Essays on Inequalities in Economics of Labor and Education

Laurenz Baertsch

TESI DOCTORAL UPF / year 2022

THESIS SUPERVISOR
Prof. Caterina Calsamiglia
Department of Economics and Business

Universitat Pompeu Fabra
Barcelona

Acknowledgments

I am very thankful for the guidance, support and patience of my advisor, Caterina Calsamiglia, throughout my Ph.D. studies. Moreoever, this thesis has greatly benefited from the detailed feedback provided by Libertad Gonzalez, Gianmarco León and Albrecht Glitz, who were always available for help. I would also like to thank Marta Araque and Laura Agustí.

I am deeply indebted to my parents, Ursula and Peter, for their continuous support throughout my studies, without which I would not have been able to reach this point, and to my brother, Marc-Andrea. I am enormously grateful to Betsy, whose relentless encouragement has enabled me to face this Ph.D.'s many challenges.

My experience would not have been the same without the friends that I have made in the last years. I would like to thank Elisa and Sampreet for accompanying me through joyful and difficult moments of life in Barcelona; Niko and Milan for unforgettable trips to nature, and Anna, Iolanda and Naïla for making me feel at home in our shared apartment. Finally, it was a pleasure to share my Ph.D. experience with Philipp, Jairo, András, Ignacio and my office mates Marco and Nico.

Laurenz Baertsch
Barcelona
December 2022

Abstract

This thesis studies inequalities in the areas of education and labor markets, as well as public policies designed to reduce these inequalities. The first chapter studies whether part-time work options in paid parental leave (PL) systems can mitigate the large earnings losses that mothers experience after child birth. Analyzing a German paid PL reform, we show that such policies can increase labor market attachment of mothers in the short run but do not affect their long run labor market outcomes. In the second chapter, I show that centralizing the admissions system of public schools can improve socio-economically disadvantaged students' access to high quality education. However, this comes at the expense of increased income-based segregation as high income students leave for the private school sector. The last chapter analyzes to what extent and how individuals in occupations that are beneficially affected by structural transformations can transmit their gains in socio-economic status to their offspring. We document that the (grand)sons of machinists, an occupation particularly demanded in the United States during the Second Industrial Revolution, held occupations with higher earnings than the (grand)sons of comparable non-machinists. We identify rural-to-urban migration and secondary education as the main channels of intergenerational transmission.

Key words: paid parental leave, child penalty, part-time work incentives, public child care, access to education, centralized assignment, academic achievement, intergenerational transmission, rural-urban migration, investment in education.

Resumen

Esta tesis estudia desigualdades en los ámbitos de la educación y los mercados laborales, así como políticas públicas diseñadas para reducir dichas desigualdades. El primer capítulo analiza si las opciones de trabajo a tiempo parcial en los sistemas de permiso parental pagado pueden mitigar las importantes pérdidas de ingresos que experimentan las madres tras el nacimiento de sus hijos. Analizando una reforma del permiso parental pagado en Alemania, mostramos que estas políticas pueden aumentar la vinculación de las madres al mercado laboral a corto plazo, pero no afectan a sus resultados en el mercado laboral a largo plazo. En el segundo capítulo, muestro que la centralización del sistema de admisión en escuelas públicas puede mejorar el acceso a la educación para estudiantes desfavorecidos socioeconómicamente. Sin embargo, esto se produce a expensas de un aumento de la segregación según ingresos, ya que los estudiantes de altos ingresos se reubican al sector de la educación privada. El último capítulo analiza hasta qué punto y de qué manera los individuos que ejercen ocupaciones que se ven beneficiadas por las transformaciones estructurales pueden transmitir sus logros socioeconómicos a su descendencia. Documentamos que los hijos (y nietos) de maquinistas, una ocupación especialmente demandada en Estados Unidos durante la Segunda Revolución Industrial, ocuparon puestos con mayores ingresos que los hijos y (nietos) de personas comparables que no eran maquinistas. Identificamos la migración del campo a la ciudad y la educación secundaria como los principales canales de esta transmisión intergeneracional.

Palabras claves: permiso parental pagado, penalización por hijo, incentivos al trabajo a tiempo parcial, guarderías públicas, acceso a la educación, asignación centralizada, rendimiento académico, transmisión intergeneracional, migración rural-urbana, inversión educativa.

Preface

This dissertation consists of three essays that study inequalities in the areas of education and labor markets, as well as the role public policies can play to mitigate them. To do so, I rely on causal inference techniques, natural experiments, economic theory and data from a variety of sources, including administrative records, surveys and newly collected historical data. I hope this work can, next to contributing to the academic literature, also provide evidence to policy debates in the corresponding areas.

The first chapter is concerned with the negative long run effects on earnings that mothers experience after giving birth to a child. In high income countries, these negative earnings effects, known as the child penalty, have become the main driver of earnings inequalities between men and women. Simultaneously, policies that increase the compatibility of work and family life, for example by promoting part-time work for parents, have been introduced. Yet, there is little empirical evidence on the effects of such pro-part-time policies. We analyze how a paid parental leave reform that increases monetary incentives for part-time work in the two years after child birth, affects maternal labor market outcomes in the short and in the long-run. To do so, we use German social security records and exploit the fact that only mothers whose child is born in or after July 2015 are eligible for the new part-time PL option in a Difference-in-Differences strategy. We find that the policy's pro-part-time incentives increase the probability that high income mothers return to work during the first year after child birth by 3pp ($\approx 15\%$), without reducing working hours of mothers who would already be working full-time in absence of the policy. Low income mothers do not choose the new part-time option, most likely due to financial constraints, and are unaffected by the reform. The policy's part-time work incentives do not impact maternal employment along the intensive margin (part-time or full-time work) in the long run, leaving the child penalty unaffected. Our analysis shows that part-time options in paid PL schemes can increase maternal labor market attachment directly after child birth and alleviates concerns that such options reduce working hours of mothers in the long-run ("lock-in" effect into part-time employment).

In the second chapter, I study whether *centralizing* the school admissions system improves socio-economically disadvantaged students' access to high quality education and their academic results. In search of a more efficient and equal assignment procedure of students to educational institutions, an increasing number of both developing and developed countries rely on centralized admission systems. In such centralized systems, students are assigned to schools based on the students' submitted preference lists and the schools' availability, rather than by letting schools choose which students to admit. I study the introduction of a national centralized admission system, the School Assignment System (SAS), for pre-tertiary education in Chile. For identification, I exploit its staggered implementation across the country in a (triple) difference-in-differences framework and use administrative data on all secondary school students. I find that the SAS enables low SES students to access schools of higher quality and that most of this effect is driven by municipalities in which private schools have a high market share. The mechanism behind this finding is that high SES students leave for the private sector - in which schools make admission decisions independently - and, thereby free up seats at high-quality public schools. As a result, income-based segregation increases in these municipalities. However, I do not find that these changes in access to educational quality impact the students' GPA or their retention rates up to the third grade of secondary school. The unintended increase in segregation implies that the presence of private schools needs to be taken into account when designing centralized admission systems.

The third chapter focuses on the intersection between intergenerational mobility and structural change. We try to understand to what extent, and how individuals in occupations that are beneficially affected by structural transformations can transmit their gains in socio-economic status to their (grand)children. While this question is very timely, considering for instance the current wave of automation, the analysis of recent shocks is inhibited by the time horizon required for intergenerational research. Therefore, we analyze the case of machinists whose occupation experienced a relative labor demand spike during the Second Industrial Revolution (1870-1914), resulting in higher income and job stability. To do so, we complement data from the US full count census with newly digitized data on the county-level supply of secondary education and occupation- state level earnings. Using matching and fixed effects regressions, we document that the (grand)sons of men who were machinists in 1870 held occupations with significantly higher earnings than the (grand)sons of comparable non-machinists. The higher earnings of machinists' sons mainly stemmed from parental investment in their education, but this effect is absent for those sons who were already too old to attend high school when the income of machinists started to rise. Additionally, the sons of initially rural machinists benefited from rural-to-urban migration. Our results are robust to controlling for family-fixed effects (comparing machinists to their non-machinist brothers), pre-1870 spatial sorting, and a rich set of next-door neighbor and grandparental characteristics. Our findings suggest that the transmission of unequal effects of labor market transformations to later generations can be mitigated by strengthening the public education system and enabling people to move to opportunity.

Contents

List of Figures

List of Tables

Chapter 1

REDUCING THE CHILD PENALTY BY INCENTIVIZING PART-TIME WORK?

EVIDENCE FROM A PAID PARENTAL LEAVE REFORM IN GERMANY

Joint with Malte Sandner

1.1. Introduction

Gender-based earnings inequalities in the labor market continue to exist even in today's most advanced economies (OECD 2019; EC 2022). Many factors that explain these inequalities, such as differences in the educational attainment between women and men, have largely disappeared over the last decades. However, the importance of the child penalty, i.e. the negative labor market effect mothers experience after having a child, in explaining earnings inequalities has doubled since the 1980s (Kleven et al. 2019). Simultaneously, policy makers in many countries seek to increase the compatibility of work and family life, for example through parental leave policies that promote part-time work for parents. Approximately one third of all upper-medium and high income countries (14 out of 43) surveyed in Blum et al. (2018) offer parents a part-time option in their paid parental leave (PL) system. Yet, the evidence for how labor market outcomes of mothers who take up such pro-part-time policies are affected is scarce.[1]

Concerns about potential negative effects of pro-part-time policies on maternal labor market outcomes in the long-run have been raised recently (e.g. Boneva et al. 2021). Kunze (2022) points to the risk of a "lock-in" effect into part-time work, according to which mothers who work part-time directly after child birth are less

[1] Fernández-Kranz and Rodríguez-Planas (2021) study the effect of an *unpaid* parental leave policy, which incentivizes parents with children that are younger than six to work part-time, on labor market outcomes of *all* mothers, i.e. not only those who take up the policy.

likely to change to higher-paying, full-time jobs once the child has grown older. This might be caused by pro-part-time habit formation in labor supply choices (Woittiez and Kapteyn 1998; Kubin and Prinz 2002) or by discrimination by the employer. On the other hand, one can also make the case for positive long-run labor market effects of pro-part-time policies. These could arise if mothers return to work more quickly after child birth and, thus, experience less human capital depreciation by reducing the time they are absent from their job. Additionally, employers might reward parents who return to work earlier by granting pay raises or promotions more easily in the long run.[2] Thus, the long-run labor market effects of pro-part-time policies are theoretically ambiguous.

In this paper we analyze how increased incentives for part-time work directly after child birth (instead of working full-time or not working at all) affect maternal labor market outcomes. To this end we exploit a reform of the paid PL system in Germany in 2015. With the objective of increasing the compatibility of work and family life for parents, the reform gives them the option to choose between a new part-time scheme, called *Parental Benefit Plus* (hereafter *PB+*; German: *Elterngeld Plus*) and the already existing *Parental Benefit* (hereafter *PB*; German: *Elterngeld*). We study the policy's effects on maternal labor market outcomes in the short run (the first 24 months in which parents are eligible for PL benefits), as well as in the long run (up to 4.5 years after child birth).[3] Additionally, we assess whether there exist complementarities between the availability of public child care and the take-up of the new part-time paid PL scheme, since the latter provides incentives for mothers to return to work earlier after child birth.

Under the already existing (old) scheme, PB, each parent is eligible for benefit payments for 12 months after child birth.[4] This benefit amount positively depends on income before child birth. Parents can work up to 30 hours per week while receiving PL benefits, however, the benefit amount decreases in post-child birth income. Compared to the old scheme, the new PB+ incentivizes part-time work by paying a higher *total* benefit amount (i.e. over the entire benefit period) to parents who work part-time in the 24 months after child birth. However, since parents can take PB+ for 24 months (instead of 12 months under PB), the new scheme pays most of these parents less each *month* than under the pre-existing PB to limit the increase in the total benefit amount. The combination of a higher *total*, but a lower *monthly* benefit amount in the benefit structure of the new policy leads to an income threshold in the take-up of the policy as explained below. Furthermore, the

[2] Tô (2018) conceptualizes the timing of return to work in a signaling model.

[3] We do not analyze the policy's effects on paternal outcomes, since we do not observe fathers in our data. Our study period is limited to 4.5 years after child birth, because the latest child birth cohorts in our sample enter the period of the COIVD-19 pandemic thereafter, making it impossible to distinguish between effects caused by the policy and those caused by the pandemic.

[4] Two extra months are offered if both parents take paid PL for at least two months.

introduction of PB+ coincides with a change in the *unpaid* PL legislation, which allows parents to take two instead of one year of unpaid PL between the child's third and eight birthday leaving the total number unpaid PL years unchanged.[5]

Only parents of children born on or after July 1^{st} 2015 are entitled to choose between PB and PB+. Parents of children born before this threshold date are eligible for PB only. We compare parents whose child is born in the two months after (treated) to those whose child is born in the two months before (control) the threshold date using Difference-in-Differences models.[6] Additionally, we net out seasonality in parental characteristics by including parents who have a child in the years before the reform, i.e. from 2011-2014. The fact that the policy passed the German parliament less than 9 months before the implementation date makes it unlikely that parents sorted across the threshold in anticipation of the policy.[7] In line with this interpretation, we do not find systematic differences in characteristics at conception between parents who give birth before and after the threshold date.

Our analysis is based on German social security records, from which we obtain detailed employment histories of roughly 400'000 mothers who gave birth between 2011-2015. These data are particularly suitable for our analysis for two reasons: first, social security records are ideal to analyze the subset of mothers to whom the policy is particularly attractive, namely mothers that are employed before child birth. Second, since these data allow us to draw on the universe of employed women in Germany, the sample size is sufficiently large (20'000 births per month) to identify the policy effect based on child births occurring in a few months around the threshold month. This ensures that mothers in the treated and control groups are comparable. Additionally, we combine these data with information on the local availability of public child care.

As theoretical framework for the mothers' short-run labor supply decisions in response to the policy, we use a simple two-period labor-leisure (child care) model. This framework captures the policy's changes in the benefit structure and offers the following predictions:[8] first, only mothers with sufficiently high income prior to child birth take up the new scheme. This is mainly due to the fact that PB+

[5]While all parents are entitled to three years of unpaid PL until the child's eighth birthday, parents who are eligible for PB+ can also choose to use two out of three years (previously one out of three) of unpaid parental leave between the child's third and eighth birthday. We are not able to precisely disentangle the effect of the change in the unpaid PL regulation from the one in the paid PL regulation. However, this does not affect our ability to rule out a lock-in effect, since it implies that our zero effect on full-time employment is a lower bound of the true effect size.

[6]Note that we do not use a Regression-Discontinuity-Design (RDD) since the child's date of birth is measured with error in our data. This introduces too much noise in the running variable to implement an RDD.

[7]To exclude the possibility of parental anticipatory sorting and taking into account that the children's birth date in German social security data relies on a proxy, we exclude births in the months directly preceding and following the implementation date, i.e. June and July 2015.

[8]This simple framework abstracts from inter-temporal externalities of labor choice, such as the labor decision in period one affecting wages or child quality in later periods, and intra-household dynamics.

replaces a *lower* fraction of pre-child birth earnings each month compared to the alternative scheme. Only mothers with sufficiently high pre-child birth income can compensate this lower replacement rate. Second, mothers that choose PB+ are expected to smooth their labor supply over the first two years. More precisely, PB+ changes the benefit structure such that working becomes relatively cheaper in the first year (by lowering the replacement rate) and more expensive in the second year (by partially replacing pre-birth income). Thus, parents are expected to increase their labor supply in the first and lower it in the second year with respect to the control group. Third, parents return to work earlier after child birth as a result of increasing employment (at the extensive margin) in the first year of benefit reception.[9]

Germany makes for a particularly suitable setting to study how increased incentives to part-time work affect mothers. Compared to similarly developed countries, in Germany $i)$ the maternal child penalty is relatively large ($\approx 60\%$ 10 years after giving birth according to Kleven et al. (2019)), *ii)* the incidence of maternal part-time work is particularly high, as many mothers that worked full-time before child birth work only part-time thereafter (OECD 2019) and *iii)* attitudes towards gender roles are relatively conservative. This is exemplified by fathers being little involved in child care and mothers being expected not to work or at most work part-time while her child is in school age or younger (Boneva et al. 2021).

In line with our theoretical prediction of an income threshold, we find that only high income mothers (i.e. upper 60% of the pre-child birth income distribution) take up the policy and are 3 pp ($\approx 15\%$ compared to the sample mean) more likely to be employed in the first year after child birth. These mothers (compliers) return to work during the first year instead of in later years as a result of the policy. Although PB+ monetarily incentivizes mothers to reduce their working hours during the first two years after child birth (relative to their pre-child birth labor supply), we do not observe a reduction in working hours for any subsample of mothers. The policy does not affect maternal labor supply during the second year after child birth. This implies that high income mothers do not smooth their labor supply, work more and have higher earnings during the first two years after child birth as a result of the policy. Thus, the policy increases labor market attachment and reduces the child penalty of high income mothers in the short-run. Low income mothers do not take up the new part-time paid PL option and are unaffected by the policy.

Furthermore, a higher local availability of affordable public child care does

[9] Recent research rationalizes the fact that some mothers return to employment before exhausting the paid PL period through a signaling model in which the timing of return to work provides employers with private information about future labor market choices and productivity (Tô 2018). In such a model, PB+ lowers the costs of early return to work.

not facilitate the observed return to work of mothers during the first year after child birth. Potential explanations are that *i)* high income mothers, who choose the part-time paid PL option, can afford more expensive private child care options or *ii)* that fathers, who are incentivized to work less by the part-time paid PL reform, engage more in child care.[10]

We do not find that the policy's pro-part-time incentives lead to a lock-in effect into part-time employment in the long run. While a lock-in effect would be observed as a shift in the probability from full-time to part-time employment, we do not find changes in employment along the intensive margin for high income mothers (compliers) or low income mothers.[11] The long run effect on monthly earnings is close to zero, yet significantly negative in some months, due to a temporary drop in the probability of being employed around four years after child birth, which we attribute to the change in the *unpaid* PL legislation rather than the paid PL's pro-part-time incentives.[12] In line with this interpretation, this temporary drop in employment is statistically insignificant among mothers who choose the part-time paid PL option (i.e. high income mothers). The policy's pro-part-time incentives do not affect the child penalty in the long run.[13]

To corroborate our findings, we conduct a series of robustness checks. Placebo tests, in which we consecutively define one of the years from 2011-2014 as the treatment year (i.e. before the actual implementation year), show that mothers do not adjust their labor supply in response to these "fake" policies. Our results are also robust to reducing the birth month window from two months to one month on either side of the threshold date and to clustering the standard errors at the week of birth-level (instead of the birth county).

Our results offer important insights for the design of paid PL schemes (see section 1.7 for a detailed discussion). First, monetarily incentivizing part-time-work can increase maternal labor market attachment through earlier return to work without harming long-run labor market outcomes through a lock-in effect into part-time employment. Second, the German paid PL reform we analyze shows that the government can achieve these effects without increasing public spending on paid PL benefits, since mothers who work while receiving paid PL benefits

[10]Unfortunately, since we do not have data on non-public child care providers or on the fathers' labor supply decisions, we are not able to distinguish between these explanations.

[11]This is also true once the employment protection period ends three years after child birth. Since employers are legally obliged to allow mothers to return to their pre-child birth job, only labor supply choices after the employment protection period can be considered "final".

[12]This change allows parents to take two of the three years of unpaid PL (instead of one out of three) between the child's third and eighth birthday (see section 2.2 for more details).

[13]Monthly earnings show a temporary drop around four years after child birth. Since this effect coincides with the temporary drop in the employment probability, we attribute it to the change in the *unpaid* PL legislation previously mentioned. Apart from this temporary drop, the point estimates of the policy's effect on monthly earnings are close to zero and statistically insignificant. Thus, the newly introduced pro-part-time incentives in the paid PL system do not have a statistically significant effect on monthly earnings and the child penalty.

receive a lower total benefit transfer than mothers who stay at home.[14] Third, the fact that only high income mothers choose the part-time paid PL option highlights the importance of taking the benefit structure's impact on different subgroups into account when designing policies.[15]

This paper makes various contributions to the literature. Numerous papers have investigated the impact of (changes in) paid PL policies on maternal labor supply (see Olivetti and Petrongolo 2017 and Rossin-Slater 2017 for excellent review articles). These papers analyze changes in the duration (Lalive and Zweimüller 2009), in the benefit amount (Asai 2015) or in both combined (Schönberg and Ludsteck 2014, Kluve and Schmitz 2018).[16] The role of part-time options in paid PL systems is almost absent from this literature. The only exception is Joseph et al. (2013), who study a 6-months part-time option in the French paid PL system using a survey of 3000 mothers with self-reported labor market outcomes. They find that it negatively impacts wages of high-income mothers two years after child birth, potentially explained by the fact that mothers can return to their pre-child birth job during an employment protection period of three years. Our analysis differs form the aforementioned paper in that we *i)* use large-scale social security records allowing for a cleaner identification of the causal effect, *ii)* measure employment outcomes after the employment protection period ends and *iv)* study a reform with a significantly longer benefit period and larger, income-dependent benefit payments.

Another strand of the literature studies the effects of (subsidized) public child care on maternal employment. This literature focuses on changes in the availability of public child care through changes in prices or capacity and finds positive effects (Givord and Marbot 2015; Nollenberger and Rodriguez-Planas 2015; Ravazzini 2018; Nix and Andresen 2019) or null effects (e.g. Havnes and Mogstad 2011) on maternal employment.[17] Instead, we study the *complementarities* between the regional availability of public child care and paid PL. To the best of our knowledge, the only paper with a similar approach is Girsberger et al. (2021), who find that the introduction of the first paid PL scheme in Switzerland positively affects fertility in the long-run but does not impact maternal employment. We contribute to this literature by studying a setting in which the paid PL reform incentivizes employment, while the introduction of the *first* paid PL system in the aforementioned paper reduces employment in the short run.

[14] Additionally, mothers who return to work earlier generate additional tax revenue.

[15] Since official documents do not state that PB+ is designed to specifically target high income mothers, the exclusive take-up among this group of mothers is likely unintended.

[16] In summary, this strand of the literature finds that paid PL affects maternal labor supply negatively in the short-run, as labor income is substituted with PL benefits, while the effect is zero in the longerrun. Schönberg and Ludsteck (2014) show that negative long-run effects do emerge if the employment protection period is shorter than the paid PL period

[17] Brewer et al. (2022) show that expanding public child care from half-day to full-day, rather than the availability of free half-day child care, positively affects maternal employment.

The paper is structured as follows: in section 2.2 we provide details about the institutional setting and the changes to the paid PL system in Germany introduced with PB+. Our theoretical framework for the parents' short-run labor supply and its predictions are presented in section 1.3. In section 2.3 we describe the empirical strategy and the data we use in more detail. The results are presented in section 2.4, followed by a discussion of the policy implications in section 1.7. We conclude in section 2.6.

1.2. Institutional setting

1.2.1. The child penalty in Germany

Recent studies have documented the existence and the extent of the child penalty in a large number of countries. Kleven et al. (2019) estimate the child penalty to be 61% five to ten years after child birth in Germany, situating it among the highest when compared to other industrialized countries. Scandinavian countries are found to have the lowest child penalties (21% and 26% in Denmark and Sweden), followed by English-speaking countries (31% and 44% in the United States and the United Kingdom respectively) and German-speaking countries (51% and 61% in Austria and Germany respectively).

The magnitude of the child penalty is determined both by the extensive margin of maternal employment, i.e. whether mothers return to employment after giving birth, and the intensive margin, i.e. how many hours these mothers work. In the case of Germany, the intensive margin explains the majority of the child penalty according to Kleven et al. (2019). Figure A1 shows that the share of mothers who work part-time is among the highest in Germany when compared to other OECD countries. Furthermore, figure A2 suggests that the high levels of part-time work among mothers in Germany indeed arise after child birth: when comparing the incidence of part-time work among women aged 25-29 (proxy for before child birth) and women aged 40-44 (proxy for after child birth), mothers in Germany experience the largest *increase* of part-time work among similarly developed countries.[18]

Recent research offers various explanations for the large magnitude of part-time work in Germany. Cultural norms play an important role. For example, mothers of young children in Germany *believe* that friends and family want them to stay at home, or at most work part-time (Boneva et al. 2021). There is evidence that these beliefs are accurate, as more than 60% of the population in Germany state that

[18]Age groups as a proxy for before/after child birth are motivated by the fact that we are not aware of a data set that offers consistent information across a set of countries on part-time incidence and information on (past) child births.

Table 1.1: Family policies for the treatment and and the control group

		Control	Treatment
Employment protection		3 years	3 years
Unpaid	child's age ≤ 8	3 years	3 years
PL	$3 <$ child's age ≤ 8	1 year	2 years
Paid PL		PB	PB or PB+

Explanation: the employment protection period refers to the period after child birth during which parents can return to their pre-child birth employer. Each parent is entitled a total of three years of unpaid PL, which can be taken until the child's eight birthday. The control (treatment) group can take 1 (2) of these years between the child's third and eight birthday. PB is the only paid PL option available to the control group. The treatment group can choose between PB and PB+ (see table 1.2 for a comparison).

mothers with children under school age or in school should stay at home instead of working (Kleven et al. 2019). Moreover, Boneva et al. (2021) point to the limited availability of affordable child care as a constraint on maternal (full-time) employment. In this paper, we analyze whether public policies can affect maternal labor supply in the long run through monetary incentives to return to (part-time) work earlier after child birth.

1.2.2. Family policies in Germany

Similar to other high income countries, Germany has a set of family policies that aim at making child care and work more compatible for parents. In general, this set of policies consists of three types: first, an employment protection period sets a maximum period directly following the birth of their child, during which parents can choose to leave their job and return to an equivalent job in terms of responsibilities and pay at the same employer. Second, unpaid PL policies allow parents to reduce their working hours or to go on employment leave for some time while their child is younger than a certain age. The employer has to be previously notified about the length and timing of these periods of absence or reductions in working hours. Third, paid PL policies entitle parents to receive government funded benefit transfers for some months immediately following child birth. In most countries, these transfers are made on a monthly basis and their amount depends on pre-child birth labor market income.[19] Importantly, if parents wish to receive paid PL benefits after child birth and adjust their labor supply (i.e. work

[19] All OECD countries - with the exception of the United States - have passed national legislation that offers mothers paid leave for some time around child birth. In the United States, federal legislation entitles mothers who are employed at companies with at least 50 employees to 12 weeks of *unpaid* leave after child birth since 1993 (FMLA). Some states go beyond the federal legislation and offer mothers additional paid and/or unpaid maternal leave.

less or stay at home completely), they have to use their available unpaid PL time for these labor supply adjustments.

For our analysis it is important to understand the set of family policies in place in Germany before the implementation of PB+ (see tables tables 1.1 and 1.2 for a summary). An employment protection period allows parents to return to their pre-child birth employer during three years following child birth. Parents can take unpaid PL for a total duration of three years until their child is eight years old. However, for births prior to the introduction of PB+ (i.e. the control group), only one of these three years could be taken while the child is between three and eight years old.

Prior to the paid PL reform that we analyze, *Parental Benefit* (PB) was the only available paid PL scheme in Germany. PB entitles parents to a maximum of 14 months of benefit payments per birth. These months can be shared between the parents as long as each parent takes at least two but at most 12 months of PL. Thus, if only one parent takes *PB*, the maximum duration is reduced to 12 months.[20] The benefit amount is proportional to the average labor market income during the 12 months before child birth.[21]

More precisely, under *PB* the monthly benefit amount is calculated as follows:

$$benefit(inc^{pre}, inc^{post}) = (inc^{pre} - inc^{post}) \cdot r, \tag{1.1}$$

where inc^{pre} and inc^{post} are the net labor market income prior to child birth (average over 12 months) and after child birth, respectively. The replacement rate r is 65% for the large majority of recipients, however, it rises to 100% for low-income recipients. The benefit amount is limited from below at 300€ for unemployed or low-income parents and from above at 1800€ for high-income parents. Parents are allowed to work a maximum of 30 hours per week while receiving PB, however, any post-child birth labor income is subtracted from the earnings base with which the benefit amount is calculated as equation 1.1 reveals. Thus, *PB* disincentivizes work after child birth. Thus, of any additional Euro earned after child birth, a parent effectively only retains $(1 - r)$ Euros.

The design of PB shapes the maternal labor supply as illustrated in figure A3. Most mother stay at home while receiving PB payments. While only 20% of mothers return to work during the first year after child birth, this number quickly rises to 50-60% in the subsequent 2-3 months. Furthermore, the share of mothers who work part-time doubles from 20% at conception to 40% (75% conditional on

[20]This rule is inspired by Scandinavian PL schemes that intend to incentivize *paternal* PL take-up

[21]Since prior to 2007 a flat amount, independent of pre-birth income was paid, the *PB* reform in 2007 left low-income parents with lower monthly benefit payments while high-income parents received higher benefit payments. Additionally, the benefit duration was reduced from 24 months to 12-14 months. Kluve and Schmitz (2018) and Raute (2019) analyze the effects of the 2007 *PB* reform on fertility and labor market outcomes, respectively.

being employed) two or more years after child birth. Mothers who return to work two or more years after child birth mostly enter part-time work and the share of mothers that works full-time remains constant.

1.2.3. Parental Benefit Plus: the reform

A new paid PL option, called *Parental Benefit Plus* (PB+, German: Elterngeld-Plus) was added to the already existing *PB* in July 2015. The main goal of PB+ is to increase the compatibility of work and family life, the latter mostly referring to child care duties, in a gender-neutral way.[22] Given that more than 80% of mothers stay at home during the first year after child birth, the policy is designed to facilitate their earlier return to work during the first year.[23] Furthermore, the new scheme monetarily incentivizes working part-time (i.e. at most 30h per week) during the first two years after child birth. With this modification, mothers whose child was born on or after 1^{st} July 2015 have the option to choose between the old and the new scheme (see table 1.2 for a comparison of the two schemes).[24] In the period that we study, roughly 20% of mothers choose PB+ (DESTATIS 2019).

PB+ introduces two changes to the paid PL system: first, it doubles the maximum benefit duration from 12 months to 24 months. The additional two months given to couples in which both parents take paid PL for at least two months is also doubled to four months. The maximum duration was increased to prevent couples from running out of paid PL eligibility in case both parents take paid PL in the first 7 months after child birth. Second, the reform partially removes the disincentives to work while taking paid PL by changing the calculation of the benefit amount in the following way:

$$benefit(inc^{pre}, inc^{post}) = \left\{ \begin{array}{cc} \frac{1}{2} \cdot inc^{pre} \cdot r & inc^{post} \leq \frac{1}{2} \cdot inc^{pre} \\ (inc^{pre} - inc^{post}) \cdot r & inc^{post} > \frac{1}{2} \cdot inc^{pre} \end{array} \right\}$$

As long as $inc^{post} <= \frac{1}{2} \cdot inc^{pre}$ (case 1), the benefit amount is independent of inc^{post}. However, since the maximum length is doubled only half of the amount $inc^{pre} \cdot r$ is paid as benefit, i.e. the replacement rate r is halved. If $inc^{post} > \frac{1}{2} \cdot inc^{pre}$ (case 2), the benefit amount is calculated like under *PB*. In both cases 1 and 2, the *total* benefit amount received under *PB+* is at least as high as under *PB*, while the

[22] See https://www.bmfsfj.de/bmfsfj/themen/familie/familienleistungen/elterngeld/elterngeld-73752?view= for more information available in German (as of 22 September 2022).

[23] In the case of fathers the policy intends to achieve a reduction of working hours such that they can allocate more time to child care during this period. We abstract from the policy's effect on fathers, since fathers are not covered in our data and paternal PB+ take-up is low.

[24] Parents can combine both. In practice, parents who take PB+ for at least one month, take it for 19.2 months on average, leaving only 2.5 months for PB until reaching the maximum benefit duration (DESTATIS)

Table 1.2: Comparison of paid PL schemes

	PB (old)	PB+ (new)	
Duration	12 months	24 months	
Benefit calculation	$(inc^{pre} - inc^{post}) \cdot r$	$inc^{pre} \cdot \frac{1}{2} \cdot r$	if $inc^{post} \leq \frac{1}{2} \cdot inc^{pre}$
		$(inc^{pre} - inc^{post}) \cdot r$	if $inc^{post} > \frac{1}{2} \cdot inc^{pre}$
	Monthly amount:	PB $\geq$ PB+	
	Total amount:	PB $\leq$ PB+	

Explanation: paid PL benefits can be received for twice as long under PB+ compared to PB, i.e. 24 instead of 12 months. The paid PL benefit under each scheme is calculated as reported in the second row.

monthly benefit amount is weakly lower in the new compared to the old scheme.[25] The possibility to receive a higher total benefit amount while receiving a lower monthly benefit amount, makes PB+ particularly attractive for mothers who are not financially constrained as we show in our theoretical framework in section 1.3.

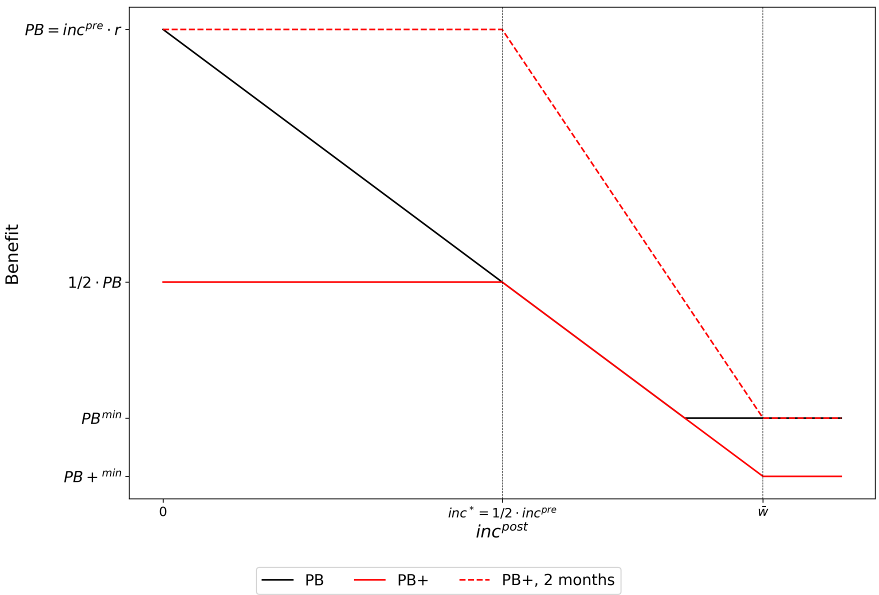
Figure 1.1 illustrates the differences in the calculation of the benefit amount between PB and PB+ for a given level of pre-child birth income graphically. The black line shows that the monthly benefit amount (vertical axis) decreases in post-child birth income (horizontal axis) until reaching the minimum benefit amount of 300€ ($PB{+}^{min}$). Under PB+, if the parents' post-child birth income is at most 50% of their income before child birth, the monthly benefit amount is independent of their income after child birth. However, the replacement rate is halved to limit the increase in the total benefit amount. In the area where parents earn more than 50% of their pre-child birth income, PB+ and PB result in the same monthly benefit amount, the only difference being the reduction in the lower limit of the benefit (150€ instead of 300€). Considering that the duration of PB+ doubles compared to PB, the red dotted line illustrates that the *total* benefit amount (scale on the right) under PB+ can be substantially larger under PB+ relative to PB.

The introduction of PB+ coincides with a change in the *unpaid* parental leave legislation. While all parents are entitled to three years of unpaid PL until the child's eighth birthday, parents who are eligible for PB+ can also choose to use two out of three years (previously one out of three) of unpaid parental leave between the child's third and eighth birthday. This change in the unpaid PL policy might additionally reduce maternal labor supply after the child's third birthday.[26] We

[25]Note that working such that $inc^{post} = \frac{1}{2} \cdot inc^{pre}$ maximizes the total *PB+* payoff

[26]We are not able to precisely disentangle the effect of the change in the unpaid PL regulation from the one in

Figure 1.1: Monthly benefit payment by scheme



The black and the red solid lines show the monthly benefit amount as a function of post-child birth labor income for a given level of income before child birth under PB and PB+, respectively. The red dotted lines represents twice the monthly benefit amount under PB+, illustrating the fact that PB+ can be taken for twice as long as PB (i.e. one PB month equals two PB+ months).

discuss how this change affects the interpretation of our results when presenting our findings in section 2.4. The employment protection period remains unchanged over the whole period of analysis. Tables 1.1 and 1.2 show a summary of the family policies available to the control and the treated groups and the differences between the new part-time paid PL option and the pre-existing scheme.

Given the complexity of the changes in the paid PL benefit structure, an important question is how parents understand these modifications. Both in government sources and on third-party websites that offer advice about the German parental leave system, the policy's modification are explained in rather general terms and with a few examples for benefit calculations.[27] The general explanations highlight that *i)* parents can receive PB+ for twice as long as PB, *ii)* if parents do not work, the *monthly* benefit is halved, and *iii)* if parents work, the *monthly* benefit under PB+ can be as high as under PB, implying that parents can receive a substantially higher *total* benefit amount due to the longer duration. The general recommendation is that PB+ is beneficial for mothers who would like to return to work earlier, in particular those who would like to work part-time (up to 30 hours per week as defined under PB+). Additionally, several example calculations of the benefit calculation highlight the fact that PB+ pays a (weakly) lower *monthly* but a (weakly) higher *total* benefit amount if parents work (see table A1). In sum, the vast majority of parents are likely familiar with these simplified ideas behind the policy rather than with its exact details.[28]

1.3. Theoretical framework: short run labor supply under PB and PB+

To study the mothers' paid PL choice, i.e. whether to choose PB or PB+, we use a simple 2-period labor-leisure model as our conceptual framework. This theoretical framework models maternal labor supply decisions in the first two years after child birth, which correspond to the period in which mothers in the treatment group are eligible for PB+. The model accurately captures the benefit structure of the paid PL schemes, however, it abstracts from aspects such as intra-household dynamics and spillovers of labor choices in period one on future wages or child

the paid PL regulation. However, this does not affect our ability to rule out a lock-in effect, since it implies that our zero effect on full-time employment is a lower bound of the true effect size, i.e. it might be positive in absence of the change in the unpaid PL regulation.

[27] This paragraph is based on Bundesministerium für Familie (2020) and https://www.elterngeld.net.

[28] Note that the official explanations are a simplification of the actual benefit calculation. For example, while the official explanations stress part-time employment, the actual benefit calculation depends on the relation of an individual's post-child birth to pre-child birth earnings.

quality. In this framework, a representative agent faces the following optimization problem:

$$\max_{c_1,c_2,l_1,l_2} U(c_1,k_1) + \beta \cdot U(c_2,k_2) \tag{1.2}$$

$$s.t. \quad c_1 + \frac{c_2}{(1+r)} = \overbrace{(1-\tau_1)\cdot w \cdot l_1 + \gamma_1 \cdot y_0}^{\text{Period 1}} + \frac{\overbrace{(1-\tau_2)\cdot w \cdot l_2 + \gamma_2 \cdot y_0}^{\text{Period 2}}}{1+r}$$

$$k_t = 1 - l_t$$

The agent chooses consumption c_t and hours worked l_t (and thereby implicitly child care k_t) in periods 1 and 2. She faces a standard inter-temporal budget constraint augmented by two terms that capture the paid PL scheme in the two periods $t \in \{1,2\}$: $\gamma_t \cdot y_0 - \tau_t \cdot w \cdot l_t$, where the first term is the pre-child birth income dependent amount that a parent receives irrespective of post-birth income and the second term is the amount by which the benefit amount is reduced if a parent works post-child birth under PB (and under some conditions also under PB+). In this specification, γ_t and τ_t represent the fraction of the pre-child birth earnings that are replaced by the PL scheme and the fraction of post-child birth earnings that are subtracted from the former amount, respectively. Pre-child birth income y_0 is exogenously given and independent of preferences over labor choices or the wage rate.

We derive the predictions of the theoretical framework by solving the agent's problem with an additively separable utility function of the form $U(c_t,k_t) = \log(c_t) + \log(k_t)$. To do so, we solve for the optimal labor supply choices in both periods as a function of the exogenous policy parameters, namely γ_t and τ_t, as well as pre-child birth income y_0 (see figure A4).[29] We infer the agent's preferred paid PL scheme based on the associated utility level (see figure A5).[30] Details are provided in section 1.A.3 in the appendix.

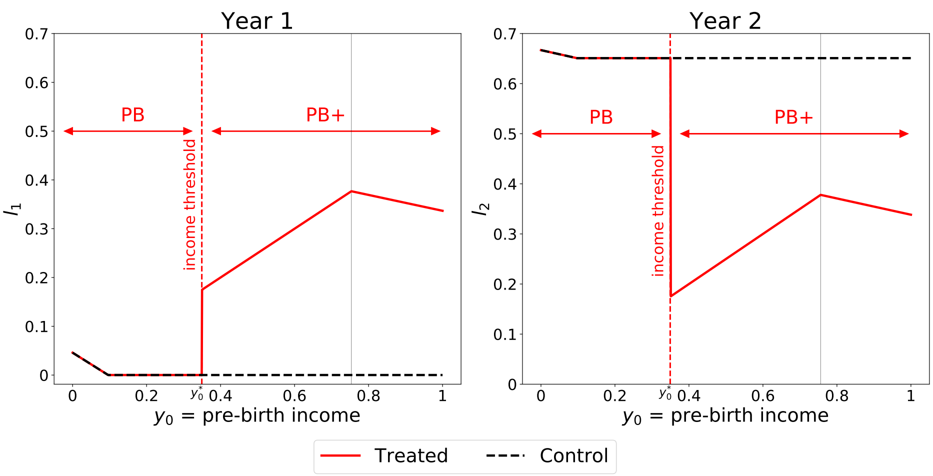
Figure 1.2 illustrates the choices of labor supply in periods one and two on the vertical axis as a function of pre-child birth labor income on the horizontal axis for the control (black dashed) and treatment (red solid) groups. These optimal labor supply choices allow us to derive the following three predictions: first, there exists an income threshold since only mothers with sufficiently high income prior to child birth, i.e. $y_0 > y_0^*$, choose PB+.[31] This income threshold is explained by

[29] The exogenous policy parameters are reported in table A2.

[30] In reality parents can choose combinations of the three alternatives, from which we abstract in this model. Official statistics show that mothers who take at least one month of PB+, receive PL benefits for 19 months on average, implying that the vast majority of benefit months are PB+ and leaving room for at most three months of PB (DESTATIS).

[31] As the more detailed figure A6 illustrates, an agent with pre-child birth income between $l_{0,t}^{high}$ and $l_{0,t}^{low}$

Figure 1.2: Labor supply choices as predicted by 2-period model



The figures on the left and right show the labor supply choices as a function of pre-child birth income in the first and second year after child birth, respectively. The red (black) line represents mothers who are (not) eligible for PB+.

the fact that the lower replacement rate under $PB+$, which is halved relative to PB, can only be afforded by individuals that are not financially constrained prior to child birth.[32] Second, since the changes to the benefit calculation introduced with $PB+$ make working less costly in the first and more costly in the second period compared to PB, PB+ takers smooth their labor supply across periods, i.e. $l_1^{treated} > l_1^{control}$ and $l_2^{treated} < l_2^{control}$. Third, compliers return to work earlier after child birth than individuals in the control group. This follows directly follow from labor supply smoothing as it raises the individual's labor supply in the first period above a level of zero.

is indifferent between $PB+^{high}$ and $PB+^{low}$ and chooses $PB+^{low}$ for a higher level of pre-child birth income. However, in our empirical analysis we abstract from the distinction between $PB+^{high}$ and $PB+^{low}$.

[32]For a smaller fraction of mothers with $inc^{post} > \frac{1}{2} \cdot inc^{pre}$, the income threshold is explained by the fact that PB+ is only attractive if the part of the benefit scheme that depends on pre-child birth income in the second year ($\gamma_2 \cdot y_0$, where $\gamma_2 = 0.65$) is sufficiently high, such that it compensates for the "tax" τ_t on post-child birth income ($\tau_2 \cdot w \cdot l_2$, where $\tau_2 = 0.65$).

1.4. Empirical strategy and data

1.4.1. Empirical strategy

To estimate the causal effect of the policy on maternal labor market outcomes we exploit the fact that only parents whose child is born on or after July 1^{st} 2015 are eligible for *PB+*. While these parents have the option of choosing between *PB+* or *PB* (or a combination of the two), parents whose child is born before the implementation date are only entitled to *PB*. This institutional feature allows us to estimate a series of Difference-in-Differences (DD) models for outcomes p months after child birth.[33] In these DD models we compare outcomes of mothers whose child was born in the months after vs. before the threshold date in the year of the reform (2015) vs. the pre-reform years (2011-2014). The interaction of these two differences captures the policy's treatment effect. Intuitively, the treatment effect equals the difference in outcomes between mothers with children born in August - September 2015 and those with children born in April - May 2015 net of the average seasonal difference in outcomes between these two groups of birth months in previous birth years.[34] We restrict our data to two birth months on each side of the implementation month. Since PL take-up is not observed in our data we estimate the policy's Intention-to-Treat effect and interpret adjustments in maternal labor supply during the first two years after child birth as a proxy for the take-up of PB+.

In our baseline specification we estimate the policy effect p months after child birth in p separate regressions of the following form:

$$\begin{aligned} y_{i,p} = \alpha_p + \beta_p \cdot (treatYear_i \times treatMonths_i) \\ + \delta_p \cdot treatMonths_i + \phi_{j(i),p} + \phi_{c(i),p} + \boldsymbol{\gamma_p} \cdot \boldsymbol{X_i} + \epsilon_{i,p} \end{aligned} \quad (1.3)$$

In equation 1.3, $y_{i,p}$ stands for individual i's outcome of interest p months after child birth, $treatYear_i$ and $treatMonths_i$ are dummies for births that occur in the reform year (2015) and in the post-implementation months (August and September in each year), respectively. The $treatMonths_i$ indicator, in combination with the inclusion of four pre-reform child birth cohorts from 2011-2014, enables us to net out seasonality in parental characteristics across months. We estimate the policy's effect within 401 German counties and five birth years (2011 - 2015) by including

[33]The reform's implementation setting, in which a threshold date determines whether an individual belongs to the control or the treatment group, is usually exploited within a Regression-Discontinuity-Design (RDD). However, in the IEB data the date of birth is proxied by the date on which the mother goes on maternity leave, which starts six weeks before birth on average (Müller and Strauch (2017); see section 1.4.2). This measurement error in the date of birth makes a DD strategy preferable to an RDD.

[34]A similar approach has been used previously, for example in Schönberg and Ludsteck (2014).

birth-county and birth-year fixed effects, namely $\phi_{c(i),p}$ and $\phi_{j(i),p}$. The vector $\boldsymbol{X}_i$ controls for additional individual-level characteristics, such as pre-child birth income, part-time work, age, the job's skill-level and industry fixed effects. All of these control variables are measured at conception. The standard errors are clustered at the county-level throughout the analysis.

The identifying assumption of our empirical strategy is that, in absence of the policy, the differences in outcomes between mothers with children born in August - September of 2015 and April - May of 2015 would not have been different from the ones observed for these same groups (August - September and April - May) in the previous years (2011 - 2014). While this assumption is not verifiable by definition, we show that there are no trends in differential outcomes between the two groups of mothers with children born in the years prior to the implementation year (see placebo tests in section 1.6).[35] Additionally, the characteristics of mothers in the treatment and in the control groups are required to be balanced at baseline (i.e. conception) to ensure comparability. Any difference in such characteristics that exists already before child birth, for example a higher probability of working part-time, is likely to reappear after child birth and would mistakenly be interpreted as a treatment effect. For this reason we show that the there are virtually no statistically significant differences in observable characteristics at conception by estimating our baseline specification with a set of employment-related and personal characteristics as independent variable (see figure A14).

A potential concern for the validity of the causal relationships estimated in this paper is whether the enactment of the law could be anticipated by parents. The German parliament passed the law on November 7^{th} 2014, implying that parents could be certain about the availability of the new part-time option under PB+ eight months before its implementation. Importantly, this is less than the nine months congestion period and even in a fertile couple three to six months are needed for conception (González 2013, Raute 2019). Thus, anticipation is very unlikely to play a role in explaining the observed effects. The balance in characteristics at conception discussed in the previous paragraph supports this interpretation (see figure A14).

1.4.2. Data

The main data source in this project are the Integrated Employment Biographies (IEB) which are provided by the German Institute for Labor Market Research and based on social security records. These data consist of the entire employment history of all social security covered employees in Germany excluding public

[35]Note that this exercise corresponds to the test for trends in outcomes between the treated and the control group prior to treatment assignment (i.e. pre-trends) in a classical Difference-in-Differences setting.

employees and the self-employed. These administrative records include highly detailed information on the gross daily wage, the start and end of an employment contract, the employer, the industry, the type of contract and the skill-level.[36] From these data we first select all mothers who give birth between the years 2008 and 2018. Since births are not directly recorded in the IEB we follow Müller and Strauch (2017) and identify all women who experience a maternity leave-related employment interruption as mothers.[37] This approach identifies women who go on mandatory maternity leave six weeks before the child's expected date of birth in 89% of the cases. The (expected) date of birth is then taken to be six weeks after the start of maternity leave.

We only consider first-time mothers who give birth to their first child during 2011-2015 (as observed in our sample) and are employed at conception. These mothers are particularly likely to take up PB+, since the latter incentivizes mothers to return to work in the months after child birth. In contrast, mothers who are already unemployed before giving birth are unlikely to respond to the policy, since they are only entitled to the minimum benefit amount (300€ and 150€ under PB and PB+, respectively). Similarly, mothers who have already given birth to a child prior to the child birth that we observe are likely less responsive to the introduction of PB+, since they are already less attached to the labor market (i.e. many of them already work part-time before the subsequent child birth). Additionally, we restrict the sample to mothers who are between 20 and 38 years old, since the birth identification procedure described above works most reliably in this age group.

Data on the availability of public child care for 0-2 year olds at the county (Kreis) level for the years 2011 - 2015 come from administrative records of the Statistical Offices of the German provinces. Child care availability is measured as the fraction of available slots for children aged 0-2 in the number of children in the same age group in a given county. Public child care for children under three years old is highly subsidized such that parents only cover 14% of the total operating costs on average. The fees paid by parents depend negatively/positively on family size/income and range from 0 to 600€ (Sandner et al. 2020). These data are merged to the data described above based on the year of child birth and the mother's county of residence at conception. We then use this information to assess whether their exist complementarities between the availability of affordable, public

[36]Unfortunately, these data do not record some relevant information at the personal and/or household level, such as whether an individual is married or cohabits with a partner.

[37]This approach makes use of the IEB's information on employment interruptions and their underlying reasons which employers are obliged to notify. Müller and Strauch (2017) flag all women who experience an employment interruption "due to entitlement to other compensation by the statutory health insurance provider (value 51 of 'grund' [variable])" as *potential* mothers. This reason for employment interruption can be due to maternity allowances, which are paid during paid maternity leave to actual mothers, and sickness allowance. The authors use further restrictions on the age of potential mothers and the lengths of the employment interruption to disentangle the two underlying reasons for employment interruptions

child care and the take-up of the part-time work option in the paid PL system in Germany. As shown in figure A7 there is large variation in the availability of public child care across counties, ranging from 5% to 45%.

Our final sample consists of roughly 380'000 mothers who give birth to a child in the months of interest (May, April, August, September) in the years 2011 - 2015, corresponding to 20'000 births per month. Descriptive statistics of the full sample are presented in table A3. Mothers are on average close to 30 years old at conception and 23% have a university degree. Furthermore, 27% work part-time prior to birth and the average gross monthly labor income equals 2500€.[38] These numbers closely match official statistics on the characteristics of mothers prior to birth and, thereby, corroborate the validity to our strategy of identifying mothers.

1.5. Results

1.5.1. All mothers

The policy's effects on labor market outcomes of all mothers are reported in figure 1.3. The three panels report the Intention-to-Treat effects for employment, part-time work (less than 20 hours per week) and full-time work (20 hours or more per week). The absence of pre-trends, i.e. significant differences in labor market outcomes between the treated and the control group before the take-up of child birth, validates that mothers in the treated and the control groups are indeed comparable. Furthermore, it also suggests that parents do not strategically alter their labor supply prior to child birth.

The first panel in figure 1.3 shows that the reform leads to an increase in employment of approximately 1-2pp ($\approx$ 5-10%) in the first year after child birth, i.e. when treated mothers are eligible for both PB and PB+. Since there is no positive employment effect in subsequent years, these mothers choose PB+ to return to work earlier, i.e. they would have returned to work at a later point in absence of the policy.[39] In the absence of individual-level data on paid PL usage, we interpret the employment effect in the first year as a proxy of PB+ take-up throughout the analysis.[40] The policy does not affect the probability of employment

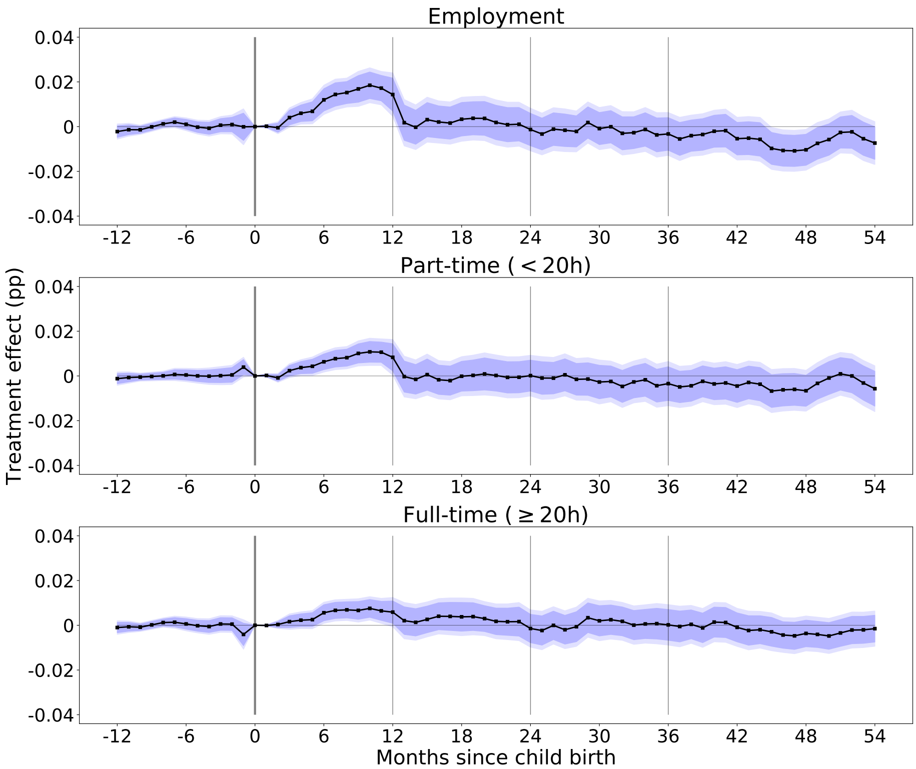
[38]Note that the benefit amount in the German paid PL system is calculated based on the net rather than the gross monthly income. A significant share of mothers change the taxation class (e.g. switch classes with their partner), which determines the tax rate in Germany, in the year before birth to increase the benefit amount they receive. Since we do not know which mothers indeed change their taxation class, we cannot infer the benefit amount they receive after child birth.

[39]Roughly two thirds of the employment effect in the first year are explained by transitions to part-time work while the rest comes from increased full-time employment as shown in the second and third panels.

[40]Note that interpreting the employment effect in the first year as a proxy for PB+ take-up only captures mothers that take-up PB+ *and* adjust their labor supply (compared to the control group). Since mothers can also choose PB+ without changing their labor supply, our proxy of PB+ take-up is lower than the take-up

along the intensive margin in the second year. This implies that, in contrast to the predictions of our theoretical framework, mothers do not smooth their labor supply across the two first years after child birth. As a result, the policy increases maternal labor supply and earnings during the first two years after child birth.

Figure 1.3: Employment outcomes of all mothers



The figures above report the DD coefficient of separate regressions for outcomes p month after child birth as specified in regression 1.3. The outcome variables in the first, second and third panels are a dummy for individual is employed, part-time employed and full-time employed, respectively. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals. See the corresponding table A4.

In line with the effects on employment, monthly earnings increase by 1.5 log points in the first year after child birth and are unaffected in the following year as shown in the first panel of figure A8. Thus, the policy reduces the child penalty in the first year. The second panel of the same figure illustrates that employer continuity, i.e. the probability of working for the same employer after as before child birth, increases by almost 2pp in the first year as well. This means that almost

reported in official statistics (DESTATIS 2019).

all mothers who return to work earlier as a result of the policy, do so by returning to their pre-child birth employer. Employer continuity is unaffected in the second year.

The absence of labor supply smoothing is most likely explained by a combination of the framing of the policy and the pre-existing patterns in maternal labor supply. More precisely, the policy is framed as an option for mothers who want to return to the labor market earlier after child birth, in particular during the first year after child birth, while less attention is given to labor supply adjustments thereafter. Given the complexity of the benefit amount calculation, this might result in the observed pattern in which mothers return to the labor market earlier in the first year and leave their labor supply unchanged during the second year. Additionally, roughly two thirds of women employed in the second year after child birth worked in part-time jobs in the pre-reform years (see figure A3). Thus, already before PB+ was implemented, a large fraction of mothers adjusted their labor supply in the second year as in line with the general recommendations the new part-time scheme, which emphasize that the policy is particularly beneficial in combination with part-time work.

After the employment protection period ends (i.e. three years after child birth), treated mothers are temporarily 1.5pp less likely to be employed as depicted in panel 1 of figure 1.3. We attribute this effect to the change in the unpaid PL legislation (i.e. 2 instead of 1 year of unpaid PL from the child's 3^{rd} to 8^{th} birthday) rather than to the pro-part-time incentives in the paid PL system for two reasons:[41] first, parents who go on employment leave using their unpaid PL time are recorded as unemployed in our data. Second, among high income mothers in West Germany, who are particularly likely to make use of PB+, the temporary drop in employment around year 4 is no longer statistically significant (see section figure 1.4 and section 1.5.2). This suggests that the temporary drop in employment is driven by a different subsample than the one that responds to pro-part-time incentives.

Panels 2 and 3 of figure 1.3 show how the reform affects employment at the intensive margin in the long run. For most months after the third year, the point estimates are negative, however, close to zero and not statistically significant. In a small number of months the coefficients are almost statistically significant at a 10% significance level. However, those are precisely the months that coincide with the temporary drop in employment in panel 1. Thus, these slightly negative estimates for part-time and full-time work are most likely the result of the change in the unpaid PL legislation. The fact that the effect on full-time employment is a precise null effect among the sample of high-income mothers in West Germany,

[41] In Germany, parents can take three years of unpaid parental leave until the child is 8 years old. Unpaid parental leave can be taken in 3 parts and a maximum of two years can be taken between the child's third and eight birthday. Parents need to accompany paid with unpaid parental leave in order to modify their working hours.

for which the temporary drop in unemployment is less pronounced (see figure 1.4), corroborates this interpretation. Based on this evidence, we conclude that the reform's pro-part-time incentives do not affect employment at the intensive margin in the long run.

In absence of any labor supply adjustments along the intensive margin, the temporary drop in employment around 4 years after child birth translates into a temporary reduction in labor market earnings of just below 0.1 log points as shown in figure A8. Thus, in the sample of all mothers the child penalty is temporarilly slightly increased in the long-run. Following the explanation in the previous paragraphs, this is most likely caused by the reform's change in the unpaid PL legislation rather than by its pro-part-time incentives. Figure A8 also shows that the employer continuity after the third year turns is unaffected in the long run. We do not find evidence for the policy affecting other labor market outcomes, such as job quality, or fertility.

1.5.2. Heterogeneity

In this section we focus on mothers that are particularly likely to make use of PB+. In absence of individual-level information on the take-up of PB+, we rely on the hypotheses of our theoretical framework, which identifies pre-child birth income as an important determinant for PB+ take-up (see section 1.3), and on the existing literature on maternal labor supply in Germany. The latter points to two heterogeneities that might be relevant for the take-up of PB+: first, there are important cultural differences between East and West Germany that result in different maternal labor supply patterns in these regions (Boelmann et al. 2020). These might lead to a differential take-up of PB+ in East and West Germany. Second, mothers perceive a lack of child care availability as a constraint on maternal employment (Boneva et al. 2021), suggesting that the take-up of PB+ might be facilitated by a higher local provision of affordable public child care.

High and low income mothers in West Germany

To empirically examine whether the theoretically predicted income threshold exists, we interact the main effect in the baseline specification with monthly pre-child birth earnings in table 1.3. By binning the earnings into five categories (quintiles) we are able to assess the interaction of the policy's main effect over the distribution of monthly earnings. The first two columns in table 1.3 report the resulting coefficients for employment in the first year after child birth (months 6 and 9 months). The first five rows show that the there is no employment effect for the bottom 40% of the income distribution (rows 1 and 2) and that the upper 60% entirely drive the employment effect observed in the main sample. As a result, the

Table 1.3: Take-up of PB+

Outcome:				
Employment in month	6	9	6	9
treatMonths=1 × treatYear=1	-0.002	0.002	0.016***	0.020***
	(0.005)	(0.007)	(0.003)	(0.003)
treatMonths=1 × treatYear=1 × earnings (Q2)	0.006	0.002		
	(0.007)	(0.008)		
treatMonths=1 × treatYear=1 × earnings (Q3)	0.023***	0.021**		
	(0.007)	(0.009)		
treatMonths=1 × treatYear=1 × earnings (Q4)	0.020***	0.023***		
	(0.007)	(0.008)		
treatMonths=1 × treatYear=1 × earnings (Q5)	0.019**	0.028***		
	(0.007)	(0.009)		
treatMonths=1 × treatYear=1 × east=1			-0.020***	-0.016**
			(0.006)	(0.007)
Mean of outcome	0.097	0.137	0.097	0.137
SD of outcome	0.30	0.34	0.30	0.34
Controls	Yes	Yes	Yes	Yes
County FEs	No	No	Yes	Yes
Sample size	380717	380717	380717	380717
Number of clusters	401	401	401	401

Note: the table above reports the coefficients of interest of the baseline model (equation 1.3) fully interacted with *i)* pre-child birth earnings (measured at conception and binned into quintiles) in rows 1 - 5 and *ii)* a dummy for residing in East Germany in rows 1 and 6. Row 1 represents the reference category. The outcome variables are employment 6 months after child birth in columns 1 and 3 as well as 9 months after child birth in columns two and four. Standard errors are clustered at the county level (401) and reported in parentheses. Descriptive statistics are reported at the bottom of the panel. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

point estimates in the upper 60% range from 1.8 - 2.8 pp and are slightly larger than the effects observed in the full sample. Moreover, within the upper 60% of the income distribution the effect is homogeneous. The positive employment effect in the first year being a proxy for PB+ take-up, this empirically confirms the existence of an income threshold.

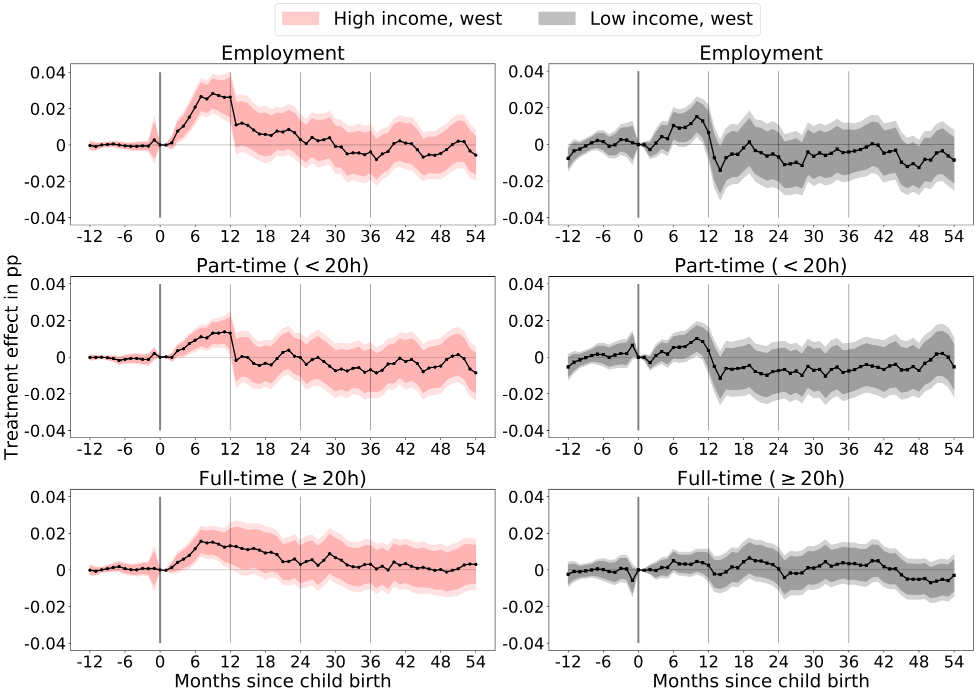
In the last two columns of table 1.3 we interact the policy's main effect with a dummy for whether a mother resides in East Germany at conception. We do so to understand whether there are differences in the policy take-up between East and West Germany. The first row shows that the policy increases employment in West Germany (reference category) during the first year by almost 2pp. This effect in West Germany is slightly higher than the one observed in the whole sample, which is due to the fact that mothers in East Germany do not make use of the policy. As can be seen in the last row of table 1.3 the interaction of the policy's main effect with the for East Germany dummy (1.9-1.6pp), reduces the effect in the first row to 0. The reason is likely that East German mothers tend to return to work earlier after child birth than their West German counterparts (Boelmann et al. 2020). PB+ partially closes this gap in the first year after child birth.

Motivated by this analysis, we analyze the full range of employment outcomes for high and low income mothers in West Germany in the panels on the left and right hand side of figure 1.4, respectively.[42] The first panel on the left illustrates that the employment effect in the first year is substantially larger (2.5 - 3pp) among high income mothers when compared to the sample of all mothers. This is due to the fact that low income mothers barely take up PB+ as can be seen in the first panel on the right. The temporary drop in the employment probability after the third year is less pronounced for both high and low income mothers. The absence of a negative temporary effect on employment for high income mothers in the long run supports our interpretation that the drop in employment probability for all mothers (see figure 1.3) is not driven by the pro-part-time incentives of the paid PL reform. Given that the long-run effects on employment along the extensive margin are similar among both high and low income mothers, i.e. mothers who take up PB+ and those who do not, they are most likely due to the change in the *unpaid* PL legislation, which makes the timing for employment leave more flexible for all mothers irrespective of whether they take PB+ or not.

To assess whether pro-part-time incentives lead to a lock in effect into part-time employment we analyze the policy's effects on part-time and full-time employment in the panels in the second and third rows of figure 1.4. While we would observe such a lock-in effect as a higher probability of working part-time and a lower probability of working full-time, we do not find any statistically significant effect

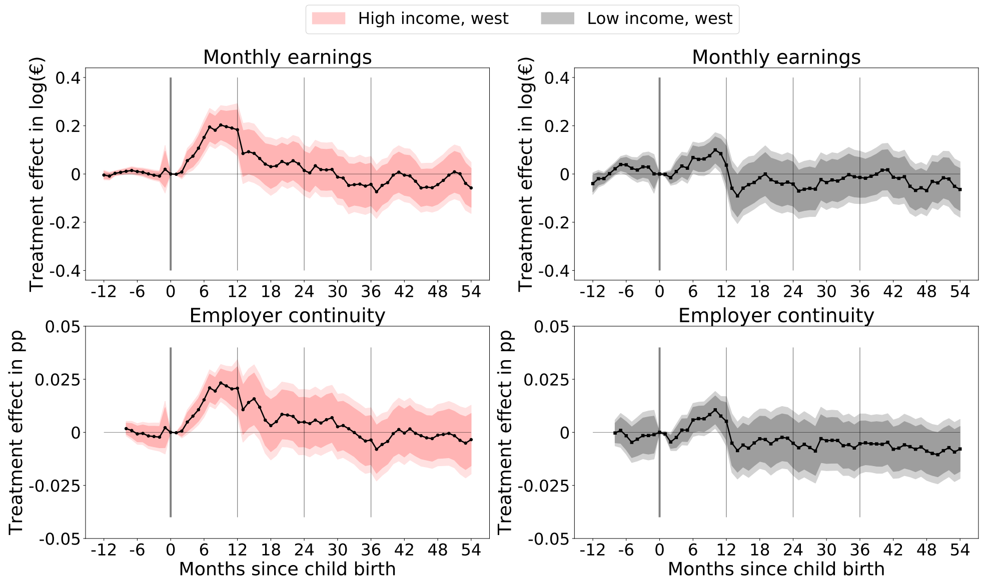
[42]To ensure equal sample sizes in both sub samples, we define high and low income mothers as being above or below the 50^{th} percentile of the pre-child birth income distribution, respectively.

Figure 1.4: Employment outcomes for high and low income mothers in West Germany



The figures above report the DD coefficient of separate regressions for outcomes p months after child birth as specified in regression 1.3. The outcome variables in the first, second and third panels are a dummy for individual is employed, part-time employed and full-time employed, respectively. High and low income mothers are defined as being in above and below the 50^{th} percentile in the pre-child birth income distribution, respectively. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals. Also see the corresponding tables A5 and A6.

Figure 1.5: Additional employment outcomes of high and low income mothers in West Germany



The figures above report the DD coefficient of separate regressions for outcomes p months after child birth as specified in regression 1.3. The outcome variables in the first and second panel are monthly earnings and employer continuity, i.e. a dummy that equals 1 if a mother works for the same employer after child birth as at conception. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals.

on part-time or full-time employment among high or low income mothers after the first year. Focusing on high income mothers, who strongly respond to the policy's pro-part-time incentives, the second panel shows that the point estimates on part-time employment after the third year closely match the ones observed for employment at the extensive margin (panel 1). This suggests that high income mothers who go on unpaid PL during that period do so by taking time off their part-time work. Furthermore, the effect on full-time employment among high income mothers after the third year is a precisely estimated zero (panel 3, left). Given that some full-time employed mothers might go on *unpaid* PL (i.e. unemployment) during this period, our estimate represents a lower bound for the true effect of pro-part-time incentives shortly after child birth on full-time employment. These results allow us to rule out that pro-part-time incentives lead to a lock-in effect into part-time employment in the case of the policy that we analyze.

We do not find a consistent, statistically significant effect of the policy on

Table 1.4: PB+ take-up and public child care availability

Outcome: *Employment in month*	3	6	9	12	15	18	21	24	27
treatMonths=1 × treatYear=1	0.033***	0.040***	0.054***	0.066***	0.002	-0.005	0.002	0.001	0.007
	(0.011)	(0.015)	(0.016)	(0.023)	(0.032)	(0.029)	(0.022)	(0.021)	(0.025)
treatMonths=1 × treatYear=1 × child care	-0.001*	-0.001	-0.001	-0.001	0.000	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of outcome	0.051	0.096	0.141	0.310	0.573	0.640	0.649	0.661	0.666
SD of outcome	0.22	0.29	0.35	0.46	0.49	0.48	0.48	0.47	0.47
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	151743	151743	151743	151743	151743	151743	151743	151743	151743
Number of clusters	400	400	400	400	400	400	400	400	400

Note: the table above reports the coefficients of interest for employment as outcomes p months after child birth (in columns) as specified in regression 1.3. Three birth months on each side of the implementation month are included. The first row reports the OLS estimates of the DD model's main effect. The second row reports the main effect's interaction with the child care availability, which is measured at conception and standardized to mean = 0 and standard deviation = 1. Standard errors are clustered at the county level (400) and reported in parentheses. Descriptive statistics are reported at the bottom of each panel. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

monthly earnings for high or for low income mothers in the long run as illustrated in the first panel of figure 1.5. The point estimates after the third year range from 0 to -0.05 log points, however, are not statistically significant in any month. For both high and low income mothers, the earnings pattern closely matches the policy's effect on employment along the extensive margin, which is most likely driven by a more flexible unpaid PL legislation (see discussion of figure 1.4). This corroborates our interpretation according to which the reform's pro-part-time incentives do not affect the child penalty up to 4.5 years after child birth. The high and low income mothers' employer continuity is not significantly affected by the policy in the long run either.

Complementarities with public child care

Previous research documents that German mothers identify the insufficient availability of child care as a major constraint for maternal employment (Boneva et al. 2021). Therefore, we examine whether the policy's effect is larger in counties with a higher availability of public child care. In Germany, public child care is inexpensive by international standards due to high subsidies, however, the availability is limited. This makes it reasonable to proxy the likelihood of being able to access public child care by the availability of child care slots, disregarding financial considerations.

To assess our prediction of complementarities between PB+ take-up and public child care availability, we interact the baseline treatment effect with the public child care availability at the county level (see section 1.4.2 for details on measurement). We focus on the subsample of high income mothers in this exercise since low income mothers do not take up PB+ because of financial constraints.

Table 1.4 reports the resulting coefficients. Surprisingly, we do not find evidence for complementarities between the maternal employment effect, i.e. the take-up of the policy, and the availability of public child care. This result is robust to using more flexible specifications, such as interacting the main effect with a squared term of public child care. The absence of complementarities between the availability of public child care and employment in the first years after child birth seems to contrast previous survey evidence in the literature (Boneva et al. 2021). These differences can potentially be explained by the fact that only high income mothers return to work earlier as a result of the policy. Being less financially constrained, these mothers might have better access to other child care options, such as private providers. Additionally, fathers, who are incentivized to work less by PB+, might (partially) compensate for the absence of mothers by engaging more in child care duties at home.[43] Taken together, we do not find that the availability of public child care fosters employment for high income mothers in the first year after child birth.

1.6. Robustness checks

In this section we present a series of robustness checks to corroborate our findings and, in particular, their causal interpretation.

Placebo tests If the observed results are indeed caused by the introduction of PB+, we should not observe any statistically signifcant treatment effect in other years. To examine whether this is the case we perform placebo tests, in which we estimate our baseline specification (equation 1.3) and sequentially define births in the pre-reform years 2011 - 2014 as treated. To ensure that the control group is not affected by the policy, parents who give birth in the actual reform year are dropped from the sample when estimating the placebo regressions (2011-2014 treated). Figure A9 shows the estimates of the true implementation year in panel 1 and of the placebo treatments in panels 2-5. It is evident that the positive employment impact is only present when the treatment status is assigned using the true implementation year. For all placebo treatments, the coefficients in the first year is close to zero and statistically insignificant.

Month of birth window around implementation date In our baseline results we restrict the sample of analysis to two birth months, excluding the ones immediately

[43]Unfortunately, we are not able to distinguish between these two potential explanations due to the lack of data on paternal labor supply and non-public child care options.

preceding/following the implementation month. Since the policy was introduced on 1 July, our sample of analysis includes the birth months April, May, August and September. In this way we ensure that parents on either side of the threshold are more comparable than if we included all births, given the seasonal differences in characteristics of parents by birth month (Buckles and Hungerman 2013). As a robustness check, we test whether our main findings hold when further restricting birth month window to one month on either side of the threshold. Figures A10 to A13 show that the effects on post-child birth employment patterns under this alternative specification are very similar.

Characteristics at conception The causal interpretation we give to our findings requires that the introduction of the policy was not anticipated by parents. If it was not anticipated, there should not be any differences in parental characteristics, measured at conception, between parents whose child is born before or after the implementation date (net of seasonality). We test whether such differences exist by estimating our baseline model (as specified in equation 1.3 but excluding controls) in which we take a set of parental characteristics at conception as outcomes. Figure A14 reports the corresponding results. The Difference-in-Differences model's main effects for each characteristic is reported in rows. The differences in characteristics at conception are small in magnitude (< 0.01 sd) and statistically insignificant at a 5% significance level, which makes anticipatory timing of births unlikely and corroborates the causal interpretation of our results.

1.7. Discussion

In this section we discuss to what extent the PB+ reform (i.e. the pro-part-time incentives and a longer benefit duration) has achieved its goals and what policy recommendations can be drawn from our analysis more broadly. The policy's goal is to increase the compatibility of work and family life in a gender-neutral way.[44] This means that the policy intends to *i)* encourage mothers to return to work earlier after child birth such that their time off work is reduced, *ii)* make fathers work less such that they can allocate more time to child care duties and *iii)* encourage couples to share child care duties more evenly.

Our analysis shows that the policy clearly achieves its first goal of encouraging mothers to return to work earlier as the share of mothers who return to work during the first after child birth, instead of later years, increases by approximately 10% on average. However, it is not clear whether this higher maternal labor market

[44] See https://www.bmfsfj.de/bmfsfj/themen/familie/familienleistungen/elterngeld/elterngeld-73752?view= for more information available in German (as of 22 September 2022).

attachment in the short run is desirable from a policy perspective as it does not translate into labor market gains for mothers (i.e. a reduction of the child penalty) in the long run. To answer this question, our study of maternal labor market effects would need to be complemented with additional analyses. An assessment of the policy's effects on child development, which depend on the quality of child care with which the mother's care is substituted, would be needed. Prior research suggests that the period in which PB+ incentivizes mothers to return to work (i.e. the first year after child birth) is particularly important for child development (e.g. Heckman 2008; Cunha and Heckman 2007). Additionally, the evidence concerning the effects of early childhood education on children from high-income households, which are precisely the ones that are affected by PB+, is mixed (e.g. Drange and Havnes 2019; Fort et al. 2019).

Another important factor for the assessment of the policy's overall effects, is whether PB+ encourages fathers to reduce their labor supply and to contribute more to child care duties, in particular in couples in which mothers return to work earlier as a result of the policy. The absence of complementarities between the availability of public child care and the take-up of PB+ among mothers could be explained by a higher fraction of fathers taking care of the child as mothers return to work earlier. However, according to official statistics, five times more mothers than fathers choose PB+ in 2015, suggesting that paternal child care could account for at most a fraction of the increase in the supply of non-maternal child care (DESTATIS 2019).[45] Alternatively, private child care centers or informal agreements with friends and family could compensate the lower maternal child care involvement. As our analysis is limited by the lack of data on child development, paternal labor supply and alternative child care arrangements, further research is needed to provide a comprehensive assessment of PB+.

Furthermore, as we show theoretically and empirically, the benefit structure introduced with PB+ encourages only high income mothers to take up the policy. Since PB+ pays a lower monthly benefit amount compared to the alternative PB for most mothers, the new scheme is unattractive for low income mothers. As official documents do not make any reference to specifically targeting high income mothers, this is most likely an unintended feature of the policy and emphasizes that the effects of incentive structures on different subgroups need to be taken into account when designing policies.

In terms of public finances, the part-time paid PL option does not increase public expenditure on paid PL benefits for mothers. This is explained by the fact that in Germany, as in most other countries, the total amount of paid PL benefits a mother receives if she works after child birth is lower than the amount she receives if she does not work. Based on back-of-the-envelope calculations, we estimate

[45]According to DESTATIS (2019), 71'000 and 13'000 mothers and fathers choose PB+ in 2015, respectively.

that mothers in the control group receive $\approx$ 9800€ in benefit payments through PB in total, while eligible mothers who take up the part-time scheme receive a total benefit amount of $\approx$ 9200€.[46] This shows that governments can - if deemed desirable - shorten the time mothers spend away from work after child birth without increasing public expenditure.

1.8. Conclusion

In this paper we analyze how monetary incentives to part-time work within a paid PL scheme affect maternal labor market outcomes up to 4.5 years after child birth. To do so, we study a German paid parental leave reform which allows eligible mothers to choose a new part-time paid PL option that monetarily incentivizes mothers to work at reduced hours during the first 24 months after child birth.

We find that only high income mothers (compliers) choose the new part-time paid PL option, while low-income mothers are unaffected by the new scheme due to financial constraints and the design of the benefit amount calculation. High income mothers return to the labor market earlier after child birth as a result of the policy, which reduces their child penalty during the first year after child birth. We do not find that the policy's pro-part-time incentives affect maternal employment along the intensive margin (part-time or full-time work) or the child penalty in the long run. Thus, our analysis alleviates concerns of a lock-in effect into part-time employment, according to which mothers stick to part-time employment in the long-run if incentivized to work at reduced hours directly after child birth (e.g. Kunze 2022; Joseph et al. 2013).

The policy achieved its primary goal of encouraging mothers to return to work earlier after child birth. It did so without increasing public expenditure on paid PL benefits for mothers. However, the fact that returning to work earlier does not translate into labor market gains for mothers in the long run raises the question of whether this effect is desirable from a policy perspective. While facilitating the return to work for mothers, irrespective of their long run labor market outcomes, can be considered beneficial in itself, it could also lead to detrimental effects on child development as mothers are less involved in child care. Additionally, a complete assessment of the policy would also require an analysis of its effects on paternal involvement in child care and the division of child care duties within couples. We see our analysis as a first step towards understanding these overall effects of pro-part-time options in paid PL schemes.

[46] These numbers are based on our estimated take-up of the policy (proxied by the employment effect in the first year) and the labor income after compared to before child birth. Furthermore, we assume that mothers receive paid PL benefits for the maximum number of available months under each scheme. Our estimates match official statistics provided by the Federal Statistical Office (DESTATIS 2019).

Our study raises additional questions for future research. The paid PL reform we study disproportionately affects high income mothers. Studying paid PL systems in countries that employ different incentive structures to foster part-time employment could create a broader understanding of the effects of pro-part-time policies for mothers with other characteristics. Moreover, incentives for part-time work during *unpaid* PL, for which parents are typically eligible when the child is already older, might have a different effect on (long-run) labor market outcomes.[47] Since, in contrast to our setting, a large fraction of mothers already work when taking up unpaid part-time PL, such policies are arguably more likely to lead to *reductions* in working hours in the short run. Consequently, the long-run effects potentially differ from the ones we find.

[47]For example, Spain currently offers parents to reduce working hours to part-time employment while their child is less than 12 years old.

1.A. Appendix

1.A.1. Background: maternal employment in Germany

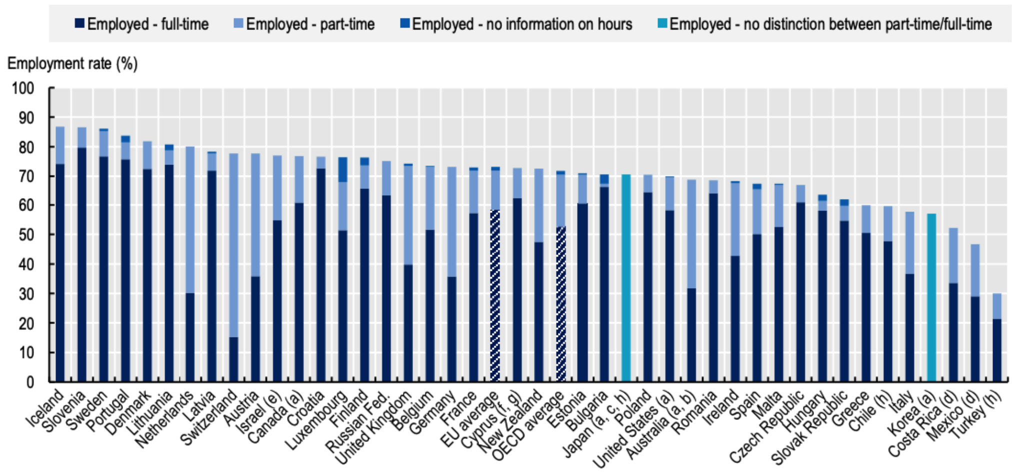
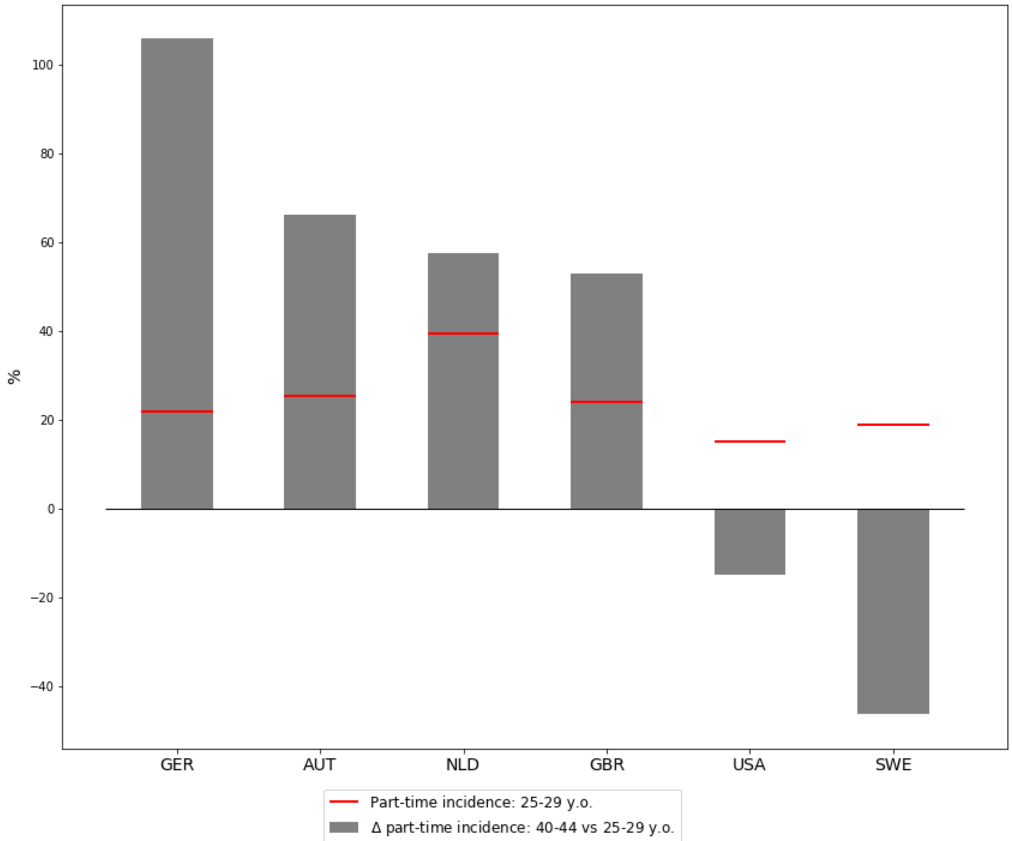


Figure A1: Part-time work across countries (OECD Family Data Base)

Figure A2: Part-time work across countries: before and after birth (OECD Family Data Base)



The figure above shows the incidence of part-time work among women aged 25-29 (proxy for before child birth) in red as well as the change in part-time incidence between women aged 40-44 (proxy for after child birth) and those aged 25-29 in grey. Age groups are used as a proxy for before/after child birth due to the absence of consistent cross-country data on part-time incidence and child birth status. Source: OECD Employment Database

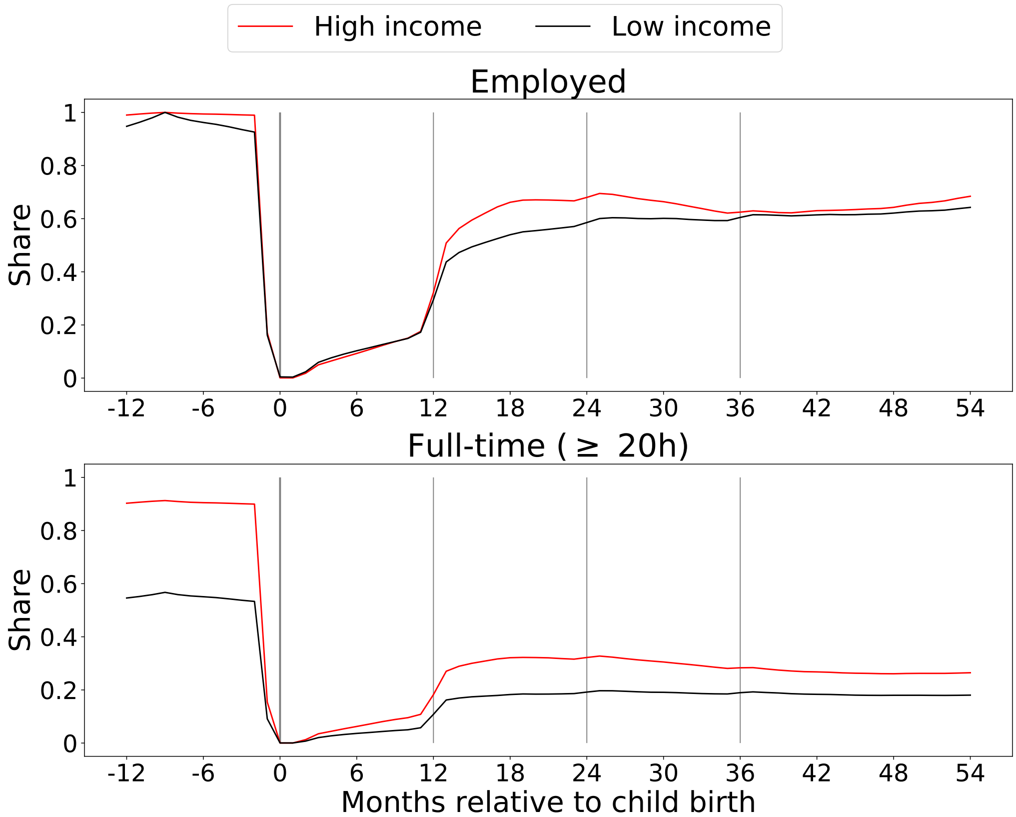
1.A.2. Institutional setting: additional notes

Table A1: Examples of benefit calculations in official documents

	PB			PB+	
Post-birth income	Monthly	Total	Limit	Monthly	Total
0	1300	15,600	650	650	15,600
500	975	11,700	650	650	15,600
1200	520	6,240	650	520	12,480

Note: The examples above are calculated for a parent with a net income prior to child birth of 2000€. It shows the monthly and total benefit amount paid under PB and PB+ in columns 2 - 3 and 5 - 6, respectively. The limit in column 3 refers to the maximum monthly benefit paid under PB+ which equals 50% of the monthly benefit under PB without post-child birth earnings.

Figure A3: Parental labor supply choices before PB+



The figure above shows employment patterns for mothers from one year before child birth to 4.5 years after child birth. The data correspond to full sample of mothers excluding the reform year (2015). The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively.

1.A.3. A 2-period model: solution

This section provides more details on the solution of the 2-period labor-leisure model outlined in section 1.3. In our model a representative agent faces the following optimization problem:

$$\max_{c_1,c_2,l_1,l_2} U(c_1,k_1) + \beta \cdot U(c_2,k_2) \tag{1.4}$$

$$s.t. \quad c_1 + \frac{c_2}{(1+r)} = (1-\tau_1)w \cdot l_1 + b_1 + \frac{(1-\tau_2)w \cdot l_2 + b_2}{1+r}$$

$$k_t = 1 - l_t$$

$$b_t = \tau_t \cdot w \cdot l_0$$

Additionally, PB+ imposes the following conditions on post-child birth labor supply (see below for details):

$$\text{Under } PB+^{high}: \qquad l_1 > 0.5 \cdot l_0 \quad \text{and} \quad l_2 > 0.5 \cdot l_0 \tag{1.5}$$

$$\text{Under } PB+^{low}: \qquad l_1 \leq 0.5 \cdot l_0 \quad \text{and} \quad l_2 \leq 0.5 \cdot l_0 \tag{1.6}$$

The agent chooses consumption c_t and hours worked l_t (and thereby implicitly child care k_t) in periods 1 and 2. She faces a standard inter-temporal budget constraint augmented by several terms that model the paid PL scheme: $b_t - (\tau_t \cdot w \cdot l_t) = \gamma_t \cdot w \cdot l_0 - (\tau_t \cdot w \cdot l_t)$, where b_t is the pre-birth income dependent amount that a parent receives irrespective of post-birth income and the second term is the amount by which the benefit amount is reduced if a parent works post-child birth under PB (and under some conditions also under PB+). In this specification, γ_t governs the fraction of the pre-birth earnings that are replaced by the PL scheme and τ_t defines the fraction of the post-birth earnings that are subtracted from the former amount. The policy parameters thagt correspond to each paid PL option are reported in table A2.

	First year		Second year		
Scheme	γ_1	τ_1	γ_2	τ_2	Condition
PB	0.65	0.65	0	0	
$PB+^{low}$	0.65/2	0	0.65/2	0	$l_1, l_2 \leq \frac{1}{2} \cdot l_0$
$PB+^{high}$	0.65	0.65	0.65	0.65	$l_1, l_2 > \frac{1}{2} \cdot l_0$

Table A2: Policy parameters

The Kuhn Tucker conditions (equations 1.5 and 1.6) are derived from the inequalities imposed by the benefit amount calculation under PB+. The resulting first-order-conditions of the maximization problem are as follows:

$$\frac{U_k(k_1) + \lambda_2}{U_c(c_1)} = w \cdot (1 - \tau_1) \tag{1.7}$$

$$\frac{U_k(k_2) + \lambda_3}{U_c(c_2)} = w \cdot (1 - \tau_2) \tag{1.8}$$

$$\frac{U_c(c_1)}{U_c(c_2)} = \beta \cdot (1 + r) \tag{1.9}$$

$$\frac{U_k(k_1) + \lambda_2}{U_k(k_2) + \lambda_3} = \frac{(1 - \tau_1)}{(1 - \tau_2)}\beta \cdot (1 + r) \tag{1.10}$$

$$-\lambda_1 \cdot (c_1 + \frac{c_2}{(1 + r)} - (1 - \tau_1) \cdot w \cdot l_1 - b_1 - \frac{(1 - \tau_2) \cdot w \cdot l_2 + b_2}{1 + r}) = 0 \tag{1.11}$$

$$-\lambda_2 \cdot (l_1 - 0.5 \cdot l_0) = 0 \tag{1.12}$$

$$-\lambda_3 \cdot (l_2 - 0.5 \cdot l_0) = 0 \tag{1.13}$$

The Lagrange multipliers λ_1, λ_2 and λ_3 in equations 1.7 to 1.13 correspond to the intertemporal budget constraint and the constraints on labor supply in periods one and two, respectively. Assuming logarithmic utility, i.e. $U(c_t, k_t) = ln(c_t) + ln(k_t)$, we can solve for the optimal labor supply choices in periods 1 and 2, i.e. l_1^* and l_2^*, and their partial derivatives with respect to all policy parameters of interest:

$$l_1^*(\tau_t, \gamma_t, l_0, r, \beta) = 1 - \frac{1}{2 + 2\beta} - \frac{1 - \tau_2}{(2 + 2\beta)(1 + r)(1 - \tau_1)} - \frac{\gamma_1 \cdot l_0 + \frac{\gamma_2 \cdot l_0}{1+r}}{(2 + 2 \cdot \beta)(1 - \tau_1)} \tag{1.14}$$

$$l_2^*(\tau_t, \gamma_t, l_0, r, \beta) = 1 - \frac{1 - \tau_1}{1 - \tau_2} \cdot \beta(1 + r) \cdot (1 - l_1^*) \tag{1.15}$$

$PB+^{low}$ and $PB+^{high}$ can only be chosen if $l_1, l_2 \leq \frac{1}{2} \cdot l_0$ and $l_1, l_2 > \frac{1}{2} \cdot l_0$, respectively. This gives rise to the following threshold values $\overline{l_{0,1}}^j$ for $t \in \{1, 2\}$ and $j \in \{$PB+^{low}, $PB+^{high}\}$:

$$l_1^* = 0.5 \cdot l_0 : \qquad \Rightarrow \qquad l_0 = \frac{1 + 2\beta - \frac{1-\tau_2}{(1+r)(1-\tau_1)}}{1 + \beta + \frac{\gamma_1 + \frac{\gamma_2}{1+r}}{1-\tau_1}} = \overline{l_{0,1}}^j \tag{1.16}$$

$$l_2^* = 0.5 \cdot l_0 : \qquad \Rightarrow \qquad l_0 = \frac{2 - \beta(1 + r)(\frac{1-\tau_1}{1-\tau_2} - \frac{1}{1+r})}{1 + \beta(1 + \frac{1+r}{1-\tau_2}(\gamma_1 + \frac{\gamma_2}{1+r}))} = \overline{l_{0,2}}^j \tag{1.17}$$

The optimal labor supply choices l_1^* and l_2^* together with the threshold values $\overline{l_{0,1}^j}$, which define the areas in which optimal labor supply is constrained, are illustrated in figure A4.

To solve for the agent's choice between PB, $PB+^{low}$ and $PB+^{high}$, we compute the utility level resulting from each of these options (see figure A5). The optimal policy choice depends on the agent's pre-child birth labor supply, which equals pre-child birth income since wages are constant in all periods:

$$\underset{PB,\ PB+^{low},\ PB+^{high}}{\text{argmax}} \quad Problem \quad (1.4)$$

$$= \left\{ \begin{array}{cc} PB & \text{if } l_0 <= \overline{l_{0,t}}^{high} \\ PB+^{high} \text{ or } PB+^{low} & \text{if } \overline{l_{0,t}}^{high} < l_0 <= \overline{l_{0,t}}^{low} \\ PB+^{low} & \text{if } \overline{l_{0,t}}^{low} < l_0 \end{array} \right\} \tag{1.18}$$

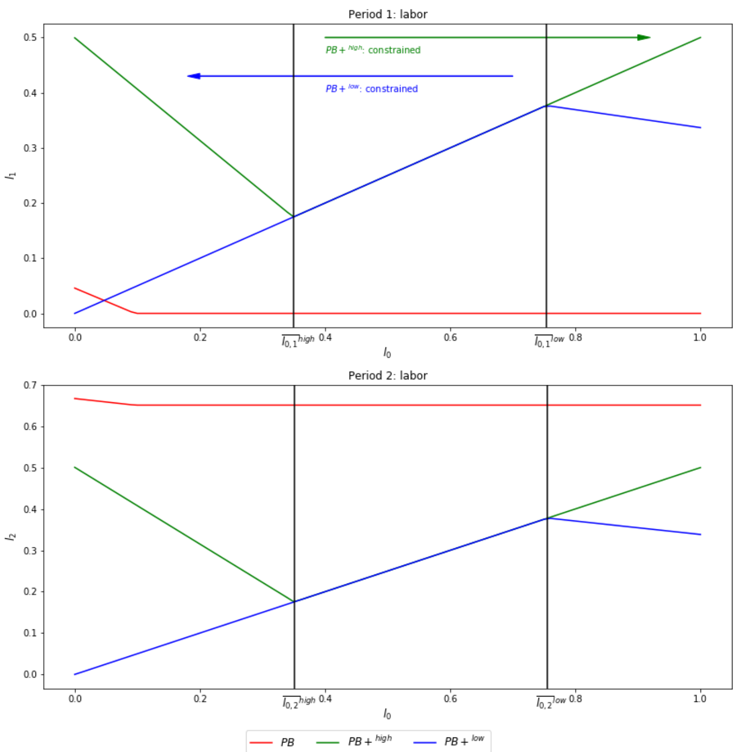


Figure A4: Labor choices by policy options

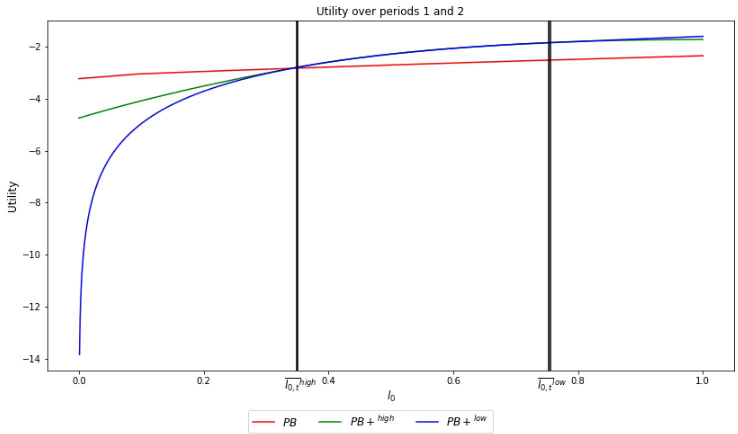


Figure A5: Utiliy by policy options

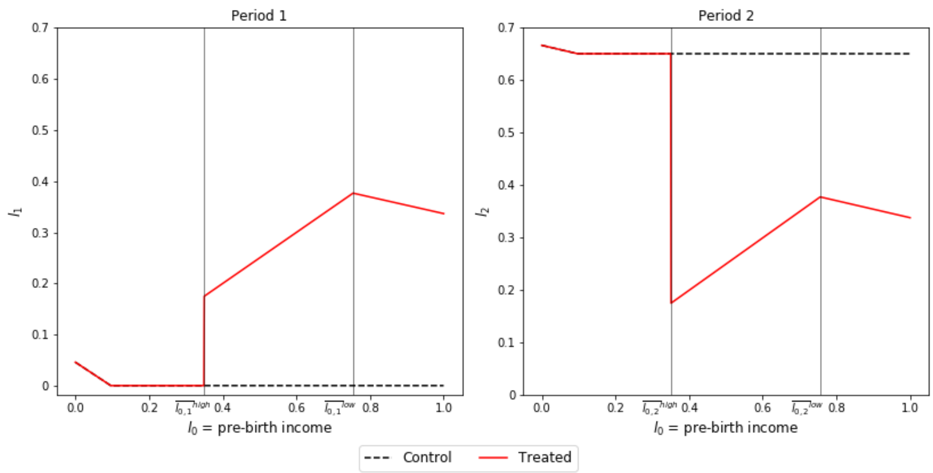
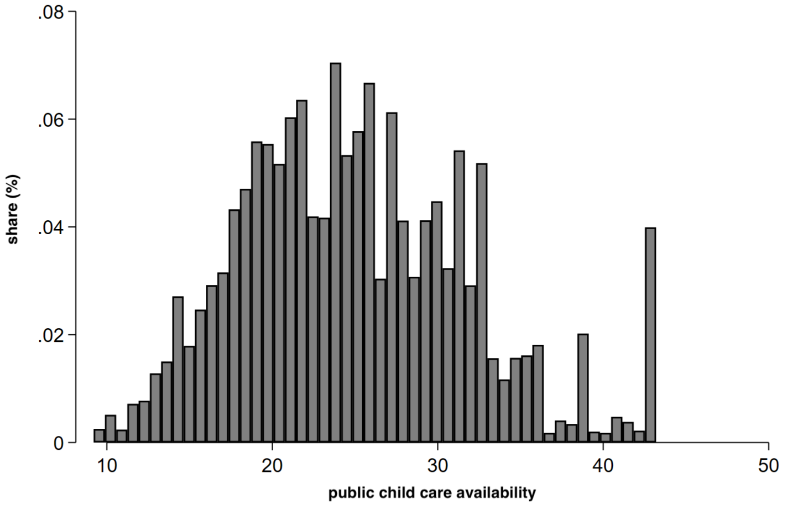


Figure A6: Labor supply choices as predicted by 2-period model

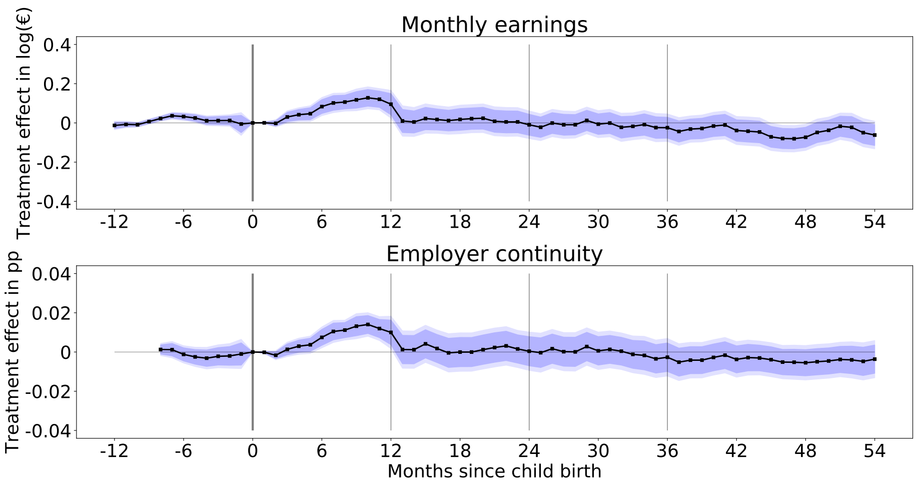
1.A.4. Additional results: figures

Figure A7: Public child care availability by county



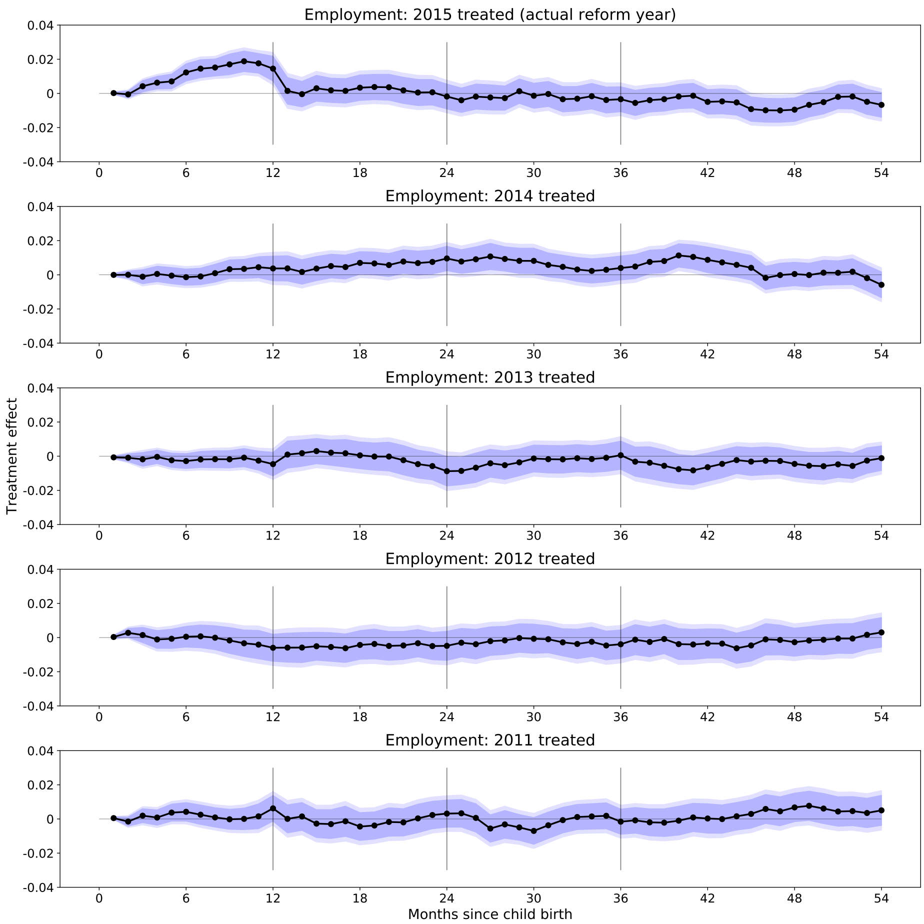
The figure above shows the distribution of public child care availability for 0-2 year old children at the county level (Kreis). The data have been trimmed at 98^{th} percentile.

Figure A8: Additional employment outcomes of all mothers



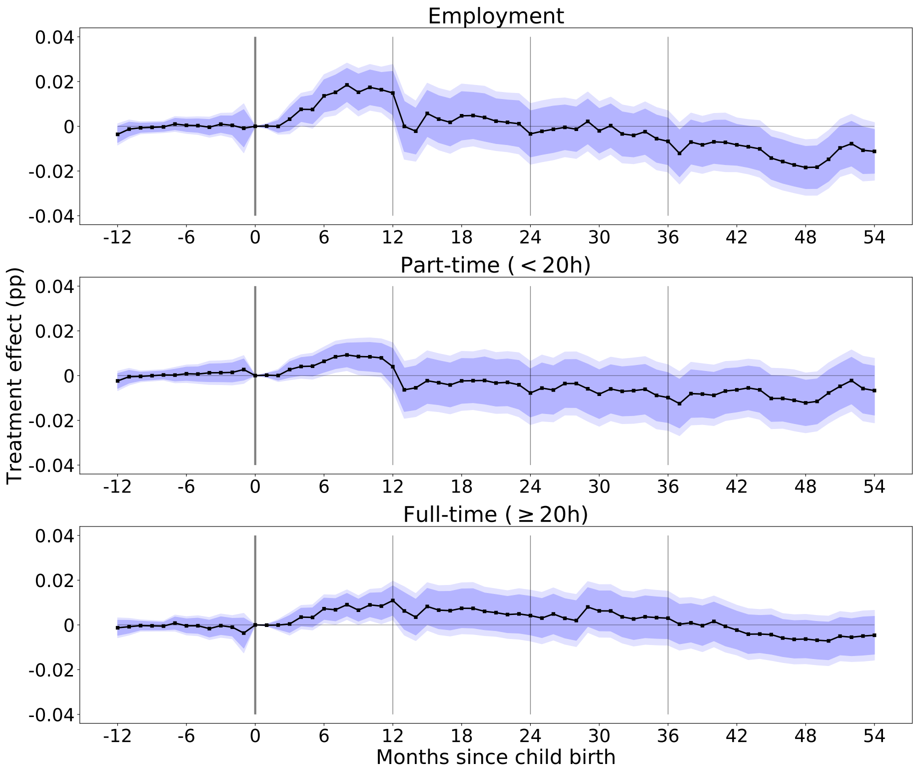
The figures above report the DD coefficient of separate regressions for outcomes p months after child birth as specified in regression 1.3. The outcome in the first panel is the monthly earnings (log). The outcome in panel two is a dummy taking a value of 1 if an individual is employed at the same employer as before child birth and 0 otherwise. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals.

Figure A9: Placebo tests for all mothers



The figures above report the DD coefficient of separate regressions for employment as outcome p month after child birth as specified in regression 1.3. Two birth months on each side of the implementation month are included. In the first panel, the actual reform year (2015) is specified as treatment variable, while the years 2011 to 2014 are used for treatment assignment as a placebo test in panels 2-5. The actual reform year (2015) is excluded from the sample in panels 2-5. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals.

Figure A10: Employment outcomes - one month birth window



Note: the figures above report the DD coefficient of separate regressions for outcomes p month after child birth as specified in regression 1.3. One birth month on each side of the implementation month are included. The outcome variable in the first panel is a dummy for whether an individual is working. The outcome in panel two (three) is a dummy for whether an individual has a part-time (full-time) work contract, i.e. less than (more than) 20 hours of work per week. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals.

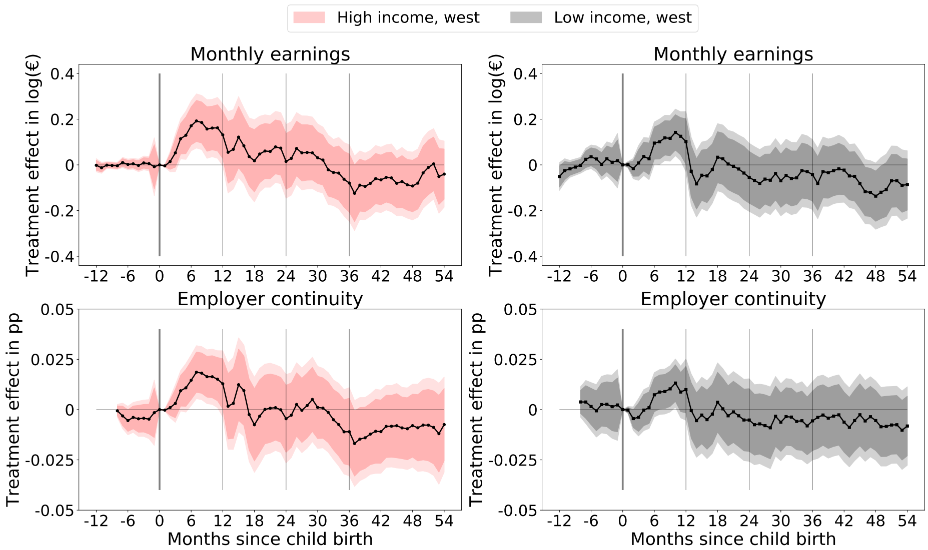
Figure A11: Employment outcomes - one month birth window (cont.)

Note: the figures above report the DD coefficient of separate regressions for outcomes p month after child birth as specified in regression 1.3. One birth month on each side of the implementation month are included. The outcome variables are monthly earnings and employer continuity, respectively. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals.

Figure A12: Employment outcomes by income - one month birth window

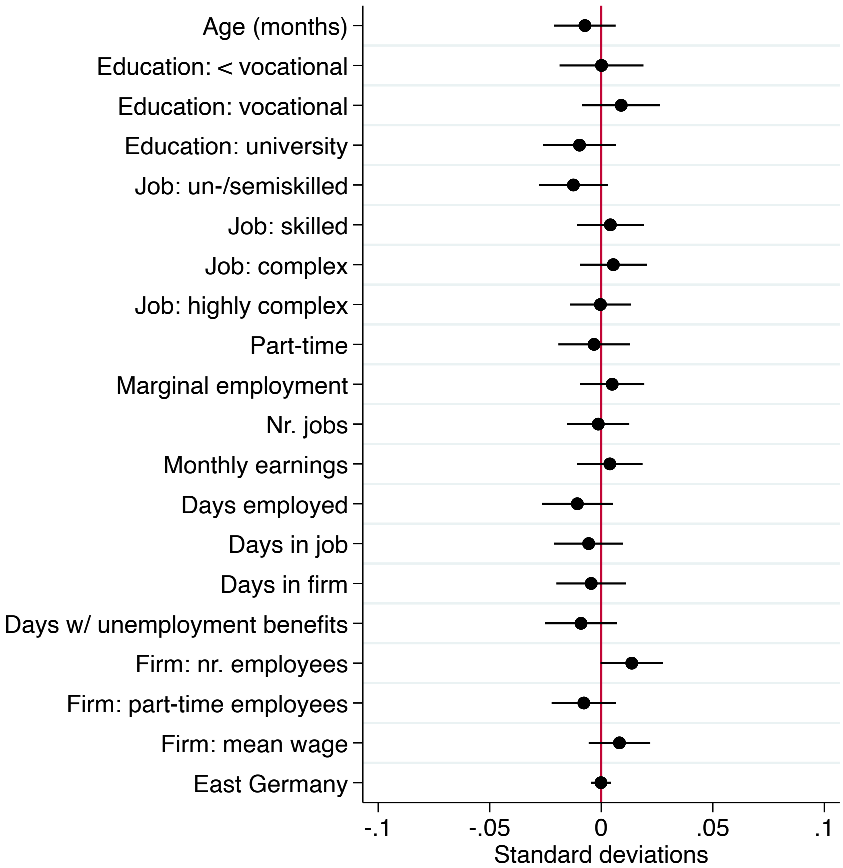
Note: the figures above report the DD coefficient of separate regressions for outcomes p month after child birth as specified in regression 1.3. One birth month on each side of the implementation month are included. The outcome variable in the first panel is a dummy for whether an individual is working. The outcome in panel two (three) is a dummy for whether an individual has a part-time (full-time) work contract, i.e. less than (more than) 20 hours of work per week. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals.

Figure A13: Employment outcomes by income - one month birth window (cont.)



Note: the figures above report the DD coefficient of separate regressions for outcomes p month after child birth as specified in regression 1.3. One birth month on each side of the implementation month are included. The outcome variables are monthly earnings and employer continuity, respectively. The three last vertical lines mark the maximum duration of PB (12-14 months), of PB+ (24-28 months) and the end of the employment protection period (36 months), respectively. Standard errors are clustered at the county level (401). The shaded areas represent the coefficients' 95% and 99% confidence intervals.

Figure A14: Characteristics at conception - all mothers



Note: Each row in the figure above represents a separate estimate of the Difference-in-Differences model's main effect (as specified in equation 1.3, excluding controls) with the variable labeled on the left hand side as dependent variable. The dependent variables are standardized s.t. mean = 0 and sd = 1. Two birth months on each side of the implementation month are included. Standard errors are clustered at the county level (401). The bars represent the coefficients' 95% confidence intervals.

1.A.5. Additional results tables

Table A3: Descriptive statistics of all mothers

Characteristics at conception	mean
age (months)	29.44
part-time	0.26
monthly earnings (€, gross)	2,541.09
Education:	
< vocational	0.06
vocational	0.71
university	0.23
Skill level (job):	
un-/semiskilled	0.09
skilled	0.64
complex	0.12
highly complex	0.15
Employment history:	
days employed	2,523.50
days in job	1,266.76
days in firm	1,305.99
days w/ unemployment benefits	109.09
Firm characteristics:	
nr. employees	87.36
part-time employees	0.33
female employees	0.66
mean wage (€, daily)	93.69
Observations	380,717

Note: The table above reports mean values in observable characteristics, measured at conception for the full sample of analysis

Table A4: Employment outcomes for all mothers

Outcome:									
Employment in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.012***	0.014***	0.003	-0.001	-0.001	-0.003	-0.005	-0.010***	-0.007**
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Mean of outcome	0.097	0.308	0.601	0.634	0.633	0.614	0.623	0.632	0.664
SD of outcome	0.30	0.46	0.49	0.48	0.48	0.49	0.48	0.48	0.47
Part-time in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.006***	0.008***	-0.000	0.000	-0.003	-0.003	-0.005	-0.007*	-0.006
	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Mean of outcome	0.048	0.163	0.349	0.377	0.385	0.378	0.397	0.411	0.441
SD of outcome	0.21	0.37	0.48	0.48	0.49	0.48	0.49	0.49	0.50
Full-time in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.006***	0.006**	0.004	-0.001	0.002	0.000	-0.001	-0.004	-0.002
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Mean of outcome	0.049	0.144	0.251	0.257	0.248	0.236	0.225	0.220	0.222
SD of outcome	0.22	0.35	0.43	0.44	0.43	0.42	0.42	0.41	0.42
Earnings in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.083***	0.095***	0.018	-0.010	-0.006	-0.024	-0.039	-0.073***	-0.062**
	(0.018)	(0.027)	(0.028)	(0.027)	(0.027)	(0.027)	(0.028)	(0.026)	(0.027)
Mean of outcome	0.691	2.225	4.408	4.680	4.686	4.568	4.627	4.713	4.944
SD of outcome	2.13	3.36	3.62	3.58	3.59	3.63	3.62	3.61	3.54
Employer continuity in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.007***	0.010***	-0.000	0.000	0.001	-0.003	-0.004	-0.005	-0.004
	(0.002)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
Mean of outcome	0.071	0.222	0.409	0.403	0.376	0.342	0.323	0.308	0.309
SD of outcome	0.26	0.42	0.49	0.49	0.48	0.47	0.47	0.46	0.46
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	380717	380717	380717	380717	380717	380717	380717	380717	380717
Number of clusters	401	401	401	401	401	401	401	401	401

Note: the table above reports the DD models' main effect on outcomes p months after child birth (in columns) as specified in regression 1.3. The panels (from top to bottom) show the following outcomes: a dummy for employment, for part-time work (less than 20 hours per week), for full-time work (20 hours per week or more), monthly earnings and employer continuity (a dummy for working for the same employer after child birth as at conception). Standard errors are clustered at the county level (401) and reported in parentheses. Descriptive statistics are reported at the bottom of each panel. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Employment outcomes for high income mothers in West Germany

Outcome:									
Employment in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.021***	0.026***	0.006	0.003	-0.001	-0.004	0.001	-0.005	-0.006
	(0.004)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Mean of outcome	0.096	0.310	0.640	0.662	0.647	0.606	0.616	0.630	0.675
SD of outcome	0.29	0.46	0.48	0.47	0.48	0.49	0.49	0.48	0.47
Part-time in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.009***	0.013***	-0.003	-0.000	-0.007	-0.007	-0.002	-0.005	-0.009
	(0.002)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Mean of outcome	0.032	0.140	0.342	0.360	0.360	0.341	0.365	0.386	0.427
SD of outcome	0.18	0.35	0.47	0.48	0.48	0.47	0.48	0.49	0.49
Full-time in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.011***	0.013***	0.009	0.003	0.007	0.003	0.003	0.000	0.003
	(0.003)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Mean of outcome	0.064	0.170	0.298	0.302	0.286	0.265	0.250	0.244	0.249
SD of outcome	0.24	0.38	0.46	0.46	0.45	0.44	0.43	0.43	0.43
Earnings in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.152***	0.183***	0.030	0.015	-0.013	-0.043	-0.005	-0.044	-0.058
	(0.028)	(0.042)	(0.046)	(0.040)	(0.044)	(0.045)	(0.043)	(0.041)	(0.041)
Mean of outcome	0.733	2.359	4.883	5.081	4.975	4.686	4.746	4.862	5.202
SD of outcome	2.27	3.54	3.70	3.67	3.71	3.80	3.78	3.76	3.64
Employer continuity in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.015***	0.021***	0.003	0.005	0.003	-0.004	-0.000	-0.001	-0.003
	(0.003)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Mean of outcome	0.078	0.245	0.482	0.478	0.446	0.398	0.382	0.373	0.384
SD of outcome	0.27	0.43	0.50	0.50	0.50	0.49	0.49	0.48	0.49
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	152368	152368	152368	152368	152368	152368	152368	152368	152368
Number of clusters	401	401	401	401	401	401	401	401	401

Note: the table above reports the DD models' main effect on outcomes p months after child birth (in columns) as specified in regression 1.3. The panels (from top to bottom) show the following outcomes: a dummy for employment, for part-time work (less than 20 hours per week), for full-time work (20 hours per week or more), monthly earnings and employer continuity (a dummy for working for the same employer after child birth as at conception). Standard errors are clustered at the county level (401) and reported in parentheses. Descriptive statistics are reported at the bottom of each panel. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Employment outcomes for low income mothers in West Germany

Outcome:									
Employment in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.010***	0.006	-0.002	-0.007	-0.006	-0.004	-0.005	-0.013**	-0.009
	(0.004)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Mean of outcome	0.108	0.274	0.493	0.541	0.560	0.569	0.586	0.595	0.618
SD of outcome	0.31	0.45	0.50	0.50	0.50	0.50	0.49	0.49	0.49
Part-time in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.005*	0.004	-0.006	-0.007	-0.007	-0.008	-0.007	-0.007	-0.005
	(0.003)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Mean of outcome	0.070	0.182	0.344	0.381	0.399	0.406	0.427	0.440	0.463
SD of outcome	0.26	0.39	0.48	0.49	0.49	0.49	0.49	0.50	0.50
Full-time in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.005**	0.003	0.004	0.001	0.001	0.003	0.002	-0.005	-0.003
	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Mean of outcome	0.038	0.092	0.147	0.158	0.160	0.162	0.157	0.153	0.153
SD of outcome	0.19	0.29	0.35	0.36	0.37	0.37	0.36	0.36	0.36
Earnings in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.068***	0.037	-0.016	-0.042	-0.035	-0.014	-0.013	-0.069*	-0.064
	(0.025)	(0.040)	(0.044)	(0.046)	(0.043)	(0.042)	(0.042)	(0.040)	(0.044)
Mean of outcome	0.722	1.845	3.351	3.724	3.890	3.995	4.118	4.198	4.354
SD of outcome	2.08	3.01	3.41	3.44	3.45	3.47	3.46	3.46	3.42
Employer continuity in month	6	12	18	24	30	36	42	48	54
treatMonths=1 × treatYear=1	0.006*	0.005	-0.003	-0.005	-0.004	-0.005	-0.008	-0.009*	-0.008
	(0.003)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Mean of outcome	0.070	0.170	0.283	0.287	0.274	0.260	0.244	0.228	0.222
SD of outcome	0.26	0.38	0.45	0.45	0.45	0.44	0.43	0.42	0.42
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	152368	152368	152368	152368	152368	152368	152368	152368	152368
Number of clusters	401	401	401	401	401	401	401	401	401

Note: the table above reports the DD models' main effect on outcomes p months after child birth (in columns) as specified in regression 1.3. The panels (from top to bottom) show the following outcomes: a dummy for employment, for part-time work (less than 20 hours per week), for full-time work (20 hours per week or more), monthly earnings and employer continuity (a dummy for working for the same employer after child birth as at conception). Standard errors are clustered at the county level (401) and reported in parentheses. Descriptive statistics are reported at the bottom of each panel. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 2

CENTRALIZED ADMISSION, ACCESS TO EDUCATION AND ACADEMIC ACHIEVEMENT

EVIDENCE FROM A SCHOOL ADMISSION REFORM IN CHILE

2.1. Introduction

Concerns regarding socio-economic inequalities and equality of opportunity, particularly at a young age, have received increased attention in the past decade, both among policy makers and in the academic debate (Chetty and Hendren 2018a). The role of education has been identified as one of the key determinants for economic social mobility in a wide range of settings (J. J. Heckman and Karapakula 2019, Laliberté 2021a). At the same time the access to schools that provide high quality education is unequally distributed in many countries around the world (OECD 2018). An increasing number of countries have resorted to centralized admission systems to address the unequal access to education, among other issues. In these systems, students submit their preferences over available schools. Thereafter, an algorithm assigns students to schools taking the preferences of all applicants into account. In many of these systems, students and their families can access information about available schools in a centralized and systematic way and have reduced (time-based) application costs. Currently, more than 40 countries around the world use centralized assignment systems, mostly for admission to tertiary education (C. Neilson 2021).

In this paper I analyze whether a centralized admission system for pre-tertiary education in Chile, the *Sistema de Admisión Escolar* (School Assignment System, SAS hereafter), improves socio-economically disadvantaged students' access to secondary schools in terms of educational quality. In particular, I analyze whether the SAS reduces the *school quality gap*, i.e. the average difference in the quality of schools attended by low vs. high socio-economic status (SES) students.[1] I then proceed to study whether these changes in the school quality gap also impact the students' academic achievement in later grades. I focus on first grade secondary school students and their academic results since *i)* an analogous analysis at the primary education level is impossible due to data limitations (see section 2.3.3), and *ii)* almost 85% of applications to secondary schools are submitted for the first grade (see figure 2.3).

Various characteristics of the education system in Chile make it a particularly interesting setting to study these questions. First, returns to education and segregation are high by international standards (OECD 2015, OECD 2019). Second, the local school markets, here defined at the municipality level, are heterogeneous in characteristics, such as the level of competition between schools and the proportion of different school types (public, voucher, private). Since the SAS is a national reform, its interactions with these characteristics can be analyzed. Lastly, the SAS is implemented in a context in which school choice, under which students choose the school they wish to apply to instead of being assigned to their neighborhood school, has existed for many decades. This allows me to isolate its effects from the impact of school choice itself (Campos and Kearns 2021).

The SAS changes the admission system in various ways. Prior to the SAS students needed to apply to each school *individually* and schools decided themselves about admission decisions. In the SAS, students apply to schools in a *centralized* way, i.e. by submitting a preference list over available schools on an online platform, and admission decisions are *coordinated* through the use of an algorithm (see section 2.A.1). Importantly, only public and voucher (i.e. semi-private) schools are part of the SAS, while admissions to the private sector are unaffected by the SAS. Since an additional school can be added to the preference list with a few clicks and additional information about each school is provided on the same platform, this results in a decrease in (time-based) costs and an increase in information for all students during the application process.

There are several reasons to expect that low SES students benefit relatively more from a centralized assignment mechanism than high SES students, although I am not able to disentangle them empirically. First, given that low SES students are less informed about the characteristics of schools that are available to them, an increase in information, even if it is available to all students, can benefit low SES students relatively more (Allende et al. 2019). Second, the previous literature has shown that low SES students underestimate the returns to education leading

[1]Throughout the entire analysis each school's quality is kept fixed at the level prior to the implementation of the SAS (see section 2.3.3). This is due to the lack of data on standardized test scores in the post-implementation period. These tests could not be administered as a result of social unrest and COVID-19 in the years 2019 and 2020, respectively.

to under-investment in education (Jensen 2010). Thus, low SES students can also be expected to be more sensitive to (time-based) cost reductions in the application process. In contrast, there is mixed evidence on whether students value school quality (value added) or the quality of their peers more when choosing schools (Hastings and Weinstein 2008, Abdulkadiroğlu et al. 2020). Whether the SAS affects the school quality gap is, thus, an empirical question.[2]

The empirical strategy exploits the fact that the SAS was implemented in 15 regions over four years. In the main part of the empirical analysis I employ a staggered triple Difference-in-Differences (DDD) model to estimate the effect of the SAS on the gap in outcomes between low and high SES students. Not (yet) treated municipalities serve as a control group. This estimation strategy relies on the assumption that the gap in outcomes between low and high SES students would have been constant in absence of the treatment. To corroborate the validity of this estimation strategy, I show that the trends are indeed parallel before the SAS is implemented.

I find that the SAS enables low SES students to access schools of higher quality (measured in value added) in areas with a high provision of private schools. This is driven by positive effects for low SES students rather than negative effects for high SES students. The explanation behind these results is that, in areas with a high provision of private schools, high SES students are more likely to transition to the private sector, thereby freeing up space for low SES students at high quality non-private schools. Since private schools do not participate in the SAS, they represent an outside option for students who are not financially constrained (Calsamiglia and Güell 2018, Kutscher et al. 2020). Exploiting the design of the SAS, I show that high SES students who have a guaranteed secondary school seat for the following year - because their current primary school also offers secondary education - are not more likely to move to the private sector. This corroborates the interpretation that high SES students leave the non-private sector either in anticipation of an inflow of low SES students or because they are not satisfied with their assignment result of the SAS.

Once the structure of the local school market is taken into account, the SAS does not impact the school quality gap. Even in areas in which school competition is high, i.e. where many schools are available, the evidence for an effect on the school quality gap is limited. This points to limited roles for increases in information and decreases in application costs in the context of pre-tertiary education in Chile. A potential explanation for this result is that, already before the SAS, students were well-informed about available schools and application costs were low.[3]

[2]Abdulkadiroğlu et al. (2017) find that coordinating admission decisions increases allocative efficiency, in particular for students who are most likely to remain unassigned in the main round of the uncoordinated admissions process. These students tend to be from areas with higher income.

[3]Indeed, information about schools, such as different measures of school quality, was already available before

To analyze whether low SES students benefit academically from attending higher quality schools as a result of the SAS, I analyze its impact on the students' GPA two years after the implementation and on their grade progression from second to third grade of secondary school. On average low-SES students have a 0.6 points lower GPA (on a scale of 1-7) and are 12.5pp less likely to proceed to third grade of secondary school (84.9% and 72.4% for high and low SES students, respectively). I do not find that the SAS affects these gaps in academic outcomes.

There are various potential explanations for why improved access to higher quality schools does not impact the gap in academic achievement between low and high SES students in the context of this study. First, school value added is a composite measure of quality that, among other factors, depends on peer quality. Since high SES students leave for the private sector, peer quality in the (ex-ante) higher quality schools, to which low SES students have gained access via the SAS, might have decreased. However, I do not find that the policy impacted the students' peer quality as measured by the peers' prior performance on standardized tests at the end of primary school.[4] Second, if low SES students have different schooling needs than high SES students, the standard value-added measure used in this analysis might not accurately reflect school quality for these two groups of students (Loviglio 2020). Third, it is possible that the period of analysis is too short and effects might show up for academic outcomes in the long-run, e.g. in graduation or university entrance rates.

I perform a series of robustness checks to corroborate these findings. First, I use an alternative treatment definition that exploits detailed school-level data on vacancies in the SAS. Based on these data I compute the share of vacant seats in each municipality, which I take as a continuous measure of treatment intensity at the municipality level. I argue that this measure is exogenous because whether a given secondary school has a high share of vacancies or not depends, by and large, on whether this school also imparts primary education (see section 2.A.3 for more details). The results with this continuous treatment measure are qualitatively and quantitatively similar to the baseline findings. Second, to assess whether the baseline results depend on the school quality measure employed, I use the official (categorical) quality measure that is shown to applicants at the moment of application on the SAS platform. The results from this specification are less statistically significant but qualitatively the same.[5] Third, I include municipality-

---

the SAS. However, with the SAS the information is provided on the same platform that students use to apply, which arguably makes it more likely that students make use of that information.

[4] Unfortunately, I am not able to examine how school value added is affected by the introduction of the SAS, due to the lack of standardized test score data in the years after the SAS.

[5] This is likely due to the fact that the official measure is a categorical variable with only four levels. Thus, a lot of information about school quality is lost when compared to the continuous measure I use in the baseline specification.

level trends and show that the findings are unaffected.

This paper contributes to various strands of the literature. First, there is a large literature that studies how to improve access to education for low SES students. Allende et al. (2019) study the role of increased information about available primary schools during the application process in Chile and find that treated students attend schools of higher quality according to various measures, such as average test scores or value added. The evidence on how decreases in costs of applications affect application and enrollment decisions is rather limited. One exception is Knight and Schiff (2022) who find that the Common Applications platform, which allows students to apply to multiple participating colleges in the United States at once, leads to a more racially diverse and higher income student body. In contrast, I study these two aspects, i.e. a uniform reduction in time-based costs and an increase in information, jointly. Empirical studies exploiting implementations or changes to centralized assignment systems as natural experiments remain relatively scarce. Exceptions with a focus on effects on educational equity are Terrier et al. (2021) and Mello (forthcoming).[6] The most closely related paper is Mello (forthcoming) which finds that the introduction of a centralized assignment mechanism for colleges in Brazil crowds out low SES students from the least competitive degrees since they are less geographically mobile than high SES students.

I contribute to this literature by studying a context in which geographical mobility can be excluded as relevant margin of adjustment since the overwhelming majority of students attend a secondary school in their municipality of residence. Furthermore, since the SAS is implemented nationally, the structure of the affected schooling markets are very heterogeneous. This allows me to study how a centralized assignment mechanism interacts with the structure of the local schooling market, such as the availability of different school types (public, voucher and private) and the degree of competition between schools.

This paper also relates to the literature that studies the effects of segregation on academic achievement. The relation between the two is theoretically ambiguous. Being surrounded by better peers (e.g. in terms of achievement or higher SES) is often found to be beneficial for students (e.g. Garlick 2018, Booij et al. 2016). On the other hand, Calsamiglia and Loviglio (2019) find that having better peers can harm students because teachers take the class as a reference group in internal evaluations. Similarly, Denning et al. (2021) show that, conditional on ability, a student's *rank* has an impact on academic and later-life outcomes. In contrast, ability-based segregation (or tracking) might be beneficial for all students if teachers can teach a more homogeneous class better (e.g. Duflo et al. 2011). The effect of segregation on academic achievement (e.g. repetition rates, GPA, standardized test scores,..) has been analyzed in various settings. In the case of desegregation plans

[6] Kutscher et al. (2020) analyze how the SAS affects SES-based segregation in schools.

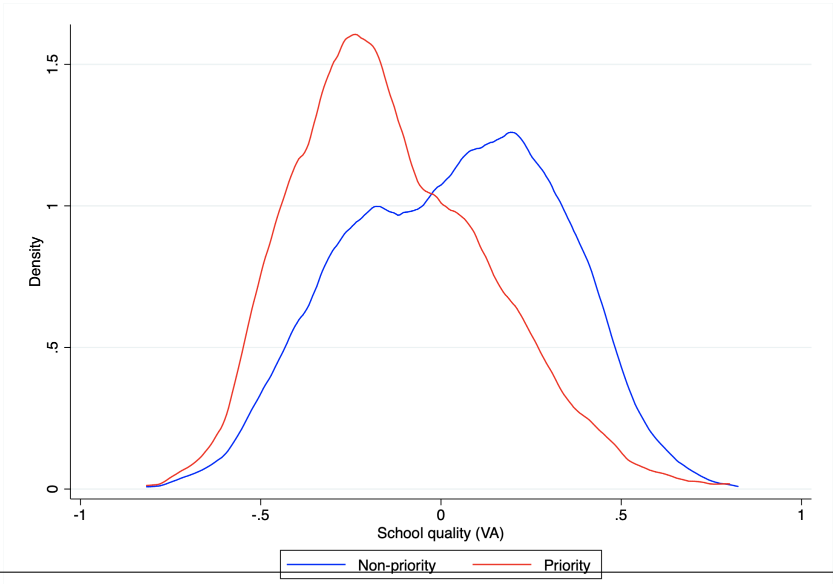
in the United States, studies in general find that decreases in segregation close the achievement gap between black and white students (Lutz 2011, Billings et al. 2013). Arguably the most similar paper in this literature is Hsieh and Urquiola (2006) who study the introduction of vouchers that allow students to attend voucher schools (mostly) free of charge. They find this policy to increase segregation by SES due to high SES students leaving the public sector. Contrary to what I find, this increase in segregation harms public schools, which have a high share of low SES students, in terms of academic achievement. I show that a similar increase in segregation does not necessarily lead to worse academic outcomes in the short run.

The paper is structured in the following way: context on the Chilean education system and details on the SAS are provided in section 2.2. The empirical strategy and the data are explained in section 2.3. In section 2.4 I present and discuss the results. My concluding remarks follow in section 2.6.

2.2. Institutional setting

2.2.1. Pre-tertiary education in Chile

Figure 2.1: School quality by enrolled students' socio-economic status



Note: the figure above shows the distribution of secondary school quality, measured as value added, that children from low (priority) and high income (non-priority) households attend

Pre-tertiary education in Chile is divided into eight years of primary education and additional four years of secondary education.[7] Since the 1980s the pre-tertiary education market is characterized by a high degree of privatization. Three types of schools exist: public, voucher and private schools which represent 45%, 45% and 10% of student enrollment, respectively. Thus, more than half of all students attend a non-public school. Since the 1980s public schools are administered at the local level, i.e. by their respective municipalities. Voucher institutions have to be accredited by the Ministry of Education. Upon approval, voucher institutions receive funding based on the number of enrolled students and their attendance. Apart from this main funding source, a complex system with a large number of subsidies based on school characteristics, such as the student composition (depending on SES) and academic performance, has been put in place over the last decades. [8] Additionally, voucher schools can charge tuition fees. Private schools serve majoritarily high SES students, do not receive any state subsidies and are financed through tuition fees. Private schools offer higher quality instruction, measured in terms of value added, than voucher schools and public schools (see figure 2.6). However, it is important to note that the voucher sector is itself heterogeneous in terms of quality and size.

Free school choice, under which students can apply to any school irrespective of their location of residence, exists in Chile since the 1980s. It was introduced to improve the quality of schools via the demand side, assuming that families choose high quality schools, thereby driving low quality schools out of the market. In practice, various papers have shown negative side effects of this system, leading to more unequal outcomes for low SES compared to high SES students (e.g. Hsieh and Urquiola 2006).

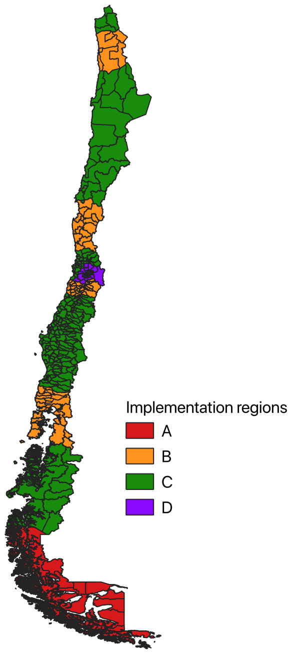
2.2.2. The reform: Sistema de Admisión Escolar

The School assignment system (SAS) was approved in 2015 as part of a larger education reform know as the Ley de Inclusión Escolar (LIE). The reform was designed to "improve the quality of the Chilean education system and to equalize the conditions such that all schools that receive state support [...] can provide high

[7] Secondary education is composed of two tracks: an academic track, which grants access to tertiary education upon graduation, and a technical track, which mainly leads to vocational training. Approximately 65% and 35% of secondary school students are enrolled in the academic and technical tracks respectively. In more recent years an increasing number of students pursues further studies at the tertiary education level upon completion of the technical track.

[8] The most notable modification of the recent years occurred in 2008 when voucher schools started receiving additional funding for each enrolled low SES student. Low SES status is assigned to students by the MINEDUC and revised annually based on various criteria defined in the Law Nr. 20.248. Characteristics that are taken into account are being part of social security programs for vulnerable families (e.g. *Chile Solidario*, *Programa de Ingreso Etico Familiar*), being part of the bottom third in the *Social Registry of Households*, or a combination of household income, parental education, the municipal poverty rate and rurality. The share of priority students is approx. 50% in the sample used in this project, i.e. excluding private school students.

Figure 2.2: SAS implementation: rolled out over four years in 15 regions



Note: The figure above shows the implementation scheme of the SAS in the four regions over the years 2016 - 2019. The implementation started in region A for the admissions 2016 (academic year 2017). Adding one region by year, the implementation was completed by adding region D in 2019 (academic year 2020).

quality education" (MINEDUC 2017). Importantly, while the SAS is implemented in four different regions over four years, the other parts of the LIE took effect in all regions in 2016. This feature of the LIE allows me to disentangle the effects of the SAS from the other aspects of the reform (see below for more details).

The SAS was implemented in order to establish a "transparent and non-discriminatory" admissions procedure that "allows parents to choose the [...] (school) that they like most for their children" (MINEDUC 2017). To apply to public and voucher schools under the SAS students need to use an online platform, to which students log in with their credentials.[9] On this platform students can see available schools either in the form of a list or on a map and access a wide range of information about each school. The information includes an official school quality measure, past test-score results, tuition fee, the educational plan, teaching staff, extracurricular activities as well as photos of the school facilities. While much of this information was already available before the SAS, it had to be accessed on different websites. Thus, it is likely that students are better informed about their choice at the moment of application under the SAS than before its availability. Students then submit a preference list over the available schools. The time-based cost of applying to an additional school is very low since the latter simply needs to be added to the preference list, as compared to physically going to the school to hand in the application documents prior to the SAS. The assignment algorithm is based on the Deferred-Acceptance algorithm and does not take into account the location of residence or the prior academic achievement (see appendix section 2.A.1 for more details).

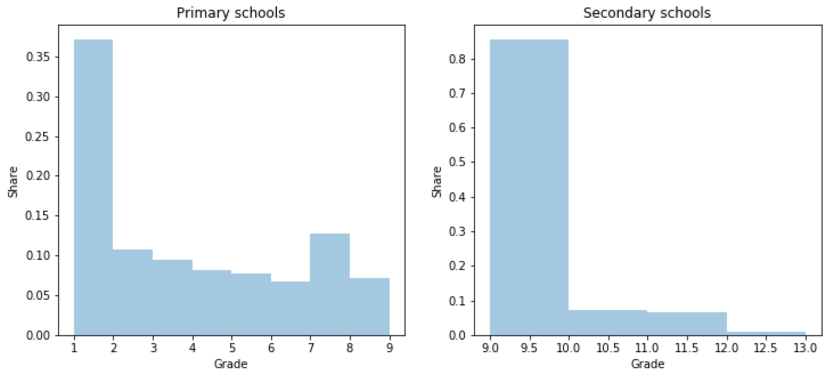
The implementation of the SAS was staggered across 15 regions over four years. Figure 2.2 shows how the implementation was rolled out starting in 2016. The staggered implementation design was likely chosen to gain experience with the administration of the SAS before its implementation in the capital, Santiago de Chile. As table 2.7 shows, the implementation regions are relatively heterogeneous in various characteristics such as, the number of students, the share of low SES students, enrollment patterns, the number of secondary schools and the number of municipalities within each implementation region.[10] Additionally, the implementation was staggered at the grade level. In the first implementation year in each region only the 1st and 7th grade of primary school and the 1st grade of secondary school are integrated into the SAS, while the rest of the grades followed in the second year.[11]

[9]The SAS can be accessed under the following link: https://www.sistemadeadmisionescolar.cl/

[10]At the start of the observation period, Chile was organized in 15 regions, 56 provinces and 346 municipalities. A 16^{th} region (Ñuble) was created in 2018. Throughout this paper, the old territorial organization with 15 regions is used.

[11]I abstract from this additional source of exogenous variation by restricting the empirical analysis to students

Figure 2.3: SAS usage by grade



Note: The figures above show the distribution of student applications made via the SAS by grade for primary schools (left) and secondary schools (right).

The second component of the LIE requires all schools that receive state funding to gradually reduce tuition fees and to finally abolish them within the following ten years. In particular, schools are compensated by a gradual increase in subsidies and a bonus subsidy for schools that declare themselves tuition-free and non-profit. The objective of this component, also referred to as the end of copayment, was to make K12 education financially accessible to everyone. The end of copayment should not affect priority students because they were already exonerated from paying tuition fees previously. Since this reform component was implemented in all regions at once in 2016 its potential effect should not confound the effect of the SAS implementation, which was introduced in a regionally staggered way. As shown in table 2.14, using data at the school level I do not find any impact of the SAS on the probability that schools charge tuition fees.

The third component of the LIE prohibits any school that receives state-funding from making profit. In practice, this change obliges the for-profit schools within the voucher sector to change their statute to that of a non-profit organization within the following ten years and to reinvest any profit resulting from its activity, for example in its infrastructure or personnel. The objective of this policy, referred to as the end of profit, was to eliminate profit maximization from the education market, recognizing that it might not be conducive to the objective of providing high quality education for a large number of students. Previously, voucher schools could declare themselves as for-profit organizations and dispose of their profit as

who are in the first grade of secondary school (and their later academic outcomes), since 85% of all secondary school applications are submitted for the first grade.

they wished. Financial incentives in the form of an additional subsidy are provided to schools that declare themselves non-profit. Identically to the end of copayment, this reform component also took effect in all regions in 2016.

It is important to note that private schools are not affected by these policies since they do not receive any state funding.

2.3. Empirical strategy and data

2.3.1. Empirical strategy

To analyze the effect of the SAS on the outcome variables of interest I exploit its regionally staggered implementation and use a triple Difference-in-Differences (DDD) specification. This amounts to comparing the gap between low SES and high SES students in regions where the SAS is already implemented to regions in which it is not yet in place. The exogenous variation across time and regions allows me to disentangle the causal effect of the SAS from potential confounders, such as other legal changes in the education system or broader trends in the outcome variable. The underlying assumption is that - in the absence of the treatment - the gap in the outcome variables between low SES and high SES students in the treated regions would evolve in the same way as in the control regions (Olden and Møen 2022). I assess the plausibility of the parallel trends assumption by analyzing whether the trends in the outcome variables are parallel across regions before the SAS is implemented (see figure 2.4). Additionally, legal changes introduced at the the start of the SAS implementation in 2016, namely the end of profit and the end of copayment, could bias the estimated effect if schools comply with those legal requirements differentially across implementation regions. For this reason I show that the compliance with these laws does not coincide with the staggered implementation scheme of the SAS. The baseline estimation model is specified as follows:

$$\begin{aligned} y_{ijt} = \alpha + \beta \cdot (SAS_{mt} \times lowSES_i) \\ + (\phi_m \times \phi_t) + (\phi_m \times lowSES_i) + (\phi_t \times lowSES_i) \\ + \delta \cdot lowSES_i + \phi_t + \phi_m + \epsilon_{imt}, \end{aligned} \tag{2.1}$$

where y_{ijt} is the outcome of interest of student i at school j at time t, SAS_{mt} is a dummy for whether municipality m is treated at time t (interaction of treatment status dummy and post-reform dummy), $lowSES_i$ indicates whether student i is a priority student, ϕ_m is a set of roughly 300 municipality fixed effects and ϕ_t is a set of year fixed effects.

Additionally, to examine changes at the municipal and/or school level I estimate a simple Difference-in-Differences (DD) model, comparing treated to untreated municipalities (schools). Analogously to the DDD case, the identifying assumption is that the differences in the outcome variable between treated and untreated municipalities would have remained constant in absence of the treatment.

$$y_{mt} = \alpha + \beta \cdot SAS_{mt} + \phi_t + \phi_m + \epsilon_{mt} \tag{2.2}$$

2.3.2. Data

The data for this study come from administrative records of the Chilean Ministry of Education (MINEDUC), which provides individual and school-level panel data of all students and schools in Chile. In these data I observe each students' enrollment decision, academic performance (e.g. gpa, grade progression,...) and school preference list submitted via the SAS. These data sources also provide rich information at the school-level, such as available seats (vacancies) in the SAS and the level of tuition fees.

The *Quality in Education Agency* provide individual-level test scores for primary and secondary school as well as surveys on household characteristics answered by the parents. These data are used to estimate each school's educational quality (value added) and to identify low-SES students (see section 2.3.3 for details).

The sample is restricted to the years 2014 - 2019, allowing me to assess whether the trends in the outcome variables are parallel in the pre-reform period as well as the effect of the reform thereafter. Furthermore, only students in the first grade of primary school, i.e. 9th grade overall, are included in the final sample for the following reasons: first, information on the students' SES status come from surveys on household characteristics answered by the students' parents before their children enter secondary school, i.e. in the final grades of primary school. Since there is no equivalent information before students enter primary school, access to primary education cannot be analyzed with the same methodology. Second, approximately 85% of all preference submissions at the secondary school level are made for the first grade as shown in figure 2.3. Thus, the impact of the SAS can be expected to be largest in first grade of secondary school. Additionally, the first region in which the SAS is implemented in 2016 is also dropped because the official school quality variable was not yet available for applicants at that time.[12] These restrictions result in a sample of approximately 160'000 students per year.

[12]The official quality measure, published yearly by the *Quality in Education Agency*, is a categorical variable with four levels and part of the information that is provided about all schools in the SAS.

2.3.3. Key variables

SES status - It is crucial to have a reliable indicator of each students' SES status for the purpose of this paper. I follow previous research that takes the mothers' educational attainment as a proxy for a student's SES status (McLanahan 2004, Kutscher et al. 2020). More precisely, I classify students whose mothers have less than a high school degree as low SES. Summary statistics by SES status are reported in table 2.1.

School quality (VA) - The official quality measure, which students see when applying via the SAS, is a *discrete* variable coded in four levels and is, thus, relatively insensitive to quality differences across schools. To obtain a more precise *continuous* measure, I estimate each school's value added using the following standard regression specification:

$$s_{it}^{10th} = \alpha + q_{jt} + s_{it}^{8th} + \gamma \cdot X_{it} + \epsilon_{it}, \tag{2.3}$$

where s_{it}^{10th} and s_{it}^{8th} are student i's test scores in 10_{th} (second grade of secondary school) and s^{8th} grade (last grade of primary school) respectively, q_{jt} is a school fixed effect and X_{it} is a vector of student and household characteristics. X_{it} contains the student's gender, the mother's and father's educational attainment, household income, and a dummy for rural schools. In this regression q_{jt} captures school j's contribution to each student's score in $10th$ grade, conditional on past achievement and individual characteristics. Angrist et al. (2020) show that value added models that control for past achievement deliver reliable estimates of the causal effect of school quality. Figure 2.7 shows that the estimated value added closely matches the official quality measure.

Competition - To measure school competition at the municipality level I rely on the Herfindahl-Hirschmann-Index (HHI), which is widely used to measure market concentration. The HHI at the municipality level is defined as:

$$HHI_m = \sum_{j=1}^{J} \sqrt{\left(\frac{\#students_j}{\#students_m}\right)^2}, \tag{2.4}$$

where $\#students_j$ and $\#students_m$ are the number of students at school j and in municipality m respectively. Thus, $HHI = 1$ corresponds to a market in which a single school has a monopoly and $HHI = 0$ to the case where atomistic schools compete for students.

Segregation - The Duncan Index was first introduced in the sociological literature to study gender-based segregation across occupations and has thereafter been applied in many papers in economics (O. D. Duncan and B. Duncan 1955, Cutler et al. 2008). To measure segregation by SES at the municipality level, the Duncan Index is defined as follows:

$$DI_m = \frac{1}{2} \cdot \sum_{j=1}^{J} \frac{h_{jm}}{H_m} - \frac{l_{jm}}{L_m}, \tag{2.5}$$

where l_{jm} (h_{jm}) is the number of low SES (high SES) students at school j in municipality m and L_m (H_m) is the total number of low (high) SES students in municipality m. Thus, $DI \in [0, 1]$, where $DI = 1$ corresponds to a municipality with a perfectly segregated school market, while $DI = 0$ if the school market is not segregated at all.

2.4. Results

2.4.1. Descriptive statistics

Table 2.1 reports descriptive statistics of the final sample by SES status according to the definition used in this paper. Low SES students attend schools of 0.17 standard deviations lower quality and almost exclusively attend public or voucher schools. They score significantly worse in terms of academic achievement across a variety of measures, such as grade progression, grades, standardized test scores and attendance. In terms of socio-economic background, low SES students come from families with ≈ 1000 USD PPP less household income than high SES students on average. The absolute level of household income roughly corresponds to the level of the Chilean minimum wage, emphasizing the disadvantaged economic situation. While 80% of high SES students' fathers have a high school degree, this is only true for 28% of fathers in the case of low SES students.[13]

2.4.2. Main results

Figure 2.4 shows the effect of the SAS implementation on the difference in quality between schools attended by high and low SES students relative to the region-specific implementation year. Importantly, there does not seem to be any effect in the years prior to the implementation, lending credibility to the identifying assumption of parallel trends between high and low SES students. After its implementation, the SAS leads to a reduction in the school quality gap of $\approx$ 0.01 - 0.02 standard deviations (sd).

[13]By definition of my measure of low SES students, all (none of the) mothers of high (low) SES students have a high school degree.

Table 2.1: Obervables of high vs. low SES students before implementation

Variable	High SES	Low SES	Diff
School quality (VA)	0.03	-0.14	-0.170***
Public	0.27	0.51	0.237***
Voucher	0.62	0.49	-0.125***
Private	0.11	0.00	-0.112***
Academic track	0.70	0.46	-0.248***
Passing 2nd grade (secondary)	0.88	0.78	-0.100***
GPA	5.32	4.84	-0.487***
Attendance	90.09	84.73	-5.357***
Reading 8th grade	0.15	-0.29	-0.434***
Mathematics 8th grade	0.21	-0.41	-0.624***
Household income (monthly, USD PPP)	1687.66	674.33	-1013.329***
Father has HS degree	0.80	0.28	-0.518***
Age	14.38	14.63	0.252***
Female	0.49	0.50	0.005

The effects of the SAS on the school quality gap between low and high SES students are reported in table 2.2. On average low SES students attend a school that has 0.17 sd lower value added during the study period. Column (1) shows that the SAS reduces this gap by 5% on average. Columns (2) to (4) explore heterogeneities of the SAS introduction with the local school market structure, of which I examine two characteristics at the municipal level: the share of students enrolled in each school type (public, voucher and private) and the school market concentration proxied by the Herfindahl-Hirschmann-Index (HHI). Column (2) shows that the effect of the SAS on school quality is significantly lower in areas with higher public and voucher enrolment. The estimates show that a 1 sd increase in private school enrolment (which is equivalent to a 1 sd decrease in public *or* voucher school enrolment *separately* all else equal), increases the effect of the SAS by 0.035 sd. This is equivalent to closing the school quality gap by $\approx 21\%$. As column (3) reports, the effect of the SAS is 0.009 sd higher in municipalities that are 1 sd more competitive ($\approx 5\%$ of school quality gap). The estimates in column (4) show that, once the model is fully interacted, the positive main effect of the SAS is driven by positive effects in municipalities with higher private school enrollment (rows 2 and 3) and higher competition among schools (row 4). Furthermore, the point estimates of the interactions are very similar when interacted one at a time or jointly (columns 2 and 3 vs. 4). Table 2.8 shows that the change in the school quality gap in municipalities with a high private enrollment share is driven by low SES students being able to access higher value added schools rather than high SES

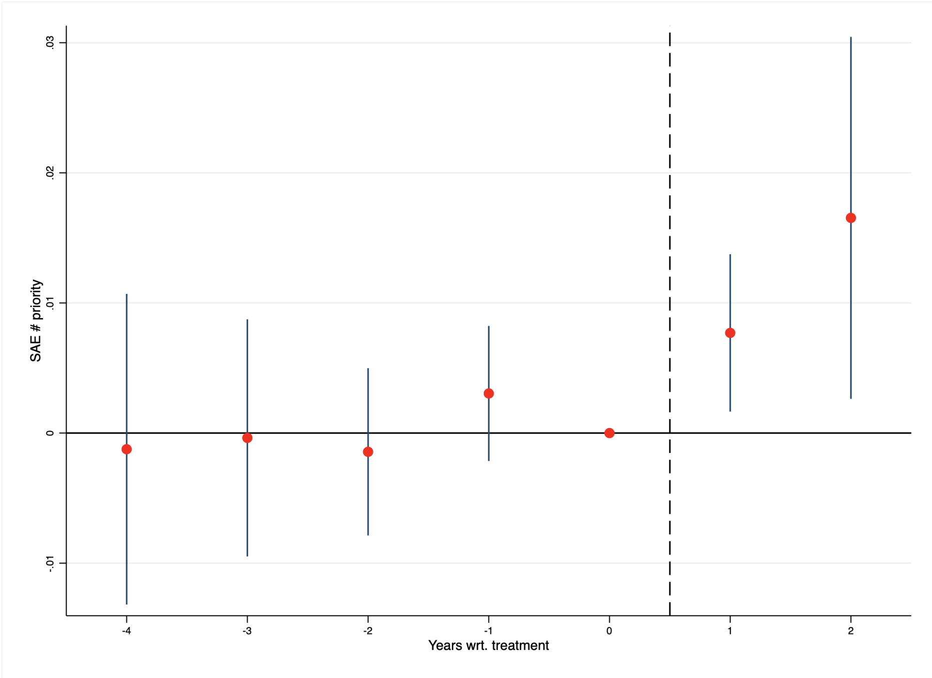


Figure 2.4: Effect of SAS on school quality

The coefficients above show the treatment effect of the SAS, relative to its the implementation year in each regioin, as identified in our baseline specification (see equation 2.1) in the sample of all students. The outcome variable is school quality measured as its value added (see equation 2.3) . The standard errors are clustered at the municipality level. The bars around the point estimates represent the coefficients 95% confidence interval.

downgrading to lower value added schools. In contrast, the reduction in the school quality gap in school districts with high school competition appears to be driven by high-SES students accessing schools of lower quality.

Turning to academic outcomes, the effects on GPA (at the end of second grade of secondary school) and grade progression (enrolling in third grade) are reported in table 2.3. The descriptive statistics show that low-SES students on average have a 0.6 points lower GPA at the end of second grade. To estimate the effect of the SAS on academic outcomes the empirical analysis follows the same strategy as in the previous table, i.e. the main effect of the SAS and its complementarities with the local school market structure are studied. In the case of the GPA, the main effect in column (1) is slightly negative which is driven by the municipalities' school type structure (columns 2 and 4). The negative interaction coefficient of voucher schools suggests that the SAS closes the GPA gap by 0.145 points for each standard deviation increase in a municipality's private school enrollment share. To assess whether this relatively small impact on the GPA translates into changes in the grade progression, I analyze enrollment in third grade as an outcome in columns 5 - 8.

Table 2.2: Effect of SAS on school quality: main effect and market structure heterogeneity.

	Dependent: school quality			
	(1)	(2)	(3)	(4)
SAS=1 × lowSES=1	0.008**	0.009**	0.001	0.004
	(0.004)	(0.004)	(0.003)	(0.003)
SAS=1 × lowSES=1 × public enrol.		-0.035**		-0.027*
		(0.015)		(0.015)
SAS=1 × lowSES=1 × voucher enrol.		-0.034**		-0.031**
		(0.014)		(0.014)
SAS=1 × lowSES=1 × HHI			-0.009**	-0.010*
			(0.003)	(0.005)
Mean SES-gap in outcome	-0.168	-0.168	-0.168	-0.168
SD of SES-gap in outcome	0.39	0.39	0.39	0.39
Municipality FEs	Yes	Yes	Yes	Yes
Sample size	965321	965321	965321	965321
Number of clusters	301	301	301	301

Note: the outcome variable in all columns is the quality (value added) of the school attended by student i in year t. The school quality is fixed at the 2016 academic year and standardized s.t. mean = 0 and standard deviation = 1. The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and its interactions with schooling market characteristics at the municipality level (rows 2-4). These interactions are standardized s.t. mean = 0 and standard deviation = 1. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

On average low-SES students are roughly 13pp less likely to enroll in third grade - conditional on attending the first grade of secondary school. Furthermore, columns 5 - 8 of table 2.3 reveal that the gap in third grade enrollment is not affected by the SAS as all coefficients are insignificant. The yearly effects on academic outcomes relative to the implementation period are shown in figures 2.9 and 2.10.

To corroborate these findings I exploit an institutional feature of the SAS: for each student enrolled in grade g the SAS reserves a seat in grade $g + 1$ at the same establishment for the following year. This ensures that students with a regular grade progression do not need to apply via the SAS each year if they wish to stay at the same establishment. Consequently, students in their last year of primary school do not need to apply via the SAS if they would like to enroll in the first grade of secondary school at the *same* establishment. This feature of the SAS is illustrated in figure 2.5, which shows the fraction of last-grade primary school students who use the SAS to submit preference lists to apply for the first grade of

Table 2.3: Effect of SAS on academic performance

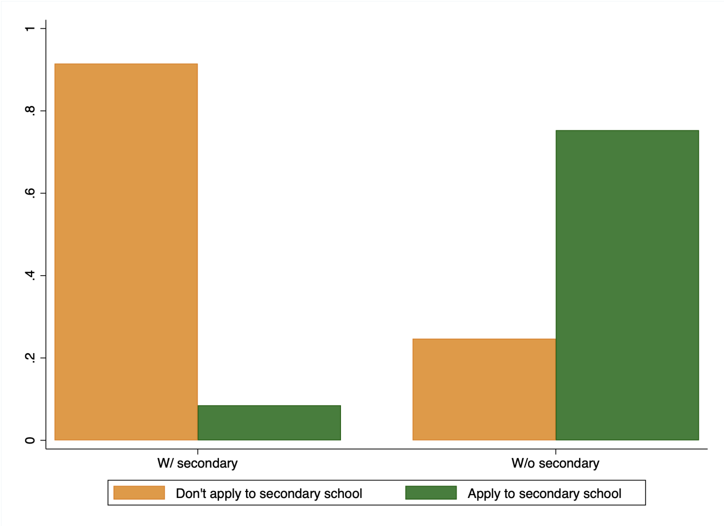
	Dep.: GPA in 2nd grade				Dep.: enrolled in 3rd grade			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SAS=1 × lowSES=1	-0.033*	0.011	0.016	0.031	-0.008	-0.006	-0.002	0.001
	(0.020)	(0.033)	(0.037)	(0.038)	(0.005)	(0.008)	(0.008)	(0.008)
SAS=1 × lowSES=1 × public enrol.		-0.110		-0.130		-0.017		-0.019
		(0.087)		(0.095)		(0.020)		(0.023)
SAS=1 × lowSES=1 × voucher enrol.		-0.145*		-0.145*		-0.017		-0.014
		(0.084)		(0.086)		(0.020)		(0.020)
SAS=1 × lowSES=1 × HHI			0.042	0.038			0.004	0.013
			(0.035)	(0.047)			(0.008)	(0.012)
Mean SES-gap in outcome	-0.644	-0.644	-0.644	-0.644	-0.125	-0.125	-0.125	-0.125
SD of SES-gap in outcome	2.50	2.50	2.50	2.50	0.57	0.57	0.57	0.57
Municipality FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	965321	965321	965321	965321	965321	965321	965321	965321
Number of clusters	301	301	301	301	301	301	301	301

Note: The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and its interactions with schooling market characteristics at the municipality level (rows 2-4). These interactions are standardized s.t. mean = 0 and standard deviation = 1. The outcome variables in columns 1-4 and 5-8 are student i's GPA at the end of second grade and a dummy for enrolling in 3rd grade of secondary school respectively. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

secondary school. The sample is split into establishments that impart both primary and secondary education (left side) and establishments that offer only primary education (right side). The figure shows that while less than 10% of last-grade primary school students at establishments *with* secondary education apply via the SAS, this number is almost 80% for students at establishments *without* secondary education. Thus, I expect the effect of the SAS to be more pronounced among students who, in the year in which they apply to the first grade of secondary school, are enrolled at a primary school without secondary education.

To exploit this feature, I split the sample by whether the students were enrolled at a primary school *with* or *without* secondary school. Table 2.4 reports these results for the school quality attended by the students. The school quality gap between high and low SES is roughly three times as large in the subsample of students who come from schools with both primary and secondary education (-0.191 sd, column (5)) when compared to students from schools that only offer primary education (-0.066 sd, columns (1) - (4)). A comparison of columns (1) - (4) to column (5) also shows that the closing of the school quality gap in municipalities with relatively more private school enrollment is primarily driven by the subset of students who come from primary schools without secondary education at the same establishment. This can be seen in columns (2) and (4) in which the interaction of public and voucher school enrollment with the main effect of the SAS is significantly negative, while the coefficients in the subsample of students from schools with both primary and secondary education (column 5) are not statistically different from 0. These

Figure 2.5: SAS usage by availability of secondary school



Note: this figure shows the share of students in the last grade of primary school that apply to secondary school via the SAS. *W/ secondary* (*W/o secondary*) refers to primary schools that (don't) have a secondary school at the same establishment. Students who are enrolled at a primary school *with* a secondary school do not need to apply to transition from their current primary school to the secondary school at the same establishment.

findings corroborate that the observed effects on the school quality gap are due to the SAS.

Similarly, in 2.5 I analyze the same subsamples of students for the academic outcomes previously analyzed as outcome variables. For both the GPA at the end of second grade of secondary school (columns 1 - 3) and enrollment in third grade (columns 4 - 6) I analyze the full sample (columns (1) and (4)) the sample of students from primary schools *without* and *with* secondary schools (columns (2), (5) and (3), (6) respectively). If the SAS impacts academic outcomes its effect should be particularly notable in the subsample of students from primary schools without secondary schools. However, both for the GPA and grade progression there are no notable differences between the school type interactions of students from primary schools with and without secondary schools.[14]

In sum, the main results show that the SAS enables low SES students to access schools of higher quality relative to high SES students. This effect is concentrated in two subsamples, namely *i)* in municipalities with a relatively high level of private enrollment and *ii)* among students who come from schools at which they cannot proceed to secondary school. These students are more reliant on the SAS to enroll

[14]While the main effect of the SAS in column 2 of table 2.5 is positive, there is no impact on grade progression in the corresponding subsample.

Table 2.4: Effect of SAS on school quality by availability of secondary school

	Primary w/o secondary				Primary w/ secondary
Dependent: school quality	(1)	(2)	(3)	(4)	(5)
SAS=1 × lowSES=1	0.002	0.006	0.001	0.004	0.001
	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)
SAS=1 × lowSES=1 × public enrol.		-0.060**		-0.054*	-0.036
		(0.030)		(0.031)	(0.023)
SAS=1 × lowSES=1 × voucher enrol.		-0.064**		-0.061**	-0.030
		(0.030)		(0.030)	(0.021)
SAS=1 × lowSES=1 × HHI			-0.002	-0.005	-0.005
			(0.005)	(0.008)	(0.008)
Mean SES-gap in outcome	-0.066	-0.066	-0.066	-0.066	-0.191
SD of SES-gap in outcome	0.37	0.37	0.37	0.37	0.39
Municipality FEs	Yes	Yes	Yes	Yes	Yes
Sample size	319867	319867	319867	319867	474232
Number of clusters	299	299	299	299	301

Note: the outcome variable in all columns is the quality (value added) of the school attended by student i in year t. The school quality is fixed at the 2016 academic year. The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and their interactions with schooling market characteristics at the municipality level (rows 2-4). The sample is split into students that, in the year before applying to secondary school, are enrolled at a primary school *without* and *with* a secondary school at the *same* establishment (columns 1-4 and 5 respectively). Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in secondary school, since they do not have a reserved seat at the same school at *secondary* level. However, I do not find a consistent impact of the SAS on academic outcomes.

2.4.3. Mechanism

In this section I assess the mechanisms that could give rise to the findings presented so far. The narrowing in the school quality gap is observed in municipalities with higher levels of private school enrollment. It is important to note that private schools are not part of the SAS, i.e. private schools have their own admission criteria and the SAS does not assign students to them based on submitted preferences. For this reason, they can respresent an outside option for students that are not willing to apply to schools via the SAS or are not satisfied with their assignment result. Due to tuition fees that tend to be substantially higher in private than in voucher schools, this outside option is likely more relevant for high SES students, who are less financially constrained.

To test this mechanism I first analyze the impact of the SAS on segregation proxied by the Duncan Index. Table 2.9 reports the corresponding results. Since the unit of observation are municipalities I use a Difference-in-Differences estimation strategy (see section 2.3.1). Column (1) shows that the main effect of SAS has a

Table 2.5: Effect of SAS on academic outcomes by availability of secondary school

	Dep.: GPA in 2nd grade			enrolment in 3rd grade		
	All	W/o second.	W/ second.	All	W/o second.	W/ second.
SAS=1 × lowSES=1	0.031	0.108**	-0.020	0.001	0.010	0.004
	(0.038)	(0.049)	(0.072)	(0.008)	(0.012)	(0.013)
SAS=1 × lowSES=1 × public enrol.	-0.130	0.020	-0.106	-0.019	-0.017	-0.036
	(0.095)	(0.109)	(0.130)	(0.023)	(0.038)	(0.029)
SAS=1 × lowSES=1 × voucher enrol.	-0.145*	-0.100	-0.049	-0.014	-0.037	-0.019
	(0.086)	(0.099)	(0.116)	(0.020)	(0.034)	(0.026)
SAS=1 × lowSES=1 × HHI	0.038	0.021	0.043	0.013	-0.002	0.018
	(0.047)	(0.056)	(0.073)	(0.012)	(0.019)	(0.015)
Mean SES-gap in outcome	-0.644	-0.345	-0.864	-0.125	-0.073	-0.154
SD of SES-gap in outcome	2.50	2.36	2.60	0.57	0.58	0.55
Municipality FEs	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	965321	319867	474232	965321	319867	474232
Number of clusters	301	299	301	301	299	301

Note: the outcome variables are a dummy for passing 2nd grade (columns 1-3) and the students' GPA in second grade (columns 4-6). The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and their interactions with schooling market characteristics at the municipality level (rows 2-4). Columns 1 and 4 are based on the full sample. In columns 2 and 5 (3 and 6) the sample is split into students that, in the year before applying to secondary school, are enrolled at a primary school *without* (*with*) a secondary school at the *same* establishment. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

significantly negative coefficient i.e. it lowers segregation. In columns (2) and (3) I separately interact municipality-level market structure characteristics. Column (4) reports the estimates of the joint interactions. This last column shows that the effect on segregation is zero on average but that segregation increases/decreases in municipalities with high/low private school enrollment. This means that, in areas with a relatively high provision of private schooling high and low SES students increasingly go to separate schools as a result of the SAS.

To provide further evidence on the relevant mechanism I assess whether the SAS affects the likelihood of switching from a *public* primary school to a *private* secondary school. Due to financial costs that are associated with switching to a private secondary school, I expect high SES students to move to the private sector disproportionately more often than low SES students with the introduction of the SAS. Additionally, switches to private schools can be expected to happen relatively more frequently for students who do not have a reserved seat when transitioning from primary school to secondary school. This is the case for students who attend a primary school that does not offer secondary education at the same establishment.

Table 2.6 reports the results of this analysis. In these regressions the sample is restricted to students who attend either a public or a voucher school in the last grade of primary school. The outcome variable is a binary dummy variable for enrolling in a private secondary school. Columns (1) - (4) show that low SES students who attend a school that only offers primary education (i.e. without reserved seat at secondary level) are 0.6pp less likely to move from a non-private primary to a private secondary school than high SES students. This gap almost doubles to 1pp for students at schools that offer both primary and secondary education (i.e. with reserved seat at secondary school). Furthermore, a comparison of these two subsamples shows that the SAS only affects the differential likelihood of switching to a private secondary school in the case of primary school students who do not have a reserved seat at the secondary level at the same establishment. In this subsample of students, the gap in switching to private schools between low and high SES students further widens. Column (4) shows that this differential likelihood is $i)$ negatively affected at mean levels of enrollment patterns and school competition (row 1) and $ii)$ further widens in the municipalities' private school enrollment (rows 2 and 3). On the contrary, column (5) shows that the differential likelihood of switching to a private option for secondary education is not affected in the subsample of students from schools that impart both primary and secondary education.

I interpret these findings in the following way: primary school students from schools that do not provide secondary education necessarily have to look for secondary education at a different establishment - both before and after the introduction of the SAS. However, after the introduction of the SAS, high SES students are more likely to switch to the private sector - either because they are unsatisfied with their assignment result obtained via the SAS or because they believe to have less chances to be admitted to their desired school under the SAS. Since private schools charge tuition fees this outside option is more relevant for high SES students who are less financially constrained. There is no evidence for this mechanism in the subsample of primary school students at schools that offer both primary and secondary education since they can continue secondary education at the same establishment.

2.5. Robustness checks

In this section I perform a series of robustness checks to corroborate the validity of my findings.

Continuous treatment Within each region the SAS is implemented for all public and voucher schools at once. However, students can only be assigned to a given school if the latter has available seats (see appendix section 2.A.3 for details on how vacancies are created in the SAS). As figure 2.8 demonstrates, there is large variation in the share of available seats at the school level.[15] It can be expected that

[15]The ratio of vacancies at secondary schools is largely determined by whether the same school also offers primary education, since in that case, seats are automatically reserved for graduating primary school students. School quality plays a minor role as figure 2.8 shows. Due to this institutional feature, the extent of vacancies

Table 2.6: Effect of SAS on transitions from *public* primary to *private* secondary school

Dependent: private secondary school	Primary w/o secondary				Primary w/ secondary
	(1)	(2)	(3)	(4)	(5)
SAS=1 × lowSES=1	-0.005*	-0.004***	-0.002	-0.004**	-0.001
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
SAS=1 × lowSES=1 × public enrol.		0.044**		0.041*	0.021
		(0.020)		(0.023)	(0.028)
SAS=1 × lowSES=1 × voucher enrol.		0.041**		0.040*	0.021
		(0.019)		(0.021)	(0.024)
SAS=1 × lowSES=1 × HHI			0.004*	0.002	0.009
			(0.002)	(0.004)	(0.008)
Mean SES-gap in outcome	-0.006	-0.006	-0.006	-0.006	-0.010
SD of SES-gap in outcome	0.09	0.09	0.09	0.09	0.12
Municipality FEs	Yes	Yes	Yes	Yes	Yes
Sample size	319440	319440	319440	319440	344769
Number of clusters	299	299	299	299	294

Note: the outcome variable in all columns is a dummy for switching from *public* primary school to *private* secondary school. The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and their interactions with schooling market characteristics at the municipality level (rows 2-4). In columns 1-4 (column 5) the sample is split into students that, in the year before applying to secondary school, are enrolled at a primary school *without* (*with*) a secondary school at the *same* establishment. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the impact of the SAS is larger in municipalities with a larger fraction of vacancies. To examine to what extent the main results depend on the constant treatment definition used in the baseline specification, I define an alternative treatment as $SAS^{alt}_{t,j} = \frac{\#vacancies_j}{\#students_j}$ if the SAS is implemented in municipality m at time t and 0 otherwise. The variables $\#vacancies_m$ and $\#students_m$ are kept constant at the level prior to the implementation of the SAS. The corresponding results are similar to the ones obtained in the baseline estimation and are reported in table 2.11.

Municipality-level trends To ensure that the results are not driven by trends at the municipality level instead of by the implementation of the SAS, I include municipality-level trends in the main specification. Specifically, I add the following terms to equation 2.1: $trend_t$, $trend_t \times \phi_m$ and $trend_t \times \phi_m \times lowSES_i$, where $trend_t$ is a linear trend and the other terms are defined as before. The results are reported in table 2.12. The coefficients are qualitatively and qualitatively very similar with one exception: the point estimate of the interaction of the HHI with the main effect is slightly lower and no longer signficant. However, the role of the municipality-level enrollment patterns by school type is virtually unchanged.

can be regarded as pre-determined to the policy.

Official school quality data The baseline estimates for the effects of the SAS on access to school quality (e.g. in table 2.2) rely on my own estimations of school value added. To ensure that the baseline results are not driven by differences in my estimation of value added and the official quality information provided to students on the SAS platform I repeat the main analysis with the official quality measure as dependent variable. Note that the latter is only available as categorical variable with four levels. The results reported in table 2.13 are relatively similar to the baseline results. In particular, the school quality gap shrinks with the introduction of the SAS (column 1) and this effect is driven by municipalities with a relatively higher private school enrollment (column 2) and higher competition between schools (column 3). However, once these variables are interacted in column (4) the coefficients are less precisely estimated and not significant. Likely this is due to the fact that the official school quality measure is a categorical variable with only four levels. In comparison with the continuous measure that I use in the baseline analysis, a lot of information about school quality is lost. This is the case because the quality levels in the official measure are created by discretizing the estimated value added at predefined thresholds, namely at the 12^{th}, 35^{th} and the 85^{th}. Thus, in terms of quality there is no distinction between a school at the 40^{th} and the 80^{th} percentile in the quality distribution.

2.6. Conclusion

In this paper I study whether centralizing the applications and admissions process of students to secondary schools impacts the access to educational quality and the students' subsequent academic achievement. Analyzing the staggered introduction of the School Assignment System (SAS), a national centralized assignment mechanism for pre-tertiary education in Chile, I show that low SES students are able to access higher value added schools under the new system. Moreover, I find important interactions with the structure of the local schooling market. In particular, access to higher value added schools is concentrated in municipalities in which the private sector has a high enrollment share. This is explained by high SES students being more likely to enroll in private secondary schools, which have their own admissions policies and are not administered within the SAS. This highlights the importance of outside options in the design of centralized assignment systems (Calsamiglia and Güell 2018, Kutscher et al. 2020). Somewhat surprisingly, the evidence for the impact of SAS on educational quality is very limited in once the size of the private sector is taken into account (e.g. in municipalities where school competition is high). A potential explanation is that (time-based) application costs were low and students were well-informed already before the implementation of the SAS.

In terms of academic outcomes, I do not find consistent evidence that the observed changes in access to educational quality impact the GPA or the grade progression in the first two years of secondary school. Potential explanations for the absence of an effect are: *i)* a too short study period, *ii)* different schooling needs for low and high SES students that are not captured in the standard value added model and *iii)* unobserved changes in school value added after the implementation of the SAS driven by the outflow of high SES peers from the non-private sector.

2.A.1. The assignment algorithm

Below I describe the application and assignment process including the specific tiebrakers, that give priority to students in case of over-demand:

1. Family submits a preference list containing a minimum of one or two schools.[16] There is no maximum amount of schools that can be listed.

2. If demand is higher than supply at the grade-school level, priority is given to students based on the pre-defined characteristics listed in order below:

   a) Siblings of enrolled students.

   b) Priority students until the share of priority students at the establishment level reaches 15%.

   c) Children of permanent staff.

   d) Non-expelled ex-students of the establishment.

3. Family accepts or rejects the assignment.

4. Complementary round: if the family rejects the assignment, it can submit a new preference list over schools with available seats.

5. After the same assignment algorithm is run, families can accept/reject the assignment result of the complementary round.

It is important to note that neither the location of residence nor prior academic achievement are taken into account in the assignment mechanism.

2.A.2. Details on the official school quality measure

The official quality measure, which is computed by the *MINEDUC* and publicly available, is a categorical variable with four quality levels: insufficient, medium-low, medium and high. Importantly, the quality measure is publicly available on the SAE platform, i.e. families can check every school's quality category since 2016 for primary schools and since 2017 for secondary schools. MINEDUC constructs and publishes the Performance Category for two reasons: first, it monitors the quality provided by schools and provides technical and financial support to schools

[16]Families are required to list a minimum of one school if the student is currently enrolled at a school at which she/he could continue her/his studies or the family is applying to a rural establishment. The minimum of two schools applies to families that are entering the Chilean K12 education system or the student's school does not offer the appropriate grade, e.g. if a student needs to apply for the 1st grade of secondary school but her/his current school only offers primary education.

that fall in the lowest category (insufficient). The MINEDUC can revoke the accreditation of schools that are repeatedly categorized in the lowest category. The second goal is to inform families about the quality of available schools, such that this information can be taken into account accordingly. The measure is constructed and made available for primary and secondary schools separately (even if an establishment provide both education levels), however, no distinction is made for secondary school tracks. It is constructed by creating an index based on the students' performance on standardized evaluation and - to a lesser extent - on the students' personal and social development at each school, before being adjusted for the SES of the school's enrolled students. It is, thus, a value-added measure. However, to make it impossible to rank schools by quality, the MINEDUC discretizes the continuous measure into four quality categories according to three predefined cutoffs, namely the 12th, 35th and 85th percentile. Below the variables and the respective weights that are used to compute the index are reported:

- Learning standards (67%)
- Average scores on standardized tests
- Evolution in standardized tests
- Personal and social development (academic motivation, environment,...

2.A.3. Details on robustness checks

This section briefly explains how vacancies are generated in the SAS.

1. Before students submit their preferences, schools declare:
   - maximum capacity, which cannot be modified through the following academic year.
   - Estimate of number of students who will repeat each grade (between 0 and median number of last three years)
2. SAS computes vacancies:

$$vacancies = capacity^{max} - enrollment^{grade-1} - retained^{net}.$$

2.A.4. Additional tables

Table 2.7: Characteristics of municipalities by implementation region

Region	Students	Priority (%)	Private (%)	Voucher (%)	Public (%)	Schools	Municip.
A	733	29	9	41	50	9	3
B	583	39	4	53	42	7	72
C	605	36	5	53	42	7	177
D	1829	28	12	64	23	20	52
Total	810	34	8	57	35	9	129

Note: This table shows summary statistics of selected characteristics (columns) by the four implementation regions (A, B, C and D in columns) in the year before the SAS is implemented in the first region.

Table 2.9: Effect of SAS on segregation

	(1)	(2)	(3)	(4)
SAE=1	-0.029***	-0.014***	-0.012**	-0.006
	(0.007)	(0.005)	(0.006)	(0.006)
SAE=1 × public enrol.		-0.113***		-0.058*
		(0.039)		(0.032)
SAE=1 × voucher enrol.		-0.106***		-0.062**
		(0.039)		(0.030)
SAE=1 × HHI			-0.004	-0.011
			(0.005)	(0.008)
Mean of outcome	0.214	0.214	0.214	0.214
Standard deviation of outcome	0.16	0.16	0.16	0.16
Region FEs	Yes	Yes	Yes	Yes
Sample size	1944	1944	1944	1944
Number of clusters	324	324	324	324

Note: the outcome variable is the Duncan Index, a measure of segregation ranging from 0 (no segregation) to 1 (full segregation). The reported coefficients correspond to OLS estimates of the Difference-in-Differences model's main effect (row 1) and interactions with schooling market characteristics at the municipality level (rows 2-5). Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.8: Effect of SAS on quality by SES status

	High SES				Low SES			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent: school quality								
SAS=1	-0.003	-0.007	0.006*	0.005	0.004	0.008**	0.006**	0.009***
	(0.003)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
SAS=1 $\times$ public enrol.		-0.015		-0.019		-0.039**		-0.036**
		(0.011)		(0.012)		(0.017)		(0.018)
SAS=1 $\times$ voucher enrol.		-0.009		-0.007		-0.037**		-0.033*
		(0.011)		(0.011)		(0.017)		(0.017)
SAS=1 $\times$ HHI			0.008**	0.017***			-0.001	0.002
			(0.003)	(0.006)			(0.003)	(0.005)
Mean gap in outcome	0.037	0.037	0.037	0.037	-0.132	-0.132	-0.132	-0.132
SD of gap in outcome	0.29	0.29	0.29	0.29	0.27	0.27	0.27	0.27
Municipality FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	636721	636721	636721	636721	328600	328600	328600	328600
Number of clusters	301	301	301	301	301	301	301	301

Note: the outcome variable in all columns is the quality (value added) of the school attended by student i in year t. The reported coefficients correspond to OLS estimates of the Difference-in-Differences model's main effect (row 1) and its interactions with schooling market characteristics at the municipality level (rows 2-4). These interactions are standardized s.t. mean = 0 and standard deviation = 1. The school quality is fixed at the 2016 academic year and standardized s.t. mean = 0 and standard deviation = 1. The sample is split into high SES (columns 1 - 4) and low SES (columns 5 - 8) students. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.10: Effect of SAS on peer quality

	Dep.: peers' maths score				Dep.: peers' reading score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SAS=1 $\times$ lowSES=1	-0.004	-0.017	0.007	0.016	-0.011	0.005	0.015	0.024
	(0.026)	(0.024)	(0.022)	(0.019)	(0.021)	(0.024)	(0.025)	(0.022)
SAS=1 $\times$ lowSES=1 $\times$ public enrol.		-0.030		-0.046		-0.029		-0.000
		(0.111)		(0.117)		(0.120)		(0.137)
SAS=1 $\times$ lowSES=1 $\times$ voucher enrol.		0.028		0.041		-0.007		0.031
		(0.112)		(0.114)		(0.115)		(0.121)
SAS=1 $\times$ lowSES=1 $\times$ HHI			-0.008	0.061*			0.004	0.030
			(0.029)	(0.036)			(0.024)	(0.044)
Mean SES-gap in outcome	-0.616	-0.616	-0.616	-0.616	-0.496	-0.496	-0.496	-0.496
SD of SES-gap in outcome	1.21	1.21	1.21	1.21	1.24	1.24	1.24	1.24
Municipality FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	917857	917857	917857	917857	917971	917971	917971	917971
Number of clusters	301	301	301	301	301	301	301	301

Note: The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and its interactions with schooling market characteristics at the municipality level (rows 2-4). These interactions are standardized s.t. mean = 0 and standard deviation = 1. The outcome variables in columns 1-4 and 5-8 are the grade-level average test scores of student i's peers in mathematics and reading respectively. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.11: Alternative treatment: effects on school quality gap

	Dependent: school quality			
	(1)	(2)	(3)	(4)
lowSES=1 × SAS (cont.)	0.008*	0.011**	0.002	0.007*
	(0.005)	(0.005)	(0.004)	(0.004)
lowSES=1 × SAS (cont.) × public enrol.		-0.079***		-0.065**
		(0.029)		(0.029)
lowSES=1 × SAS (cont.) × voucher enrol.		-0.073***		-0.066**
		(0.028)		(0.027)
lowSES=1 × SAS (cont.) × HHI			-0.011***	-0.011*
			(0.004)	(0.006)
Mean SES-gap in outcome	-0.168	-0.168	-0.168	-0.168
SD of SES-gap in outcome	0.39	0.39	0.39	0.39
Municipality FEs	Yes	Yes	Yes	Yes
Sample size	965321	965321	965321	965321
Number of clusters	301	301	301	301

Note: The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and its interactions with schooling market characteristics at the municipality level (rows 2-4). These interactions are standardized s.t. mean = 0 and standard deviation = 1. The outcome variable in all columns is the quality (value added) of the school attended by student i in year t. The school quality is fixed at the 2016 academic year and standardized s.t. mean = 0 and standard deviation = 1. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.12: Municiaplity level trends: effects on school quality gap

	School quality (VA)			
	(1)	(2)	(3)	(4)
SAE=1 × priority=1	0.004	0.004	-0.001	0.000
	(0.003)	(0.004)	(0.003)	(0.003)
SAE=1 × priority=1 × public enrol.		-0.039**		-0.031*
		(0.016)		(0.018)
SAE=1 × priority=1 × voucher enrol.		-0.039**		-0.035**
		(0.016)		(0.017)
SAE=1 × priority=1 × HHI			-0.006**	-0.008
			(0.003)	(0.005)
Mean of outcome	-0.168	-0.168	-0.168	-0.168
Standard deviation of outcome	0.39	0.39	0.39	0.39
Municipality FEs	Yes	Yes	Yes	Yes
Sample size	965321	965321	965321	965321
Number of clusters	301	301	301	301

Note: the outcome variable in all columns is the quality (value added) of the school attended by student i in year t. The school quality is fixed at the 2016 academic year and standardized s.t. mean = 0 and standard deviation = 1. The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and its interactions with schooling market characteristics at the municipality level (rows 2-4). These interactions are standardized s.t. mean = 0 and standard deviation = 1. The regression additionally includes linear municipality-level trends. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.13: Official quality measure: effects on school quality gap

	Dependent: school quality (categorical)			
	(1)	(2)	(3)	(4)
SAS=1 × lowSES=1	0.018*	0.007	-0.004	0.001
	(0.010)	(0.011)	(0.011)	(0.010)
SAS=1 × lowSES=1 × public enrol.		-0.074*		-0.058
		(0.043)		(0.046)
SAS=1 × lowSES=1 × voucher enrol.		-0.054		-0.047
		(0.043)		(0.042)
SAS=1 × lowSES=1 × HHI			-0.026**	-0.014
			(0.011)	(0.017)
Mean gap in outcome	-0.408	-0.408	-0.408	-0.408
SD of gap in outcome	1.14	1.14	1.14	1.14
Municipality FEs	Yes	Yes	Yes	Yes
Sample size	965321	965321	965321	965321
Number of clusters	301	301	301	301

Note: the outcome variable in all columns is the quality (official) of the school attended by student i in year t, measured as categorical variable in four levels. The school quality is fixed at the 2016 academic year. The reported coefficients correspond to OLS estimates of the Triple Difference-in-Differences model's main effect (row 1) and its interactions with schooling market characteristics at the municipality level (rows 2-4). These interactions are standardized s.t. mean = 0 and standard deviation = 1. Summary statistics on the gap between low and high SES students in the outcome variables are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

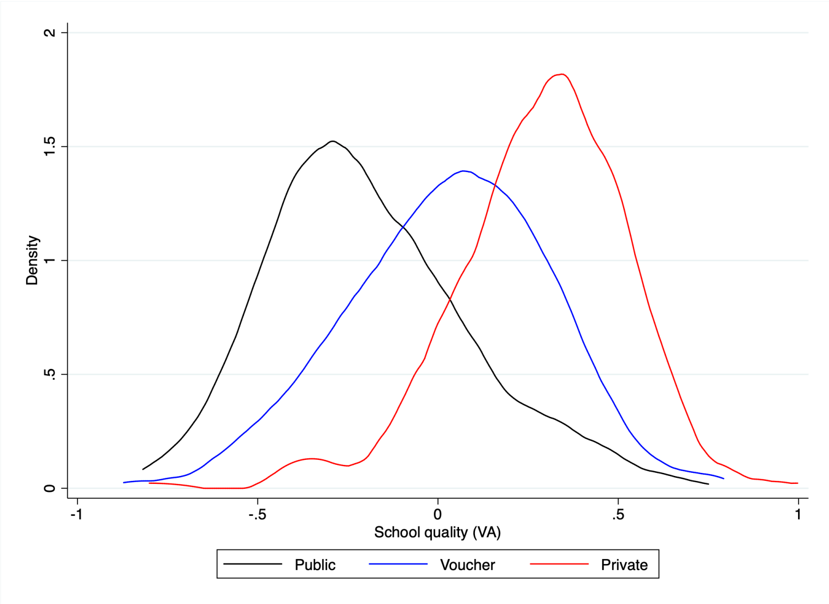
Table 2.14: Official quality measure: effects on school quality gap

	Dependent: tution fee (0/1)			
	(1)	(2)	(3)	(4)
SAS=1	0.030	0.018	0.011	0.020
	(0.020)	(0.017)	(0.017)	(0.016)
SAS=1 × public enrol.		0.015		0.006
		(0.057)		(0.059)
SAS=1 × voucher enrol.		-0.003		-0.007
		(0.061)		(0.060)
SAS=1 × HHI			-0.004	0.006
			(0.022)	(0.028)
Mean outcome	0.379	0.379	0.379	0.379
SD in outcome	0.49	0.49	0.49	0.49
Municipality FEs	Yes	Yes	Yes	Yes
Sample size	11380	11380	11380	11380
Number of clusters	301	301	301	301

Note: the outcome variable in all columns is a dummy indicating whether a school charges tuition fees in a given year. The regressions are estimated at the school-year level. The reported coefficients correspond to OLS estimates of the Difference-in-Differences model's main effect (row 1) and its interactions with schooling market characteristics at the municipality level (rows 2-4). These interactions are standardized s.t. mean = 0 and standard deviation = 1. Summary statistics of outcome variable are reported in the bottom panel. Standard errors are clustered at the municipality level and reported in parentheses. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

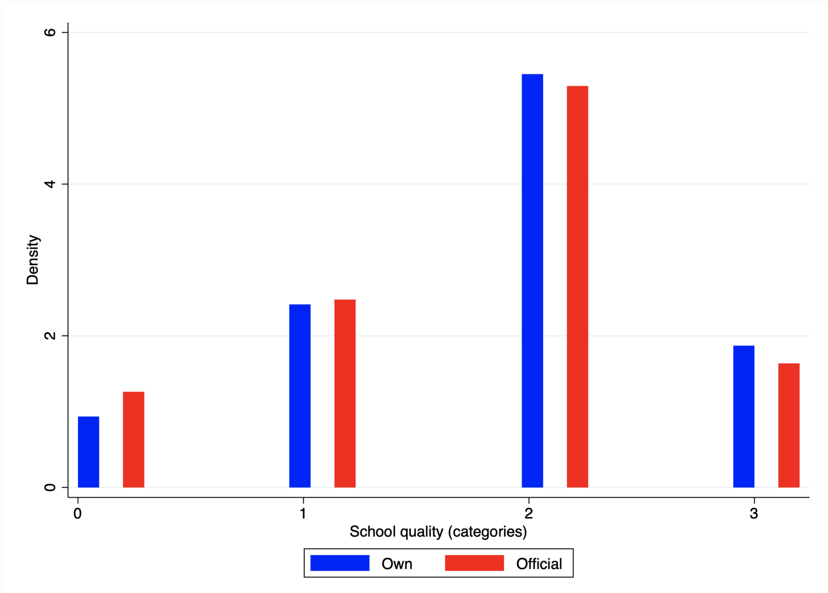
2.A.5. Additional figures

Figure 2.6: School quality (value added) by school type



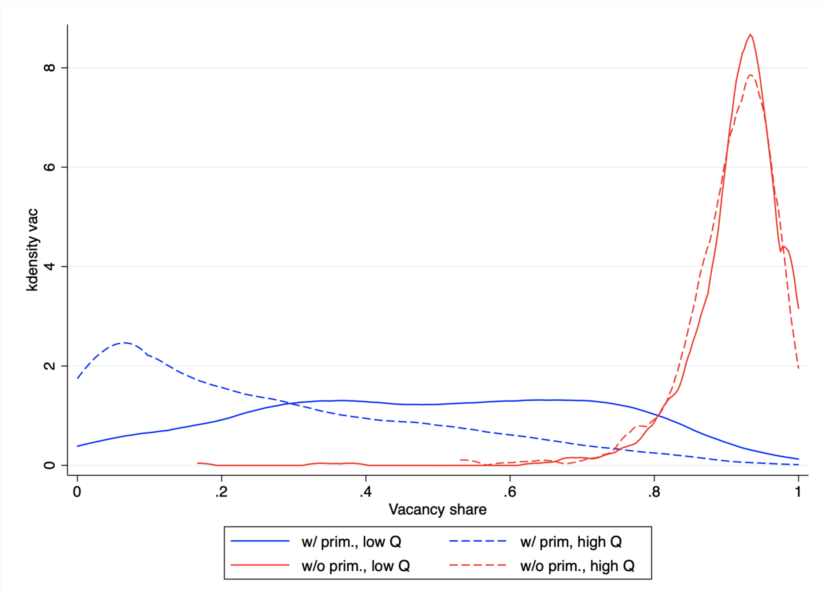
Note: the figure above shows the distribution of secondary school quality, measured as value added, by school type. Voucher schools are semi-private schools.

Figure 2.7: Comparison of quality measures: baseline (own) and official



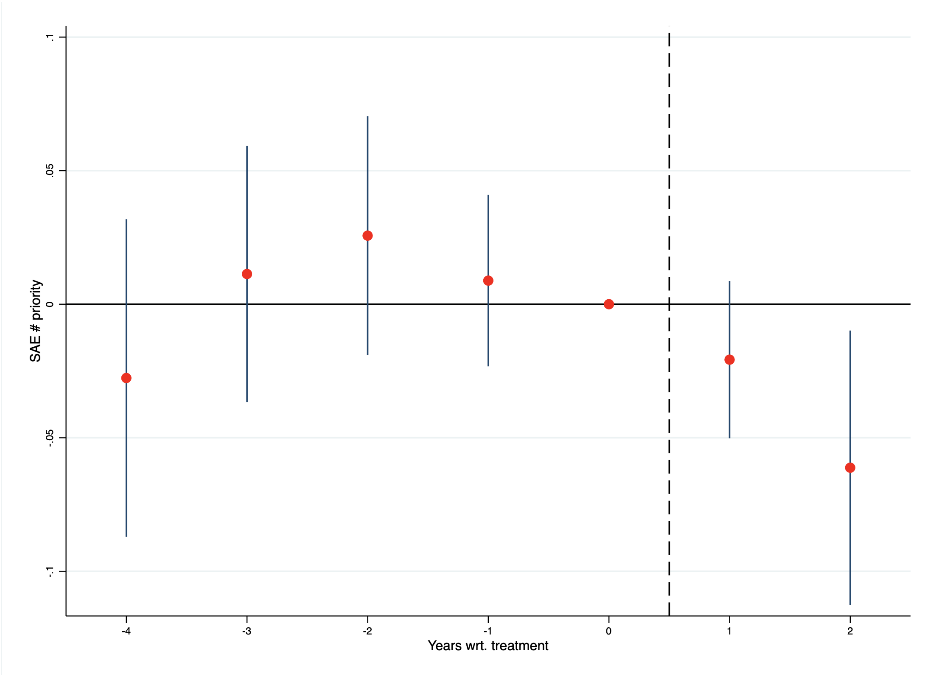
The figure above shows the continuous value added (i.e. school quality) measure used in the baseline estimations in blue and compares it to the official, discrete quality variable. The own, continuous value added measure is estimated as detailed in equation 2.3 and discretized into four categories in the same way as the official quality measure (see section 2.A.2 for more details).

Figure 2.8: Distribution of vacancies by: *i)* secondary schools with/without primary education and *ii)* quality



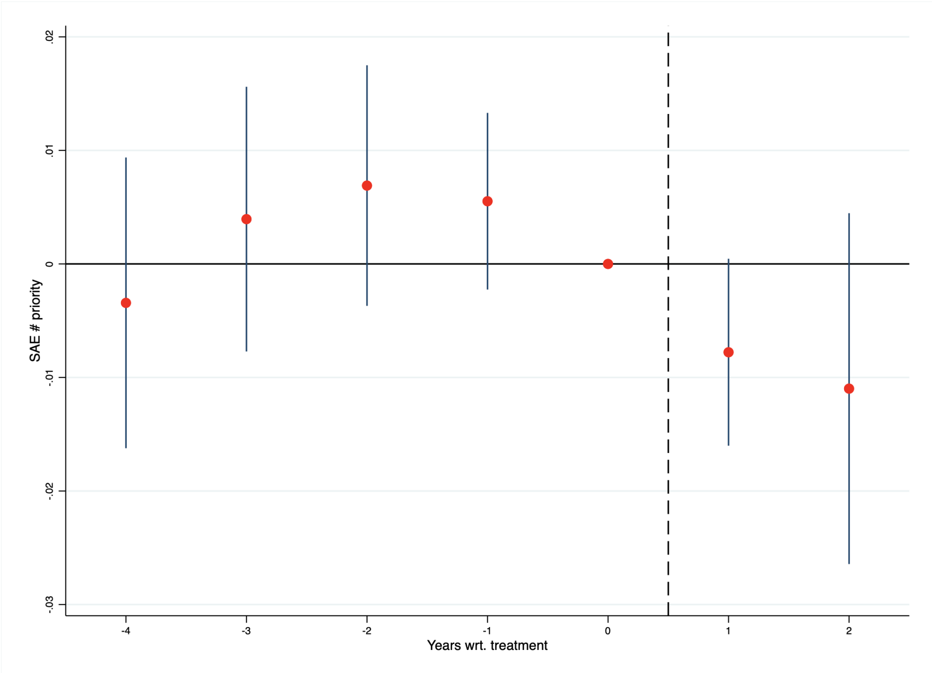
The figure above shows the distribution of vacancies at the school level for secondary schools. Vacancies are largely independent of a secondary school's quality, and determined by whether a secondary school also provides primary education. Abbreviations: *w/ prim* = secondary school that also offers primary education; *w/o prim* = secondary school that does not offer primary education; *high Q* = high quality secondary school; *low Q* = low quality secondary school.

Figure 2.9: Effect of the SAS on the GPA



The coefficients above show the treatment effect of the SAS, relative to its the implementation year in each region, as identified in our baseline specification (see equation 2.1) in the sample of all students. The outcome variable is the student i's GPA. The standard errors are clustered at the municipality level. The bars around the point estimates represent the coefficients 95% confidence interval.

Figure 2.10: Effect of the SAS on grade promotion



The coefficients above show the treatment effect of the SAS, relative to its the implementation year in each region, as identified in our baseline specification (see equation 2.1) in the sample of all students. The outcome variable is a dummy for whether student i passes the second grade of secondary school. The standard errors are clustered at the municipality level. The bars around the point estimates represent the coefficients 95% confidence interval.

Chapter 3

THE MECHANICS OF GOOD FORTUNE

ON INTERGENERATIONAL MOBILITY DURING THE SECOND INDUSTRIAL REVOLUTION

Joint with András Jagadits

3.1. Introduction

Structural transformations have a profound impact on the career and socio-economic status of most people. In particular, recent waves of robotization or trade shocks changed the structure of the labor market and the life of millions of workers depending on their industry or occupation (Acemoglu and Restrepo 2019; 2020; Autor et al. 2016; Dauth et al. 2021a; b; Graetz and Michaels 2018; Humlum 2019; Traiberman 2019). However, since these shocks are very recent and the grandchildren of affected workers have not even been born yet, we can merely speculate how the offspring of demanded tech workers or of displaced manufacturing workers might fare in the very long run. Therefore, in this paper we go back in time to study the effect of an arguably equally disrupting time period on the labor market: the Second Industrial Revolution (ca. 1870-1914). We try to understand *to what extent* and *how* members of a particularly demanded occupation - machinists[1] - could pass on their gains in socio-economic status to later generations in the United States.

To the best of our knowledge, this is the first work which documents the persistence in income gains caused by a labor market shock on the grandchildren of affected individuals, i.e. over three generations. Moreover, we also shed light on the mechanisms which underlie the documented intergenerational persistence: internal migration and increased (secondary) education. We show that these two channels

[1]Workers in charge of installing and maintaining machinery.

may account for the entire positive effect of machinists on their offspring's earnings, though their relative importance depends on initial urban status: the offspring of initially rural machinists gained from more schooling as well as from internal migration to more urban areas, whereas the channel of internal migration does not play a significant role for the sons of urban machinists.

Using the US full count census, we can overcome the main hurdle to intergenerational studies: the scarcity of available data connecting generations. We exploit this data set leveraging the strengths of two, complementary estimation methods: propensity score matching and fixed effects regression. Our empirical strategy amounts to comparing the post-1870 outcomes of machinists to those of non-machinists, who were observationally very similar to machinists before the onset of the Second Industrial Revolution. Next, we identify the offspring of these individuals and investigate their outcomes as well. In our baseline strategy, we use personal and residential characteristics from the census as controls and complement them with occupation-based education (Song et al. 2020) and novel earnings scores. These earnings scores, constructed based on U.S. Department of Labor (1900), are another contribution of this paper as we are the first ones to calculate state-specific earnings scores for a large number of occupations before 1890.

In this paper, we document that machinists could pass on their relative gains in socio-economic status to their (grand)sons. First, we find that machinists, whose occupation experienced a relative labor demand boom starting in the 1870s, enjoyed higher earnings and occupational stability, and were more likely to live in urban places after 1870. As explained in Section 3.2, the surge in demand for machinists resulted from innovations leading to mechanization and the rapid spread of factory production methods in the US. Therefore, much demanded machinists could avoid switching to lower-paying, often agricultural occupations during the volatile business cycle of the Gilded Age. Thus, besides a relative wage improvement, the identified occupational earnings gains are driven by less occupational downgrading rather than occupational upward mobility, which could be suggestive of unobserved ability. Second, the sons of machinists held occupations with 5-12 log-points higher real or nominal earnings scores than the sons of comparable non-machinists in 1900.[2] Finally, a significant positive effect is estimated on the individual- or occupation-level income of grandsons in 1940, seventy years after 1870.

Next, we shed light on the mechanisms behind the documented intergenerational transmission. For the sons of initially rural machinists, the positive earnings effect partly stems from a higher probability of living in an urban area as an adult (urban wage premium). To quantify the approximate size of earnings gains

[2]After correcting for the bias which stems for mismeasurement. See Section 3.6.4.

originating from rural-to-urban migration, we multiply the differential likelihood of urban status with an earnings score-based estimate of the rural-to-urban migration premium pertaining to the early-twentieth-century United States (Ward forthcoming).

Additionally, machinists' sons benefited from parental investment in their education irrespective of initial urban status, receiving approx. 0.35 more years of (mainly secondary) schooling. To study the role of education in explaining the earnings effect, we simply combine our years of schooling point estimate with a returns to schooling estimate in Goldin and Katz (2000). Moreover, by exploiting a newly digitized, county-level data set on high school provision, we establish that the positive earnings and especially schooling effects on the sons of machinists increased in county-level private high school provision.[3] This complementarity between a machinist father and local high school supply was especially strong when free-of-charge public high school supply was limited and private schools had a high teacher-student ratio. On the other hand, gains from private high schools decreased if these schools could be attended at a low price and put emphasis on scientific education in their curricula (e.g., mechanical drawing), the type of knowledge which the sons of machinists could more easily acquire at home. This suggests that passing on scientific knowledge in an informal way, within a family also helped machinist's sons succeed. Furthermore, the estimated positive effects on machinists' sons declined in public high school provision as well. This empirical result is consistent with financially more constrained non-machinist parents (Becker and Tomes 1979; 1986). Last, we estimate a coefficient on education which is not significantly different from zero for sons who were older than ten years in 1870, suggesting that machinists were not differentially more likely to invest in the education of their sons before 1870.[4]

Apart from heterogeneity exercises, we conduct a series of robustness checks to mitigate concerns that the identified positive effects can be explained by (the transmission of) the machinists' unobserved ability. Arguably the most convincing robustness checks are regressions containing family-fixed effects, i.e. comparing machinists to their own brothers.[5] The results from this specification, which controls for the similar environment of upbringing and inherited genes, are qualitatively and quantitatively similar to those obtained from the baseline analysis. In addition, the lack of correlation - both within and across families - between the machinist

[3]We show that this effect is driven by the medium-level tuition fee. At this cost level, education was less affordable for rival boys but not prohibitively costly for machinists.

[4]In accordance with the literature documenting dynamic complementarities in the production of human capital (see J. J. Heckman and Cunha 2007), it was arguably already too late to invest in their education when the relative earnings of machinists started to rise.

[5]This robustness test can only be conducted for the generation of machinists themselves because we run into sample size limitations for later generations.

indicator and standard (historical) proxies of unobserved ability (e.g., number of children or spousal literacy - measured in 1870) suggests that machinist fathers were not more able compared to their brothers or comparable peers.

We demonstrate that the results are not driven by occupation-state or census division-level pre-trends (e.g., changes in the employment share, probability of switching to agriculture, etc. in the 1850s and 1860s), and are insensitive to which specific occupations are the dominant "control occupations". Moreover, our preferred propensity score matching strategy eliminates initially large differences in the overwhelming majority of characteristics of wives, fathers and next-door-neighbors between machinists and non-machinists - even *without* matching on these characteristics. We also establish that similarly aged sons of younger and older machinists experienced similar positive effects, indicating the absence of early sorting into the machinist occupation by more talented individuals. Additionally, the inclusion of birth state-destination county (1870)-fixed effects makes it very unlikely that the results reflect spatial sorting prior to 1870.

Related literature This work is closely connected to the literature which examines the effect of parental labor market shocks on affected children. Exploiting layoffs, Hilger (2016) and Mörk et al. (2020) find at most very small negative effects on the education and adult earnings of affected children.[6] As both papers point out, these might be the consequence of a generous welfare state offsetting otherwise reduced parental spending on education. A more accurate comparison to our setting might come from papers that focus on less developed countries with a rather weak welfare state or low-income (financially constrained) families. These papers tend to find that changes in parental income - not necessarily induced by job loss - do matter for the offspring (see, e.g., Aizer et al. 2016; Akee et al. 2010; Dahl and Lochner 2012; Di Maio and Nisticò 2019; Løken et al. 2012; Manoli and Turner 2018). Surveying the literature, Cooper and Stewart (2017) conclude that there is «*strong evidence that income has causal effects on a wide range of children's outcomes, especially in households on low incomes*», whereas wealth shocks do not seem to have substantial effects on children either in a historical (Bleakley and Ferrie 2016) or in a modern context (Cesarini et al. 2017). Additionally, there is a large literature documenting the role of credit constraints and grants, mostly for college education.[7] Our contribution is to show that the effect of

[6]Early papers tend to exploit mass layoffs or factory closures, and find mixed effects on schooling and future earnings of children affected by parental job loss (Bratberg et al. 2008; Coelli 2011; Oreopoulos et al. 2008; Rege et al. 2011). However, Hilger (2016) argues that many early findings on large, negative effects might be driven by the assortative matching of low-quality workers and low-quality firms leading to selection into layoffs or closure. Løken (2010), exploiting the oil boom in Norway as a permanent income shock, finds no effect on children either.

[7]There is ample evidence that credit constraints and grants for schooling matter even in modern contexts and in many developed countries. The early literature is summarized in Lochner and Monge-Naranjo (2012), see

labor market shocks may persist even for the offspring in the second generation. In addition, we pin down mechanisms which lead to the documented intergenerational persistence. These are not well-understood even in the modern context and, to the best of our knowledge, have not been studied in a historical context yet.

This paper also speaks to a literature which seeks to identify the determinants of intergenerational mobility in the 19$^{\text{th}}$-20$^{\text{th}}$ century United States. Parman (2011) demonstrates that children from high-income families benefited disproportionally more from improving public high school availability in Iowa at the turn of the 20$^{\text{th}}$ century, resulting in a higher intergenerational income elasticity. However, we find a negative association between public high school supply and the relative gains of machinists' sons, in line with Solon (2004) and Olivetti and Paserman (2015). Since parents of similar socio-economic background tend to have similar preferences over education (Boneva and Rauh 2018), we believe that comparing machinist fathers to fathers in other middle-class occupations might eliminate the effect uncovered by Parman (2011) in our case. In a comparison of migrating to non-migrating brothers, Ward (forthcoming) finds that rural-urban migration was an important contributor to upward mobility in the early-twentieth-century US, particularly so for people from the poorest households. This finding is in line with our results on the importance of urban place of living for initially rural machinists' sons. Furthermore, Olivetti and Paserman (2015) and Song et al. (2020) show that industrialization was a major determinant of a relatively low intergenerational mobility around 1900. Our case study of machinists aligns well with this view and suggests highly persistent positive effects on their offspring.

By analyzing the effect of a change in occupational labor demand on machinists themselves, this work is also connected to a fast growing literature which investigates the effect of technology-induced occupational labor demand changes on affected individuals. Papers studying the impact of automation or robotization typically find that robots decrease the employment share of lower-skilled production workers and benefit workers in occupations with complementary tasks - just as early machines did to machinists (Acemoglu and Restrepo 2020; Dauth et al. 2021b; Graetz and Michaels 2018; Humlum 2019). Focusing on the automation of telephone operation, Feigenbaum and Gross (2020) find that incumbent telephone operators bore most of the losses: they were more likely to be in lower-paying occupations or left the labor force entirely after automation started. However, growth in middle-skill jobs absorbed the labor supply of later generations. Using exceptionally disaggregated Swedish data on occupations, Edin et al. (2019) show that those facing occupational decline lost about 2-5 percent of mean cumulative earnings and were less likely to remain in their starting occupations - the mirror im-

also Bettinger et al. (2019), Castleman and B. T. Long (2016), Denning et al. (2019), Fack and Grenet (2015), Hai and J. J. Heckman (2017), Lee and Seshadri (2019), Molina and Rivadeneyra (2021), Solis (2017), and Wright (2021) .

age of what we estimate in the US for machinists. Additionally, Swedish earnings losses are partly accounted for by reduced employment and increased time spent in unemployment and retraining. Our contribution to this literature lies in analyzing a different time period, mainly the Second Industrial Revolution, in detail.

The paper is structured as follows. First, Section 3.2 discusses the historical background, then, Section 3.3 addresses questions related to data sources and sample construction. Section 4 presents the empirical strategy while Section 3.5 contains the main results. Thereafter, the reader may find a battery of robustness exercises and a discussion of a non-classical measurement error in Section 3.6. Finally, Section 7 concludes.

3.2. Historical background

The machinist occupation was born in the First Industrial Revolution in the United Kingdom, but members of this occupation played an important role in innovative activities in the United States in the early nineteenth century as well (Kelly et al. 2020; Meisenzahl and Mokyr 2011; Sokoloff and Khan 1990). Nevertheless, professional engineers had taken over this inventive role by the mid-nineteenth century, even before the Second Industrial Revolution started (Hanlon 2021; Maloney and Valencia Caicedo 2020). Thus, the assembly and maintenance of industrial machinery was left as the task of most machinists (U.S. Department of Labor 1899). People could enter this occupation through the helper system, a type of informal apprenticeship. This meant initially simple operations followed by a sequence of more demanding tasks as they gained experience next to senior machinists. Additionally, the division of labor among American machinists reached a substantially higher level compared to the UK, resulting in a relatively lower skill requirement and making a cross-country earnings comparison of machinists almost impossible (Rosenbloom 2002). In spite of reduced skill requirements in the US, machinists remained a part of the so-called ”labor aristocracy” alongside other skilled craftsmen, for instance, blacksmiths, carpenters, conductors, masons, painters or plumbers (Dawson 1979; Rosenbloom 2002).

While at-scale factory production was limited to the textile industry until the Civil War, the situation changed rapidly after the onset of the Second Industrial Revolution around 1870. Mechanization and factory production methods spread swiftly across a wide range of industries, led by steel and chemicals production, and was supercharged by the utilization of electricity and novel ways of transportation (e.g., the railway; Mokyr 1999; Rosenbloom 2002). American manufacturing harnessed steam engines with a total capacity of approx. 1.000 thousand HP in 1870. This figure exponentially increased to almost 9.000 thousand by 1900 (Rosenberg and Trajtenberg 2004). The average establishment size stagnated

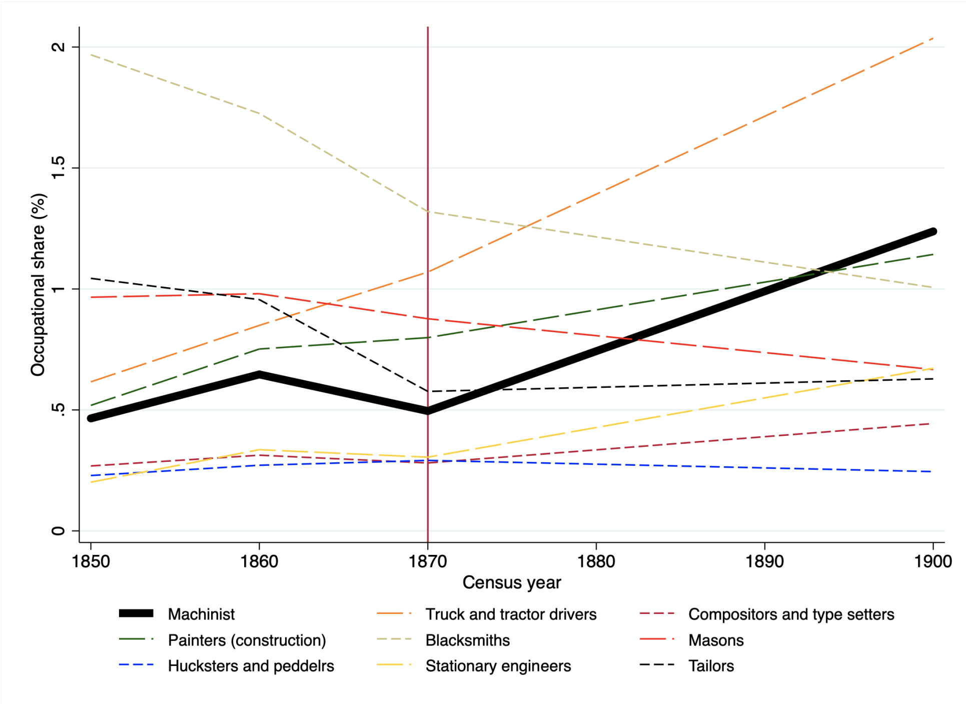


Figure 3.1: The evolution of occupational employment shares over time

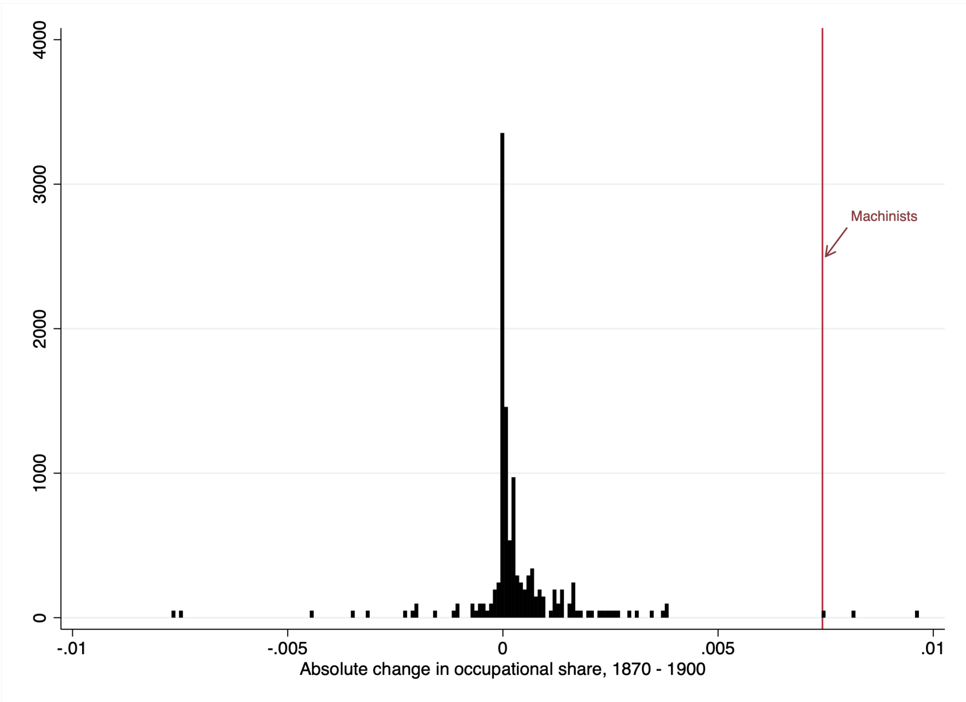


Figure 3.2: Histogram of occupational employment changes

Notes. **Figure 1**: the sample includes all males aged 16-65 who did not give a non-occupational response in the full count census in a given year. The number of workers in each occupation is divided by the total number of workers in 1850, 1860, 1870 and 1900. Harmonized occupations (1950) are used. Therefore, people classified as 'Truck and tractor drivers' were predominantly teamsters in the nineteenth century. **Figure 2**: the same sample used as in Figure 1. Only the employment of 'Mine operatives and laborers' and 'Truck and tractor drivers' grew faster than that of machinists out of the narrowly defined (i.e., not 'not elsewhere classified') occupations (see Section 3.3.1).

between 1849 and 1869, but it experienced a historically unprecedented growth in the 1870s and 1880s as production concentrated in factories (O'Brien 1988).

As assembling, setting up and maintaining the machinery were the main tasks of machinists, they were found in a wide range of mechanized sectors by the end of the 19th century (U.S. Department of Labor 1899): brush, buttonmold, canned corn, cigarette, faucet, female shoes, ingrain carpet, needle or teaspoon production, etc.. The sudden need for expertise to handle machines in these sectors led to a fast rising demand for machinists. The change in their employment share, which could barely outpace the growth of the labor force and was similar to that of some other craftsmen prior to 1870, experienced a steep acceleration (see Figure 3.1). As a result, their number almost doubled between 1870 and 1880, and a five-fold increase is registered in the full count census between 1870 and 1900. The US population merely doubled in these three decades. Thus, the expansion of the machinist occupation surpassed practically any other major group of craftsmen.

Despite the outstanding growth in their number, machinists did not experience a relative earnings decline. On the contrary, their relative earnings increased compared to most occupations from the early 1870s to the 1880s, and relative earnings gains seem to have disappeared only by the end of the century to some extent (see Table 3.1 and Section 3.5).[8] Taken together, the substantial employment expansion and relative earnings growth are consistent with a positive labor demand shock induced by the Second Industrial Revolution - relative to most other middle-skilled occupations.

3.3. Data

The main data sources for this work are various waves of the US full count census between 1850 and 1940 (Ruggles et al. 2021). This data set is complemented with i) novel, state- and time-varying earnings scores pre-1900 (Section 3.3.2); ii) newly digitized measures of county-level high school provision around 1880 (Appendix 3.C.2); iii) the occupational education rank of Song et al. (2020); and iv) some development-related county characteristics from the NHGIS (Manson et al. 2021).

[8]One potential cause behind the disappearance of earnings gains as measured by occupational earnings scores is the following. While the machinist occupation was growing, it started to employ relatively more young, less experienced workers. Thus, a declining average experience level might have pushed the occupational earnings level down.

Table 3.1: Occupational earnings (1850-1892; in 1890 dollars)

	Yearly earnings score					Growth (%)			Growth (Massachusetts)	
Occupation	1850	1860	1870-72	1879-1881	1890-92	1850-1872	1872-1880	1872-1892	1872-1880	1872-1892
Blacksmith	453	462	541	427	523	19	-21	-3		1
Bricklayer		457	684	671	895		-2	31	-25	13
Cabinetmaker			400	430	487		7	22	-12	14
Carpenter	376	389	478	422	492	27	-12	3	-7	14
Locomotive engineer	568	542	654	758	874	15	16	34	6	49
Locomotive fireman	330	310	356	367	488	8	3	37	19	31
Machinist	**414**	**430**	**445**	**473**	**530**	**8**	**6**	**19**	**11**	**22**
Mason	398	459	580	535	734	46	-8	27	5	37
Painter	455	417	447	528	460	-2	18	3	3	12
Pattern maker	407	435	544	474	618	34	-13	14	3	44
Plasterer	429	414	613	625	766	43	2	25		
Shoemaker			456	380	454		-17	0		
Stone cutter		438	733	640	858		-13	17	-24	9
Teamster	364	290	344	369	447	-6	7	30	12	45
Watchman	269	270	290	288	362	8	-1	25	7	16

Note: the data source is U.S. Department of Labor (1900). Occupations are not harmonized. Earnings are converted to 1890 dollars using inflation values from measuringworth.com. Every yearly earnings score is constructed as follows. First, all state-year daily wage observations are collected which are based on at least ten individuals. For 1870-1872, 1879-1881 and 1890-1892, we take the state-year observation with the largest number of individuals. Second, the conversion of daily wage rates to yearly earnings is described in Appendix 3.C.2. Finally, the values presented are the weighted averages of state-level scores. The weights are the number of individuals who contributed to the average wage calculation in every state. The last two columns contain only observations from Massachusetts.

3.3.1. Linking historical censuses

Analyzing intergenerational mobility necessitates linking individuals over time across distinct waves of the full count census. In this paper, we start out with the census conducted in 1870 to find the fathers (first generation - G1), whose offspring we follow in later decades and whose male parent (i.e. the grandfather - G0) we find in earlier decades in subsequent parts of this analysis.[9]

A few major restrictions are made on the 1870 full father (G1) sample. Exclusively fathers who were between 20 and 40 years old are included for two reasons. First, teenager workers tend to have transient occupations (Papageorgiou 2014). Second, relatively old workers did not live with their kids anymore (the only way to identify family relationships) and were often not alive in 1900, the year chosen for the analysis of their long-run outcomes.[10] Furthermore, we exclude every individual with a non-occupational response or outlier wealth (personal property or real estate value above the 99th percentile). Individuals who held an agricultural occupation (farmer, farm manager/foreman/laborer), reported certain apprenticeship, or their harmonized occupation was a type of ”not elsewhere classified” (e.g., ’Clerical and kindred workers (n.e.c.)’) are also omitted. These restrictions are important because farmers had completely different characteristics compared to non-agricultural workers. Additionally, apprenticeships could obviously not be the final occupation of young adults. Finally, loosely classified occupations make the use of occupational education ranks or earnings scores less reliable if not impossible.

[9]The paper is limited to the analysis of male observations since the surname change of women upon marriage makes their linking over time impossible.

[10]The 1890 census records were burnt in a fire.

As a next step, fathers are linked to their own 1900 observation. An individual is considered linked if at least one of the two conservative linking methods offered by Abramitzky et al. (2020) yields a match.[11] These linking methods have a particularly low false positive ratio (Bailey et al. 2020). Thus, we can avoid erroneously linking observations between two different people which helps us reduce the attenuation bias at the expense of a reduced sample size. Importantly, this linking rule is used for *every* linking in the entire paper.

In 1870, we can identify sons (G2) who lived with their father and link them to 1900 and 1940, separately. Exclusively sons who were at most 20 years old in 1870 are included. Then, we link these sons between 1870-1900 and 1870-1910, and find their kids in the respective end year in order to identify grandsons (G3). As a final step, we link grandsons found in 1900/1910 to 1940.

In Section 3.6.1 and 3.6.3, we use the characteristics of grandfathers (G0) in 1860. To do so, we link fathers back to 1850 and 1860. If a grandfather is only found in 1850 (e.g., because he already lived separately from the father in 1860), we link him forward to 1860 in order to obtain grandfathers' characteristics from the exact same year.

3.3.2. Occupational earnings scores for the late nineteenth century

One of our contributions is providing novel, state-specific earnings score estimates for the late-nineteenth-century United States. There are at least three reasons why these measures are crucial for this project. First, the traditional approach used in the literature - generating occupational income scores based on income reported in the 1940 census and using them in earlier decades - has been shown to perform more poorly the earlier it is applied prior to 1940 (Inwood et al. 2019; Saavedra and Twinam 2020). Especially for periods when relative wages are changing rapidly, Inwood et al. (2019) recommend constructing earnings scores based on data from the studied time period, even if the sample might not be representative. Second, a considerable share of education received was informal in the 19th century (e.g., apprenticeships; see Goldin and Katz 2008; Kelly et al. 2020; Meisenzahl and Mokyr 2011). Therefore, while we can control for the (formal) education percentile rank devised by Song et al. (2020), we might not be able to capture the full difference in occupational human capital across occupations with this measure. However, earnings scores combined with the educational rank might very well capture the actual level of human capital implied by the sum of formal and informal education. Third, even the labor market of the north-eastern part of the United States (New England,

[11] The conservative linking methods provided by Abramitzky et al. (2020) require matches be unique by name and birthplace within a five-year age band.

Middle Atlantic, East-North Central), where most of the machinists lived, was not integrated until the 1880s and the difference between the north-eastern and Pacific (or southern) regions persisted even longer (Rosenbloom 1996; 1998). Kaboski and Logan (2011) also find spatially-varying returns to education in the United States in the early twentieth century. Consequently, applying the same earnings score to a certain occupation all over the United States could lead to inaccurate conclusions. To the best of our knowledge, all existing earnings scores data sets for the late nineteenth century provide a single score for each occupation and pertain to the last decade of the 19th century (Preston and Haines 1991; Sobek 1996). Hence, we proceed to construct our own measure of state-specific occupational earnings for the 1870s and 1880s.

In this section, we outline the main steps of calculating these earnings scores. The interested reader can find detailed information and the discussion of the underlying assumptions in Appendix 3.C.2. The source of our occupational earnings information is U.S. Department of Labor (1900). For many occupations,[12] we digitized the average daily wage found in 1870-72 (the 1872 score), 1879-81 (the 1880 score), 1890-92 (the 1892 score) in every state. In case of multiple observations within a three-year period, we digitized the daily wage which was calculated based on the largest number of observations. Then, daily wages were converted to yearly earnings scores and 1890 dollars. In this way, the earnings scores could be calculated for many large, low- and medium-skilled occupations. The income of high-skilled occupations (e.g., lawyers or physicians) was imputed by combining the earnings scores provided by Sobek (1996) with our own earnings scores.

The previously described steps provide *nominal* earnings scores. However, it is well-known that the costs of living differed significantly between urban and rural areas, and across states (Koffsky 1949; Stecker 1937). Hence, we also calculated *real* earnings scores adjusting for these price differences following Collins and Wanamaker (2014) (see Appendix 3.C.2 for more details).

3.3.3. Summary statistics

Machinists were not the "representative agents" of the US economy. As it can clearly be seen from Table 3.2, most of their observables differed from the rest of the population. Machinist fathers in our analysis were slightly younger, more educated, less wealthy, more likely to be immigrants (especially of English ancestry) and lived in more urban, larger places than non-machinists in 1870. Since they were concentrated in the New England and Middle Atlantic census divisions, one might want to disentangle the effect of spatial distribution from other causes

[12] Besides machinists, the focus was on occupations i) which are in the control group in a large number in 1870 following propensity score matching, and ii) which played a large role in the economy later (i.e., important possible occupations for fathers or sons in 1900).

Table 3.2: Summary statistics of fathers (G1 in 1870)

Variable	Mean (*non-machinists*)	Difference (machinists - non-machinists) *Raw difference*	*Conditional on state-fixed effects*
Age (in years)	34,1	-0,5 [0.076] ***	-0,5 [0.089] ***
Literate (Yes=1)	0,92	0,05 [0.0082] ***	0,04 [0.0075] ***
Education rank of occupation (Song et al. 2020)	50,4	4,3 [1.182] ***	4,5 [0.968] ***
Value of real estates (in 1870 dollars)	793,4	-128,0 [32.218] ***	-74,9 [14.424] ***
Value of personal property (livestock, jewels, bonds, etc.; in 1870 dollars)	350,9	-87,7 [15.703] ***	-85,9 [26.635] ***
Both parents native born (Yes=1)	0,58	-0,10 [0.025] ***	-0.10 [0.0229] ***
Both parents foreign born (Yes=1)	0,37	0,08 [0.0226]***	0,07 [0.0218] ***
Immigrant - UK or Ireland (Yes=1)	0,16	0,12 [0.0181] ***	0,1 [0.0173] ***
Immigrant - Germany (Yes=1)	0,15	-0,04 [0.0128] ***	-0,02 [0.009] *
Urban place of living (Yes=1)	0,44	0,34 [0.0195] ***	0,26 [0.0285] ***
Population of place of living	75532	33543 [14867] **	27692 [13249] **
New England (Yes=1)	0,13	0,16 [0.0731] **	-
Middle Atlantic (Yes=1)	0,32	0,04 [0.051]	-
East-north Central (Yes=1)	0,28	-0,1 [0.0354] **	-
West-north Central (Yes=1)	0,09	-0,04 [0.0220] *	-
South (Yes=1)	0,15	-0,05 [0.0249] **	-
West and Pacific (Yes=1)	0,03	-0,01 [0.0119]	-

Note: robust standard errors clustered at the state level (1870) in brackets. The summary statistics presented pertain to the final, total sample used in Table 3.5 and C6. The raw difference between means of machinists and non-machinists is the coefficient on the machinist dummy in an OLS regression with a constant and the dummy. This OLS regression also includes state-fixed effects (1870) in the last column. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of significant difference. Therefore, differences in means are also presented after netting out state-fixed effects. Nonetheless, machinists seem to exhibit similar, though somewhat smaller differences in characteristics within states.

3.4. Empirical strategy

In this section, we describe our two, complementary empirical strategies: propensity score matching and fixed effects regressions.

3.4.1. Propensity score matching

Our primary empirical strategy is propensity score matching on many observable characteristics of fathers in 1870 (Austin 2011; Ho et al. 2007; Leuven and Sianesi 2003; Rosenbaum and Rubin 1983). This estimation strategy amounts to estimating each individual's probability of being a machinist in a logit regression as a first step. For every machinist father, the five non-machinist fathers with the closest estimated probability are chosen as control observations with replacement.[13] Then, we compare the outcomes of machinist fathers and of their offspring to the outcomes of matched control fathers (and of their offspring) in the resulting sample. The relatively small share of machinists in the full sample implies that there are many potential control observations, making our setting particularly well-suited for matching. The aim of matching is to reduce the correlation between the machinist

[13] Additionally, we use a caliper of 0.01 and restrict the analysis to the common support of machinist and non-machinist fathers. This never results in losing more than ten treated observations in the main analysis. In a few analyses of later generations, we use ten instead of five neighbors because of the small sample size but this change is always duly noted.

dummy, which indicates if a father was a machinist in 1870, and the control variables. The full list of these control variables is shown in Appendix 3.C.2. In short, we include i) personal characteristics (e.g., age, literacy, proxies of migration background, education rank of occupation, etc.); ii) place of living characteristics (e.g., urban dummy, state-fixed effects, measures of county-level industrialization, etc.); and iii) state-occupational level features constructed pre-1870 (e.g., probability of job switching or migration). Importantly, 1870 was the last historical census wave in which detailed information was collected on personal wealth: the value of real estates and personal property (the contemporary dollar value of all stocks, bonds, mortgages, notes, livestock, plate, jewels, and furniture owned by the respondent), separately.[14] Interactions and squares of many background characteristics are also included to match the distribution of these covariates more closely (Ho et al. 2007; Imai et al. 2008).

The main advantage of matching is that by reducing the correlation between the explanatory variable of interest and observables, such as personal wealth or urban status, we considerably reduce the influence of correlated unobservables. For example, the wealth proxies are most likely correlated with individual talent and family heritage, or the urban status can capture many urban (dis)amenities. Furthermore, matching diminishes our own discretion over how to control for a given background characteristic (Ho et al. 2007).[15]

The main limitation of using matching in our setting is that the full count census does not provide individual-level information on earnings and education before 1940. To overcome this lack of data, occupation-based characteristics are used. For education, the occupational education percentile rank of Song et al. (2020) is included. This is a percentile rank (0-100) based on the average occupational years of (formal) schooling in a person's birth cohort.[16] For income, which is probably more volatile over time than the education requirement of most occupations, we use our own state-level real earnings score constructed for 1870-72. The latter is exclusively included in the analyses of income-related outcomes because its inclusion reduces the sample size along with the precision of the estimation without significantly changing the coefficients on non-pecuniary outcome variables. In fact,

[14] Wealth at a young age is an even better predictor of future wealth than parental wealth, and a good proxy for intergenerational correlation in savings behaviour and additional transfers from parents (Boserup et al. 2018).

[15] The application of propensity score matching in this paper is mostly immune to the criticism of King and Nielsen (2019) for several reasons: i) contrary to their claim that matching often increases imbalance compared to the unmatched sample, we transparently show that matching decreases it in our application; ii) the large sample makes the "propensity score matching paradox" less likely to appear; and iii) even though a caliper is used, the number of unmatched and, consequently, dropped machinists is always one-digit.

[16] For fathers and sons, we use the earliest available birth cohort around 1880 whose percentile rank is based on detailed years of schooling data and not merely on literacy. For grandsons, we use the percentile of the birth cohort around 1900.

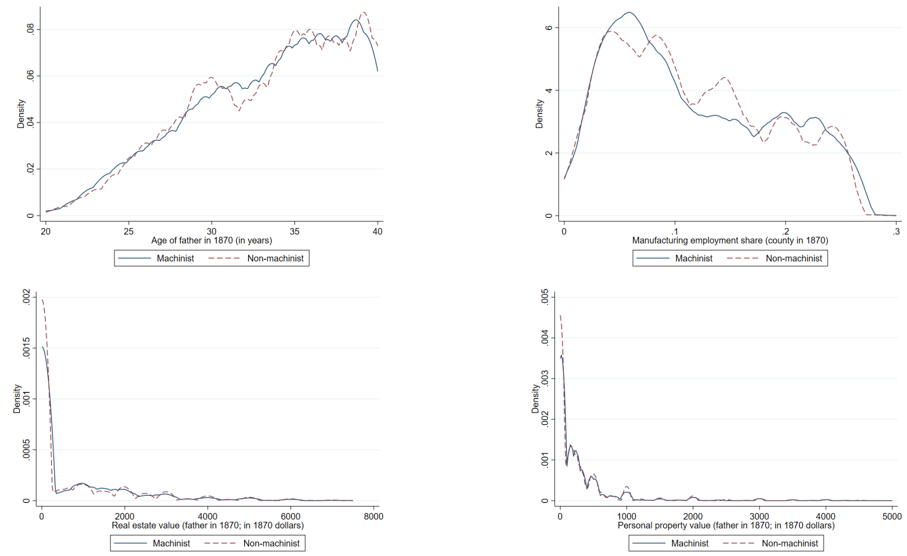


Figure 3.3: The histogram of some continuous characteristics after matching

Notes: the figures are created after the propensity score matching which generates Table 3.5. They depict the density of certain continuous control variables for machinists and matched non-machinists using weights obtained from the matching process.

the real earnings score tends to be somewhat *lower* for machinists than for matched non-machinist control observations when it is not included in the list of control variables.[17]

In practice, propensity score matching works well in this setting and the correlation between observables and the machinist dummy, which is highly significant for most cases (Table 3.2), vanishes. Apart from similar means, the whole distribution of control covariates is closely matched (see Figure 3.3). However, the mean of a small subset of variables remains significantly different in some cases. The typical example is urban status: while machinists tend to be significantly *more* urban compared to the full sample, they are somewhat *less* urban in the matched one. Nevertheless, the standardized difference lies below 10% (the upper bar for tolerable difference - Austin 2011) even in this case.[18] To avoid any bias from such residual differences, we include every control variable (their main effects)

[17]Machinists have an average score of $500, while the matched (unmatched) control average is $530 ($582) in Table 3.5. Notice that this imbalance works *against* our findings.

[18]In Table 3.3, the difference (machinist minus non-machinist) between the probability of urban place of living in 1870 is 34% before matching and -4% after matching. In this particular application, the mean (median) standardized bias is 15.4 (9.2) before matching and 2.2 (1.3) after matching.

which has a significantly different mean at 5% in a regression after matching.[19] The other reason for running a regression on the matched sample instead of reporting the immediate outcome of matching is to construct clustered, more conservative standard errors at the state level.

The occupations with the most matched control observations are presented in Table C1. While the role of carpenters, truck and tractor drivers, and shoemakers is relatively large, none of them exceeds 10% of the control observations. We also show in Section 3.6 that their omission does not affect the results in any meaningful way. It must be emphasized that we use harmonized occupational codes provided by IPUMS as it is usually done in the literature. Therefore, the category ’Truck and tractor drivers’ mainly consists of teamsters, draymen and hackmen in the 19th century.

3.4.2. Fixed effects regression

Despite the appealing features of propensity score matching, it precludes the inclusion of numerous fixed effects for two reasons: the algorithm occasionally does not converge when including county or county-urban status-fixed effects, and the small size of the matched subsample makes the estimation of fixed effects very imprecise. Another problem with matching is that it does not allow for weighting, so the sample cannot be weighted to make it representative of the US population (more details in Section 3.6.2). To address these issues, we also present some results using fixed effects regressions.

In our fixed effects regressions, exactly the same baseline controls are included as in matching in addition to county-fixed effects (1870).[20] Therefore, the offspring of machinists are compared to the offspring of non-machinists who lived in the exact same county in 1870 and had similar paternal (G1) observables. To bring this analysis in spirit closer to matching, fathers whose occupation is below the 25th or above the 85th educational rank percentile are omitted from the analysis (the rank of machinists is the 55th). In this way, the very low-skilled (e.g., lumbermen or miners) and high-skilled (e.g., architects or lawyers) fathers are not in the sample so that we can focus on the ”middle class”. Another advantage of fixed effects regressions is that they allow us to precisely estimate interaction terms between the machinist dummy and other variables as well.

[19]We are aware of the ”balance test fallacy” coined by Ho et al. (2007) and Imai et al. (2008), who discourage researchers to use the significance of difference between means as a balancing threshold. However, we find in practice that the inclusion of significantly different (p-value below five percent) characteristics matters to a very limited extent and the inclusion of non-significantly different variables does not have any effect on the estimation.

[20]We use the *reghdfe* package in Stata by Correia (2016).

Formally, the regression specification takes the following form:

$$y_{s,f,c,1900} = \beta \cdot Machinist_{f,1870} + \gamma \cdot x'_{f,1870} + \delta_{c,1870} + \epsilon_{s,f,c,1900} \qquad (3.1)$$

where $y_{s,f,c,1900}$ represents an outcome variable for son s of father f measured in 1900 (e.g., a binary variable if the son held an agricultural occupation). The explanatory variable of interest is $Machinist_{f,1870}$, which equals one if the father was a machinist in 1870. County-fixed effects ($\delta_{c,1870}$) and all paternal baseline controls ($x_{f,1870}$) are also included. Reassuringly, the effects on main outcomes estimated by propensity score matching and fixed effects regressions tend to be quantitatively and qualitatively very similar.

In order to get a consistent estimate of β, the error term, $\epsilon_{s,f,c,1900}$, must be uncorrelated with the machinist dummy conditional on our predetermined controls. Thus, the main concern about the validity of the empirical strategy is that particularly talented fathers sorted into the machinist occupation before 1870 in an unobserved way, causing omitted variable bias. To alleviate this concern, we present many heterogeneity and robustness checks in Sections 3.5.2 and 3.6. These empirical exercises suggest that (a within-family intergenerational transmission of) unobserved ability is not driving our results.

3.5. Main results

In the first part of this section, key results establishing the gains of the machinist occupation post-1870 and the intergenerational transmission between machinists and their (grand)sons are presented. To elaborate on mechanisms of transmission, we conduct some heterogeneity exercises in the second part.

3.5.1. Long-term effects and intergenerational transmission

Fathers (G1) between 1870 and 1900 Table 3.3 contains the main, non-pecuniary outcomes for our linked 1870-1900 father sample using propensity score matching. The first column shows that machinists were 8.7 percentage points (0.2 standard deviation) less likely to switch their occupation. This coefficient can be decomposed into switching to different types of jobs. In particular, roughly one-third of the total effect stemmed from a lower likelihood of switching to an agricultural job (Column 2), while the rest can be attributed to a less likely change for another non-agricultural occupation (Column 3). We interpret the lower likelihood of leaving the initial occupation as the first sign of a beneficial effect on machinists post-1870. Namely, there is an extensive literature which documents the large costs of occupation switching in many contexts (e.g., Artuç et al. 2010; Cortes and Gallipoli 2018; Dix-Carneiro 2014; Kambourov and Manovskii 2009; Sanders and

Table 3.3: Main outcomes - fathers (G1; 1870-1900)

	Occupational change [(1) = (2)+(3)]			Migration (Yes=1)		Place of living (1900 - Yes=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any occupation (Yes=1)	Agricultural	Non-agricultural	Within-state	Across states	Higher population than in 1870	Urban
Machinist (G1)	-0.087***	-0.033***	-0.054***	0.006	0.011	0.058***	0.073***
	(0.010)	(0.006)	(0.012)	(0.007)	(0.008)	(0.012)	(0.010)
Mean of outcome	0.77	0.19	0.57	0.20	0.37	0.46	0.50
Standard deviation of outcome	0.42	0.39	0.49	0.40	0.48	0.50	0.50
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	18811	18811	18811	18811	18811	18811	18811
Number of clusters	50	50	50	50	50	50	50

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. All specifications are weighted by weights obtained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 3902 matched machinist fathers. The outcome variable is a binary variable which equals one if the father changed occupation (Col. 1), changed occupation and the new occupation is agricultural (Col. 2 - farmer, farm manager/foreman/laborer) or non-agricultural (Col. 3), migrated within-state across counties (Col. 4) or across states (Col. 5), his place of residence fell into a larger *SIZEPL* category in 1900 than in 1870 (Col. 6), he lived in an urban place in 1900 (Col. 7). Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Taber 2012; Traiberman 2019). This literature suggests that machinists lost less lifetime earnings caused by the costly accumulation of occupation- or task-specific human capital due to their lower likelihood of changing their occupation.

Internal migration has been established as a pre-eminent way to upward mobility in the studied time period (J. Long and Ferrie 2007; 2013; Ward forthcoming). However, no evidence is found on a differential probability of migration within or across states (Columns 4-5). We further elaborate on migration *destinations* in Table C2. First, we decompose the insignificant migration differential and find that machinist fathers tended to migrate significantly more (less) to urban (rural) places. Second, we also establish that initially rural machinists were particularly more likely to move to urban areas and initially urban machinists were less likely to migrate to rural areas. These effects can clearly be seen in Columns 6-7 of Table 3.3 as well: machinist fathers lived in more populous and more urban places by 1900 (both effects stronger than 0.1 standard deviation). The urban environment could provide them and their offspring with better opportunities in a period when urbanization and growth were tightly intertwined.

Next, we direct our attention to analyze the effect on occupational earnings scores. In Columns 1-2 of Table 3.4, we assume that fathers held the same occupation and lived in the same place in 1880 as in 1870. We do so because an additional linking to 1880 would come at the expense of a large sample size reduction. The coefficients suggest that machinist fathers experienced a relative increase of 8-9 log-points in their earnings score.[21] This finding is unsurprising since it is documented in Table 3.1 that the relative wage of the machinist occupation increased compared to most other occupations in this time period. While the magnitude of the effect is substantial (0.25 s.d.), we treat it as an *upper* bound on the actual effect because control fathers could switch their occupation or place of living in

[21] The same coefficient on the nominal and real earnings score is mechanical. Since we assume that fathers do not change their occupation, state and urban status between 1870 and 1880, only the nominal wage change of the given occupation matters in this calculation.

order to reduce the relative earnings gap. Therefore, we also linked fathers between 1870 and 1880 (instead of 1900) and, thus, allowed for occupation and place of living change in Table C3. As expected, the estimated earnings effect is somewhat smaller (6-7 log-points) but still significantly positive.

In the last four columns of Table 3.4, we use the occupation and state of living of fathers in 1900 to construct outcome variables. The first observation is that both the nominal and real earnings score gains expectedly declined compared to 1880. The second observation is that using the widely-used earnings scores (Preston and Haines 1991; Sobek 1996) results in a larger coefficient compared to our own nominal score.[22] We suspect that this discrepancy partly stems from the treatment of agricultural workers. In particular, the ratio between the score of farm laborers and other laborers is substantially lower in Sobek (1996) or Preston and Haines (1991) than in the case of our scores. Knowing that machinists were significantly less likely to switch to agricultural occupations, assigning lower scores to agricultural jobs amplifies the relative earnings gains of machinists. We believe that our scores might be more accurate since Alston and Hatton (1991) or Hatton and Williamson (1991) show that a large part of the gap in nominal earnings between farm and common laborers can be explained by more in-kind benefits (especially the value of accommodation) for the former group. As explained in Appendix 3.C.2, we calculate farm laborers' remuneration based on daily wages *without* accommodation which brings the ratio between the earnings of farm and common laborers close to those reported in Alston and Hatton (1991) and Hatton and Williamson (1991), and takes into account the monetary value of accommodation. Nevertheless, the 3.5-8 log-points higher nominal earnings scores do not account for the fact that machinist fathers were more likely to reside in more populous, urban places in 1900 - implying higher consumer prices. When these differences in cost of living are adjusted for, the estimated positive effect becomes insignificant (Column 5). In other words, the real gains of initially machinist fathers were arbitraged away in the (very) long run.

We further investigate the effect on earnings scores in Table C4, focusing on individuals who changed their occupation between 1870 and 1900. The main takeaway of this table is that, besides the increasing relative wage of the machinist occupation, the relative earnings gain of initially machinist fathers was the result of a six percentage points lower likelihood of switching to an occupation with considerably lower earnings rather than differential upward mobility. In our interpretation, non-machinists lost their occupations more frequently in the turbulent times of the Gilded Age, when recurrent busts in the aftermath of panics characterized an overall robust growth. As breadwinners of their family, they had to find an

[22]Preston and Haines (1991) do not provide an earnings score for owner-occupier farmers and calculate earnings scores based on an urban sample in the Cost of Living survey. Sobek (1996) instead calculates an unweighted average of all distinct earnings scores for every occupation.

Table 3.4: Measures of economic status (medium- and long-run) - fathers (G1)

	State-occupation in 1870		State-occupation in 1900			
	(1) State-level nominal log-score (1880)	(2) State-level real log-score (1880)	(3) Sobek log-score	(4) State-level nominal log-score (1892)	(5) State-level real log-score (1892)	(6) Preston-Haines log-score
Machinist (G1)	0.085*	0.085*	0.067***	0.034***	0.019	0.084***
	(0.048)	(0.048)	(0.009)	(0.011)	(0.012)	(0.012)
Mean of outcome	6.11	6.25	6.21	6.25	6.37	6.49
Standard deviation of outcome	0.33	0.34	0.56	0.44	0.43	0.37
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	16124	16124	11573	11573	11573	9895
Number of clusters	32	32	47	47	47	50

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1870 in Col. 1-2; 1900 in Col. 3-6) level. All specifications are weighted by weights obtained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 3820 (Col. 1-2), 2669 (Col. 3-5) and 2334 (Col. 6) matched machinist fathers. The outcome variable is the state-level nominal and real log-score (Col. 1 and 2.) merged to the state and occupation of fathers in 1870; the Sobek, state-level nominal and real, and Preston-Haines log-score (Col. 3-6) merged to the state and occupation of fathers in 1900. Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

alternative, potentially lower-paying (agricultural) occupation in the absence of generous unemployment benefits. This interpretation is consistent with Boone and Wilse-Samson (2019) who show that movement to farms served as a source of migratory insurance during the Great Depression. Moreover, the fact that machinists were not more likely to switch to managerial jobs or becoming proprietors (Column 6) supports our claim that the improved outcomes for themselves and their offspring were not the result of unobserved talent.

The outcomes of sons (G2) The next question we answer is if the benefits of fathers could be transmitted to their sons. The main, non-pecuniary outcomes are presented in Table 3.5. Similarly to their fathers, sons were significantly less likely to hold an agricultural occupation (Column 1). Furthermore, they held occupations which had significantly higher education ranks (almost +0.1 s.d.). Whereas the latter finding simply suggests that machinists' sons held occupations with on average more educated peers, we can estimate *individual-level* schooling using the 1940 census. We linked sons between 1870 and 1940 to this end. The results in Table C5 show that machinists' sons had indeed 0.21 years more schooling. The effect is mainly the result of a 3.6 percentage points (+0.1 s.d.) higher likelihood of having some secondary education, meanwhile the effect on university education is a tightly estimated zero. The secondary school coefficient should be treated as a *lower* bound on the actual effect since the beneficial effect of education on longevity could lead to endogenous attrition. Thus, as sons were at least seventy years old in 1940, the less educated control sons might have been more likely to pass away before 1940, leading to a downward bias in the estimated coefficient.

In line with the higher level of educational attainment, we find a significantly higher probability of long-distance migration for sons between 1870 and 1900 (+0.1 s.d. - Column 4 in Table 3.5; Malamud and Wozniak 2012; Rosenbloom and Sundstrom 2003; Wozniak 2010). This foreshadows our findings on higher earnings because the migration premium increased in distance in this time period (Ward

Table 3.5: Main outcomes - sons (G2; 1900)

	Occupational characteristics		Migration (Yes=1)		Place of living			Personal characteristics		
	(1) Agricultural occ. (Yes=1)	(2) Education rank	(3) Within-state	(4) Across states	(5) Urban (Yes=1)	(6) Higher population than in 1870 (Yes=1)	(7) Manuf. emp. per capita (county)	(8) # of children	(9) Married (Yes=1)	(10) Owning house (Yes=1)
Machinist (G1)	-0.028***	2.403***	-0.009	0.043***	0.045***	0.033**	0.004	-0.080**	0.005	0.000
	(0.007)	(0.745)	(0.011)	(0.012)	(0.011)	(0.016)	(0.003)	(0.038)	(0.010)	(0.014)
Mean of outcome	0.19	49.68	0.30	0.31	0.55	0.52	0.09	1.70	0.77	0.43
Standard deviation of outcome	0.39	28.25	0.46	0.46	0.50	0.50	0.07	1.86	0.42	0.50
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	8745	8745	8745	8745	8745	8745	8745	8745	8745	8745
Number of clusters	45	45	45	45	45	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. All specifications are weighted by weights obtained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 1842 matched machinist sons. The outcome variable is an agricultural occupation indicator (Col. 1 - farmer, farm manager/foreman/laborer), the education rank of occupation (Col. 2), a binary variable which equals one if the son migrated within-state across counties (Col. 3) or across states (Col. 4) between 1870 and 1900, a binary variable which equals one if the son lived in an urban place in 1900 (Col. 5) or his place of residence fell into a larger *SIZEPL* category in 1900 than in 1870 (Col. 6), manufacturing employment per capita (as % of total county population; Col. 7), the number of children in the household (Col. 8), a marriage status (Col. 9) and house ownership (Col. 10) indicator. Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Measures of economic status - sons (G2; 1900)

	(1) Sobek log-score	(2) State-level nominal log-score	(3) State-level real log-score	(4) State-level real score (level)	(5) Preston-Haines log-score
Machinist (G1)	0.070***	0.042***	0.032***	15.208**	0.069***
	(0.014)	(0.011)	(0.010)	(6.452)	(0.011)
Mean of outcome	6.23	6.23	6.35	628.89	6.45
Standard deviation of outcome	0.53	0.43	0.41	294.05	0.39
Unbalanced controls	Yes	Yes	Yes	Yes	Yes
Sample size	6812	6812	6812	6812	6687
Number of clusters	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. All specifications are weighted by weights obtained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 1548 (1514 in Col. 5) matched machinist sons. The outcome variable is the Sobek, state-level nominal and real log-score (Col. 1-3), the state-level real score in levels (Col. 4) and the Preston-Haines log-score (Col. 5). Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

forthcoming). The effect on a higher probability of urban homes and larger cities persists, though it slowly starts to fade away compared to the first generation.[23] In fact, the difference in the manufacturing employment share of the county of living in 1900 is insignificant. This suggests a certain convergence in the type of place of living across the sons of machinists and non-machinists. Finally, we uncover some evidence that the more educated sons of machinists had fewer kids, perhaps because they faced higher opportunity costs of raising children (Ager et al. 2020). The effect on marriage probability and house ownership is insignificant.

The pattern of the earnings effect for sons is similar to the paternal one: the well-known nominal scores having a more positive coefficient than our own score, and a diminished coefficient once across-state and rural-urban price differences are accounted for (Table 3.6).

The outcomes of grandsons (G3) Table 3.7 documents the main, non-pecuniary outcomes for machinists' grandsons. The set of possible outcomes is richer thanks to the increased data collection effort in the 1940 census. First, we learn that even the grandsons were less likely to be engaged in an agricultural occupation, they

[23]For instance, the positive effect of an urban place of living drops by 60% in magnitude, from 1.5 to 0.9 standard deviations.

Table 3.7: Main outcomes - grandsons (G3; 1940)

	Occupational characteristics				Education		Personal characteristics		
	(1) Agricultural occ. (Yes=1)	(2) Weeks worked	(3) Self-employed (Yes=1)	(4) Education rank	(5) # of grades completed	(6) More than primary education (Yes=1)	(7) # of children	(8) Married (Yes=1)	(9) Owned a house (Yes=1)
Machinist (G1)	-0.024**	0.865**	0.013	1.902***	0.312***	0.047**	-0.002	0.016	0.023
	(0.011)	(0.408)	(0.014)	(0.666)	(0.113)	(0.020)	(0.061)	(0.012)	(0.022)
Mean of outcome	0.10	44.29	0.22	31.95	9.90	0.54	1.52	0.87	0.53
Standard deviation of outcome	0.30	14.33	0.42	19.04	3.26	0.50	1.66	0.33	0.50
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	6383	6383	6383	6383	6383	6383	6383	6383	6383
Number of clusters	49	49	49	49	49	49	49	49	49

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1940) level. All specifications are weighted by weights obtained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 969 matched machinist grandsons. We use ten matched control observations instead of five owing to the small number of machinist grandsons. The outcome variable is an agricultural occupation indicator (Col. 1 - farmer, farm manager/foreman/laborer), the number of weeks worked (Col. 2), a self-employed status indicator (Col. 3), education rank of occupation (Col. 4), highest grade of schooling (Col. 5 - winsorized at the 99th percentile in the final sample), a binary variable which equals one if the highest grade of schooling is at least 9 years (Col. 6), number of children in the household (Col. 7), a marriage status (Col. 8) and house ownership (Col. 9) indicator. Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8: Measures of income - grandsons (G3; 1940)

	Self-employed & wage workers				Wage workers			
	(1) Log-wage (nominal)	(2) Log-wage (real)	(3) Wage (level)	(4) Non-wage income (Yes=1)	(5) Log-wage (nominal)	(6) Log-wage (real)	(7) Wage (level)	(8) Non-wage income (Yes=1)
Machinist (G1)	0.068**	0.061*	112.774**	-0.009	0.082**	0.074**	117.019**	-0.026**
	(0.034)	(0.034)	(52.069)	(0.012)	(0.036)	(0.037)	(55.541)	(0.011)
Mean of outcome	7.26	0.20	1855.92	0.30	7.21	0.15	1771.97	0.17
Standard deviation of outcome	0.82	0.81	1278.30	0.46	0.82	0.81	1207.89	0.38
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	6212	6212	6212	6212	5244	5244	5244	5244
Number of clusters	49	49	49	49	49	49	49	49

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1940) level. All specifications are weighted by weights obtained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 908 (746 in Col. 5-8) matched machinist grandsons. We use ten matched control observations instead of five owing to the small number of machinist grandsons. The outcome variable is the log of reported nominal wage (Col. 1 and 5 - winsorized at the 95th percentile in the final sample), the log of reported real wage (Col. 2 and 6 - winsorized at the 95th percentile in the final sample), the level of reported nominal wage (Col. 3 and 7 - winsorized at the 95th percentile in the final sample), and a meaningful non-wage income indicator (Col. 4 and 8. - more than $50). The sample includes wage workers as well as self-employed people reporting non-zero wage in Columns 1-4, while it is restricted to wage earners in Columns 5-8. The imputation of self-employed income is described in Appendix 3.C.2. Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

worked more weeks, but did not have a differential likelihood of self-employment (Columns 1-3). The availability of individual-level educational attainment allows us to compare the magnitude of the effect on the occupational education rank and on the highest grade of individual-level schooling (Columns 4 and 5). Reassuringly, both variables imply a very similar, positive magnitude: 0.1 standard deviation. This comparison corroborates our entire analysis because we seem to approximate actual education very closely with occupation-level average education scores. We also see that machinist grandsons were almost five percentage points more likely to have completed at least primary school. However, we do not find any significant effect on the number of children, marriage probability and house ownership indicator (though the signs of the coefficients are in the expected direction).

The final piece of main results concerns the income of grandsons (Table 3.8). The first four columns include wage earner as well as self-employed grandsons. As self-employed individuals did not report their income, we impute it following the best practice in the literature (see Appendix 3.C.2). The results are clear: both the nominal and the real wage effect are positive and significant. This conclusion becomes even stronger when we focus exclusively on wage earners whose wages do not require imputation (Columns 5-8).

In conclusion, we document large and significant gains for the sons and grandsons of machinists in terms of education- and income-related outcomes even after seventy years. The implied limited level of intergenerational mobility[24] is consistent with a large literature which demonstrates that intergenerational mobility was indeed low and declined at the turn of the twentieth century in the United States (J. Long and Ferrie 2013; Olivetti and Paserman 2015; Song et al. 2020; Ward 2019). Therefore, the initial gains of machinist fathers dissipated slowly over time and generations.

3.5.2. Mechanisms behind the intergenerational transmission

The goal of this section is to understand the mechanism behind the intergenerational transmission of the improved socio-economic status of machinist fathers to their offspring. We focus on the transmission from fathers to sons since the small sample size for grandsons does not let us draw robust conclusions.

Secondary education as a pathway to upward mobility Education meant at most primary schooling for the overwhelming majority of young people in the late-nineteenth-century United States: merely nine percent of American youth had high school diploma even in 1910. This share only moderately increased from the 1870s until the start of the so-called High School Movement in the 1900s. In the studied time period, high schools were mostly attended by the children of the (upper)-middle class. To a lesser extent, farmers or manual workers also sent their offspring to study as they saw high school education as a way out of a rural life and physical toil for their children. Rural areas maintained mostly private high schools and only cities could afford to finance public high schools. Private secondary schools regularly charged a tuition fee and non-residents were expected to pay a boarding fee (cost of accommodation) as well, meanwhile public institutions normally did not demand any payment. Nevertheless, the role of public schools remained inferior to private institutions until the 1890s. Therefore, in the absence of strictly implemented compulsory schooling laws for secondary schooling, it mainly depended on their parents' income and preferences if the sons of machinists and their peers received post-primary education (e.g., Goldin 1998; Goldin and Katz 2000; 2008; Lingwall 2010; Tyack 1974).

We documented earlier that machinist fathers experienced occupational stability and higher earnings in the period when most sons in the sample reached high school age around 1880. Thus, they could afford to educate their sons more easily. Indeed, we present evidence consistent with a complementarity between local private secondary school provision and parental income. Additionally, it has been demonstrated that parents of similar socio-economic status tend to have similar

[24]Around 65% of the earnings gains of fathers (Column 1 in Table C3) were transmitted to their sons (Column 2 in Table 3.6), which is consistent with the values reported in Ward (2019).

Table 3.9: Heterogeneity by the level of private tuition fee - sons (G2; 1900)

	Education rank						Other outcomes (medium fee)	
	(1) Full sample	(2) Low tuition fee (¡25th percentile)	(3) Medium tuition fee (25th-75th percentile)	(4) High tuition fee (¿75th percentile)	(5) Medium tuition fee (below median public HS)	(6) Medium tuition fee (teacher-pupil ratio¿0.04)	(7) Max. primary education (% in occupation)	(8) Sobek log-score
Machinist (G1)	4.207***	6.454***	4.000***	3.143	4.369**	1.477	-2.067**	0.091***
	(0.928)	(1.491)	(1.270)	(1.900)	(2.074)	(1.625)	(0.869)	(0.024)
Private high school (%) x Machinist (G1)	0.893*	-2.225	3.574***	0.292	4.803***	4.368***	-2.145***	0.035*
	(0.512)	(1.543)	(0.973)	(0.854)	(1.278)	(1.191)	(0.794)	(0.020)
Manufacturing emp. (%) x Machinist (G1)	0.058	0.094	0.526	-0.131	0.510	1.979**	-0.390	-0.016
	(0.491)	(0.978)	(0.637)	(1.124)	(1.088)	(0.926)	(0.477)	(0.013)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-fixed effects (1870)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	39570	9871	19742	9957	6455	13511	19742	19181
Number of clusters	45	45	45	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. None of the specifications is weighted. The sample includes all sons who were not older than ten years in 1870, and whose father held an occupation between the 24.7th and 84.7th education rank percentiles in 1870. The sample is additionally restricted to sons who lived in a county in 1870 i) with private tuition fee below the 25th percentile (Col. 2) / between the 25th and 75th percentiles (Col. 3 and 5-8) / above the 75th percentile (Col. 4); ii) with below median public high school share (male public high school students as % of 14-20 year-old males in the county in 1880 - Col. 5); iii) with an average teacher-pupil ratio above 0.04 in private schools. The outcome variable is the education rank of occupation (Col. 1-6), the share of workers who had at most primary education in the son's occupation (Col. 7) and the Sobek log-score (Col. 8). The share of private high school students and of manufacturing employment (as % of county population in 1870) are winsorized at the 99th percentile and standardized. Baseline controls are described in Appendix 3.C.2. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

preferences over the schooling of their kids (Boneva and Rauh 2018). Thus, we do not expect these (unobserved) preferences to drive the findings. Moreover, our subsequent findings are inconsistent with preferences for more schooling at every tuition fee (cost of education) level.[25]

First, the effect of private high school provision is studied by interacting the machinist main effect with the share of boys who attended high school in the the county. We assume that sons still lived in the county where they were located in 1870 when they reached high school age. In addition, only those sons are included who were not older than ten years in 1870, so that they were not too old to benefit from secondary education and reached high school age around 1880 - the year which our schooling measure corresponds to. Every specification includes an interaction with the county-level share of manufacturing employment as well, so that we can avoid that the results are driven by the known negative association between high schooling and industrialization (Goldin and Katz 1999). Both the high school provision and industrialization proxy are standardized in the full sample. This means that the coefficients can be interpreted as the effect of one standard deviation increase in the given variable.

The results are presented in Table 3.9. When the tuition fee was neither too cheap (so the main cost of schooling was the foregone wage and practically everyone could attend high school; Column 2) nor prohibitively expensive even for machinists (Column 4), the sons of machinists benefited from the increased availability of private high schools. At mean private high school provision, a machinist son had an occupation with a four percentiles higher education rank. If he instead grew up in a county with a one standard deviation lower high school provision, the entire positive effect might have vanished (Column 3). To strengthen our increased parental investment interpretation, we show that the identified positive coefficient on the interaction term is driven by counties which had low public high school provision, i.e. boys could mainly pursue secondary education at private schools as public high school provision was very limited (Column 5). Additionally, Column 6 establishes that the coefficient on the interaction term is particularly large

[25]The entire schooling data collection and preparation process is described in Appendix 3.C.2.

across counties with high-quality private secondary schools (high teacher-pupil ratio - Card and Krueger 1992; Chetty et al. 2014). The last two columns show that the complementarity between a machinist father (income effect) and local private high school provision also manifests itself for other relevant outcomes. Better private high school provision at medium tuition fee level led to machinists' sons having fewer people with at most primary education in their occupation (Song et al. 2020) and a larger increase in their earnings score.

Second, we study the effect of public secondary education supply in large cities (at least 7.500 inhabitants in 1880). In the standard model of Becker and Tomes (1986), the effect of a higher income of machinist fathers allowing their sons to stay in school longer should be diminishing in the expansion of the mostly free of charge, public high school system. This happens if parents faced restrictions to borrowing or savings and public schooling purely substituted for private schooling.[26] The test of this hypothesis is presented in Table 3.10. At mean public high school provision, an urban machinist's son had a three percentage points higher occupational education rank compared to sons of non-machinists. However, half of this relative gain was lost in counties with one standard deviation higher public high school provision (e.g., Akron, OH, Hartford, CT or Richmond, VA). Compared to these places, the gains of machinists' sons were three times larger in cities with one standard deviation below the mean (e.g., Indianapolis, IN, Jersey City, NJ or Joliet, IL). Columns 2 and 3 show that the other two outcomes of interest were influenced by expanding public secondary schools in a similar way. Column 4 establishes that the expansion of public schools particularly mattered under medium private tuition fee, in line with the previous analysis of private high schools. Exclusively urban sons were included in the estimation so far, even though Goldin and Katz (2008) write that township public schools sometimes educated the youth of the urban center as well as those of nearby rural communities. Therefore, the sample is expanded with rural sons within the county of large cities in Column 5. The interaction coefficient becomes somewhat smaller, suggesting that the effect is driven by the urban subsample who grew up in the physical proximity of schools. In Column 6, cities with more than 100.000 inhabitants are excluded from the sample which makes the interaction term even larger in magnitude.

The extent of local public high schooling was influenced by other factors than the level of industrialization (high opportunity cost of staying in school in industrialized counties) as well. Wealthier, more equal and stable communities tended to be associated with a more abundant public high school supply. Using proxies following Goldin and Katz (1999), we demonstrate that our interaction with public schooling does not capture, for instance, the beneficial effect of wealthier

[26] Goldin and Katz (2008) report that the tuition fee itself was on average 5% of the gross earnings of skilled workers. The boarding fee could double or triple the costs. The recent empirical evidence on credit constraints is discussed in the introduction.

Table 3.10: Heterogeneity by the supply of public schooling - sons (G2; 1900)

	Full urban sample			Dependent variable: education rank				
	(1) Education rank	(2) Max. primary education (% in occupation)	(3) Sobek log-score	(4) Medium tuition fee	(5) Rural and urban county population	(6) No large cities	(7) Wealth proxies	(8) Inequality and old population
Machinist (G1)	3.033***	-1.770***	0.029	4.557***	3.179***	2.724*	2.804**	1.991*
	(1.068)	(0.610)	(0.018)	(1.640)	(0.793)	(1.396)	(1.168)	(1.060)
Public high school (%) x Machinist (G1)	-1.644***	1.031***	-0.022*	-2.587***	-1.302**	-2.379***	-1.475***	-1.940***
	(0.588)	(0.359)	(0.012)	(0.743)	(0.522)	(0.769)	(0.540)	(0.526)
Manufacturing emp. (%) x Machinist (G1)	0.534	-0.183	0.011	-0.296	0.703	0.438	0.819	-0.822
	(0.559)	(0.303)	(0.012)	(0.783)	(0.511)	(0.826)	(0.674)	(0.972)
Agricultural production per agric. worker x Machinist (G1)							2.234	
							(1.496)	
Wealth per capita x Machinist (G1)							-0.613	
							(1.061)	
Top 1% share of wealth x Machinist (G1)								4.781**
								(1.863)
Elderly population (%; above 65 y.o.) x Machinist (G1)								2.239*
								(1.317)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-fixed effects (1870)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	19393	19393	18767	9901	26555	12424	19393	19393
Number of clusters	45	45	45	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. None of the specifications is weighted. The sample includes all sons who were not older than ten years in 1870, whose father held an occupation between the 24.7th and 84.7th education rank percentiles in 1870, and who lived in places classified as urban in 1870. Additionally, the sample is restricted to sons in counties with medium tuition fee (Column 4 - see Table 3.9), is expanded to include rural sons as well within a county (Column 5), and does not include sons in cities with more than 100.000 inhabitants in 1870 (Column 6). The outcome variable is the education rank of occupation (Col. 1 and 4-8), the share of workers who had at most primary education in the son's occupation (Col. 2) and the Sobek log-score (Col. 3). The share of public high school students (as % of 14-20 years old males in the county in 1880), the share of manufacturing employment (as % of county population in 1870), the agricultural production per agricultural worker (the estimated yearly, county-level agricultural production is from Manson et al. (2021), while the number of agricultural workers is from the 1870 full count census - agricultural workers are farmers, farm managers, foremen and laborers), the wealth per capita (calculated as the total wealth in a county - the sum of real estates and personal property - divided by county population in 1870), the top 1% share of wealth (calculated as the county-level wealth - sum of real estate and personal property - share of the richest one percent in 1870; only males who were above 16 years old), and share of elderly people (share of people older than 65 years in the 1870 full count census) are winsorized at the 99th percentile and standardized. Baseline controls are described in Appendix 3.C.2. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

residents who, in turn, were willing to invest in public schools. In Column 7, two wealth proxies are included, but the coefficient of interest remains practically unaffected. We use the wealth share of the top 1% of residents as a proxy for wealth inequality and the share of elderly people to capture the stability of the local community in Column 8. Interestingly, the machinist effect seems to increase in local wealth inequality. We suspect that this this effect is attributable to the presence of wealthy factory owners who utilized modern, mechanized production methods in their establishments, requiring the intensive involvement of machinists. Alternatively, intergenerational mobility might have simply been lower in counties with higher concentration of wealth (Chetty and Hendren 2018c; Chetty et al. 2014), which could make the catch-up of non-machinists' sons more difficult. Nonetheless, the interaction with public high schooling becomes even more negative in this column.[27] We conclude that the provision of public high schooling could dampen the difference between the offspring of machinists and non-machinists, in line with Becker and Tomes (1986). This result also suggests that machinist families did not have particularly strong preferences for education, since otherwise they could have sent their sons to college using money saved from substituting private with free public high school education or, simply, let their sons stay in public high school longer.

Third, being able to decipher blueprints, having some elementary knowledge of algebra or chemistry, and mechanical drawing skills were all valuable on the labor market in the late nineteenth century (Goldin and Katz 2000; 2008). The sons

[27] The highly significant interaction term in Column 3 of Table 3.9 also survives the inclusion of these control interactions.

Table 3.11: Information channel - sons (G2; 1900)

	Full sample	Low tuition fee (below city median)	Low tuition fee & population ¿ 5.000 (city in 1870)		
	(1)	(2)	(3)	(4)	(5)
	Education rank	Education rank	Education rank	Max. primary education (% in occupation)	Sobek log-score
Machinist (G1)	4.502***	5.691***	3.846**	-2.277**	0.059*
	(0.787)	(0.894)	(1.449)	(0.979)	(0.032)
Technical education (% of HS students) x Machinist (G1)	-0.643	-1.852**	-3.489**	1.616*	-0.049**
	(0.652)	(0.781)	(1.458)	(0.945)	(0.019)
Manufacturing emp. (%) x Machinist (G1)	-0.065	-0.621	0.862	-0.479	0.013
	(0.539)	(0.888)	(1.499)	(0.868)	(0.028)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County-fixed effects (1870)	Yes	Yes	Yes	Yes	Yes
Sample size	34475	20746	8327	8327	8070
Number of clusters	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. None of the specifications is weighted. The sample includes all sons who were not older than ten years in 1870 and whose father held an occupation between the 24.7th and 84.7th education rank percentiles in 1870. The sample is additionally restricted to sons who lived in 1870 i) in a county with private tuition fee below the median of cities (places with more than 5.000 inhabitants in 1870 - Col. 2-5) and ii) in places with more than 5.000 inhabitants in 1870 (Col. 5). The outcome variable is the education rank of occupation (Col. 1-3), the share of workers who had at most primary education in the son's occupation (Col. 4) and the Sobek log-score (Col. 5). The share of technical education (% of private high school students - institutions for secondary instruction or preparatory schools - whose school had a chemical laboratory or taught mechanical drawing) and of manufacturing employment (as % of county population in 1870) are winsorized at the 99th percentile and standardized. Baseline controls are described in Appendix 3.C.2. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of machinists could easily learn many of these skills from their fathers, thereby gaining some advantage outside formal schooling - which we call the information channel. However, schools increasingly started to incorporate scientific subjects into their curriculum which may have decreased the benefits of machinists' sons. To test this hypothesis, we use the Reports of the Commissioner of Education. These volumes contain relevant information - if the given school taught mechanical drawing or had a chemical laboratory - on two types of private high schools: institutions for secondary instruction and preparatory schools. We calculate the share of high school students whose school replied with a yes to any of the two questions. The underlying assumption is that these institutions put an emphasis on technical education in their curriculum. The interaction between technical education at school and a machinist father is estimated in Table 3.11. In line with our hypothesis, offering technical education decreased the relative gains accruing to machinists' sons, but only if private high schools were accessible to "rival" boys too (relatively low tuition fee; Column 2). Moreover, this effect is particularly strong in cities, where the benefits of technical skills could be reaped in manufacturing production, as opposed to rural areas and is also present for other potential outcomes (Columns 3-5).

A decomposition of gains in earnings A simple goal is set in this subsection: understanding and quantifying to what extent the earnings effects of machinists' sons are driven by rural-urban differences and education.

First, we split the nominal earnings effect between sons who resided in villages (settlements with less than 5.000 inhabitants) and in cities in 1870. The first column of Table 3.12 shows that rural machinists' sons had significantly larger earnings gains: they had on average 4.7 log-points higher nominal earnings scores relative to city-dweller machinists' sons. A possible explanation is that the sons of rural machinists, being more educated than their peers, migrated to urban places more intensively or they migrated with their father to these areas and, thereby, had access to better paying urban occupations (see Tables 3.3 and C2). Machinists' sons in cities, on the other hand, could have experienced a relative urban premium only if initially urban non-machinists' sons would have left cities for rural areas.[28]

In line with the previous interpretation, Column 2 shows that the entire higher probability of urban place of living effect can be attributed to sons of initially rural machinists. We can use the differential probability between the offspring of rural and city-dweller machinists to calculate the differential earnings effect which can be explained by rural-urban earnings differences. Ward (forthcoming) estimates that rural-to-urban migration led to a 30 log-point increase in the log-earnings score in the early-twentieth-century United States. Assuming that this figure accurately describes the average gains of machinists' sons derived from rural-to-urban migration, we can conclude that the majority of the 4.7 log-points difference can be explained by the differential relative probability of urban status $(2.6 = 0.085 \cdot 30)$.

In Column 3, the sample is restricted to villagers' sons who lived in counties with medium level tuition fee (see Table 3.9). In line with anecdotal evidence in Goldin and Katz (2000) and Goldin and Katz (2008), we find that the (secondary) education of rural sons was indeed the pathway to urban life. At mean private high school provision, the son of a villager machinist was 12% more likely to live in an urban place three decades later than a comparable non-machinist's son. However, this effect increases by 70% when private high school provision increases by one standard deviation. We believe that this result lends support to the interpretation that machinists' sons ended up in urban places at least partly because they were more educated.

Second, we want to understand to what extent the rest of the machinist effect $(2.5 = 7.2 - 4.7)$ can be explained by returns to education. In unreported results, we establish that less than the half of the 0.21-year-longer schooling (see Column 1 in Table C5) stemmed from longer primary schooling, while the majority was the result of secondary schooling. Taking the returns to schooling estimates of Goldin and Katz (2000), we calculate that two-thirds $(1.7 = return\ to\ high\ school + return\ to\ primary\ school = 10.3\% \cdot 0.13 + 4.8\% \cdot 0.08)$ of the remaining machinist effect was the result of more years of schooling.[29] Considering that Goldin and Katz (2000) argue that their returns estimated in Iowa (1915) might be a lower bound on returns to education and that our estimated 0.21-year-longer schooling might be a lower bound too (owing to endogenous attrition), we can attribute

[28]Additionally, the magnitude of earnings losses from urban-to-rural migration was significantly smaller than gains from rural-to-urban migration (Ward forthcoming).

[29]We cannot analyze a heterogeneous years of schooling effect by initial urban status owing to the small number of sons in the 1940 sample.

Table 3.12: The urban-rural gap in the earnings effect (G2; 1900)

	State-level nominal log-score (1892)	Urban place of living (Yes=1)	
	(1)	(2)	(3)
	Full sample	Full sample	Villagers & medium tuition fee
Machinist (G1)	0.072***	0.099***	0.124***
	(0.019)	(0.023)	(0.032)
City (1870) x Machinist (G1)	-0.047*	-0.085***	
	(0.025)	(0.030)	
Private high school (%) x Machinist (G1)			0.088***
			(0.026)
Manufacturing emp. (%) x Machinist (G1)			-0.037
			(0.024)
Baseline controls	Yes	Yes	Yes
County-fixed effects (1870)	Yes	Yes	Yes
Sample size	45605	45605	6542
Number of clusters	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. None of the specifications is weighted. The sample includes all sons who were not older than ten years in 1870 and whose father held an occupation between the 24.7th and 84.7th education rank percentiles in 1870. The sample is additionally restricted to sons who lived in a place with less than 5.000 inhabitants and in a county with private tuition fee between the 25th and 75th percentiles in 1870 (see Table 3.9). The outcome variable is state-level nominal log-score (Col. 1) and an indicator for an urban place of living in 1900 (Col. 2-3). The specifications in Columns 1-2 also include a city indicator which equals to one if a son lived in a place with more than 5.000 inhabitants in 1870. Baseline controls are described in Appendix 3.C.2. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

practically the entire remaining earnings effect to returns to schooling.[30]

3.5.3. Fathers in other demanded occupations

The main results section is closed by looking at the sons of fathers in other occupations which were already present around 1870 and also received a boost from technological innovations during the Second Industrial Revolution (see Mokyr 1999).

The first such occupational group contains fathers who were chemists, engineers (mainly civil or mechanical), or telegraph operators - all white-collar jobs. Column 2 in Table 3.13 shows that their sons might have experienced even larger benefits, as measured by the education rank of occupation, than the sons of machinists. The conclusion is similar for the point estimate of the log-earnings score (Column 6), though this coefficient is imprecisely estimated, potentially owing to the small sample size.

Subsequent columns investigate the effect on the sons of two other, relatively

[30]The returns to education of Goldin and Katz (2000) combine within and across occupations gains, whereas earnings scores-based estimates can exclusively capture the latter. The estimates of Feigenbaum and Tan (2020) - those based on income scores measured before the Great Compression (Goldin and Margo 1992) - indicate that 60-70% of the effect of a year of education on individual wages is captured in the effect on occupational earnings scores (4.4% vs 2.6-3.1%; see Tables 7 and A.9 of Feigenbaum and Tan 2020).

Table 3.13: Sons of fathers in other occupations (G2; 1900)

	Education rank				Sobek log-score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Machinist (G1)	2.403***				0.041***			
	(0.745)				(0.013)			
White-collar occupation boosted by the Second Ind. Rev. (G1)		4.959***				0.055		
		(1.757)				(0.038)		
Employee of railways (G1)			-0.300				0.030	
			(1.531)				(0.021)	
Metal industry operative (G1)				-0.517				-0.051
				(1.539)				(0.040)
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	8745	1643	3317	2893	8424	1448	3020	1453
Number of clusters	45	45	43	43	45	42	44	41

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. All specifications are weighted by weights obtained from propensity score matching described in the main text. The final sample includes 1842, 200, 763, 665, 2779, 175, 746, 628 (Col. 1-8, respectively) matched sons of machinists. To mitigate the imprecision caused by the small number of treated observations, ten controls are chosen for the sons of white-collar workers instead of the usual five. The outcome variable is the education rank of occupation (Col. 1-4) and the Sobek log-score (Col. 5-8). Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. The Sobek score merged to the occupation of fathers (1870) is also included in the propensity score matching in Columns 5-8. White-collar occupations boosted by the Second Industrial Revolution are: chemists, engineers (IPUMS's harmonized *OCC1950* code between 41 and 49), and telegraph and telephone operators. Railway employees are: brakemen, locomotive engineers, locomotive firemen and switchmen. Metal industry operatives are: filers, furnacemen, heaters, grinders, polishers and smeltermen. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

lower-skilled groups of workers: employees of the railways (for instance, locomotive engineers or firemen) and operatives of the metal industry (smeltermen, heaters, etc.). Interestingly, we do not find evidence on any significant effect on their sons using our baseline matching estimation. While the explanation of the missing effect is beyond the scope of this work, we suspect that the labor market competition stemming from masses of low-skilled, European immigrants might have affected these lower-skilled workers more severely. Thus, the labor supply could more easily match the rising demand in these occupations.

3.6. Robustness checks

We discuss the robustness of our main findings below, implementing modifications in our baseline matching or regression estimations. In most cases, we concentrate on the effect on the two crucial outcomes of sons for the sake of brevity (the education rank and urban place of living indicator in 1900). However, we deviate from these outcomes in a few specifications owing to sample size considerations and, instead, look at some outcomes of fathers.

3.6.1. Robustness checks using matching

Occupational employment pre-trends Facing occupational choice in their teenager years, fathers could elicit information about the future of certain occupations from their employment growth. For instance, the employment share of sailors was on a constant decline after the spread of steamships, indicating a gloomy future for prospective sailors. If machinists followed a relatively faster employment growth path compared to baseline control occupations, the identified positive effects could be the result of better foresight (and correlated talent) of machinist fathers, or simply the result of pre-trends leading to better occupation-level outcomes even in the absence of the Second Industrial Revolution. To assess this potential bias, we constructed the changes in the employment share of occupations at the census division level for the 1850s and 1860s (see Appendix 3.C.2). These two measures are also included in the propensity score matching implemented in Columns 1 and 2 of Table C8. In comparison with Columns 2 and 5 of Table 3.5, both coefficients (insignificantly) increase in magnitude. Consequently, differential employment growth trends in the decades when fathers chose their occupation cannot explain our findings.

Manufacturing control occupations One might be concerned that workers of the manufacturing sector might have been more open-minded to modernity than people employed in more traditional sectors and predisposed to benefit from the overarching industrial and urban transformation in the late-nineteenth-century US. If this was the case, our matching estimation would be upward biased as the baseline control group contains many workers outside of manufacturing as well (e.g., carpenters or teamsters). Therefore, the sample is restricted to fathers who were employed in durable or non-durable manufacturing in Columns 3 and 4 in Table C8. However, this restriction causes no meaningful change in the coefficients of interest.

Maternal observables As more than 95% of mothers were not active on the labor market in our sample in 1870, we cannot use occupation-based measures of their socio-economic status. However, next to maternal age, we constructed an indicator variable if the mother was native born and if she was literate. Our baseline matching strategy in Section 3.5 balances our sample on maternal age and nativity even without including them as controls, but it is significantly more likely that a machinist's son had a more literate mother (2.6 percentage points difference - which amounts to a 9.6% standardized difference). We assess if more educated mothers drive our results in Columns 5 and 6 (Table C8), where we match on the three maternal observables as well. Once again, our findings are not affected by this change in the baseline specification.

Influential control occupations A particular concern could be an influential role played by the largest control occupations (Table C1). The interpretation of our findings would be profoundly different if the results were driven by a certain small group of control occupations. Therefore, we exclude fathers employed in the three largest control occupations - exclusively these three have a larger than five percent share among matched controls - from the pool of potential control individuals in last two columns of Table C8. The omission of these occupations, which provide approximately one-quarter of the control individuals in the baseline matching, does not influence the results in any significant way.[31]

The role of next-door neighbors The important effect of the neighborhood where kids grow up is well-established both in current and historical US context (see e.g., Abramitzky et al. 2021; Chetty and Hendren 2018b; c; Chetty et al. 2014; Durlauf 2004; Galster 2012; Ward 2020). In our (subsequent) regression analysis, at most county-fixed effects can be included to capture the effect of growing up in the same neighborhood. However, within-county residential segregation along ethnic (Eriksson and Ward 2019) or other socio-economic lines calls into question whether neighborhoods should be defined at the county-level.

To demonstrate that our results are not driven by machinists residing in more prosperous neighborhoods, we exploit the fact that next-door neighbors can be identified in the full count census. We construct the average value of personal property, real estate value, occupational education rank, literacy and foreign-born status of the closest household heads in 1870 (see Appendix 3.C.2 for details). Reassuringly, our baseline matching strategy balances on these initially significantly different characteristics even without their inclusion (e.g., in the estimation in Table 3.5 - not reported). Thus, machinists tend to have very similar neighbors compared to matched control observations. Therefore, we believe that omitted differences in neighborhood quality cannot drive the findings.

The role of grandparental (G0) characteristics Grandfathers (G0) could influence our results and their interpretation in many ways. For instance, grandfathers with better foresight could nudge fathers to choose an occupation that was expected to be prosperous or to leave agriculture. Additionally, if machinists had significantly richer or more educated parents, this could introduce a more mundane form of omitted variable bias into the empirical analysis. All these reasons make the linking of fathers (G1) to their fathers (grandfathers; G0) important. In the resulting sample, we can assess the difference in coefficients with and without controlling for a large number of grandparental observables measured in 1860. Before doing so, we acknowledge that our sample might be selected since the parents of most foreign-born individuals did not live in the United States and some grandfathers

[31]While the omission of the largest control occupations does not matter for our results, if control occupations experienced an employment decline or rise in 1870-1900 does matter. Restricting control occupations only to those which experienced an increasing (decreasing) employment share in these decades would result in different coefficients: 1.6 (4.3) for the educational rank and 0.034 (0.084) for the urban status indicator. In our baseline matching strategy, the average employment share change of the matched control group is approximately zero (unreported results).

might have died before 1860. However, the resemblance of coefficients estimated in Tables 3.3 and C9 suggests that the degree of this selection is not severe.[32]

First, we investigate how well the baseline matching strategy performs *without* explicitly balancing the sample on grandparental observables. The fact that the age, wealth (both real estate and personal property), urban status, population of place of living, and steel and iron industry dummy of grandfathers are significantly different before, but not significantly different after matching lends credibility to our estimation strategy. Furthermore, even when the difference cannot be eliminated in the case of certain remaining variables, it shrinks substantially. For instance, non-machinist grandfathers are fifteen percentage points more likely to have an agricultural occupation initially. This gap is reduced to five percentage points with a p-value of 1%. Nonetheless, there are several variables which are still highly significantly different, the most prominent one being the indicator variable of a machinist grandfather.

Second, Table C9 reports the results with and without controlling for grandparental characteristics (see the notes below the table for the full list). It can be observed that the inclusion of these G0 background variables in the matching procedure does not change the results. Consequently, we can conclude that the main findings are not driven by grandparental observables.

3.6.2. Robustness checks using regressions

As a validation step before presenting the full set of robustness checks with fixed effects regressions, the baseline results for sons are estimated using these regressions instead of matching. The comparison of Tables 3.5 and 3.6 to Tables C6 and C7 reveals that the two estimation methods produce very similar coefficients which are not significantly different from each other.

Spatial sorting before 1870 Even though we can include county-fixed effects in our regressions, individuals who resided in a certain county in 1870 might have still been different in their migration history. Ideally, people who were born in a given county should not be compared to people who migrated there. Since the Second Industrial Revolution does not have a well-defined starting date, it could be the case that, when only county-fixed (1870) effects are used, in-migrated machinists with a good instinct to spot places with a growth potential are compared to locals who happened to be born there.[33]

[32]The sole qualitatively different result is long-distance migration. Unlike the baseline analysis, where it is insignificantly positive, the coefficient becomes significantly positive at 5% in the new sample.

[33]Klein and Crafts (2020) argue that in the early-twentieth-century United States «*technological progress accelerated at this time but its progress was quite erratic and the development of new technologies and industrial locations was unpredictable.*»

Therefore, more detailed fixed effects are specified to tackle the possible spatial sorting prior to 1870. To do so, we generate fixed effects combining state of birth (country of birth for the foreign-born), county of living in 1870, an urban status indicator in 1870, and an indicator variable for above median age of the father. For instance, if a 28 year-old machinist was born in South Carolina, but then moved to the rural part of Erie county (NY), we are going to compare him to individuals with exactly the same migration history and below median age. Consequently, we will cease to compare individuals to all other locals in 1870. The underlying assumption is that individuals sharing the same migration history had very similar information and keenness to migrate. While the coefficients in Table C10 (Columns 1-2) somewhat decrease compared to Table C6, a large part of this insignificant difference is attributable to a slightly different, reduced sample.[34] This sample size reduction is the result of our narrowly defined fixed effects as we lose observations in less densely populated, rural areas or with a peculiar migration history. Finally, we can conclude that spatial sorting preceding the 1870s does not drive our results.

Additional state-occupation level pre-trends Our baseline matching strategy contains merely two occupation-state level characteristics (probability of migration and occupation change in the 1860s) because the matching algorithm would not convergence if many more were added. This limitation is simply the result of the occupation-based "treatment". However, many other similar variables can be included in fixed effects regressions. To this end, we calculated the two aforementioned variables for the 1850s, and added the average change in the urban status indicator and the probability of switching to an agricultural occupation for every occupation in the 1850s and 1860s (see Appendix 3.C.2). The absence of any significant change after the inclusion of these control variables in Table C10 shows that the results are not outcomes of spatially-varying, occupation-level pre-trends.

Weighting for a representative sample The implementation of propensity score matching does not allow us to use any kind of weights. However, it is a well-known issue in the literature using the full count census that linking across different census waves might engender a non-representative sample. Therefore, we calculated the widely used inverse proportional weights to make the sample representative of the US population around 1870 (see Appendix 3.C.3 for the details), then applied them in Columns 5-6 of Table C10. One can clearly see that our regression estimation without weighting produces coefficients very close to these new estimates. Therefore, we believe that our results accurately reflect the US population at the onset of the Second the Industrial Revolution.

[34]Results with the new sample but without the new fixed effects are available upon request.

Restricting the set of control occupations The baseline regression estimation includes all fathers whose occupation is above the 25^{th} but below the 85^{th} educational rank percentile. In the last robustness exercise reported in Table C10, we further restrict the sample of fathers to the 45^{th}-65^{th} educational rank percentiles. No significant change ensues aside from a slight drop in the coefficients.

Old and young fathers/sons Before occupations start to grow rapidly or are about to decline, there is much uncertainty about their future. More forward-looking and able individuals might have anticipated the eventual rise of machinists and took up this occupation early on. This type of sorting would imply that more positive effects should be observed for the sons of older machinists. Table C12 presents a comparison of the effect on sons depending on the age of the father. The age of sons is restricted between 0 and 5 in 1870 because otherwise older fathers have substantially older kids who, in turn, grew up in different years. Reassuringly, we do not find any significant difference between the sons of older and younger machinists when the sample is split by the age of the median machinist father.

Dynamic complementarity in the production of human capital is a well-established finding in the literature of education economics (see Caucutt and Lochner 2020; J. J. Heckman and Cunha 2007; Lee and Seshadri 2019). This implies that those sons of machinists who were relatively old in 1870 should have experienced a relatively smaller increase in their level of education compared to the younger ones because they lacked complementary education investments during their early childhood. We investigate this question in the last column of Table C12. Confirming the theoretical prediction, machinists' sons who were older than ten years around the onset of the Second Industrial Revolution did not enjoy any gains in education (proxied by the education rank) in comparison with sons of similar, non-machinist workers.

3.6.3. Grandfather-fixed effects

Our arguably most important robustness checks are regressions in which grandfather-fixed effects are included. In other words, we compare machinists to their non-machinist brother(s). In this way, we can eliminate concerns related to machinists growing up in more advantaged families (unobservables not captured by the job, place of living or wealth of the grandfather) or inheriting a particular genetics, which helps them succeed in life (see Mogstad and Torsvik (2021) for a recent survey on this topic). To eliminate within-family differences in talent across siblings, we still control for many of their personal characteristics in 1870: county of living, education rank, literacy or wealth. Our regression specification thus takes the following form:

Table 3.14: Within-family estimation - fathers (G1; 1870-1900)

	Occupational change (Yes=1)				Other outcomes		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Baseline	20-40 y.o.	Unrestricted sample	Urban in 1900 (Yes=1)	Migration (within-state; Yes=1)	Migration (across states; Yes=1)
Machinist (G1)	-0.107***	-0.138**	-0.133*	-0.204***	0.116**	-0.027	0.066
	(0.037)	(0.062)	(0.079)	(0.041)	(0.050)	(0.043)	(0.049)
Grandfather (G0)-fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
County-fixed effects (1870)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	22799	22799	17793	69221	22799	22799	22799
R^2	0.24	0.64	0.65	0.67	0.72	0.62	0.65

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are multiway clustered at the grandfather-county (1900) level. None of the specifications is weighted. The sample includes all fathers who held an occupation between the 24.7th and 84.7th education rank percentiles in 1870 - except for Column 4 which includes all fathers irrespective of the education rank of their occupation. In every column, the age of included fathers is between 16 and 50 years (inclusive) - except for Column 3 where the age is restricted between 20 and 40 years (inclusive). The outcome variable is a binary variable which equals one if i) the father changed occupation between 1870 and 1900 (Col. 1-4), ii) the father lived in an urban place in 1900 (Col. 5); iii) the father migrated within-state across counties (Col. 6) or across states (Col. 7). Personal controls included in the regressions are (all measured in 1870): the education rank of occupation, urban status and literacy indicator, age (in years), value of real estate and personal property, number of inhabitants in the place of living and a farmer-farm manager-farm foreman indicator. The interactions of the urban indicator, size of place of living, two wealth measures, education rank and age are also included. The squared size of place of living, wealth measures and age are included as well. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

$$y_{f,c,g,1900} = \beta \cdot Machinist_{f,1870} + \gamma \cdot x'_{f,1870} + \delta_{c,1870} + \kappa_g + \epsilon_{f,c,g,1900} \quad (3.2)$$

where the fixed effect for grandfather g of father f appears as a new control variable (κ_g). Unfortunately, we can only apply this estimation strategy for fathers' outcomes due to sample size limitations. Moreover, the baseline sample must be extended in two ways even for fathers. First, we include all fathers who were between 16 and 50 years old (originally 20-40). Second, loosely defined occupations are not omitted anymore (e.g., *Clerical and kindred workers (n.e.c.)*). Nevertheless, the baseline regression sample restriction is still implemented and we only include fathers whose occupation had an occupational education rank between 25$^{\text{th}}$ and 85$^{\text{th}}$ percentiles, thereby excluding farm laborers, fishermen but even high-skilled individuals such as bookkeepers or physicians.

The results of this estimation are shown in Table 3.14. The comparison of Columns 1 and 2 shows that the inclusion of grandfather-fixed effects does not significantly change the coefficient of interest in spite of a forty percentage-point increase in the R^2. This suggests that machinist fathers were significantly less likely to change their occupation even compared to their non-machinist brothers. In Column 3, the age of fathers is restricted to the original 20-40 range. The point estimate is practically unchanged but less precisely estimated owing to the sample size reduction. Next, we include all brothers irrespective of their education rank in Column 4. This produces an even larger coefficient than the initially estimated one in Column 2. Other outcomes of fathers are presented in Columns 5-7. The same conclusion can be drawn quantitatively and qualitatively as before (see Tables 3.3 and C9): a substantial positive likelihood of living in an urban place and (if anything) a positive probability to migrate across states.

The within-family estimation can greatly reduce the role of certain confounding unobservables, but it cannot entirely eliminate differences stemming from the different ability of brothers. In our previous analysis, we already made two steps to reduce their role. First, analogously to Feigenbaum and Tan (2020), who restrict their sample to small years of education differences between twins, brothers holding occupations with the lowest and highest education ranks were excluded. The underlying assumption is that brothers with more similar education ranks are more likely to be similar in terms of unobservables as well. Second, the included personal characteristics (for instance, the two wealth measures, the education rank of occupation or literacy dummy) should already capture a certain degree of differences in ability. To further reduce the likelihood that the results are driven by unobserved ability, we borrow from the literature which estimates returns to schooling using twins (e.g., Ashenfelter and Rouse 1998; Feigenbaum and Tan 2020). They argue that some observable variables - marriage status,[35] spousal education, number of kids, etc. - are correlated with ability. In Table C11, we demonstrate that none of these variables are correlated with the machinist dummy. Perhaps even more importantly, specifications *without* grandfather-fixed effects show no significant association either.

3.6.4. Correcting measurement error and magnitude comparison

It is well-known that the misreporting of binary independent variables produces a non-classical measurement error in regression estimations because the measurement error is mechanically negatively correlated with the correctly measured value (see e.g., Aigner 1973; Bingley and Martinello 2017; Dupraz and Ferrara 2021). Consequently, the OLS estimate is a lower bound on the consistent coefficient normally. The relationship between the correct coefficient and inconsistent OLS estimate is the following:

$$plim\ \hat{\beta}_{OLS} = \beta \cdot (1 - p - q) \tag{3.3}$$

where β is the consistent coefficient, p is the share of false positives (among fathers classified as machinists, $p\%$ were incorrectly classified as one), and q is the share of false negatives (among fathers classified as non-machinists, $q\%$ were actually machinists).

In our case, q can be set equal to zero owing to the small share of machinists in the whole sample. A non-negligible p can be the result of two, distinct measurement errors. First, a machinist observation might be linked to a non-machinist one when we link across census waves. For conservative linking methods used in this paper, Bailey et al. (2020) estimate a false positive ratio of 10-15%. Second, even

[35] In the absence of a separate census question on marriage status in 1870, a father is imputed to be married if the age of the spouse is known.

if we could perfectly link individuals to their own observations over time, the misreporting of occupations can cause measurement error. Ward (2019) shows that around one-third of respondents misreported their occupation in the full count census, relying on a census re-enumeration in Saint Louis in 1880.[36] Therefore, we believe that assuming $p \approx 40\%$ might capture the true extent of false positives.

Using the previously introduced formula, one can see that the OLS coefficient is assumed to be downward biased by a factor of 0.6 (=1-0.4). Under this assumption, the consistently estimated effects are around 66.67% larger than the earlier OLS estimates. This implies that a machinist's son had on average a four percentiles higher education rank (Col. 2 of Table 3.5), 0.35 years more of schooling (Col. 1 of Table C5), and a seven log-points higher nominal earnings score (Col. 2 of Table 3.6) than a son of a comparable but non-machinist father.

3.7. Conclusion

In this paper, we investigate to what extent and how winners of structural transformations can transmit their gains in socio-economic status to their offspring. Combining full count census data with newly digitized data sources, we establish that machinists, whose occupation experienced a relative labor demand spike in the United States during the Second Industrial Revolution, experienced relatively higher income and job stability. Relying on propensity score matching and fixed effects regressions, we document that the (grand)sons of machinists were significantly better-off in terms of earnings-related outcomes than (grand)sons of observationally similar non-machinists. In addition, the main contribution of this work is pinning down the mechanism which underlies the documented intergenerational transmission. We find that the sons of rural machinists benefited from rural-to-urban migration and parental investment in their education, while the sons of urban machinists mostly gained from the latter channel. A wide range of robustness checks show that the results are unlikely to be driven by the (transmitted) unobserved ability of machinist fathers.

In conclusion, the main mechanisms behind intergenerational mobility seem to have changed little over more than a century: the opportunities offered by high-quality urban neighborhoods (see Chetty and Hendren 2018b; c; Chetty et al. 2014; Durlauf 2004; Galster 2012; Laliberté 2021b) and by high educational attainment guarantee a higher socio-economic status in the age of telegraphs as well as of smartphones. We also show that expanding public schools could equally well reduce inequality stemming from financially constrained parents in the past as nowadays (Dobbie and Fryer 2011; Duflo 2001; Lucas and Mbiti 2012; C. A.

[36] If a reported machinist was more than 66.67% likely to actually hold the machinist occupation, the magnitude of the adjustment factor declines along with p.

Neilson and Zimmerman 2014; Wantchekon et al. 2015). Taken together, our results suggest that the effects of current transformations in the labor market, such as automation, might be passed on to later generations, but to a lesser extent due to today's considerably more expanded public education and unemployment benefit system (allowing for less occupational downgrading) - especially if people are allowed to move to places offering better economic prospects.

3.C. Appendix

3.C.1. Additional tables

Table C1: Top control occupations

Top control occupations (*OCC1950*)	**% of all control observations**
Carpenters	9,92%
Truck and tractor drivers	7,68%
Shoemakers	6,16%
Painters (construction)	4,17%
Blacksmiths	4,15%
Masons	3,34%
Hucksters and peddlers	2,94%
Stationary engineers	2,91%
Tailors	2,58%
Molders (metal)	2,36%
Bookkeepers	2,24%
Compositors and typesetters	2,10%
Meat cutters	1,99%
Stone cutters	1,86%
Clergymen	1,64%

Note: the results presented in this table pertain to the propensity score matching in Table 3.5.

Table C2: Migration destination decomposition - fathers (G1; 1900)

	Migration (within and across states) [(1)=(2)+(3)]			Urban destination [(2)=(4)+(5)]		Rural destination [(3)=(6)+(7)]	
	(1) Any destination	(2) Urban destination	(3) Rural destination	(4) Urban in 1870	(5) Rural in 1870	(6) Urban in 1870	(7) Rural in 1870
Machinist (G1)	0.017	0.037***	-0.020***	0.016	0.021***	-0.013*	-0.007**
	(0.011)	(0.009)	(0.007)	(0.011)	(0.004)	(0.006)	(0.003)
Mean of outcome	0.58	0.28	0.30	0.15	0.13	0.13	0.17
Standard deviation of outcome	0.49	0.45	0.46	0.36	0.33	0.33	0.38
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	18811	18811	18811	18811	18811	18811	18811
Number of clusters	50	50	50	50	50	50	50

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. All specifications are weighted by weights gained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The outcome variable is a binary variable which equals one if the father migrated between 1870 and 1900 (across or within states; Col. 1), if he migrated and was found in an urban (Col. 2) or rural (Col. 3) place of living in 1900, if he migrated to an urban destination by 1900 and lived in an urban (Col. 4) or rural (Col. 5) place of living in 1870, if he migrated to a rural destination by 1900 and lived in an urban (Col. 5) or rural (Col. 6) place of living in 1870. Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C3: Measures of economic status - fathers (G1; 1880)

	(1) State-level nominal log-score (1880)	(2) State-level real log-score (1880)
Machinist (G1)	0.065**	0.057**
	(0.025)	(0.023)
Mean of outcome	6.07	6.20
Standard deviation of outcome	0.42	0.42
Unbalanced controls	Yes	Yes
Sample size	19120	19120
Number of clusters	47	47

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1880) level. All specifications are weighted by weights gained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 4428 matched machinist fathers. The outcome variable is the state-level nominal and real log-score (Col. 1 and 2). Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C4: Measures of occupational income and mobility - occupation switcher fathers (G1; 1870-1900)

	Earnings scores (state-occupation in 1900)					Occupational mobility measures (1900)	
	(1) Sobek log-score	(2) State-level nominal log-score (1892)	(3) State-level real log-score (1892)	(4) Higher-paying occupation (+$150 or more; Yes=1)	(5) Lower-paying occupation (-$150 or less; Yes=1)	(6) Manager/official/proprietor (Yes=1)	(7) Siegel's prestige log-score
Machinist (G1)	0.036**	0.032**	0.026**	0.004	-0.060***	-0.015	0.028**
	(0.014)	(0.012)	(0.013)	(0.018)	(0.016)	(0.010)	(0.012)
Mean of outcome	6.07	6.17	6.30	0.22	0.34	0.13	3.56
Standard deviation of outcome	0.60	0.45	0.43	0.42	0.47	0.33	0.36
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	7337	7337	7337	7337	7337	7337	7337
Number of clusters	47	47	47	47	47	47	47

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. All specifications are weighted by weights obtained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 1578 matched machinist fathers and non-machinist fathers who did not hold the same occupation in 1870 and 1900. The outcome variable is the Sobek, state-level nominal and real log-score (Col. 1-3), an indicator variable if the state-level real log-score was at least $150 higher (Col. 6) or lower (Col. 5) in 1900 than in 1870, an indicator variable if the father held a managerial/proprietor occupation (Col. 6; *OCC1950* code=290), and Siegel's prestige log-score (Col. 7). Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C5: Measures of education and wealth - sons (G2; 1940)

	(1) Highest grade completed	(2) Some primary education (years ¡ 9; Yes=1)	(3) Some secondary education (9 ¡= years ¡=12; Yes=1)	(4) Some university education (12 ¡ years; Yes=1)	(5) Owned a house (Yes=1)
Machinist (G1)	0.209**	-0.032**	0.036***	-0.004	0.018
	(0.091)	(0.012)	(0.010)	(0.009)	(0.019)
Mean of outcome	7.91	0.75	0.17	0.08	0.68
Standard deviation of outcome	3.42	0.44	0.38	0.28	0.47
Unbalanced controls	Yes	Yes	Yes	Yes	Yes
Sample size	7543	7543	7543	7543	7543
Number of clusters	49	49	49	49	49

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1940) level. All specifications are weighted by weights gained from propensity score matching described in the main text. The summary statistics reported are unweighted and pertain to the full estimation sample before matching. The final sample includes 919 matched machinist sons. We use ten matched control observations instead of five owing to the small number of machinist sons. The outcome variable is the highest grade of schooling completed (Col. 1 - winsorized at the 99th percentile), a binary variable which equals one if i) the years of schooling is below nine years (Col. 2), ii) the years of schooling is between nine and twelve years (Col. 3), or iii) the years of schooling is more than twelve years (Col. 4), and an indicator variable for house ownership (Col. 5). Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C6: Main outcomes - sons (G2; 1900)

	Occupational characteristics		Migration (Yes=1)		Place of living (1900)			Personal characteristics		
	(1) Agricultural occ. (Yes=1)	(2) Education rank	(3) Within-state	(4) Across states	(5) Urban (Yes=1)	(6) Higher population than in 1870 (Yes=1)	(7) Manuf. emp. per capita (county)	(8) # of children	(9) Married (Yes=1)	(10) Owning house (Yes=1)
Machinist (G1)	-0.025***	3.368***	-0.014	0.048***	0.044***	0.035***	0.003	-0.051	0.002	-0.001
	(0.007)	(0.648)	(0.010)	(0.011)	(0.009)	(0.012)	(0.002)	(0.040)	(0.009)	(0.011)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-fixed effects (1870)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	63857	63857	63857	63857	63857	63857	63857	63857	63857	63857
Number of clusters	45	45	45	45	45	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. None of the specifications is weighted. The sample includes all sons whose father held an occupation between the 24.7th and 84.7th education rank percentiles in 1870. The outcome variable is an agricultural occupation indicator (Col. 1 - farmer, farm manager/foreman/laborer), the education rank of occupation (Col. 2), a binary variable which equals one if the son migrated within-state across counties (Col. 3) or across states (Col. 4) between 1870 and 1900, a binary variable which equals one if the son lived in an urban place in 1900 (Col. 5) or his place of residence fell into a larger *SIZEPL* category in 1900 than in 1870 (Col. 6), manufacturing employment per capita (as % of total county population), the number of children in the household (Col. 8), a marriage status (Col. 9) and house ownership (Col. 10) indicator. Baseline controls are described in Appendix 3.C.2. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C7: Measures of economic status - sons (G2; 1900)

	(1) Sobek log-score	(2) State-level nominal log-score	(3) State-level real log-score	(4) State-level real score (level)	(5) Preston-Haines log-score
Machinist (G1)	0.060***	0.040***	0.030***	15.170**	0.075***
	(0.011)	(0.010)	(0.010)	(6.406)	(0.009)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County-fixed effects (1870)	Yes	Yes	Yes	Yes	Yes
Sample size	49268	44553	44553	44553	40331
Number of clusters	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. None of the specifications is weighted. The sample includes all sons whose father held an occupation between the 24.7th and 84.7th education rank percentiles in 1870. The outcome variable is the Sobek, state-level nominal and real log-score (Col. 1-3), the state-level real score in levels (Col. 4) and the Preston-Haines log-score (Col. 5). Baseline controls are described in Appendix 3.C.2. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C8: Robustness checks - sons (G2; 1900)

	Occupational pre-trends		Manufacturing control occs		Maternal characteristics		Top 3 control occs excluded	
	(1) Education rank	(2) Urban (Yes=1)	(3) Education rank	(4) Urban (Yes=1)	(5) Education rank	(6) Urban (Yes=1)	(7) Education rank	(8) Urban (Yes=1)
Machinist (G1)	4.133***	0.074***	2.913**	0.050***	3.153***	0.050***	2.107***	0.052***
	(1.010)	(0.015)	(1.210)	(0.015)	(0.830)	(0.013)	(0.641)	(0.012)
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	7015	7015	3111	3111	8904	8904	8570	8570
Number of clusters	45	45	45	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. All specifications are weighted by weights gained from propensity score matching described in the main text. The sample includes 1842 (Col. 1-2, 7-8), 1794 (Col. 3-4) and 1823 (Col. 5-6) matched machinist sons. The outcome variable is the education rank of occupation (every odd column) or an urban place of living indicator (every even column). Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. In Columns 1-2, changes in the employment share of father's occupation (measured in percentage points and calculated for fathers' 1870 census division) between 1850-1860 and 1860-1870 are also included in the matching process (see Appendix 3.C.2 for more details). In Columns 3-4, exclusively those fathers are included who worked in durable or non-durable manufacturing in 1870. In this specification, the matching process chooses a single control father owing to the reduction in the number of potential control occupations. In Columns 5-6, the matching process balances the sample on maternal characteristics (1870): a literacy and a native-born status indicator, and her age in 1870 (in years). In Columns 7-8, carpenter, truck & tractor driver and shoemaker fathers are excluded from the control group in matching. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C9: Main outcomes - fathers (G1; 1870-1900)

	Occupational change (Yes=1)		Agricultural occupation in 1900 (Yes=1)		Migration (across states) (Yes=1)		Urban in 1900 (Yes=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Machinist (G1)	-0.098***	-0.110***	-0.037***	-0.030***	0.038**	0.030**	0.073***	0.070***
	(0.015)	(0.016)	(0.010)	(0.009)	(0.015)	(0.012)	(0.017)	(0.016)
Unbalanced controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grandfather (G0) controls	No	Yes	No	Yes	No	Yes	No	Yes
Sample size	5429	5456	5429	5456	5429	5456	5429	5456
Number of clusters	48	49	48	49	48	49	48	49

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. All specifications are weighted by weights gained from propensity score matching described in the main text (every even column matches on grandfather controls as well). The sample includes 1120 (1116 with grandfather controls included) matched machinist fathers. The outcome variable is a binary variable which equals one if i) the father changed occupation (Col. 1-2) and the new occupation is agricultural (Col. 3-4 - farmer, farm manager/foreman/laborer), ii) he migrated across states between 1870 and 1900 (Col. 5-6) , iii) he lived in an urban place in 1900 (Col. 7-8). Unbalanced controls included in the regressions are characteristics whose mean between machinist and control fathers is still significantly different at 5% after matching. Grandfather controls are (all measured in 1860): a literacy indicator, age (measured in years), an indicator if the father lived in the same state in 1870 as the grandfather in 1860 but in a different county, an indicator if the father lived in a different state in 1870 from the grandfather in 1860, indicator variables for the grandfather holding an agricultural (farmer, farm manager/foreman/laborer) or manufacturing (durable or non-durable manufacturing) occupation, indicator variables if the grandfather worked for the railways (railroad conductor, locomotive engineer, locomotive fireman, brakeman, switchman) / in the metal industry (molder, structural metal worker, furnaceman, heater, filer, grinder, polisher, roller, tinsmith and coppersmith) / in the chemical industry (*IND1950:* Cement, concrete, gypsum and plaster products; Miscellaneous chemicals and allied products; Petroleum refining; Miscellaneous petroleum and coal products; Rubber products) / in the steel and iron industry (*IND1950:* Blast furnaces, steel works, and rolling mills; Other primary iron and steel industries; Fabricated steel products) / in machinery (*IND1950:* Agricultural machinery and tractors; Office and store machines and devices; Miscellaneous machinery; Electrical machinery, equipment, and supplies) / as a machinist, an indicator for urban status, the number of inhabitants in the place of living (*SIZEPL*), and the value of personal property and real estates. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C10: Robustness checks with regressions - sons (G2; 1900)

	Spatial sorting pre-1870		State-occupation pre-trends		Weighting		Restricted control occs	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Education rank	Urban (Yes=1)	Education rank	Urban (Yes=1)	Education rank	Urban (Yes=1)	Education rank	Urban (Yes=1)
Machinist (G1)	2.587***	0.037***	2.782***	0.035***	3.166***	0.039***	3.162***	0.039***
	(0.898)	(0.012)	(0.678)	(0.009)	(0.730)	(0.010)	(0.655)	(0.011)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-fixed effects (1870)	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Detailed fixed effects	Yes	Yes	No	No	No	No	No	No
Sample size	55770	55770	61796	61796	63857	63857	52046	52046
Number of clusters	45	45	45	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. None of the specifications is weighted - except for Columns 5-6, where we use inverse proportional weights (see Appendix 3.C.3 for details). The sample includes all sons whose father held an occupation between the 24.7th and 84.7th education rank percentiles in 1870 - except for Columns 7-8, where these cutoffs are 44.7 and 64.7, respectively. The outcome variable is the educational rank of occupation (every odd column) or a binary variable which equals one if the son lived in an urban place in 1900 (every even column). Baseline controls are described in Appendix 3.C.2. Detailed fixed effects are generated by interacting the state of birth (county for the foreign-born) indicator, the county of residence indicator (1870), an urban place of living indicator (1870), and an indicator if the father was at least 34 years old in 1870. In Columns 3-4, state-occupation level measures of migration (within and across states jointly), occupation change probability, change in urban status and the probability of switching for an agricultural occupation (farmer, farm manager/foreman/laborer) are included for the 1850s and 1860s (see Appendix 3.C.2 for details). Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C11: Ability bias - fathers (G1 in 1870)

	Having a child (Yes=1)		Number of children		Having a spouse (Yes=1)		Literate spouse (Yes=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Machinist (G1)	-0.005	0.041	-0.002	0.053	-0.005	0.018	0.003	-0.037
	(0.010)	(0.042)	(0.024)	(0.095)	(0.009)	(0.042)	(0.005)	(0.036)
Grandfather (G0)-fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
County-fixed effects (1870)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	159716	22799	159716	22799	159716	22799	90473	9581
R^2	0.38	0.73	0.41	0.76	0.43	0.77	0.41	0.76

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are multiway clustered at the grandfather-county (1900) level. None of the specifications is weighted. The sample includes all fathers who held an occupation between the 24.7th and 84.7th education rank percentiles in 1870. In every column, the age of included fathers is between 16 and 50 years (inclusive). The outcome variable is i) a binary variable which equals one if the father had at least one child in 1870 (Col. 1-2), ii) the number of children in 1870 (Col. 3-4); iii)-iv) a binary variable which equals one if the father had a spouse (Col. 5-6) and, conditional on having a wife, she was literate (Col. 7-8). Personal controls included in the regressions are (all measured in 1870): the education rank of occupation, urban status and literacy indicator, age (in years), value of real estate and personal property, number of inhabitants in the place of living and a farmer-farm manager-farm foreman indicator. The interactions of the urban indicator, size of place of living, two wealth measures, education rank and age are also included. The squared size of place of living, wealth measures and age are included as well. Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C12: Robustness checks by the age of fathers and sons - sons (G2 in 1900)

	Sons (0-5 y.o.) of young fathers (¡33 y.o.; G1)		Sons (0-5 y.o.) of old fathers (¿=33 y.o.; G1)		Old and young sons
	(1)	(2)	(3)	(4)	(5)
	Education rank	Urban (Yes=1)	Education rank	Urban (Yes=1)	Education rank
Machinist (G1)	4.176***	0.041**	3.616***	0.044**	4.001***
	(1.397)	(0.020)	(1.106)	(0.020)	(0.695)
Machinist (G1) x $\mathbb{1}$(son (G2) older than 10 y.o. in 1870)					-4.098***
					(1.323)
Baseline controls	Yes	Yes	Yes	Yes	Yes
County-fixed effects (1870)	Yes	Yes	Yes	Yes	Yes
Sample size	16338	16338	18266	18266	63857
Number of clusters	45	45	45	45	45

Note: OLS regression coefficients with standard errors in parentheses. Standard errors are clustered at the state (1900) level. None of the specifications is weighted. The sample includes all sons i) whose father held an occupation between the 24.7th and 84.7th education rank percentiles in 1870; and ii) who were not older than five years in 1870 (Col. 1-4). The outcome variable is the educational rank of occupation (Col. 1,3,5) or a binary variable which equals one if the son lived in an urban place in 1900 (Col. 2,4). Column 5 includes the son age indicator as a main effect separately. Baseline controls are described in Appendix 3.C.2. The estimation includes only fathers who were younger (older) than thirty-three years in 1870 in Columns 1-2 (3-4). Levels of significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.C.2. Details on data

Controls used in propensity score matching

The controls used in the baseline propensity score matching are the following. For every father in the census in 1870, we measure:

- Personal characteristics: age (in years), literacy (can read and write, yes=1), foreign-born dummy (yes=1), native-born dummy (yes=1), dummies for the UK (yes=1) and for Germany (yes=1) as country of birth;
- Occupational characteristics: education rank of occupation (percentile rank; Song et al. 2020),[37] occupation-state level migration and probability of occupation change between 1860 and 1870 (see Appendix 3.C.2);
- Measures of individual wealth: value of personal property and real estates (separately);
- Characteristics of place of residence: urban status (yes=1), size category of place of living,[38] dummy for living in the state of birth (yes=1) and state-fixed effects.

Moreover, we use the pairwise interactions of the following six variables: real estate, personal property, age, urban dummy, population of place of living, education rank. We also include the square of the two wealth measures, the age and the population of place of living. Finally, we include several county characteristics downloaded from the NHGIS (Manson et al. 2021; for 1870): the share of manufacturing employment (% of total population), manufacturing output per capita, manufacturing output per manufacturing wage earners, the share of steam engine-provided engine power (% of steam engine- and water-driven engine-provided total). We refer to these controls jointly as *baseline controls*.

The construction of state-level earnings scores

The source of state-level earnings data is the *Fifteenth Annual Report of the Commissioner of Labor* which reports daily average wages for US states and other countries mainly for years in the second half of the 19th century. We sought to find a close match for every occupation i) which has a large role as control occupation for machinist fathers in 1870, or ii) which is a common occupation across fathers or sons in 1900.

[37]We use the first available rank which is constructed for those born around 1880.

[38]We converted the original *SIZEPL* variable into actual population numbers using the midpoint of every interval. The first and last categories are defined using half the length of the second and penultimate intervals, respectively.

We checked all relevant state-occupation pairs for 1870-72, 1879-1881 and 1890-1892, and digitized every entry in which at least ten individuals were used for average wage calculation. In case of multiple entries within any of the three-year time spans for a given state, we chose the average wage which was based on the largest number of wage reporting individuals. We exclusively included entries for males. The ending years of 1872, 1881 and 1892 were chosen because they preceded the Panics of 1873 and 1893, and the Depression of 1882-85. In a few cases, we deviated from our baseline data collection strategy to improve our sample. For miners, census-based, daily average wages were used from 1889 for the 1892 income score because the number of reporting states and observations used for average wage calculation were undoubtedly superior to other publications between 1890-1892. For farm laborers, data were digitized from the *Ninety-ninth Bulletin of U.S. Department of Agriculture (Wages of Farm Labor).* Daily wages were digitized *without* board (accommodation) to reduce the gap in in-kind compensation between agricultural and manufacturing laborers (Alston and Hatton 1991, Hatton and Williamson 1991). For the 1892 score, the number of occupations available is increased in our sample by using the publication titled *The slums of Baltimore, Chicago, New York, and Philadelphia: prepared in compliance with a joint resolution of the Congress of the United States (1892).* We take mostly occupations in services[39] for Maryland, Illinois, New York and Pennsylvania.

Next, daily wages were converted to yearly earnings following Sobek (1996). We assumed 245 days of work for the majority of occupations, 225 days for building trades (bricklayers, cabinetmakers, carpenters, masons, painters, plasterers) and farm laborers, 270 days for clerical occupations (bookkeepers, clerks, telegraph operators).[40] For farm labor, we assumed 30 days of harvest wages and 195 (=225-30) days of non-harvest wages.[41] We multiplied the yearly earnings of farm labor by the ratio of farmer-to-farm labor score in Sobek (1996) to compute earnings scores for owner-occupier farmers.

The main limitation of our earnings score is that the earnings of high-skilled workers, for instance, lawyers or physicians, cannot be observed. To solve this problem, the following imputation procedure is set up. First, we took the earnings

[39]These occupations are: barbers, bartenders, watchmen, policemen, detectives, agents (n.e.c), clerks, long-shoremen, hucksters, salesmen (n.e.c). We included an observation if the average wage could be calculated using at least ten individuals.

[40]The slum report provides weekly wages. Following Sobek (1996), we assumed 45 weeks worked apart from clerical jobs, where 48 weeks are assumed.

[41]Unlike Sobek (1996), we did not assume 245 days of work for farm laborers because it gave rise to a tendency of *nominal* farm laborer wages surpassing laborer wages. This would be inconsistent with existing evidence (Alston and Hatton 1991; Hatton and Williamson 1991). Our ratio between farm laborer to laborer nominal earnings scores is really close to the estimates found in the literature which takes into account the pecuniary value of in-kind remuneration as well. Moreover, the similar length of (un)employment spells between farm workers and workers in building trades is also consistent with Engerman and Goldin (1991).

scores of the fifteen occupations (TOP15) for which we have the most state-year level observations.[42] Afterwards, we calculated the earnings scores of missing, predominantly white-collar occupations[43] by multiplying our earnings scores for the available TOP15 occupations with the ratio of the Sobek score of the missing occupation and each of the TOP15 occupations. Then, we took the unweighted average of the implied earnings scores which constitutes our earnings score for missing occupations. We calculated this average only if at least eight of the fifteen (more than half) occupations were available for the case of a given state-year pair in order to reduce measurement error. The main assumption underlying this imputation procedure is that the ratio of earnings scores found in Sobek (1996) around 1890 is the same across states (for our 1892 score), or the same across states and time (for our 1872 and 1880 scores). Reassuringly, Katz and Margo (2014) find that the skilled artisans-to-clerks earnings ratio remained stable between the 1840s and 1880s. The debate if the earnings of higher-skilled workers differed across states more or less than the earnings of production workers or craftsmen has not been settled yet (see Goldin 1998; Rosenbloom 1990; 1996; 2002; Sundstrom and Rosenbloom 1993). Therefore, applying the Sobek score ratio-implied premia for earlier decades might not introduce a large measurement error since most of our TOP15 benchmark occupations are classified as artisans/craftsmen and we mostly impute the wages of white-collar workers.[44]

Another empirical barrier is that some harmonized occupations have many potential matches in our earnings score data. For instance, we had to aggregate the earnings score of miners of coal, iron or zinc into a single score for miners. The affected occupations are brickmasons (bricklayers and masons), railroad conductors (freight, passenger or not specified), miners (coal, iron, lead and zinc), spinners and weavers (cotton or woolen goods). Our state-year level earnings score for these harmonized occupations is defined as the observation-weighted average earnings score (of ”subcategories”).[45] As a last step, missing earnings scores were imputed

[42]These occupations are: blacksmiths, boilermakers, cabinetmakers, carpenters, compositors, engineers (locomotive), firemen (locomotive), laborers (n.e.c.), machinists, molders, painters, pattern makers, plumbers, stone cutters and teamsters.

[43]These occupations are: operatives (n.e.c.), managers, physicians, lawyers, meat cutters, clergymen, pharmacists, policemen, insurance agents, foremen (n.e.c.), teachers (n.e.c.), craftsmen (n.e.c.), fishermen, engineers (civil and mechanical separately), accountants, chemists, draftsmen, editors, funeral directors, musicians, ship officers, stenographers, real estate agents, janitors, waiters, gardeners and sailors. For 1872 and 1880, the list also includes barbers, bartenders, agents (n.e.c.) and hucksters.

[44]For 1872, we need an additional step because there are no data on the wages of miners, shoemakers and tailors who play an important part in the control group of machinists. To impute their wages, we follow our procedure described in the main text with one exception. Instead of using the ratio of Sobek scores, we calculate our observation-weighted, US-level earnings score in 1880 for the TOP15 occupations as well as for miners, shoemakers and tailors. We use the ratio of these earnings scores to implement the imputation procedure in order to diminish the potential effect of the Second Industrial Revolution on relative wages over time.

[45]Additionally, we included furnacemen in foundries or in the gas industry as furnacemen, and lumbermen

with the unweighted average of states within a given census division whenever it was possible.

The estimates of measuringworth.com were used to convert all earnings scores into 1890 dollars. The 1872/1880 earnings scores were multiplied by 0.75/0.89.

The conversion of nominal earnings scores to real scores requires state-level and urban-rural price differences. Nominal earnings (1872, 1880 and 1892) were deflated by the state-level price index of Haines (1989), and nominal wages (1940) by the cost of living measures reported in Stecker (1937). As Haines (1989) and Stecker (1937) do not contain information on all states, we use the price index of a neighboring state in the case of missing values (the actual pairs are available upon request).[46] We inflate earnings scores in places with less than 25.000 inhabitants by 1.192 (1872, 1880 and 1892 - Hatton and Williamson 1991) and by 1.205 (1940 - Williamson and Lindert 1980) to account for urban-rural price differences. In doing so, we follow the best practice in earlier literature (e.g., Collins and Wanamaker 2014).

The construction of schooling supply measures

The source of our high school supply proxies are different *Reports of the Commissioner of Education*. We followed a distinct data collection strategy for private and public high schools.

Private high schools We refer to institutions for secondary instruction, preparatory schools, commercial and business colleges (excluding evening schooling), preparatory departments of colleges and universities, and schools of science as private high school.

First, all available data on private high schools were digitized from the 1880 Report. If a school was reported as not replying to the query of the Commissioner's office, we tried to find it in the 1882 Report. Different types of schools were expected to report different data, so the following pieces of information could be digitized:

- Institutions for secondary instruction: number of teachers and students (split by gender), tuition fee, dummy whether mechanical drawing is taught, dummy if they had a chemical laboratory;

can be lumber handlers, lumber pilers or wood choppers as well. Two of the different "subcategories" of furnacemen or lumbermen never coincided within a state-year cell. Thus, there was no need to calculate observation-weighted averages.

[46] Stecker (1937) reports cost of living for more than one city in some states. We calculated the unweighted average of cost of living in cities within those states.

- Preparatory schools: number of teachers and students,[47] tuition fee and dummy if they had a chemical laboratory;
- Commercial and business colleges: number of teachers and students[48] (split by gender), and tuition fee;
- Preparatory departments of colleges and universities: number of teachers and students (split by gender);
- Schools of science: number of teachers and students (split by gender).

If tuition fees were not reported for the entire scholastic year (but for a term or month), a 40-week (10-month) long scholastic year was assumed which was the most common length. The children of residents sometimes did not have to pay the tuition fee. In such cases, the tuition fee is set equal to zero. In the next step, schools were matched to counties (1870) one-by-one using their reported location (post office).

Public high schools The data collection process for public high schools is more complex. While detailed statistics were reported for private high schools starting from the 1870s, no school-level information is available on public high schools until 1890. Moreover, the year of establishment is solely recorded in the Reports published in the mid-1900s.

To circumvent these data limitations, we adopted the following data collection strategy. First, we restricted our attention to schools in cities which had a population of 7.500 in 1880 since municipality/school name changes between 1890 and the mid-1900s would be an insurmountable barrier to data collection considering the number of public schools. Then, we turned to Reports of the mid-1900s for the list of public high schools which were established in these cities until 1880. Next, all available data were digitized on these high schools in the 1890/91 Report. If a school did not report despite being established pre-1880, we searched for it in the 1892/93 Report. For high schools which existed in 1890 but had no establishment year, we searched the web to gather information about their establishment year. As a result of this process, we obtained information on the number of teachers and students (split by gender) in public high schools around 1890. We believe that this value should be strongly positively correlated with its counterpart in 1880 since high school completion rates started their rapid increase only after the turn of the century (Goldin 1998; Goldin and Katz 2008). One might also argue that in the

[47]Preparatory schools are not included in our high school student shares since we do not know the exact number of male students.

[48]We digitized the number of students in day education if it was available. Otherwise, the missing value was imputed with the number of all students including evening schooling.

largest, fastest growing cities schools might have been split between 1880 and 1890 and, consequently, we underestimate the true extent of high school provision. Nonetheless, we show in the relevant analysis that our results our robust to the omission of these metropolises.

Imputation of missing values Before the creation of the final measures of schooling supply, missing values for the six school types had to be imputed. We followed the same procedure for all of them (except for public high schools - see the last paragraph of this section). First, if the number of all students was missing, we used the unweighted average of the same type of schools within-state (if there were less than ten such schools, then within-census division). The number of male students was imputed using the unweighted share of males in the same type of schools within-state (if there were less than ten such schools, then within-census division) and multiplying it by the (imputed) number of all students. The number of teachers was imputed similarly - the unweighted average of the same type of schools within-state (if there were less than ten such schools, then within-census division). Finally, a missing tuition fee was imputed as the number of students-weighted tuition fee within-state (if there were less than ten schools of the underlying type, then within-census division).

The aggregation of school-level measures to the county level amounts to a simple summation of the number of students and teachers, and taking the weighted average (by number of students) in case of the tuition fee. The five different private school types were pooled together before summation. The share of private and public high school students was calculated as the number of male students divided by the number of males aged 14-20 in a given county in 1880. The teacher-pupil ratio is defined as the student-weighted ratio of teachers to all students (male and female) at each school. The share of students having technical education (at institutions for secondary instruction or preparatory schools) was constructed as follows. All students who were at a school which offered mechanical drawing or had a chemical laboratory were indicated as having technical education. The sum of these students is divided by the total number of students at the county level.

For public high schools, the strong dependence of school size on local population necessitated a different imputation strategy. First, we ran the following regression:

$$y_{c,s} = \beta \cdot Population_{c,1880} + \gamma \cdot Population^2_{c,1880} + f_{state(c)} + \epsilon_{c,s} \tag{3.4}$$

where $y_{c,s}$ is the number of students or teachers in city c and public high school s. City population and its squared form (population figures are from the Report of the Commissioner of Education in 1880), and the state-fixed effects produce an $R^2 \approx 0.5$. This model is used to impute the missing number of students and teachers if a public high school already existed before 1880. To split the number of students by gender, the average gender ratio is used within state - if at least ten public high schools have non-missing data -, otherwise the average of public high schools in the census division.

Construction of other variables

Occupational employment growth until 1870 The full count censuses of 1850, 1860 and 1870 are used to compute changes in employment shares in the period preceding the Second Industrial Revolution. In all three years, we dropped individuals who were not between 16 and 65 years old and gave a non-occupational response (*OCC1950* codes larger than 978). Then, we calculated the share of every harmonized occupation for each census division. Last, we created the differences between 1850-60 and 1860-70, and merged them to fathers in 1870 based on their occupation and census division.

Occupation-state level measures in the pre-period We computed several occupation-state level measures based on the 1850, 1860 and 1870 full count censuses. To do so, the census was first restricted to individuals between 16 and 40 years old. We assigned to every state the occupational level i) probability of changing occupation; ii) probability of migration (changing county or state), iii) average change in the urban status dummy, iv) probability of having agricultural occupation at the end of the decade - based on individuals who at the beginning of the decade (1850s and 1860s) lived in the given state.

Imputing self-employed income We followed the literature in imputing the income of self-employed individuals in the 1940 full count census (see e.g., Collins and Wanamaker 2017; Ward forthcoming). As a first step, a sample of male self-employed workers was created in the 1960 5% census. We calculated the ratio between the total income and wage income for these individuals. Finally, the wage of self-employed individuals with non-zero reported wage in 1940 was inflated by the median of the calculated ratio (1.89). The main assumption of this imputation is that the ratio remained constant between 1940 and 1960.

For self-employed individuals, who reported zero wage earned in 1940, we use the median total income obtained from the 1960 census after a conversion from 1960 to 1940 dollars and conditional on reporting more than 50 weeks worked. We calculated the median separately for the agricultural (*OCC1950:* 100, 123, 810, 820, 830, 840) and non-agricultural self-employed.

Characteristics of next-door neighbors To calculate observable measures for next-door neighbors, we first took every household head from the 1870 census. This data set is sorted, so neighbors appear next to each other. To every household head we assigned its ten closest neighbors, i.e. the five household heads right before and after a given person. Afterwards, the real estate and personal property values were winsorized at the 1st and 99th percentiles. Finally, the average of observable neighbor characteristics was computed and they were assigned to the 1870 full count census. We use the occupational education rank estimated for the 1880 (earliest) cohort by Song et al. (2020) for all neighbors. Literacy (foreign-born status) is measured as a dummy which is set equal to one if a given neighbor could read and write (was born outside the US).

3.C.3. Inverse proportional weights

To create inverse proportional weights for the sons' sample, sons were linked between 1870 and 1900 with the two conservative linking methods developed by Abramitzky et al. (2020). First, we merged the full count census of 1900 to the crosswalk, keeping matched as well as unmatched observations. Next, we also merged this data set with the 1870 full count census. If an observation could not be matched with any of the two conservative linking methods, we considered it unmatched and generated a variable which was set to zero for this case (one otherwise). Then, we used this binary variable as an outcome of a probit regression on age bins (following the code provided by Abramitzky et al. 2020), an urban place of living indicator, the population size category of place of living (*SIZEPL*) and census division-fixed effects - all measured in 1900. Finally, the inverse proportional weight for every single matched observation was calculated based on the following formula: $(1 - \hat{p})/\hat{p}$, where $\hat{p}$ is the predicted probability of a successful match. We set the weight equal to zero for observations which were unmatched.

Bibliography

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A. Pathak (Dec. 2017). "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match". In: *American Economic Review* 107.12, pp. 3635–89.

Abdulkadiroğlu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters (May 2020). "Do Parents Value School Effectiveness?" In: *American Economic Review* 110.5, pp. 1502–39.

Abramitzky, Ran, Leah Boustan, Elisa Jacome, and Santiago Perez (Feb. 2021). "Intergenerational Mobility of Immigrants in the United States over Two Centuries". In: *American Economic Review* 111.2, pp. 580–608.

Abramitzky, Ran, Leah Boustan, and Myera Rashid (2020). *Census Linking Project: Version 1.0 [dataset]. 2020.*

Acemoglu, Daron and Pascual Restrepo (Spring 2019). "Automation and New Tasks: How Technology Displaces and Reinstates Labor". In: *Journal of Economic Perspectives* 33.2, pp. 3–30.

— (2020). "Robots and Jobs: Evidence from US Labor Markets". In: *Journal of Political Economy* 128.6, pp. 2188–2244.

Ager, Philipp, Benedikt Herz, and Markus Brueckner (Oct. 2020). "Structural Change and the Fertility Transition". In: *The Review of Economics and Statistics* 102.4, pp. 806–822.

Aigner, Dennis J. (1973). "Regression with a binary independent variable subject to errors of observation". In: *Journal of Econometrics* 1.1, pp. 49–59.

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney (Apr. 2016). "The Long-Run Impact of Cash Transfers to Poor Families". In: *American Economic Review* 106.4, pp. 935–971.

Akee, Randall K. Q., William E. Copeland, Gordon Keeler, Adrian Angold, and E. Jane Costello (Jan. 2010). "Parents' Incomes and Children's Outcomes: A Quasi-experiment Using Transfer Payments from Casino Profits". In: *American Economic Journal: Applied Economics* 2.1, pp. 86–115.

Allende, Claudia, Francisco Gallego, and Christopher Neilson (July 2019). *Approximating the Equilibrium Effects of Informed School Choice*. Working Papers 628. Princeton University, Department of Economics, Industrial Relations Section.

Alston, Lee J. and Timothy J. Hatton (Mar. 1991). “The Earnings Gap Between Agricultural and Manufacturing Laborers, 1925-1941”. In: *The Journal of Economic History* 51.1, pp. 83–99.

Angrist, Joshua, Peter Hull, Parag A. Pathak, and Christopher R. Walters (Dec. 2020). *Simple and Credible Value-Added Estimation Using Centralized School Assignment*. NBER Working Papers 28241. National Bureau of Economic Research, Inc.

Artuç, Erhan, Shubham Chaudhuri, and John McLaren (June 2010). “Trade Shocks and Labor Adjustment: A Structural Empirical Approach”. In: *American Economic Review* 100.3, pp. 1008–45.

Asai, Yukiko (2015). “Parental leave reforms and the employment of new mothers: Quasi-experimental evidence from Japan”. In: *Labour Economics* 36.C, pp. 72–83.

Ashenfelter, Orley and Cecilia Rouse (1998). “Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins”. In: *The Quarterly Journal of Economics* 113.1, pp. 253–284.

Austin, Peter C. (2011). “An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies”. In: *Multivariate Behavioral Research* 46.3, pp. 399–424.

Autor, David H., David Dorn, and Gordon H. Hanson (Oct. 2016). “The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade”. In: *Annual Review of Economics* 8.1, pp. 205–240.

Bailey, Martha J., Connor Cole, Morgan Henderson, and Catherine Massey (Dec. 2020). “How Well Do Automated Linking Methods Perform? Lessons from US Historical Data”. In: *Journal of Economic Literature* 58.4, pp. 997–1044.

Becker, Gary and Nigel Tomes (1979). “An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility”. In: *Journal of Political Economy* 87.6, pp. 1153–89.

— (1986). “Human Capital and the Rise and Fall of Families”. In: *Journal of Labor Economics* 4.3, S1–39.

Bettinger, Eric, Oded Gurantz, Laura Kawano, Bruce Sacerdote, and Michael Stevens (Feb. 2019). “The Long-Run Impacts of Financial Aid: Evidence from California’s Cal Grant”. In: *American Economic Journal: Economic Policy* 11.1, pp. 64–94.

Billings, Stephen B., David J. Deming, and Jonah Rockoff (Sept. 2013). “ School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg *”. In: *The Quarterly Journal of Economics* 129.1, pp. 435–476.

Bingley, Paul and Alessandro Martinello (2017). “Measurement Error in Income and Schooling and the Bias of Linear Estimators”. In: *Journal of Labor Economics* 35.4, pp. 1117–1148.

Bleakley, Hoyt and Joseph Ferrie (2016). “Shocking Behavior: Random Wealth in Antebellum Georgia and Human Capital Across Generations”. In: *The Quarterly Journal of Economics* 131.3, pp. 1455–1495.

Blum, Sonja, Alison Koslowski, Alexandra Macht, and Peter Moss (2018). *14th International Review of Leave Policies and Related Research 2018*. Report. International Network on Leave Policies and Research.

Boelmann, Barbara, Anna Raute, and Uta SchÃ¶nberg (Oct. 2020). *Wind of Change? Cultural Determinants of Maternal Labor Supply*. CReAM Discussion Paper Series 2020. Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.

Boneva, Teodora, Katja Kaufmann, and Christopher Rauh (2021). *Maternal Labor Supply: Perceived Returns, Constraints, and Social Norms*. IZA Discussion Papers 14348. Institute of Labor Economics (IZA).

Boneva, Teodora and Christopher Rauh (Mar. 2018). “Parental Beliefs about Returns to Educational Investments—The Later the Better?” In: *Journal of the European Economic Association* 16.6, pp. 1669–1711.

Booij, Adam S, Edwin Leuven, and Hessel Oosterbeek (Sept. 2016). “Ability Peer Effects in University: Evidence from a Randomized Experiment”. In: *The Review of Economic Studies* 84.2, pp. 547–578.

Boone, Christopher and Laurence Wilse-Samson (2019). “Farm Mechanization and Rural Migration in the Great Depression”. Unpublished manuscript.

Boserup, Simon Halphen, Wojciech Kopczuk, and Claus Thustrup Kreiner (July 2018). “Born with a Silver Spoon? Danish Evidence on Wealth Inequality in Childhood”. In: *Economic Journal* 128.612, pp. 514–544.

Bratberg, Espen, Øivind Anti Nilsen, and Kjell Vaage (Aug. 2008). “Job losses and child outcomes”. In: *Labour Economics* 15.4, pp. 591–603.

Brewer, Mike, Sarah Cattan, Claire Crawford, and Birgitta Rabe (2022). “Does more free childcare help parents work more?” In: *Labour Economics* 74.C, S0927537121001354.

Buckles, Kasey S. and Daniel M. Hungerman (July 2013). “Season of Birth and Later Outcomes: Old Questions, New Answers”. In: *The Review of Economics and Statistics* 95.3, pp. 711–724.

Bundesministerium für Familie Senioren, Frauen und Jugend (2020). “Elterngeld, ElterngeldPlus und Elternzeit: das Bundeselterngeld und Elternzeit Gesetz”. In.

Calsamiglia, Caterina and Maia Güell (2018). “Priorities in school choice: The case of the Boston mechanism in Barcelona”. In: *Journal of Public Economics* 163, pp. 20–36.

Calsamiglia, Caterina and Annalisa Loviglio (2019). “Grading on a curve: When having good peers is not good”. In: *Economics of Education Review* 73, p. 101916.

Campos, Christopher and Caitlin Kearns (2021). "The Impacts of Neighborhood School Choice: Evidence from Los Angeles' Zones of Choice". In.

Card, David and Alan B Krueger (Feb. 1992). "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States". In: *Journal of Political Economy* 100.1, pp. 1–40.

Castleman, Benjamin L. and Bridget Terry Long (2016). "Looking beyond Enrollment: The Causal Effect of Need-Based Grants on College Access, Persistence, and Graduation". In: *Journal of Labor Economics* 34.4, pp. 1023–1073.

Caucutt, Elizabeth M. and Lance Lochner (2020). "Early and Late Human Capital Investments, Borrowing Constraints, and the Family". In: *Journal of Political Economy* 128.3, pp. 1065–1147.

Cesarini, David, Erik Lindqvist, Matthew J. Notowidigdo, and Robert Östling (Dec. 2017). "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries". In: *American Economic Review* 107.12, pp. 3917–3946.

Chetty, Raj and Nathaniel Hendren (2018a). "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects". In: *The Quarterly Journal of Economics* 133.3, pp. 1107–1162.

— (2018b). "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects". In: *The Quarterly Journal of Economics* 133.3, pp. 1107–1162.

— (2018c). "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates". In: *The Quarterly Journal of Economics* 133.3, pp. 1163–1228.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014). "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States". In: *The Quarterly Journal of Economics* 129.4, pp. 1553–1623.

Coelli, Michael B. (Jan. 2011). "Parental job loss and the education enrollment of youth". In: *Labour Economics* 18.1, pp. 25–35.

Collins, William J. and Marianne H. Wanamaker (Jan. 2014). "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data". In: *American Economic Journal: Applied Economics* 6.1, pp. 220–252.

— (May 2017). *African American Intergenerational Economic Mobility Since 1880*. NBER Working Papers 23395. National Bureau of Economic Research, Inc.

Cooper, Kerris and Kitty Stewart (July 2017). *Does Money Affect Children's Outcomes? An update*. CASE Papers 203. Centre for Analysis of Social Exclusion, LSE.

Correia, Sergio (2016). *Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator*. Tech. rep. Working Paper.

Cortes, Guido Matias and Giovanni Gallipoli (2018). “The Costs of Occupational Mobility: An Aggregate Analysis”. In: *Journal of the European Economic Association* 16.2, pp. 275–315.

Cunha, Flavio and Heckman (May 2007). “The Technology of Skill Formation”. In: *American Economic Review* 97.2, pp. 31–47.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor (2008). “When are ghettos bad? Lessons from immigrant segregation in the United States”. In: *Journal of Urban Economics* 63.3, pp. 759–774.

Dahl, Gordon B. and Lance Lochner (Aug. 2012). “The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit”. In: *American Economic Review* 102.5, pp. 1927–1956.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum (2021a). “Adjusting to Globalization in Germany”. In: *Journal of Labor Economics* 39.1, pp. 263–302.

Dauth, Wolfgang, Sebastian Findeisen, Jens Suedekum, and Nicole Woessner (May 2021b). “The Adjustment of Labor Markets to Robots”. In: *Journal of the European Economic Association*.

Dawson, Andrew (1979). “The paradox of dynamic technological change and the labor aristocracy in the United States, 1880–1914”. In: *Labor History* 20.3, pp. 325–351.

Denning, Jeffrey T., Benjamin M. Marx, and Lesley J. Turner (July 2019). “ProPelled: The Effects of Grants on Graduation, Earnings, and Welfare”. In: *American Economic Journal: Applied Economics* 11.3, pp. 193–224.

Denning, Jeffrey T., Richard Murphy, and Felix Weinhardt (Oct. 2021). “Class Rank and Long-Run Outcomes”. In: *The Review of Economics and Statistics*, pp. 1–45.

DESTATIS, German Federal Statistical Office (2019). “Statistik zum Elterngeld: Beendete Leistungsbezüge für im Jahr 2015 geborene Kinder”. In.

Di Maio, Michele and Roberto Nisticò (2019). “The effect of parental job loss on child school dropout: Evidence from the Occupied Palestinian Territories”. In: *Journal of Development Economics* 141.C.

Dix-Carneiro, Rafael (2014). “Trade Liberalization and Labor Market Dynamics”. In: *Econometrica* 82.3, pp. 825–885.

Dobbie, Will and Roland G. Fryer (July 2011). “Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children’s Zone”. In: *American Economic Journal: Applied Economics* 3.3, pp. 158–187.

Drange, Nina and Tarjei Havnes (2019). “Early Childcare and Cognitive Development: Evidence from an Assignment Lottery”. In: *Journal of Labor Economics* 37.2, pp. 581–620.

Duflo, Esther (Sept. 2001). “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment”. In: *American Economic Review* 91.4, pp. 795–813.

Duflo, Esther, Pascaline Dupas, and Michael Kremer (Aug. 2011). “Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya”. In: *American Economic Review* 101.5, pp. 1739–74.

Duncan, Otis Dudley and Beverly Duncan (1955). “A methodological analysis of segregation indexes”. In: *American sociological review*.

Dupraz, Yannick and Andreas Ferrara (2021). *Fatherless: The Long-Term Effects of Losing a Father in the U.S. Civil War*. CAGE Online Working Paper Series 538. Competitive Advantage in the Global Economy (CAGE).

Durlauf, Steven (2004). “Neighborhood effects”. In: *Handbook of Regional and Urban Economics*. Ed. by J. V. Henderson and J. F. Thisse. 1st ed. Vol. 4. Elsevier. Chap. 50, pp. 2173–2242.

EC (2022). “Report on Gener Equality in the EU”. In.

Edin, Per-Anders, Tiernan Evans, Georg Graetz, Sofia Hernnäs, and Guy Michaels (June 2019). *Individual Consequences of Occupational Decline*. CEPR Discussion Papers 13808. C.E.P.R. Discussion Papers.

Engerman, Stanley and Claudia Goldin (Jan. 1991). *Seasonality in Nineteenth Century Labor Markets*. NBER Historical Working Papers 0020. National Bureau of Economic Research, Inc.

Eriksson, Katherine and Zachary Ward (Dec. 2019). “The Residential Segregation of Immigrants in the United States from 1850 to 1940”. In: *The Journal of Economic History* 79.4, pp. 989–1026.

Fack, Gabrielle and Julien Grenet (Apr. 2015). “Improving College Access and Success for Low-Income Students: Evidence from a Large Need-Based Grant Program”. In: *American Economic Journal: Applied Economics* 7.2, pp. 1–34.

Feigenbaum, James and Daniel P. Gross (Nov. 2020). *Automation and the Fate of Young Workers: Evidence from Telephone Operation in the Early 20th Century*. NBER Working Papers 28061. National Bureau of Economic Research, Inc.

Feigenbaum, James and Hui Ren Tan (Dec. 2020). “The Return to Education in the Mid-Twentieth Century: Evidence from Twins”. In: *The Journal of Economic History* 80.4, pp. 1101–1142.

Fernández-Kranz, Daniel and Núria Rodríguez-Planas (2021). “Too family friendly? The consequences of parent part-time working rights”. In: *Journal of Public Economics* 197, p. 104407.

Fort, Margherita, Andrea Ichino, and Giulio Zanella (2019). “Cognitive and Noncognitive Costs of Day Care at Age 0–2 for Children in Advantaged Families”. In: *Journal of Political Economy* 0.0, pp. 000–000.

Galster, George C. (2012). “The Mechanism(s) of Neighbourhood Effects: Theory, Evidence, and Policy Implications”. In: *Neighbourhood Effects Research: New*

Perspectives. Ed. by Maarten van Ham, David Manley, Nick Bailey, Ludi Simpson, and Duncan Maclennan. Springer Netherlands.

Garlick, Robert (July 2018). "Academic Peer Effects with Different Group Assignment Policies: Residential Tracking versus Random Assignment". In: *American Economic Journal: Applied Economics* 10.3, pp. 345–69.

Girsberger, Esther Mirjam, Lena Hassani Nezhad, Kalaivani Karunanethy, and Rafael Lalive (2021). *Mothers at Work: How Mandating Paid Maternity Leave Affects Employment, Earnings and Fertility*. IZA Discussion Papers 14605. Institute of Labor Economics (IZA).

Givord, Pauline and Claire Marbot (2015). "Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies". In: *Labour Economics* 36.C, pp. 99–111.

Goldin, Claudia (June 1998). "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century". In: *The Journal of Economic History* 58.2, pp. 345–374.

Goldin, Claudia and Lawrence F. Katz (1999). "Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910-1940". In: *The Journal of Interdisciplinary History* 29.4, pp. 683–723.

— (Sept. 2000). "Education and Income in the Early Twentieth Century: Evidence from the Prairies". In: *The Journal of Economic History* 60.3, pp. 782–818.

— (2008). *The Race Between Education and Technology.* Belknap Press for Harvard University Press.

Goldin, Claudia and Robert A. Margo (1992). "The Great Compression: The Wage Structure in the United States at Mid-Century". In: *The Quarterly Journal of Economics* 107.1, pp. 1–34.

González, Libertad (Aug. 2013). "The Effect of a Universal Child Benefit on Conceptions, Abortions, and Early Maternal Labor Supply". In: *American Economic Journal: Economic Policy* 5.3, pp. 160–88.

Graetz, Georg and Guy Michaels (Dec. 2018). "Robots at Work". In: *The Review of Economics and Statistics* 100.5, pp. 753–768.

Hai, Rong and James J. Heckman (Apr. 2017). "Inequality in Human Capital and Endogenous Credit Constraints". In: *Review of Economic Dynamics* 25, pp. 4–36.

Haines, Michael R. (1989). "A State and Local Consumer Price Index for the United States in 1890". In: *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 22.3, pp. 97–105.

Hanlon, Walker (July 2021). "The Rise of the Engineer: Inventing the Professional Inventor During the Industrial Revolution". Unpublished manuscript.

Hastings, Justine S. and Jeffrey M. Weinstein (Nov. 2008). "Information, School Choice, and Academic Achievement: Evidence from Two Experiments*". In: *The Quarterly Journal of Economics* 123.4, pp. 1373–1414.

Hatton, Timothy J. and Jeffrey G. Williamson (Oct. 1991). “Wage gaps between farm and city: Michigan in the 1890s”. In: *Explorations in Economic History* 28.4, pp. 381–408.

Havnes, Tarjei and Magne Mogstad (2011). “Money for nothing? Universal child care and maternal employment”. In: *Journal of Public Economics* 95.11, pp. 1455–1465.

Heckman (2008). “SCHOOLS, SKILLS, AND SYNAPSES”. In: *Economic Inquiry* 46.3, pp. 289–324.

Heckman, James J. and Flavio Cunha (May 2007). “The Technology of Skill Formation”. In: *American Economic Review* 97.2, pp. 31–47.

Heckman, James J. and Ganesh Karapakula (May 2019). *Intergenerational and Intragenerational Externalities of the Perry Preschool Project*. Working Papers 2019-033. Human Capital and Economic Opportunity Working Group.

Hilger, Nathaniel G. (July 2016). “Parental Job Loss and Children’s Long-Term Outcomes: Evidence from 7 Million Fathers’ Layoffs”. In: *American Economic Journal: Applied Economics* 8.3, pp. 247–83.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart (2007). “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference”. In: *Political Analysis* 15.3, pp. 199–236.

Hsieh, Chang-Tai and Miguel Urquiola (Sept. 2006). “The effects of generalized school choice on achievement and stratification: Evidence from Chile’s voucher program”. In: *Journal of Public Economics* 90.8-9, pp. 1477–1503.

Humlum, Anders (Nov. 2019). “Robot Adoption and Labor Market Dynamics”. Unpublished manuscript.

Imai, Kosuke, Gary King, and Elizabeth A. Stuart (2008). “Misunderstandings between experimentalists and observationalists about causal inference”. In: *Journal of the Royal Statistical Society Series A* 171.2, pp. 481–502.

Inwood, Kris, Chris Minns, and Fraser Summerfield (2019). “Occupational income scores and immigrant assimilation. Evidence from the Canadian census”. In: *Explorations in Economic History* 72.C, pp. 114–122.

Jensen, Robert (May 2010). “The (Perceived) Returns to Education and the Demand for Schooling*”. In: *The Quarterly Journal of Economics* 125.2, pp. 515–548.

Joseph, Olivier, Ariane Pailhé, Isabelle Recotillet, and Anne Solaz (2013). “The economic impact of taking short parental leave: Evaluation of a French reform”. In: *Labour Economics* 25.C, pp. 63–75.

Kaboski, Joseph P. and Trevon D. Logan (2011). “Factor Endowments and the Returns to Skill: New Evidence from the American Past”. In: *Journal of Human Capital* 5.2, pp. 111–152.

Kambourov, Gueorgui and Iourii Manovskii (Feb. 2009). “Occupational Specificity Of Human Capital”. In: *International Economic Review* 50.1, pp. 63–115.

Katz, Lawrence F. and Robert A. Margo (Dec. 2014). "Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective". In: *Human Capital in History: The American Record*. NBER Chapters. National Bureau of Economic Research, Inc, pp. 15–57.

Kelly, Morgan, Joel Mokyr, and Cormac ó Gráda (June 2020). *The Mechanics of the Industrial Revolution*. CEPR Discussion Papers 14884. C.E.P.R. Discussion Papers.

King, Gary and Richard Nielsen (2019). "Why Propensity Scores Should Not Be Used for Matching". In: *Political Analysis* 27.4, pp. 435–454.

Klein, Alexander and Nicholas Crafts (Feb. 2020). "Agglomeration externalities and productivity growth: US cities, 1880–1930". In: *Economic History Review* 73.1, pp. 209–232.

Kleven, Henrik, Camille Landais, and Jakob Søgaard (2019). "Children and Gender Inequality: Evidence from Denmark". In: *American Economic Journal: Applied Economics* 11.4, pp. 181–209.

Kluve, Jochen and Sebastian Schmitz (2018). "Back to Work: Parental Benefits and Mothers' Labor Market Outcomes in the Medium Run". In: *ILR Review* 71.1, pp. 143–173.

Knight, Brian and Nathan Schiff (Feb. 2022). "Reducing Frictions in College Admissions: Evidence from the Common Application". In: *American Economic Journal: Economic Policy* 14.1, pp. 179–206.

Koffsky, Nathan (1949). "Farm and Urban Purchasing Power". In: *Studies in Income and Wealth, Volume 11*. NBER, pp. 151–220.

Kubin, Ingrid and Aloys Prinz (2002). "Labour supply with habit formation". In: *Economics Letters* 75.1, pp. 75–79.

Kunze, Astrid (2022). "Parental leave and maternal labor supply". In.

Kutscher, Macarena, Shanjukta Nath, and Sergio Urzua (2020). *Centralized Admission Systems and School Segregation: Evidence from a National Reform*. eng. IZA Discussion Papers 13305. Bonn.

Laliberté, Jean-William (May 2021a). "Long-Term Contextual Effects in Education: Schools and Neighborhoods". In: *American Economic Journal: Economic Policy* 13.2, pp. 336–77.

— (May 2021b). "Long-Term Contextual Effects in Education: Schools and Neighborhoods". In: *American Economic Journal: Economic Policy* 13.2, pp. 336–377.

Lalive, Rafael and Josef Zweimüller (2009). "How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments". In: *The Quarterly Journal of Economics* 124.3, pp. 1363–1402.

Lee, Sang Yoon (Tim) and Ananth Seshadri (2019). "On the Intergenerational Transmission of Economic Status". In: *Journal of Political Economy* 127.2, pp. 855–921.

Leuven, Edwin and Barbara Sianesi (Apr. 2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. Statistical Software Components, Boston College Department of Economics.

Lingwall, Jeff (2010). “Compulsory Schooling, the Family, and the ‘Foreign Element’ in the United States, 1880–1900”. Unpublished manuscript.

Lochner, Lance and Alexander Monge-Naranjo (July 2012). “Credit Constraints in Education”. In: *Annual Review of Economics* 4.1, pp. 225–256.

Løken, Katrine V. (Jan. 2010). “Family income and children’s education: Using the Norwegian oil boom as a natural experiment”. In: *Labour Economics* 17.1, pp. 118–129.

Løken, Katrine V., Magne Mogstad, and Matthew Wiswall (Apr. 2012). “What Linear Estimators Miss: The Effects of Family Income on Child Outcomes”. In: *American Economic Journal: Applied Economics* 4.2, pp. 1–35.

Long, Jason and Joseph Ferrie (2007). “The Path to Convergence: Intergenerational Occupational Mobility in Britain and the US in Three Eras*”. In: *The Economic Journal* 117.519, pp. C61–C71.

— (June 2013). “Intergenerational Occupational Mobility in Great Britain and the United States since 1850”. In: *American Economic Review* 103.4, pp. 1109–37.

Loviglio, Annalisa (2020). *Schools and Their Multiple Ways to Impact Students: A Structural Model of Skill Accumulation and Educational Choices*. working paper.

Lucas, Adrienne M. and Isaac M. Mbiti (Oct. 2012). “Access, Sorting, and Achievement: The Short-Run Effects of Free Primary Education in Kenya”. In: *American Economic Journal: Applied Economics* 4.4, pp. 226–253.

Lutz, Byron (May 2011). “The End of Court-Ordered Desegregation”. In: *American Economic Journal: Economic Policy* 3.2, pp. 130–68.

Malamud, Ofer and Abigail Wozniak (2012). “The Impact of College on Migration: Evidence from the Vietnam Generation”. In: *Journal of Human Resources* 47.4, pp. 913–950.

Maloney, William F and Felipe Valencia Caicedo (Aug. 2020). *Engineering Growth*. CEPR Discussion Papers 15144. C.E.P.R. Discussion Papers.

Manoli, Day and Nicholas Turner (May 2018). “Cash-on-Hand and College Enrollment: Evidence from Population Tax Data and the Earned Income Tax Credit”. In: *American Economic Journal: Economic Policy* 10.2, pp. 242–271.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles (2021). *IPUMS National Historical Geographic Information System: Version 16.0 [dataset]*.

McLanahan, Sara (2004). “Diverging Destinies: How Children Are Faring under the Second Demo- graphic Transition”. In: *Demography*.

Meisenzahl, Ralf R. and Joel Mokyr (Aug. 2011). “The Rate and Direction of Invention in the British Industrial Revolution: Incentives and Institutions”. In: *The Rate and Direction of Inventive Activity Revisited*. NBER Chapters. National Bureau of Economic Research, Inc, pp. 443–479.

Mello, Ursula (forthcoming). “Affirmative Action, Centralized Admissions and Access of Low-income students to Higher Education”. In: *AMERICAN ECONOMIC JOURNAL: ECONOMIC POLICY*.

MINEDUC (2017). “Ley de Inclusión Escolar: el primer gran debate de la reforma educacional”. In: *MINEDUC*.

Mogstad, Magne and Gaute Torsvik (May 2021). *Family Background, Neighborhoods and Intergenerational Mobility*. NBER Working Papers 28874. National Bureau of Economic Research, Inc.

Mokyr, Joel (1999). “The Second Industrial Revolution, 1870-1914”. In: *Storia dell’economia Mondiale*. Ed. by Valerio Castronovo. Laterza Publishing, pp. 219–245.

Molina, Teresa and Ivan Rivadeneyra (2021). “The schooling and labor market effects of eliminating university tuition in Ecuador”. In: *Journal of Public Economics* 196.C.

Mörk, Eva, Anna Sjögren, and Helena Svaleryd (2020). “Consequences of parental job loss on the family environment and on human capital formation-Evidence from workplace closures”. In: *Labour Economics* 67.C.

Müller, Dana and Katharina Strauch (2017). *Identifying mothers in administrative data*. FDZ-Methodenreport 201713 (en). Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].

Neilson, Christopher (2021). *The Rise of Centralized Assignment Mechanisms in Education Markets Around the World*. Working Papers.

Neilson, Christopher A. and Seth D. Zimmerman (2014). “The effect of school construction on test scores, school enrollment, and home prices”. In: *Journal of Public Economics* 120.C, pp. 18–31.

Nix, Emily and Martin Andresen (2019). *What Causes the Child Penalty? Evidence from Same Sex Couples and Policy Reforms*. Discussion Papers. Statistics Norway, Research Department.

Nollenberger, Natalia and Núria Rodriguez-Planas (2015). “Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain”. In: *Labour Economics* 36.C, pp. 124–136.

O’Brien, Anthony Patrick (1988). “Factory size, economies of scale, and the great merger wave of 1898-1902”. In: *The Journal of Economic History* 48.3, pp. 639–649.

OECD (2015). *Education at a Glance 2015*, p. 564.

— (2018). *Equity in Education*, p. 192.

OECD (2019). “Gender wage gap”. In.

Olden, Andreas and Jarle Møen (Mar. 2022). “The triple difference estimator”. In: *The Econometrics Journal*. utac010.

Olivetti, Claudia and M. Daniele Paserman (Aug. 2015). “In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940”. In: *American Economic Review* 105.8, pp. 2695–2724.

Olivetti, Claudia and Barbara Petrongolo (Feb. 2017). “The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries”. In: *Journal of Economic Perspectives* 31.1, pp. 205–30.

Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens (July 2008). “The Intergenerational Effects of Worker Displacement”. In: *Journal of Labor Economics* 26.3, pp. 455–483.

Papageorgiou, Theodore (2014). “Learning Your Comparative Advantages”. In: *Review of Economic Studies* 81.3, pp. 1263–1295.

Parman, John (Mar. 2011). “American Mobility and the Expansion of Public Education”. In: *The Journal of Economic History* 71.1, pp. 105–132.

Preston, Samuel H. and Michael R. Haines (May 1991). *Fatal Years: Child Mortality in Late Nineteenth-Century America*. NBER Books pres91-1. National Bureau of Economic Research, Inc.

Raute, Anna (2019). “Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits”. In: *Journal of Public Economics* 169, pp. 203–222.

Ravazzini, Laura (Dec. 2018). “Childcare and maternal part-time employment: a natural experiment using Swiss cantons”. In: *Swiss Journal of Economics and Statistics* 154.1, pp. 1–16.

Rege, Mari, Kjetil Telle, and Mark Votruba (2011). “Parental Job Loss and Children’s School Performance”. In: *Review of Economic Studies* 78.4, pp. 1462–1489.

Rosenbaum, Paul R. and Donald B. Rubin (Apr. 1983). “The central role of the propensity score in observational studies for causal effects”. In: *Biometrika* 70.1, pp. 41–55.

Rosenberg, Nathan and Manuel Trajtenberg (Mar. 2004). “A General-Purpose Technology at Work: The Corliss Steam Engine in the Late-Nineteenth-Century United States”. In: *The Journal of Economic History* 64.1, pp. 61–99.

Rosenbloom, Joshua L. (Mar. 1990). “One Market or Many? Labor Market Integration in the Late Nineteenth-Century United States”. In: *The Journal of Economic History* 50.1, pp. 85–107.

— (Sept. 1996). “Was There a National Labor Market at the End of the Nineteenth Century? New Evidence on Earnings in Manufacturing”. In: *The Journal of Economic History* 56.3, pp. 626–656.

— (1998). “The Extent of the Labor Market in the United States, 1870-1914”. In: *Social Science History* 22.3, pp. 287–318.
— (2002). *Looking for Work, Searching for Workers: American Labor Markets during Industrialization*. Cambridge University Press.
Rosenbloom, Joshua L. and William A. Sundstrom (July 2003). *The Decline and Rise of Interstate Migration in the United States: Evidence from the IPUMS, 1850-1990*. NBER Working Papers 9857. National Bureau of Economic Research, Inc.
Rossin-Slater, Maya (2017). *Maternity and Family Leave Policy*. NBER Working Papers 23069. National Bureau of Economic Research, Inc.
Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek (2021). *IPUMS USA: Version 11.0 [dataset]*.
Saavedra, Martin and Tate Twinam (2020). “A machine learning approach to improving occupational income scores”. In: *Explorations in Economic History* 75.C.
Sanders, Carl and Christopher Taber (July 2012). “Life-Cycle Wage Growth and Heterogeneous Human Capital”. In: *Annual Review of Economics* 4.1, pp. 399–425.
Sandner, Malte, Stephan L. Thomsen, and Libertad González (Oct. 2020). *Preventing Child Maltreatment: Beneficial Side Effects of Public Childcare Provision*. Working Papers 1207. Barcelona School of Economics.
Schönberg, Uta and Johannes Ludsteck (2014). “Expansions in Maternity Leave Coverage and Mothers’ Labor Market Outcomes after Childbirth”. In: *Journal of Labor Economics* 32.3, pp. 469–505.
Sobek, Matthew (1996). “Work, Status, and Income: Men in the American Occupational Structure since the Late Nineteenth Century”. In: *Social Science History* 20.2, pp. 169–207.
Sokoloff, Kenneth L. and B. Zorina Khan (June 1990). “The Democratization of Invention During Early Industrialization: Evidence from the United States, 1790-1846”. In: *The Journal of Economic History* 50.2, pp. 363–378.
Solis, Alex (2017). “Credit Access and College Enrollment”. In: *Journal of Political Economy* 125.2, pp. 562–622.
Solon, Gary (2004). “A model of intergenerational mobility variation over time and place”. In: *Generational Income Mobility in North America and Europe*. Ed. by Miles Corak. Cambridge University Press, pp. 38–47.
Song, Xi, Catherine G. Massey, Karen A. Rolf, Joseph P. Ferrie, Jonathan L. Rothbaum, and Yu Xie (Jan. 2020). “Long-term decline in intergenerational mobility in the United States since the 1850s”. In: *Proceedings of the National Academy of Sciences* 117.1, pp. 251–258.
Stecker, Margaret (1937). *Intercity differences in costs of living in March, 1935, 59 cities*. Washington: U.S. Govt. Print. Office.

Sundstrom, William A. and Joshua L. Rosenbloom (Oct. 1993). “Occupational Differences in the Dispersion of Wages and Working Hours: Labor Market Integration in the United States, 1890-1903”. In: *Explorations in Economic History* 30.4, pp. 379–408.

Terrier, Camille, Parag A. Pathak, and Kevin Ren (2021). *From Immediate Acceptance to Deferred Acceptance: Effects on School Admissions and Achievement in England*. NBER Working Papers 29600. National Bureau of Economic Research, Inc.

Tô, Linh T. (2018). *Competition and Career Advancement: The Hidden Costs of Paid Leave*. Working Papers.

Traiberman, Sharon (Dec. 2019). “Occupations and Import Competition: Evidence from Denmark”. In: *American Economic Review* 109.12, pp. 4260–4301.

Tyack, David (1974). *The one best system : a history of American urban education*. Cambridge, Mass. : Harvard University Press.

U.S. Department of Labor (1899). *Thirteenth Annual Report of the Commissioner of Labor*. Vol. 2. Washington: U.S. G.P.O.

— (1900). *Fifteenth Annual Report of the Commissioner of Labor*. Washington: U.S. G.P.O.

Wantchekon, Leonard, Marko Klašnja, and Natalija Novta (2015). “Education and Human Capital Externalities: Evidence from Colonial Benin”. In: *The Quarterly Journal of Economics* 130.2, pp. 703–757.

Ward, Zachary (Nov. 2019). *Intergenerational Mobility in American History: Accounting for Race and Measurement Error*. CEH Discussion Papers 10. Centre for Economic History, Research School of Economics, Australian National University.

— (Oct. 2020). “The Not-So-Hot Melting Pot: The Persistence of Outcomes for Descendants of the Age of Mass Migration”. In: *American Economic Journal: Applied Economics* 12.4, pp. 73–102.

— (forthcoming). “Internal Migration, Education and Upward Rank Mobility:Evidence from American History”. In: *The Journal of Human Resources*.

Williamson, Jeffrey G. and Peter H. Lindert (Nov. 1980). “Long-Term Trends in American Wealth Inequality”. In: *Modeling the Distribution and Intergenerational Transmission of Wealth*. NBER Chapters. National Bureau of Economic Research, Inc, pp. 9–94.

Woittiez, Isolde and Arie Kapteyn (1998). “Social interactions and habit formation in a model of female labour supply”. In: *Journal of Public Economics* 70.2, pp. 185–205.

Wozniak, Abigail (2010). “Are College Graduates More Responsive to Distant Labor Market Opportunities?” In: *Journal of Human Resources* 45.4, pp. 944–970.

Wright, Nicholas A. (2021). “Need-based financing policies, college decision-making, and labor market behavior: Evidence from Jamaica”. In: *Journal of Development Economics* 150.C.