0.318	(0.020)	0.279	(0.034)	44.920	(0.386)
Tre Valli	0.133	(0.021)	0.106	(0.009)	0.219	(0.012)	0.272	(0.036)	0.091	(0.049)	46.600	(0.526)
Unteres Baselbiet	0.129	(0.008)	0.146	(0.009)	0.223	(0.011)	0.315	(0.007)	0.331	(0.025)	48.419	(0.370)
Untersee	0.121	(0.013)	0.144	(0.009)	0.243	(0.011)	0.289	(0.015)	0.153	(0.042)	47.388	(0.455)
Uri	0.233	(0.014)	0.092	(0.006)	0.294	(0.009)	0.316	(0.024)	0.198	(0.039)	45.106	(0.367)
Val-de-Travers	0.151	(0.028)	0.105	(0.014)	0.275	(0.021)	0.265	(0.042)	0.247	(0.077)	43.771	(0.829)
Vevey	0.131	(0.012)	0.143	(0.011)	0.247	(0.013)	0.290	(0.011)	0.238	(0.035)	46.373	(0.510)
Viamala	0.188	(0.023)	0.084	(0.011)	0.273	(0.018)	0.263	(0.032)	0.373	(0.075)	43.223	(0.704)
Visp	0.116	(0.016)	0.136	(0.009)	0.221	(0.011)	0.331	(0.023)	0.119	(0.049)	48.208	(0.445)
Weinland	0.158	(0.017)	0.144	(0.014)	0.263	(0.018)	0.315	(0.017)	0.308	(0.058)	46.891	(0.675)
Werdenberg	0.130	(0.016)	0.131	(0.008)	0.336	(0.012)	0.259	(0.021)	0.053	(0.036)	41.950	(0.493)
Wil	0.143	(0.009)	0.158	(0.006)	0.244	(0.007)	0.313	(0.011)	0.189	(0.027)	48.510	(0.311)
Willisau	0.159	(0.011)	0.107	(0.005)	0.237	(0.006)	0.313	(0.019)	0.152	(0.030)	47.716	(0.286)
Winterthur	0.122	(0.008)	0.164	(0.008)	0.236	(0.010)	0.315	(0.007)	0.289	(0.028)	49.091	(0.347)
Yverdon	0.130	(0.014)	0.107	(0.009)	0.278	(0.014)	0.216	(0.015)	0.346	(0.047)	44.440	(0.496)
Zimmerberg	0.143	(0.010)	0.160	(0.012)	0.241	(0.013)	0.344	(0.007)	0.286	(0.032)	49.860	(0.493)
Zug	0.141	(0.009)	0.168	(0.009)	0.212	(0.010)	0.362	(0.009)	0.271	(0.026)	51.037	(0.394)
Zürcher Oberland	0.132	(0.008)	0.150	(0.008)	0.246	(0.009)	0.314	(0.007)	0.285	(0.026)	48.771	(0.335)
Zürcher Unterland	0.134	(0.011)	0.164	(0.011)	0.225	(0.013)	0.333	(0.009)	0.245	(0.035)	50.027	(0.467)
Zürich	0.099	(0.008)	0.191	(0.009)	0.265	(0.010)	0.298	(0.006)	0.224	(0.024)	49.550	(0.362)

Notes: This table shows the income mobility estimates by labor market regions (n=106). RRS indicates the rank-rank slope, *Q1Q5* is the American Dream measure, Q1Q1 is the cycle of poverty measure, *Q5Q5* is the cycle of privileges measure, IGE is the intergenerational elasticity, AUM P=25 shows the expected rank of children below the median of the income distribution. The estimates are based on 923,107 observations. Corresponding standard errors are shown in parentheses.

TABLE C.10: Income Mobility by Canton

	RRS		Q1Q5		Q1Q1		Q5Q5		IGE		AM P=25	.
Canton												
AG	0.142	(0.004)	0.150	(0.004)	0.219	(0.004)	0.324	(0.004)	0.242	(0.014)	49.171	(0.157)
AI	0.181	(0.021)	0.110	(0.009)	0.306	(0.013)	0.307	(0.032)	0.188	(0.047)	45.253	(0.604)
AR	0.182	(0.011)	0.104	(0.007)	0.294	(0.010)	0.319	(0.014)	0.254	(0.032)	45.145	(0.399)
BE	0.157	(0.003)	0.090	(0.002)	0.251	(0.002)	0.265	(0.003)	0.275	(0.008)	44.958	(0.092)
BL	0.132	(0.006)	0.130	(0.005)	0.226	(0.007)	0.305	(0.005)	0.299	(0.019)	47.855	(0.249)
BS	0.148	(0.010)	0.121	(0.008)	0.301	(0.011)	0.273	(0.008)	0.316	(0.031)	44.527	(0.404)
FR	0.148	(0.005)	0.118	(0.003)	0.226	(0.004)	0.294	(0.007)	0.224	(0.016)	46.797	(0.178)
GE	0.099	(0.006)	0.232	(0.007)	0.245	(0.007)	0.328	(0.005)	0.137	(0.017)	50.610	(0.304)
GL	0.145	(0.015)	0.126	(0.010)	0.253	(0.013)	0.288	(0.018)	0.129	(0.041)	47.287	(0.495)
GR	0.164	(0.006)	0.106	(0.004)	0.262	(0.005)	0.289	(0.008)	0.238	(0.020)	45.304	(0.209)
JU	0.158	(0.010)	0.108	(0.005)	0.252	(0.008)	0.284	(0.015)	0.206	(0.030)	44.737	(0.305)
LU	0.146	(0.004)	0.121	(0.003)	0.226	(0.003)	0.311	(0.006)	0.206	(0.013)	48.239	(0.144)
NE	0.160	(0.008)	0.125	(0.006)	0.260	(0.008)	0.285	(0.009)	0.234	(0.024)	45.031	(0.286)
NW	0.168	(0.013)	0.099	(0.007)	0.240	(0.010)	0.334	(0.017)	0.233	(0.037)	47.722	(0.421)
OW	0.168	(0.013)	0.094	(0.006)	0.259	(0.010)	0.259	(0.019)	0.213	(0.038)	47.119	(0.398)
SG	0.169	(0.004)	0.130	(0.003)	0.274	(0.004)	0.323	(0.005)	0.211	(0.012)	46.353	(0.145)
SH	0.133	(0.012)	0.131	(0.010)	0.226	(0.012)	0.307	(0.012)	0.284	(0.040)	47.443	(0.452)
SO	0.141	(0.006)	0.135	(0.005)	0.221	(0.006)	0.323	(0.007)	0.261	(0.020)	48.570	(0.223)
SZ	0.178	(0.007)	0.121	(0.005)	0.261	(0.006)	0.341	(0.009)	0.244	(0.021)	47.470	(0.252)
TG	0.121	(0.006)	0.150	(0.004)	0.237	(0.005)	0.295	(0.007)	0.218	(0.020)	47.846	(0.205)
TI	0.126	(0.006)	0.110	(0.003)	0.241	(0.004)	0.279	(0.007)	0.186	(0.015)	45.927	(0.183)
UR	0.233	(0.014)	0.092	(0.006)	0.294	(0.009)	0.316	(0.024)	0.198	(0.039)	45.106	(0.367)
VD	0.133	(0.004)	0.140	(0.004)	0.247	(0.004)	0.288	(0.004)	0.244	(0.013)	46.285	(0.170)
VS	0.141	(0.005)	0.129	(0.003)	0.229	(0.004)	0.308	(0.007)	0.227	(0.016)	47.007	(0.168)
ZG	0.141	(0.009)	0.168	(0.009)	0.212	(0.010)	0.362	(0.009)	0.271	(0.026)	51.037	(0.394)
ZH	0.128	(0.003)	0.170	(0.003)	0.240	(0.004)	0.329	(0.003)	0.253	(0.010)	49.629	(0.142)

Notes: This table shows the income mobility estimates by cantons (n=26). RRS indicates the rank-rank slope, *Q1Q5* is the American Dream measure, Q1Q1 is the cycle of poverty measure, *Q5Q5* is the cycle of privileges measure, IGE is the intergenerational elasticity, AUM P=25 shows the expected rank of children below the median of the income distribution. The estimates are based on 923,107 observations. Corresponding standard errors are shown in parentheses.

TABLE C.11: Educational Mobility by Canton

	Share Bottom 20 in HS		Child-Parent Years Edu		Child-Parent HS	
Canton						
AG	0.084	(0.005)	0.247	(0.006)	0.353	(0.009)
AI	0.077	(0.015)	0.201	(0.037)	0.414	(0.072)
AR	0.089	(0.012)	0.251	(0.023)	0.336	(0.039)
BE	0.076	(0.003)	0.263	(0.006)	0.392	(0.008)
BL	0.121	(0.010)	0.259	(0.013)	0.377	(0.017)
BS	0.171	(0.019)	0.309	(0.020)	0.435	(0.027)
FR	0.112	(0.006)	0.268	(0.011)	0.380	(0.020)
GE	0.273	(0.012)	0.265	(0.009)	0.247	(0.013)
GL	0.083	(0.016)	0.234	(0.033)	0.245	(0.051)
GR	0.096	(0.007)	0.277	(0.014)	0.410	(0.022)
JU	0.099	(0.008)	0.229	(0.015)	0.318	(0.025)
LU	0.076	(0.003)	0.234	(0.006)	0.369	(0.010)
NE	0.129	(0.009)	0.282	(0.011)	0.313	(0.016)
NW	0.081	(0.012)	0.187	(0.027)	0.281	(0.049)
OW	0.071	(0.011)	0.265	(0.029)	0.259	(0.063)
SG	0.098	(0.005)	0.221	(0.009)	0.346	(0.014)
SH	0.094	(0.017)	0.278	(0.025)	0.399	(0.033)
SO	0.100	(0.008)	0.245	(0.013)	0.411	(0.019)
SZ	0.072	(0.007)	0.263	(0.015)	0.403	(0.027)
TG	0.079	(0.005)	0.209	(0.010)	0.291	(0.015)
TI	0.149	(0.005)	0.268	(0.009)	0.316	(0.013)
UR	0.062	(0.009)	0.297	(0.023)	0.527	(0.044)
VD	0.148	(0.005)	0.294	(0.006)	0.311	(0.008)
VS	0.121	(0.006)	0.243	(0.011)	0.343	(0.018)
ZG	0.093	(0.011)	0.270	(0.013)	0.377	(0.020)
ZH	0.128	(0.006)	0.277	(0.006)	0.347	(0.009)

Notes: This table shows the educational mobility estimates by cantons (n=26). Share Bottom 20 in HS shows the share of children from the bottom quintile in the national parental income distribution that visit a high school (gymnasium), *Child-Parent Years Edu* shows the correlation in years of education between children and parents, Child-Parent HS shows how much more likely children are to visit a high school if at least one of their parents went to high school as well. Corresponding standard errors are shown in parentheses.

TABLE C.12: Educational Mobility by Labor Market Region

	Share Bottom 20 in HS		Child-Parent Years Edu		Child-Parent HS	
LaborMarket						
Aarau	0.070	(0.006)	0.240	(0.010)	0.357	(0.014)
Aaretal	0.073	(0.009)	0.242	(0.023)	0.443	(0.029)
Aigle	0.110	(0.018)	0.289	(0.026)	0.397	(0.033)
Appenzell A.Rh.	0.085	(0.012)	0.251	(0.022)	0.336	(0.038)
Appenzell I.Rh.	0.083	(0.016)	0.198	(0.040)	0.438	(0.077)
Baden	0.148	(0.021)	0.258	(0.017)	0.303	(0.024)
Basel-Stadt	0.171	(0.019)	0.309	(0.020)	0.435	(0.027)
Bellinzona	0.125	(0.013)	0.318	(0.024)	0.244	(0.031)
Bern	0.123	(0.010)	0.280	(0.011)	0.381	(0.014)
Biel/Bienne	0.125	(0.016)	0.247	(0.026)	0.363	(0.031)
Brig	0.121	(0.020)	0.202	(0.033)	0.330	(0.050)
Brugg-Zurzach	0.085	(0.013)	0.241	(0.017)	0.404	(0.026)
Burgdorf	0.085	(0.009)	0.262	(0.021)	0.332	(0.030)
Chur	0.107	(0.015)	0.254	(0.024)	0.439	(0.037)
Davos	0.090	(0.035)	0.200	(0.064)	0.362	(0.093)
Einsiedeln	0.065	(0.016)	0.174	(0.036)	0.130	(0.071)
Engiadina Bassa	0.088	(0.023)	0.276	(0.060)	0.324	(0.090)
Entlebuch	0.058	(0.008)	0.168	(0.023)	0.431	(0.050)
Erlach-Seeland	0.085	(0.013)	0.244	(0.027)	0.420	(0.039)
Freiamt	0.068	(0.010)	0.198	(0.018)	0.295	(0.029)
Fricktal	0.079	(0.012)	0.245	(0.018)	0.358	(0.031)
Genève	0.273	(0.012)	0.265	(0.009)	0.247	(0.013)
Glarner Hinterland	0.091	(0.026)	0.240	(0.058)	0.205	(0.086)
Glarner Unterland	0.078	(0.019)	0.232	(0.040)	0.265	(0.063)
Glattal-Furttal	0.108	(0.016)	0.245	(0.018)	0.380	(0.027)
Glâne-Veveyse	0.115	(0.014)	0.212	(0.030)	0.320	(0.061)
Goms	0.094	(0.029)	0.212	(0.075)	0.442	(0.106)
Grenchen	0.073	(0.020)	0.286	(0.042)	0.435	(0.060)
Gros-de-Vaud	0.097	(0.013)	0.232	(0.019)	0.261	(0.027)
Innerschwyz	0.063	(0.008)	0.260	(0.020)	0.472	(0.036)
Jura	0.099	(0.008)	0.229	(0.015)	0.318	(0.025)
Jura bernois	0.077	(0.013)	0.250	(0.024)	0.367	(0.036)
Kandertal	0.031	(0.009)	0.121	(0.042)	0.222	(0.065)
Knonaueramt	0.090	(0.021)	0.275	(0.030)	0.328	(0.036)
La Broye	0.103	(0.011)	0.235	(0.019)	0.289	(0.029)
La Chaux-de-Fonds	0.105	(0.011)	0.251	(0.018)	0.382	(0.025)
La Gruyère	0.070	(0.012)	0.206	(0.030)	0.332	(0.052)
La Sarine	0.143	(0.015)	0.298	(0.018)	0.371	(0.033)
La Vallée	0.037	(0.016)	0.255	(0.041)	0.343	(0.053)
Laufental	0.102	(0.017)	0.216	(0.027)	0.297	(0.040)
Lausanne	0.184	(0.013)	0.315	(0.011)	0.295	(0.015)
Leuk	0.113	(0.026)	0.158	(0.052)	0.462	(0.119)
Limmattal	0.151	(0.028)	0.206	(0.027)	0.315	(0.038)
Linthgebiet	0.103	(0.014)	0.186	(0.024)	0.343	(0.041)
Locarno	0.165	(0.012)	0.236	(0.017)	0.375	(0.027)
Lugano	0.158	(0.010)	0.281	(0.014)	0.305	(0.020)
Luzern	0.091	(0.007)	0.253	(0.010)	0.363	(0.014)
March	0.096	(0.015)	0.283	(0.024)	0.349	(0.043)
Martigny	0.088	(0.012)	0.278	(0.025)	0.316	(0.041)
Mendrisio	0.160	(0.014)	0.234	(0.021)	0.293	(0.033)
Mesolcina	0.112	(0.031)	0.319	(0.072)	0.481	(0.135)
Mittelbünden	0.083	(0.024)	0.334	(0.067)	0.380	(0.090)
Monthey	0.093	(0.016)	0.286	(0.030)	0.331	(0.048)
Morges	0.192	(0.019)	0.312	(0.018)	0.303	(0.024)
Murten/Morat	0.126	(0.014)	0.260	(0.026)	0.388	(0.038)
Mutschellen	0.114	(0.018)	0.282	(0.022)	0.344	(0.029)
Neuchâtel	0.159	(0.013)	0.290	(0.014)	0.287	(0.020)
Nidwalden	0.080	(0.012)	0.194	(0.026)	0.302	(0.047)
Nyon	0.246	(0.025)	0.272	(0.022)	0.260	(0.029)
Oberaargau	0.073	(0.008)	0.233	(0.021)	0.444	(0.030)
Oberengadin	0.171	(0.029)	0.350	(0.045)	0.364	(0.070)
Oberes Baselbiet	0.102	(0.013)	0.251	(0.021)	0.386	(0.028)
Oberes Emmental	0.038	(0.007)	0.186	(0.030)	0.244	(0.046)
Oberland-Ost	0.044	(0.009)	0.203	(0.027)	0.409	(0.036)
Oberthurgau	0.061	(0.009)	0.219	(0.021)	0.295	(0.032)
Olten	0.084	(0.013)	0.209	(0.023)	0.424	(0.037)
Pays d'Enhaut	0.084	(0.029)	0.208	(0.061)	0.221	(0.066)
Pfannenstiel	0.206	(0.026)	0.286	(0.020)	0.315	(0.027)
Prättigau	0.046	(0.014)	0.281	(0.042)	0.462	(0.065)
Rheintal	0.107	(0.014)	0.205	(0.024)	0.462	(0.046)
Saanen-Obersimmental	0.061	(0.012)	0.180	(0.035)	0.275	(0.054)
Sarganserland	0.101	(0.014)	0.169	(0.033)	0.232	(0.048)
Sarneraatal	0.071	(0.011)	0.263	(0.030)	0.196	(0.069)
Schaffhausen	0.094	(0.017)	0.278	(0.025)	0.399	(0.033)
Schanfigg	0.025	(0.025)	0.140	(0.129)	0.509	(0.160)
Schwarzwasser	0.045	(0.011)	0.134	(0.036)	0.398	(0.054)
Sense	0.096	(0.012)	0.260	(0.026)	0.443	(0.055)
Sierre	0.134	(0.019)	0.226	(0.032)	0.296	(0.052)
Sion	0.162	(0.015)	0.269	(0.023)	0.361	(0.036)
Solothurn	0.123	(0.016)	0.280	(0.021)	0.425	(0.030)
St.Gallen	0.117	(0.010)	0.255	(0.015)	0.369	(0.023)
Sursee-Seetal	0.082	(0.007)	0.214	(0.013)	0.366	(0.020)
Surselva	0.099	(0.015)	0.238	(0.038)	0.400	(0.058)
Thal	0.032	(0.014)	0.210	(0.045)	0.560	(0.070)
Thun	0.070	(0.007)	0.225	(0.016)	0.300	(0.023)
Thurtal	0.091	(0.009)	0.194	(0.017)	0.290	(0.023)
Toggenburg	0.069	(0.010)	0.175	(0.025)	0.209	(0.043)
Tre Valli	0.107	(0.014)	0.274	(0.028)	0.352	(0.040)
Unteres Baselbiet	0.166	(0.019)	0.257	(0.018)	0.352	(0.023)
Untersee	0.102	(0.013)	0.254	(0.022)	0.330	(0.033)
Uri	0.062	(0.009)	0.297	(0.023)	0.527	(0.044)
Val-de-Travers	0.090	(0.020)	0.221	(0.035)	0.370	(0.057)
Vevey	0.164	(0.018)	0.305	(0.019)	0.304	(0.024)
Viamala	0.079	(0.021)	0.349	(0.060)	0.329	(0.071)
Visp	0.129	(0.017)	0.148	(0.028)	0.367	(0.056)
Weinland	0.076	(0.021)	0.274	(0.039)	0.293	(0.050)
Werdenberg	0.107	(0.015)	0.249	(0.034)	0.333	(0.051)
Wil	0.065	(0.008)	0.179	(0.018)	0.279	(0.026)
Willisau	0.069	(0.006)	0.195	(0.012)	0.361	(0.024)
Winterthur	0.094	(0.013)	0.285	(0.018)	0.299	(0.023)
Yverdon	0.131	(0.016)	0.239	(0.019)	0.331	(0.028)
Zimmerberg	0.155	(0.022)	0.271	(0.022)	0.340	(0.029)
Zug	0.093	(0.011)	0.270	(0.013)	0.377	(0.020)
Zürcher Oberland	0.089	(0.012)	0.260	(0.017)	0.292	(0.021)
Zürcher Unterland	0.090	(0.017)	0.240	(0.023)	0.277	(0.032)
Zürich	0.204	(0.017)	0.292	(0.014)	0.393	(0.018)

Notes: This table shows the educational mobility estimates by labor market regions (n=106). Share Bottom 20 in HS shows the share of children from the bottom quintile in the national parental income distribution that visit a high school (gymnasium), *Child-Parent Years Edu* shows the correlation between years of education of children and parents, Child-Parent HS shows how much more likely children are to visit a high school if at least one of their parents went to high school as well. Corresponding standard errors are shown in parentheses.

TABLE C.13: Correlation with Alternative Location Specifications

	Mother Location Child 16	Child Place of Birth	Child Location in 2010
Income Mobility			
RRS	0.985	0.942	0.828
Q1Q5	0.998	0.986	0.913
Q1Q1	0.979	0.912	0.778
Q5Q5	0.981	0.838	0.836
AUM25	0.997	0.968	0.925
IGE	0.954	0.816	0.688
Educational Mobility			
YearsEdu	0.963	0.831	0.881
Share Bottom20 in HS	0.993	0.989	0.912
Parents HS when Child was	0.967	0.875	0.868

Notes: This table shows how the mobility estimates on a cantonal level are correlated when children are assigned to regions according to different rules. Mother Location 16 restricts the sample to children for which we know for sure that the mother lived in this place when the child was 16 (this is true for 75% of children). Child Place of Birth is the place where the child was born. Child Location in 2010 used the location where the child lives when adult. Correlations are weighted by cantonal population in 2010.

TABLE C.14: Robustness Regional Housing Price Index

	(1)	(2)	(3)	(4)
		Regional		
	CPI	HPI 1	HPI 2	HPI 3
Panel A:				
Rank-Rank Slope	0.141	0.117	0.120	0.127
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Constant	43.323	44.579	44.390	44.068
	(0.0601)	(0.0603)	(0.0602)	(0.0602)
Observations	923,107	923,107	923,107	923,107
Panel B:				
American Dream (Q1Q5)	0.124	0.125	0.123	0.121
	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Observations	184,628	184,628	184,628	184,628
Panel C:				
Poverty Circle (Q1Q1)	0.247	0.231	0.233	0.234
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Observations	184,628	184,628	184,628	184,628

Notes: This table shows the sensitivity of our measures to regional price indices. Column (1) uses the «Residential Property Privately Owned Apartments Price Index», Column (2) Residential Property Regional Housing Price Index. Column (3) Rented properties, rental housing units price index. Source: Swiss National Bank.

.

TABLE D.1: Descriptive Summary by Sample

	(1) Full Sample Mean	(2) Mother Mean	(3) Father. Mean	(4) Any Grandf. Mean	(5) Mat. Grandf. Mean	(6) Pat. Grandf. Mean
Share Female	0.49	0.49	0.49	0.49	0.49	0.48
Birthyear	1967.94	1977.79	1977.63	1982.13	1982.12	1982.62
Father Birthyear	1942.19	1947.32	1947.72	1953.70	1953.31	1955.05
Mother Birthyear	1942.42	1950.28	1950.05	1956.23	1956.47	1956.53
Pat. Grandf. Birthyear	1926.29	1926.35	1926.41	1927.22	1927.89	1927.37
Mat. Grandf. Birthyear	1927.60	1927.77	1927.73	1928.44	1928.49	1929.90
Pat. Grandfm. Birthyear	1926.12	1926.34	1926.42	1929.06	1928.46	1930.23
Mat. Grandfm. Birthyear	1927.46	1927.94	1927.86	1931.02	1931.49	1930.83
Child income (log res.)	0.04	0.05	0.04	0.07	0.06	0.06
Father income (log res.)	0.57	0.56	0.56	0.57	0.56	0.58
Grandfather income (log res.)	0.56	0.56	0.56	0.56	0.56	0.56
Grandpa income (log res.)	0.54	0.54	0.54	0.54	0.54	0.57
Region 1: Région Lémanique	0.15	0.15	0.15	0.16	0.15	0.16
Region 2: Mittelland	0.26	0.26	0.26	0.28	0.28	0.28
Region 3: Nordwestschweiz	0.14	0.13	0.13	0.14	0.14	0.14
Region 4: Zurich	0.17	0.17	0.17	0.17	0.16	0.17
Region 5: Ostschweiz	0.14	0.14	0.14	0.14	0.14	0.13
Region 6: Zentralschweiz	0.11	0.11	0.11	0.10	0.10	0.09
Region 7: Ticino	0.04	0.04	0.04	0.03	0.03	0.02
Education: High School	0.19	0.22	0.21	0.21	0.21	0.22
Education: Vocational	0.65	0.65	0.65	0.66	0.66	0.65
Observations	2,755,550	1,023,014	1,087,743	212,227	143,650	94,229

Notes: This table shows the means of children characteristics conditional for different samples.

TABLE D.2: Calculation of Residualized Lifetime Income

	Individual Log Income
Age	0.27365***
	(0.000275)
Age^2	-0.00516***
	(0.000007)
Age^3	0.00003***
	(0.000000)
Birthyear	0.26328***
	(0.001701)
$Birthyear^2$	-0.00007***
	(0.000000)
Constant	-254.33751***
	(1.666691)
Observations	138,517,168
R-Squared	0.05

Notes: This table shows the coefficients of the regression used to calculate the residualized lifetime income.
*p<.010; **p<0.05; ***p<0.01

TABLE D.3: Inheritability Coefficient by Canton

Canton	λ	(se)	λ_{alt}	(se)
AG	0.193	(0.005)	0.187	(0.010)
AI	0.289	(0.029)	0.234	(0.036)
AR	0.293	(0.008)	0.336	(0.020)
BE	0.212	(0.003)	0.202	(0.006)
BL	0.195	(0.013)	0.170	(0.020)
BS	0.198	(0.016)	0.152	(0.018)
FR	0.180	(0.005)	0.193	(0.011)
GE	0.253	(0.013)	0.234	(0.021)
GR	0.258	(0.006)	0.308	(0.018)
JU	0.250	(0.012)	0.230	(0.021)
LU	0.207	(0.006)	0.195	(0.010)
NE	0.197	(0.011)	0.177	(0.017)
NW	0.252	(0.02)	0.223	(0.030)
OW	0.212	(0.014)	0.224	(0.030)
SG	0.233	(0.005)	0.263	(0.011)
SH	0.162	(0.02)	0.167	(0.030)
SO	0.205	(0.008)	0.182	(0.012)
SZ	0.263	(0.006)	0.312	(0.017)
TG	0.194	(0.007)	0.203	(0.014)
TI	0.198	(0.013)	0.163	(0.017)
UR	0.224	(0.019)	0.215	(0.028)
VD	0.253	(0.006)	0.247	(0.011)
VS	0.138	(0.006)	0.136	(0.011)
ZG	0.230	(0.012)	0.225	(0.023)
ZH	0.184	(0.004)	0.185	(0.008)
GL	0.252	(0.020)	0.223	(0.030)

Notes: This table shows the inheritability coefficient λ for income by different cantons of birth of the grandfather. Standard errors are shown in parentheses. For λ_{alt}, the two-generation coefficient is only measured on the youngest cohorts (child and father).

TABLE E.1: Treatment Effects Child Log Income by Parent Income Deciles

	Dependent Variable: Log Child Income									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Parent-D1	Parent-D2	Parent-D3	Parent-D4	Parent-D5	Parent-D6	Parent-D7	Parent-D8	Parent-D9	Parent-D10
Treatment Post = 1	-0.083**	-0.055***	-0.073***	-0.060***	-0.045**	-0.037*	-0.035*	-0.009	-0.029*	-0.001
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Treatment Trans = 1	-0.013	0.032*	-0.023	-0.045**	-0.033***	0.012	-0.010	0.026	-0.049**	0.000
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	151,478	127,303	139,712	140,519	141,153	142,457	143,708	144,820	146120	147645
r2	0.0963	0.099	0.111	0.112	0.12	0.119	0.11	0.0965	0.0851	.0671

This table shows the treatment effects of the regression on Equation 6.1 for different subgroups according to parental income. Thereby, the outcome variable is the log of the child income around the age 30 to 33. Parent-D1 refers to parent decile 1 etc. The coefficient «Treatment Post» shows the difference-in-difference coefficient for the post-treatment period, the coefficient «Treatment Trans» shows the difference-in-difference coefficient for the transition period.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.2: Robustness: Treatment Assignment if less than 15″ from border

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post = 1	-0.103***	-0.086***	-0.059***	-0.050***	-0.039***
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
Treatment Transition = 1	-0.014	-0.046***	-0.019***	-0.011	-0.018
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	215,664	214,649	210,799	206,252	204,491
r2	0.129	0.149	0.158	0.137	0.105

This table shows the treatment effects for an alternative treatment-control-group assignment. In the treatment group are only municipalities less than 15 minutes away from the border. The control group consists of municipalities further than 30 minutes away from the border. Municipalities within 15 and 30 minutes from the border are excluded.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.3: Robustness: Treatment Assignment if less than 15″ from border and different control group

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post = 1	-0.086*	-0.069*	-0.055*	-0.037*	-0.033
	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)
Treatment Transition = 1	-0.018	-0.021	-0.021*	-0.048**	-0.042**
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	97,063	98,288	103,049	107,418	116,684
r2	0.106	0.127	0.142	0.125	0.093

Notes: This table shows the treatment effects for an alternative treatment-control-group assignment. In the treatment group are only municipalities less than 15 minutes away from the border. The control group consists only of municipalities further than 30 minutes away from the border, but that are—administratively—still part of the border region (in Figure 6.1 colored in the brightest green).
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.4: Robustness: Treatment Assignment by Government Definition

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post=1	-0.077***	-0.074***	-0.050***	-0.029*	-0.015
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Treatment Trans=1	-0.005	-0.040***	-0.030***	-0.007	-0.015
	(0.02)	(0.01)	(0.00)	(0.01)	(0.02)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	290,861	291,553	294,718	299,216	303,937
r2	0.126	0.144	0.153	0.135	0.104

Notes: This table shows the treatment effects for an alternative treatment-control-group assignment. In the treatment group are only municipalities administratively assigned to the border-region, while the control group only includes municipalities assigned to the non-border region. In the baseline specification, treatment assignment is strictly defined by distance to the border.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.5: Robustness: Child Municipality Assignment by Municipality of Birth

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post = 1	-0.078**	-0.080***	-0.036*	-0.028**	-0.015
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Treatment Trans = 1	0.006	-0.056**	-0.018*	-0.005	-0.004
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	283,040	284,196	288,305	293,446	299,267
r2	0.125	0.143	0.153	0.135	0.104

Notes: This table shows the treatment effects for an alternative treatment-control-group assignment. Children are assignent to the treatment and control group by municipality of birth instead of the municipality of the mother in the year 2010.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.6: Robustness: Child Municipality Assignment by Municipality in 2010

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post = 1	-0.059*	-0.053**	-0.028	-0.028*	-0.019
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Treatment Trans = 1	0.019	-0.031	-0.006	-0.010	-0.017
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	283,040	284,196	288,305	293,446	299,267
r2	0.125	0.143	0.153	0.135	0.104

Notes: This table shows the treatment effects for an alternative treatment-control-group assignment. Children are assignent to the treatment and control group by the municipality they live in 2010 instead of the municipality the mother lives in the year 2010. Here, the effect is likely attenuated because children might react to the policy by moving.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.7: Robustness: Measuring Child Income between Age 35 to 40

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post=1	-0.060**	-0.063*	-0.040*	-0.023	-0.009
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)
Treatment Trans=1	-0.026*	-0.028	-0.020	-0.019	-0.015
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	220,044	220,090	221,856	225,305	228,321
r2	0.209	0.221	0.243	0.232	0.209

Notes: This table shows the treatment effects for an alternative child income definition. Here, child income is measured over the age 35 to 40. In the main specification, child income is measured between 30 and 33.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.8: Robustness: Measuring Child Income between 30 and 38

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post=1	-0.009	-0.051*	-0.036*	-0.025	-0.013
	(0.03)	(0.02)	(0.01)	(0.02)	(0.01)
Treatment Transition=1	-0.019	-0.024*	-0.030**	-0.018	-0.006
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	169,311	169,019	170,310	173,160	175,630
r2	0.207	0.212	0.235	0.221	0 .188

Notes: This table shows the treatment effects for an alternative child income definition. Here, child income is measured over the age 30 to 38. In the main specification, child income is measured between 30 and 33.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.9: Robustness: Parent Income Rank within Assignment Group

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post =1	-0.059*	-0.042**	-0.048**	-0.025	-0.017
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Treatment Trans=1	0.012	-0.015	-0.011	-0.007	-0.018
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	296,312	295,886	296,095	296,086	296,070
r2	0.125	0.144	0.154	0.136	0.105

Notes: This table shows the treatment effect with an alternative parental income rank definition. In this specification, parental income rank is defined as within treatment and control group. In the baseline specification, parental income rank is defined on the national level. Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

TABLE E.10: Placebo Test: Fake Treatment Period

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post = 1	0.018	-0.011	0.041	-0.021	-0.013
	(0.02)	(0.02)	(0.01)	(0.03)	(0.01)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	93,877	93,583	94,136	96,309	98,443
r2	0.179	0.178	0.203	0.193	0.149

Notes: This table shows the result of a placebo test. Thereby, the period is shifted entirely into the pre-treatment period, where no effect should occur. The «fake»-post period between 1995 and 1999. The years 1990 to 1994 are assigned to the pre-treatment period.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

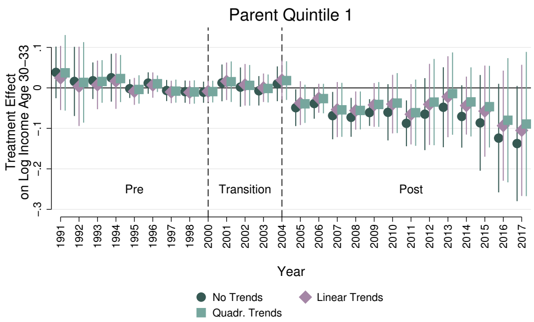
TABLE E.11: Placebo Test: Fake Treatment Groups

	Dependent Variable: Child Log Income				
Subgroup:	Parent-Q1	Parent-Q2	Parent-Q3	Parent-Q4	Parent-Q5
	(1)	(2)	(3)	(4)	(5)
Treatment Post = 1	-0.002	-0.001	0.026**	0.002	0.007
	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
Treatment Trans = 1	-0.006	0.006	0.014	0.005	-0.003
	(0.02)	(0.00)	(0.02)	(0.02)	(0.01)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Sex FE	Yes	Yes	Yes	Yes	Yes
N	109,261	106,097	98,390	91,357	82,335
r2	0.152	0.171	0.174	0.153	0.124

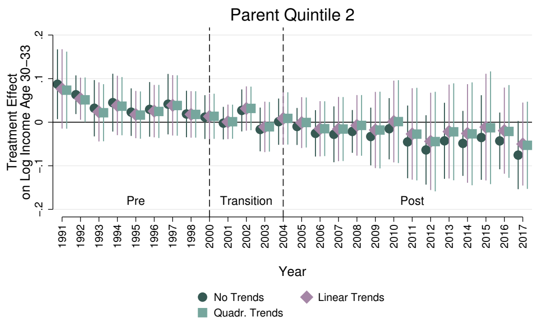
This table shows the result of a placebo test. Thereby, treatment and control groups are randomly assigned within the actual control group.
Robust standard errors clustered on regional (NUTS-2) level (*p<0.10; **p<0.05; ***p<0.01)

FIGURE E.1: Event Study Estimates on Log Child Income by Parent Income Quintiles

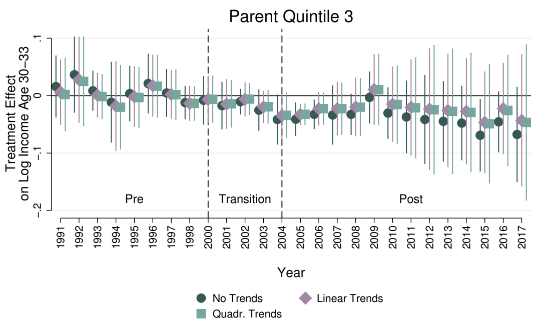
(A) Parent Q1



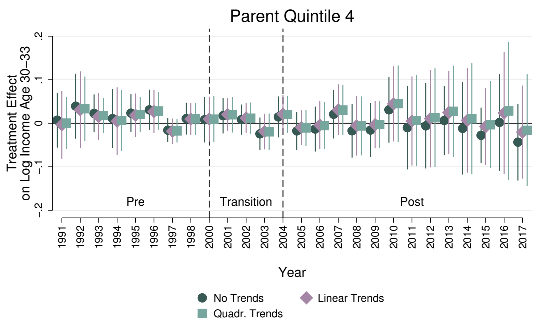
(B) Parent Q2



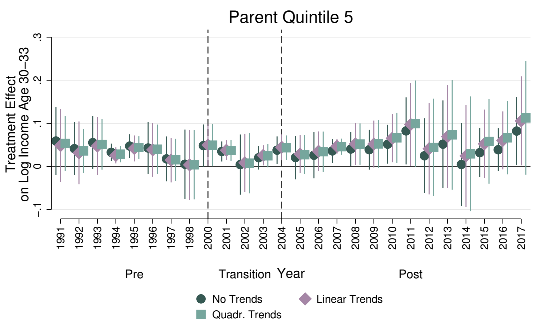
(C) Parent Q3



(D) Parent Q4

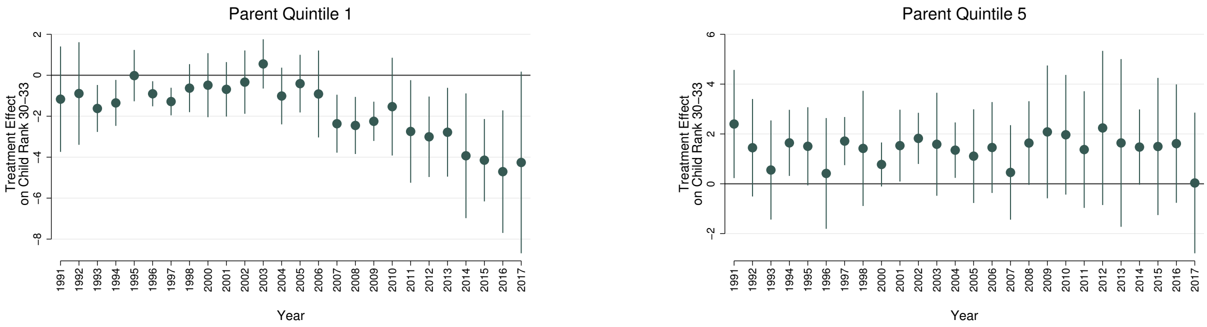


(E) Parent Q5



Notes: This figure shows the event study difference-in-difference coefficients δ_t at every year t as discribed in Equation 6.2 for children with parents from the bottom quintile (ParentQ1) and from the top quintile (ParentQ5) of the parental income distribution. The line indicates a 95% confidence interval. Standard errors are cluster on regional NUTS-2 level. The ommited year dummy is year 1999 right at the onset of the policy. The specification includes region fixed effects (NUTS-2), region specific trends, and sex. The income of the children is measured as log of the mean real income over the ages 30 to 33. Parental rank is measured as child cohort specific rank of the mean income when the child is between 12 and 22.

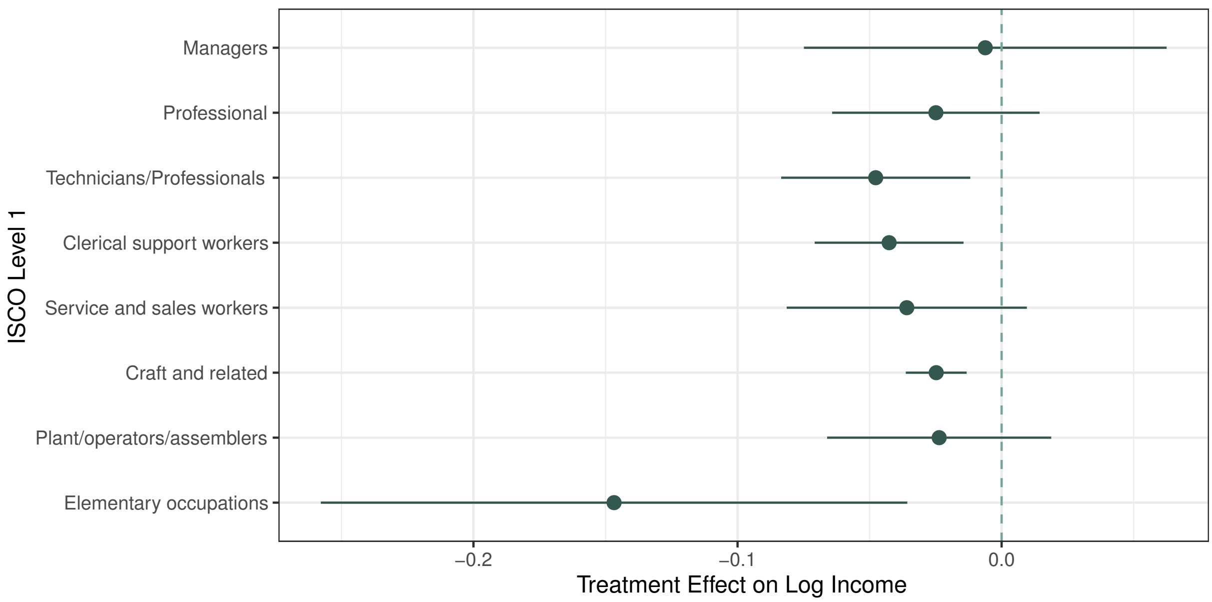
FIGURE E.2: Outcome Variable: Ranks



(A) Parent Quintile 1 (poor) (B) Parent Quintile 5 (rich)

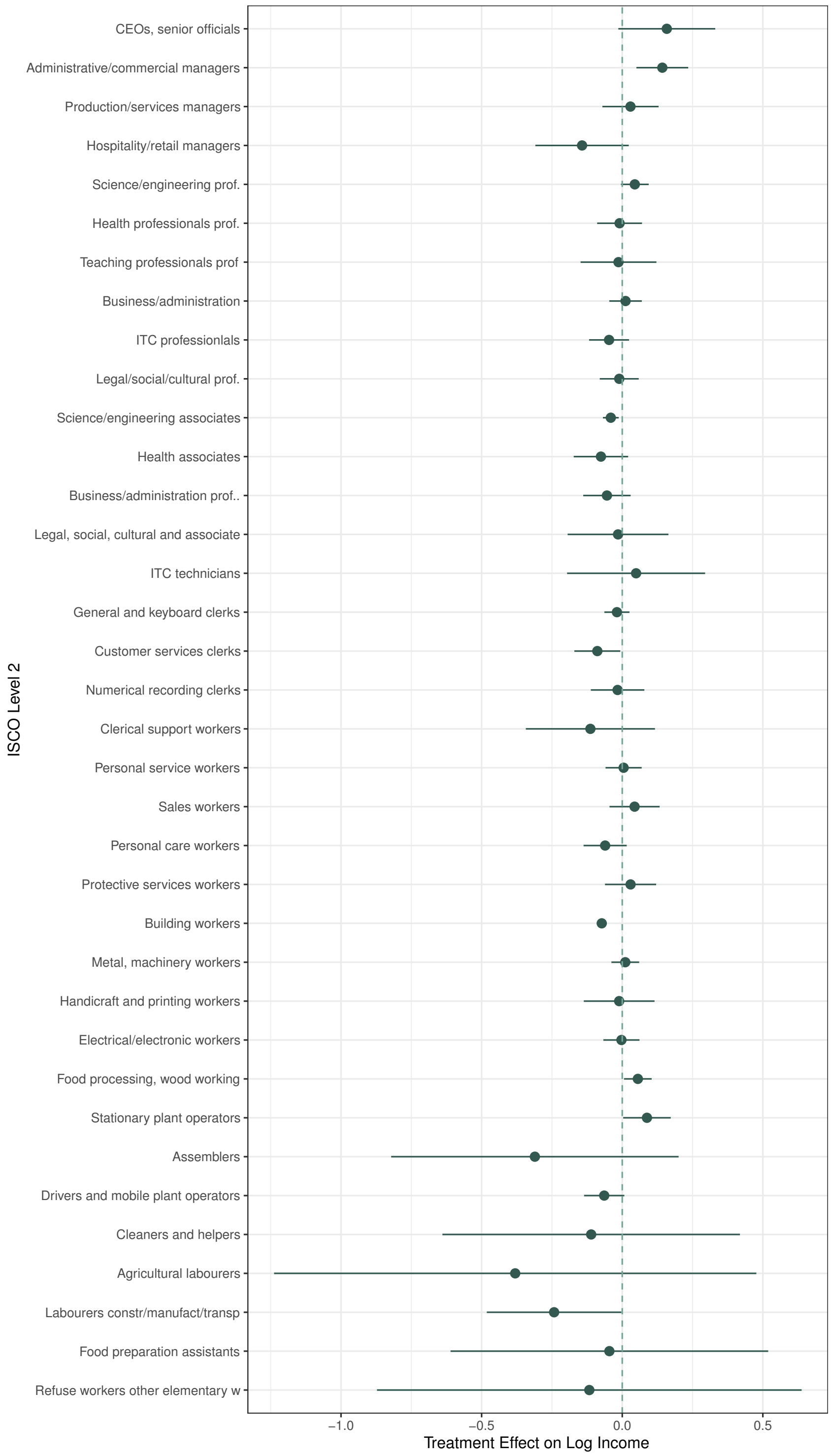
Notes: This figure show the coefficients δ_t from Equation 6.2 with cohort specific child rank as outcome variable y. Panel a shows the estimates for children with parents from the bottom quintile of the income distribution, Panel b the same for children with parents at from the top quintile.

FIGURE E.3: Treatment Effect by Learned Occupation (ISCO-08 1)



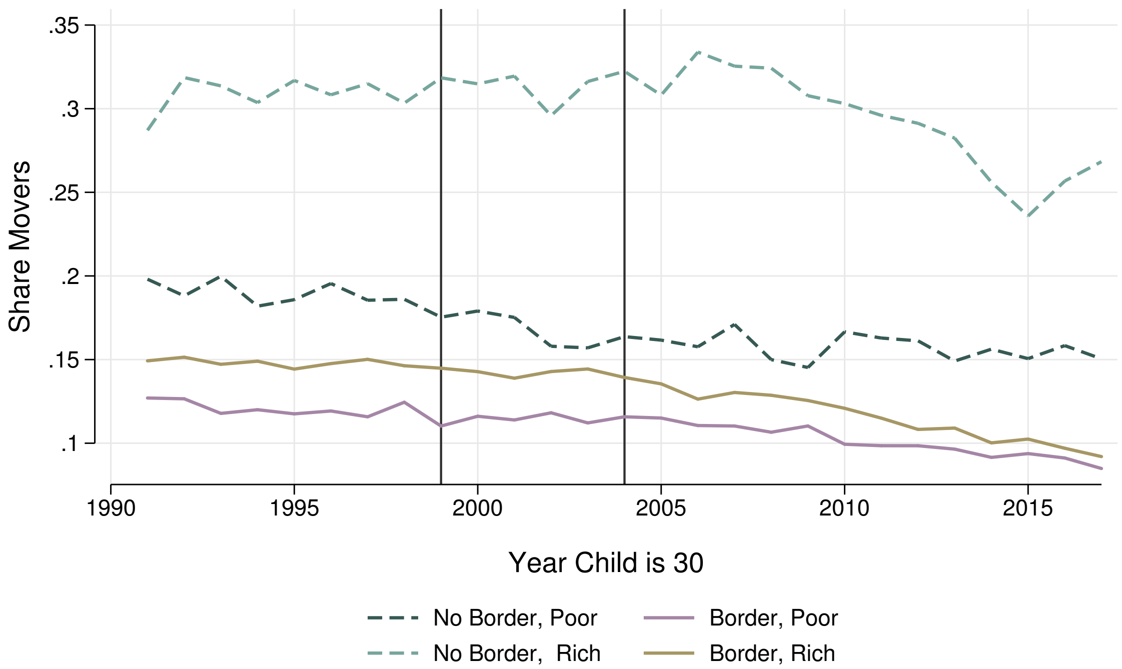
Notes: This figure shows the treatment effects in the post period on child log income according to Equation 6.1 by ISCO-08 major groups (without army and agriculture).

FIGURE E.4: Treatment Effect by Learned Occupation (ISCO 2)



Notes: This figure shows the treatment effects in the post period on child log income according to Equation 6.1 by ISCO-08 submajor groups.

FIGURE E.5: Geographic Mobility



Notes: The lines indicates the share of children born in the border region who are in the treatment region in the year 2010 and vice-versa for the non-border region. Children born in the border-region are less likely to move to the non-border region than the oppositve way

Curriculum Vitae

PATRICK MICHEL CHUARD-KELLER

Born May 22, 1986
in Brugg AG, Switzerland
Swiss national

EDUCATION

2017 - 2021	PhD Program in Economics and Finance, University of St. Gallen, Switzerland
2013 - 2015	Master of Arts in Economics, University of Zurich, Switzerland
2009 - 2013	Bachelor of Arts in Economics, University of Zurich, Switzerland
2011	Exchange Semester, Erasmus University Rotterdam, Netherlands
2008 - 2009	Studies in Biology, Swiss Federal Institute of Technology, Switzerland
2005 - 2007	Swiss Matura, Kantonsschule Baden, Switzerland
2002 - 2005	Commercial Diploma, Kantonsschule Baden, Switzerland

---

PROFESSIONAL EXPERIENCE

2017 - 2021	Doctoral Student, University of St. Gallen, Switzerland
2015 - 2017	Research Associate, Zurich University of Applied Sciences, Switzerland
2014 - 2015	Research Assistant, Swiss Economic Institute (KOF), Switzerland
2013 - 2015	Intern, Swiss National Bank, Switzerland
2013	Civil Service in Alcohol Rehab, Psychiatric Hospital Windisch, Switzerland
2008	Administrative Assistant, Oracle Software, Switzerland
2007	Military Service, Switzerland