Essays on Economic Inequality and Mobility

René Karadakic
Norwegian School of Economics

Acknowledgements

Choosing to study economics for sure was not a choice I made out of my overarching passion for the field. I simply had nothing better in mind and people with business and economics degrees seemed to do fairly well in the labour market. In my naïve mind, I thought economics was nothing more than central banking, interest rates, and a way to potentially get a job in finance later on. Little did I know at that time, but luckily throughout my academic journey I had the chance to meet many people who showed me through their passion, dedication, and knowledge for economics and the social sciences what this field truly has to offer.

Not only did many of these friends, companions, advisors, and acquaintances infect me with their interest and motivation for the social sciences, they also are the reason why I did not give up the ambition to pursue my PhD. In particular, I want to sincerely thank my three brilliant advisors. Foremost, Aline, whom I will be eternally grateful to for making the PhD not only a less lonely journey, but also for her unconditional support in all aspects during this PhD. I will always remember the times you offered guidance and resolved my self-doubts. Without you, the journey of my PhD would have been a short-lived one. Thank you also, Kjell, for all the times I could simply drop by your office and ask you for help. Your support was undoubtedly an important factor in this journey and for this I would like to extend my sincere gratitude. Furthermore, I would like to thank Alexander, you are not only a great advisor and role model but also helped me think about what is possible rather than what is not. This optimism is something truly inspiring.

Without my friends and colleagues, this journey would have also been a different one. First, thank you, Eirik, Arne, Alessandro, and Pablo, for sharing despair, jokes, and many long hours at the office during the first year. Without you, the first year would have been a lot harder and significantly less fun. Thank you to Eirik in particular for our many inspiring conversations, discussions, and sharing your immense knowledge about Norwegian economic history. I also want to thank Laura, Fanny, and Daniel for being the "best best" colleagues when it comes to socialising and telling me not to take work too seriously all the time. My gratitude also extends to Akshay and Marlis for inspiring me with your tremendous curiosity for research. Thank you to Osama for being such a centre of tranquillity and positivity in the often stressful days at the office. Furthermore, I want to thank all the other PhD students at the department of economics. The community you all cherish is the reason I always look forward to coming to the office.

Additionally, I want to thank Andreas, Katrine, and Sam for their time, constructive feedback, and support during the job market season. I always enjoyed our meetings and your input made a massive difference in this slightly terrifying and stressful time. I also want to thank my two co-authors, Ulrika and Joachim, for the many hours of pandemic Zoom calls. My gratitude also extends to the department of economics and in particular FAIR and the labour group. The support in terms of time, resources, and feedback that all of you provide is a key ingredient for the fantastic research environment here at the department. I could not have wished for a better place to pursue my PhD.

Lastly, I want to thank my family who has always supported me in my educational and career choices. To my mum for letting me make my own decisions about my education already very early in life, but still offering support if needed. To my dad for endlessly quizzing me on our weekend hikes and also to my brother for always having an open ear and a second opinion for important decisions.

René Karadakic
Bergen, January 2023

Contents

Introduction

Economic resources, such as income, food, technology, and human capital, have important implications for how well and for how long we are able to live our lives. This pattern is observable across societies, for example, by looking at the positive correlation between GDP per capita and life expectancy (see, e.g., Figure 1). But also on a less aggregate level, we see that individuals with higher access to economic resources tend to live longer (Schwandt et al., 2021; Kinge et al., 2019) and live healthier lives (Lindahl, 2005).[1] If economic resources play an important role in shaping longevity and other measures of a successful life, the distribution of such resources matters and influences the outcomes of individuals.

This thesis covers three broad topics, each touching upon a different aspect of how economic inequality, the distribution of economic resources, impacts different aspects of our lives. The first chapter looks at how inequality in income is related to mortality of individuals at different stages in their lives. Understanding the implications of economic inequality for mortality has received a lot of attention in recent years, and an extensive literature has investigated the relationship between income, mortality, and morbidity of

[1]There are many definitions besides longevity that capture success in life, but longevity is, however, one that is universally, across time and societies, seen as desirable.

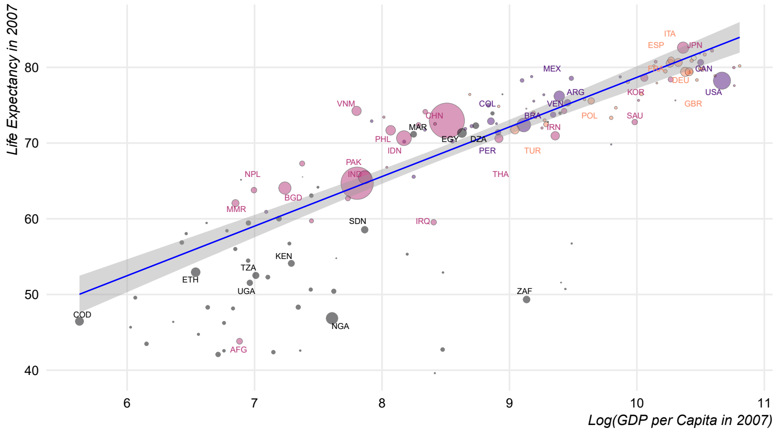


Figure 1: Association between GDP per Capita and Life-Expectancy in 2007.
Note: The figure plots the logarithm of GDP per capita on the x-axis and life-expectancy on the y-axis. The size of the bubbles indicates the population size of countries and the colour indicates the continent of the country. ISO labels are included for countries with a population above 25 Million. The blue line is fitted using a linear regression weighted by population size. Data for this figure was obtained from Bryan (2017).

both individuals and societies (Case, Lubotsky and Paxson, 2002; Currie and Schwandt, 2016*b*; Case and Deaton, 2017). The second topic also deals with a fundamental question that economists have been following for a long time, namely how the economic success of parents influences the economic success of their children and how this relationship changes over time (Becker and Tomes, 1979; Solon, 1999). The last main topic of this thesis delves into the determinants of inequality and how infrastructure and technology adoption many decades ago impacts differences in longevity between individuals up until today.

All three chapters of this thesis combine microeconometric techniques with data sources from administrative or historical records, in order to answer questions related to the consequences and determinants of socio-economic inequality. The first chapter, titled **Income Inequality and Mortality: A Norwegian Perspective**, maps the evolution of inequality in mortality in Norway over time. This chapter was part of a cross-country collaboration with research groups from eleven different OECD countries where our chapter comprised the Norwegian contribution, which had two main goals: i) providing comparable descriptive evidence on the evolution of inequality in mortality over time following the framework of Currie and Schwandt (2016*b*), and ii) focussing on aspects of inequality in income and mortality that are particularly important in the Norwegian setting.

In the US, recent evidence shows that differences in life expectancy by income increased in high-income areas, while for some disadvantaged groups, there is a downward trend in life expectancy (Chetty et al., 2016). During a similar time period, inequality in mortality at older ages has increased in the US, while inequality in mortality among younger aged individuals has declined (Currie and Schwandt, 2016*a*). Despite this recent evidence between income inequality and mortality, certain aspects remain poorly understood. First, how do increases in inequality connect to changes in mortality when there is a comprehensive social safety net in place that guarantees a living wage and free healthcare to all residents? In addition, we still do not fully understand whether closing mortality gaps at a local geographic level would close inequality in mortality also within a geographic region. A large body of literature in economics and health sciences has also established the particularly crucial role of the in-utero period and the first years in life for later life health and economic outcomes (Almond, Currie and Duque, 2018; Doyle, 2020).

Given these findings, in the first part of this chapter, using the internationally comparable framework, we use an inequality measure that is based on the poverty share of the population within a local area and then compare mortality rates for specific age groups in a particular year from the richest to the poorest five percent of Norwegian municipalities. We are therefore comparing mortality rates in more and less deprived areas of Norway over time comparably to other OECD member countries. In the second part, we then delve deeper into the evolution of infant mortality, defined as the death of a child within the first year of life, and how infant mortality at different times during the first year of life evolved for individuals from different socio-economic backgrounds. [2]

By combining rich administrative data from Norway, we are able to focus on both differences in mortality across Norwegian municipalities and individual-level differences in infant mortality where the underlying inequality measure is based on the individual income of the parents of a child. Results using the internationally comparable framework show that mortality in Norway decreased across all age groups between 1990 and 2018. Any income gradients in mortality that existed in 1990, in particular for older age groups, entirely disappeared at the geographic level by 2018. Furthermore, we show that inequality in infant mortality at the geographic level was effectively eliminated by the 1960s, while individual-level income gradients in infant mortality persisted about four decades longer. Lastly, we show that individual-level income gradients in infant mortality developed differently depending on the timing of infant death during the first year of life. By 2018, differences in infant mortality between children from

[2] Infant mortality is an extreme health outcome and an indicator for health of infants more generally. Disproportionately high rates of infant mortality in certain groups usually mask other less severe underlying health issues.

high and low-income families have been closed.

Overall, this first chapter complements a growing literature on the relationship between income and mortality in high-income countries (Chetty et al., 2016; Currie and Schwandt, 2016*b*; Schwandt et al., 2021) and points towards the importance of differences between local and individual-level inequality in the design for public policies targeted at disadvantaged groups in society.

In the second chapter of this thesis, the focus shifts to how economic resources of parents matter for the economic success of their children. Among economists and sociologists, this concept is generally known as intergenerational mobility and is often interpreted as a measure of equality of opportunity. In the paper titled **Intergenerational Mobility Trends and the Changing Role of Female Labour**, we study how the increased labour market participation of women during the second half of the twentieth century impacted the relationship between the economic status of parents and that of their children. This chapter, therefore, connects back to the main theme of economic inequality by looking at how the distribution of resources in the parent's generation is linked to the distribution of economic resources in the child's generation.

Our paper starts by estimating trends in intergenerational rank associations (IRAs) for cohorts of children born between 1951 and 1979. IRAs measure the degree of intergenerational mobility in a population, meaning the correlation in income between parents and children. The higher the IRA, the lower intergenerational mobility is, and vice versa (Dahl and DeLeire, 2008; Chetty et al., 2014*a*). Since parental income during this period is mostly a combination of paternal and maternal incomes, increased female labour market participation on both the extensive and intensive margins will influence the measurement of the IRA.

The "grand convergence" in labour market conditions between men and women, and its implications for the measurement of the IRA, is a priori unclear. On the one hand, when female labour supply increases, the relative position of a woman in the female income distribution arguably reflects her underlying skill better. All else equal, this puts upward pressure on the intergenerational persistence of income. On the other hand, however, maternal income represents a larger share of the combined parental income. If maternal income is initially less informative than paternal income with respect to the child's income potential, this would put downward pressure on measures of intergenerational income persistence such as the IRA. Due to constraints on the quality of linked survey data, it has been proven difficult in the past to estimate trends in IRAs for men and women separately and jointly (Chadwick and Solon, 2002; Björklund and Jäntti, 1997; Blanden et al., 2004), the extent to which the secular trend in female labour supply affects trends in measures of intergenerational mobility is therefore largely unexplored.

In order to overcome this data limitation, we turn to Denmark, Norway, and Sweden, three Scandinavian countries with administrative data that allow us to follow income records for individuals born between 1951 and 1979 linked to their mothers' and fathers' respective income and labour market records. Similar to large parts of the previous literature, we then estimate IRAs using OLS regressions of a child's percentile ranking in the income distribution on the parental percentile ranking, separately by the birth year of the child.[3] The resulting trend in the IRAs shows a significant increase in IRAs from birth cohorts born in the 1960s up

[3] Parental income is defined as the average income of the father's and the mother's labour earnings. Both parental and child income rankings are performed separately by the child's birth year.

until the end of the 1970s. For the period between 1951 and 1960, only modest trends have been found. Between 1961 and 1979, the IRA increased by 39

When breaking down mobility trends by the gender of the parents and children, we find that while correlations between sons and their fathers, who are not subject to increasing labour market participation, are mainly constant, all combinations involving mothers and/or daughters trend upwards over time. Conducting a similar analysis on Panel Study of Income Dynamics (PSID) data from the US, we find similar patterns, although with lagged timing, thus suggesting that this is not purely a Scandinavian phenomenon. We argue that increasing participation, rather than changing skills transmitted across generations, is the reason for the gender-specific patterns in income correlations across generations.

Our paper corroborates this in two ways. In a first step, we build a model of gender-specific mobility and latent productivity calibrated to the Scandinavian data. We use this model to better understand how different factors, such as assortative mating, skill transmission across generations, and the importance of skills for earnings, contribute to the trend in intergenerational mobility. The result of this exercise suggests that while the first two factors contribute little to the overall trend, the latter - the importance of skills for earnings - explains a large proportion of the trend. We interpret this as female earnings increasingly reflecting the earnings potential of women.

In a second step, we then empirically test this hypothesis by combining records on education, occupations, and income in order to capture latent economic status. Following previous research by Vosters and Nybom (2017); Vosters (2018); Adermon, Lindahl and Palme (2021), we argue that life-time income is not a good proxy for female economic status, especially when going back further in time. Combining information on education, occupation, and income and weighting them following the methodology laid forth in Lubotsky and Wittenberg (2006) allows us to improve upon our measure of female economic status. The corresponding results indeed suggest that rank correlations in latent economic status do not show a trend similar to the trend in father-son correlations, thereby corroborating the results of our calibrated model.

The main takeaway from this chapter is that higher income rank associations between children and parents do not necessarily need to reflect a lower degree of social mobility over time if the underlying measure changes, as in our case due to significant changes in female labour market participation.

The third and last chapter of this dissertation moves from pure consequences of economic inequality to its determinants. The chapter titled **Consequences of Rapid Structural Change - Evidence from Norwegian Hydropower Expansions** sets out to estimate the impact of childhood exposure to rapid structural change on the longevity of individuals. The setting for this is Norway during the beginning of the 20th century, when small rural communities suddenly started to industrialise due to their newly gained access to hydroelectric power. Rapid change during this time was often accompanied by negative externalities, such as poor hygienic environments, crowded housing, and the spread of infectious diseases, while at the same time offering improved economic prospects for individuals. While previous research has extensively documented the transformative effects of new technologies on local economies on various aspects of society and the economy (see e.g. Allen (2009); Lewis and Severnini (2020); Donaldson (2018)), the long-term effects, in particular for the longevity of individuals growing up during this rapid transformation of local communities, are ambiguous due to both negative and positive short-term aspects of structural change. As a consequence, evidence on the longevity effect of structural change is limited.

Choosing the Norwegian context to answer this question has two main advantages. First, Norway's industrialisation was heavily fuelled by hydroelectric power, which can be seen as local shocks to small communities that consequently shifted from predominantly agricultural production to manufacturing, thereby providing the sharp changes in the local childhood environments important for long-run health outcomes (Venneslan, 2009; Leknes and Modalsli, 2020). Second, data on hydropower adoption can be linked to individual-level census data, which in turn can be linked to novel death data by applying the new data linking methodology proposed by Abramitzky et al. (2021). The main result of this chapter indicates that growing up in rapidly changing local communities increases the longevity of individuals by ten months on average, and that this effect is significantly driven by individuals from higher socio-economic backgrounds.

In a second step, I delve deeper into the causes for this rise in longevity. Using data from historic records on municipalities, I show that employment in manufacturing, as a response to hydropower adoption, significantly rises in Norwegian municipalities, while agricultural employment stays constant. Moreover, income and wealth per taxpayer in municipalities rises significantly in response to hydroelectric power adoption. This increase in manufacturing and economic prospects leads to internal immigration from other parts of Norway and increases socio-economic inequality in hydropower communities. The rapid population growth in these local communities also puts stress on the public health environment and leads to the increased spread of common infectious diseases in the years following hydropower adoption.

Overall, these results suggest that childhood exposure to structural change and economic activity does not necessarily lead to shorter lives for individuals growing up through this transformation process. On the contrary, the increased economic resources in local communities that structurally transform lead to, on average, longer lives of individuals, despite a simultaneously deteriorating public health environment. This suggests that the economic prospects of rapid structural transformation significantly outweigh the negative aspects (e.g. spread of infectious disease) in terms of individual longevity.

This dissertation covers a broad range of topics all related to economic inequality and mobility, but what are the main takeaways from this dissertation? I think that the most important finding is that economic inequality has profound effects on many aspects of individuals' lives and, through that, on societies more generally. The first chapter documents the importance of economic inequality for mortality and shows how policies intended to benefit society as a whole might benefit certain groups earlier and more directly, even in societies that are comparably more equal in their access to resources, such as Norway. The second chapter provides important insights into the difficulty of measuring economic mobility over a longer period of time, especially when the structure of our society and the labour market changes in profound ways. It also points out that univariate measures of economic mobility perhaps lead to extreme conclusions about change in intergenerational mobility over time. The main takeaway of the last chapter is that changes to local economies decades earlier have implications for the outcomes, but also the inequality in these outcomes.

Chapter 1

Income Inequality and Mortality: A Norwegian Perspective

Abstract: While Norway has experienced income growth accompanied by a large decline in mortality during the past several decades, little is known about the distribution of these improvements in longevity across the income distribution. Using municipality-level income and mortality data, we show that the stark income gradient in infant mortality across municipalities in the 1950s mostly closed in the late 1960s. However, the income gradient in mortality for older age categories across municipalities persisted until 2010 and only flattened thereafter. Further, the infant mortality gap between rich and poor Norwegian families based on individual-level data persisted several decades longer than the gap between rich and poor municipalities and only finally closed in the early 21st century.

This chapter is written together with Aline Bütikofer and Kjell G. Salvanes. The authors gratefully acknowledge comments by Janet Currie, Sonya Krutikova, Sarah Cattan, Hannes Schwandt, James Banks, and seminar participants at the Institute for Fiscal Studies and the Norwegian School of Economics. This work was partially supported by the Research Council of Norway through project No. 275800 and through its Centres of Excellence Scheme, FAIR project No. 262675 and by NORFACE DIAL grant No. 462-16-050. Bütikofer: aline.buetikofer@nhh.no; Karadakic: rene.karadakic@nhh.no; Salvanes: kjell.salvanes@nhh.no. Address for all authors: Department of Economics, Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway.

1.1 Introduction

During the twentieth century, high-income countries experienced growth in real incomes accompanied by a historically unique decline in mortality. This negative association between income and mortality has been well documented across countries and the income gradient in mortality—the fact that relatively richer people live longer—is observed throughout the income distribution within countries (Cutler, Deaton and Lleras-Muney, 2006; Case, Lubotsky and Paxson, 2002). Various reasons exist for the gradient in mortality, including that compared to the rich, the poor have higher exposure to economic and social stress, lower use of health care services, different health behaviours and settle in different neighbourhoods. Recently, the differences in life expectancy by income even increased in many high-income countries and there is evidence that life expectancy among the most disadvantaged groups in the US is on a downward trend (see, e.g., Chetty et al., 2016). Parallel to the decrease in life expectancy, inequality in mortality at older ages has increased in the US since the 1990s, while inequality in mortality among young people has decreased (Currie and Schwandt, 2016*b*).

Nevertheless, several aspects of the relationship between income and mortality remain not well understood. For instance, a question remains whether the recent increase in inequality in mortality is also evident in the context of a comprehensive welfare state like Norway, where a social safety net guarantees a living wage and access to a public health care system that is paid for from general taxation and that guarantees free medical care to all residents.[1] Moreover, we know little about the role of a social safety

[1] Kinge et al. (2019) documents that socioeconomic status is positively correlated

net, public health measures targeted at the poorest localities, and medical innovations in the decline in the mortality gradient among the young. Lastly, we still do not understand well whether closing the mortality gap at the local level also closes the within-location differences between the rich and poor.

In this paper, we first take a closer look at the inequality in age-group and gender-specific mortality rates in Norway over the past 30 years. To date, (Case and Deaton, 2015; Currie and Schwandt, 2016*a*,*b*; Currie, Schwandt and Thuilliez, 2020) document the trends in socioeconomic inequality in mortality in the US, France and Canada. While income inequality in Norway (as proxied by the Gini coefficient) is substantially lower than in the US, Canada or France (World Inequality Database, 2020), it has been on an increasing trend over the past 30 years, similar to other Western countries (Aaberge, Atkinson and Modalsli, 2020).[2] To analyse whether similar trends in the inequality in mortality rates also exist in more equal societies such as Norway, we followed Currie and Schwandt (2016*b*) and compared inequality in mortality across different municipalities ranked by income.

Changes in social conditions and medical progress are often first visible in infant mortality because mortality in younger ages typically reacts quicker to transformations influencing the whole of society. Moreover, inequality in adult health is documented to have its roots in childhood (Case, Lubotsky and Paxson, 2002). Hence, changes in infant mortality may be important determinants of the developments in longevity and inequality in mortality. As we usually measure life expectancy at mid-life, the measure necessarily omits the health and well-being of children. Therefore, in a second

with life expectancy in Norway.

[2]Note that inequality in Norway is still among the lowest of the countries analysed in this volume.

step we examined how inequality in infant mortality changed over a period of 70 years. Our analysis period started after the invention of antibiotics and the resulting dramatic decrease in deaths from infectious disease (Easterlin, 1995). Nevertheless, we investigated a period defined by major technological, medical and pharmaceutical innovations including the innovation of respirators, the introduction of many vaccines for widespread childhood illnesses and the finding that maternal smoking and environmental toxins negatively affect unborn children.

However, such innovations do not necessarily affect individuals equally across socioeconomic groups. While disease campaigns and new pharmaceuticals lowering mortality from poverty-related diseases decrease health inequality (Bhalotra and Venkataramani, 2015; Bütikofer and Salvanes, 2020), medical improvements in recent decades, for example in cancer treatment, are shown to increase the take-up of affluent groups first, and thereby serve to reinforce health inequalities (Glied and Lleras-Muney, 2008). Moreover, public health interventions, in particular information campaigns, may have a differential effect across socioeconomic groups if richer people are responding faster to advice (de Walque, 2010; Aizer and Stroud, 2010; Kjellsson, Gerdtham and Lyttkens, 2011). Hence, studying whether and how inequalities in mortality are changing over time is central to understanding the reasons behind the current mortality gradient and its future development.

Although the neighbourhood can have large effects on children's long-term outcomes (see, e.g., Chetty and Hendren, 2018), mortality—particularly infant mortality—may not only depend on the local area with its health care offerings, but also on the family and its health behaviour. Therefore, we leveraged the Norwegian register data and focused on both municipality- and individual-level income and thereby considered the infant mortality gradient across both municipalities (1951–2018) and households (1967–2018). This allowed us to distinguish between innovations improving and levelling out health care access or public health measures both across and within municipalities.

We present four key findings. First, we find that when using an internationally comparable framework, mortality decreased in Norway across all age categories between 1990 and 2018. While there were small income gradients in mortality in 1990 for some older age categories at the local level, mortality decreased more in disadvantaged municipalities and there was no difference in age-specific mortality rates across high- and low-poverty areas in Norway in 2018. Second, we show that infant mortality decreased in Norway between 1951 and 2018. While there were large disparities in infant mortality rates across rich and poor municipalities in 1951, these effectively closed by the end of the 1960s. This period of convergence in the mortality rate between rich and poor areas coincided with the rollout of better health care access for all and the expansion of the Norwegian welfare state. Third, we show that closing the infant mortality gap between the rich and the poor at the individual level took about 40 years longer than at the municipality level. That is, infants from the 10% richest families had a substantially lower mortality rate than those from the 10% poorest families until about 2010. Subsequently, the Norwegian infant mortality gap has closed at both the local and individual level.

There are several reasons why these inequalities in infant mortality across municipalities closed much earlier than across individuals. For example, the decrease in inequality across local areas in Norway coincided with the introduction of comprehensive social welfare institutions, universal access to infant health care, and extensive disease control and vaccination campaigns (Pekkarinen, Salvanes and Sarvimäki, 2017; Bütikofer, Løken and Salvanes, 2019; Bütikofer and Salvanes, 2020). Finally, we show that absolute mortality in the first few days of life and the mortality gap between the richest and the poorest Norwegian families changed prior to the 1980s, whereas mortality at 2–12 months post-birth changed dramatically in the 1990s and the income gradient for this age category completely vanished after 2010. Importantly, high-income families always led the sharp decreases in infant mortality. This suggests that advantaged groups can leverage the various health innovations more rapidly and adapt their behaviour more quickly to new knowledge.

This paper complements the growing literature on the relationships between income and mortality in high-income countries (Currie and Schwandt, 2016*b*; Chetty, Hendren and Katz, 2016). First, we document—in an internationally comparable framework—that mortality has been decreasing in all Norwegian municipalities over the past 30 years and that inequality in mortality among older people has fallen. Second, we study trends in the mortality gradient over 70 years to shed light on whether the advent of the Norwegian welfare state was able to close the existing mortality gaps between the rich and poor. Moreover, we expand the literature by analysing local- and individual-level inequality in infant mortality to better understand whether future policies should target disadvantaged areas or disadvantaged families more directly.

The remainder of the paper is structured as follows. Section 1.2 describes the data. Section 1.3 provides the results for the internationally comparable framework. Section 1.4 presents the empirical findings on infant mortality. Section 1.5 concludes.

1.2 Data

To analyse the relationship between income inequality and mortality, we used various Norwegian register data. These population-wide panel data are primarily provided by Statistics Norway (SSB) and the Norwegian Public Health Institute (FHI). The SSB data include several registries for population, education, earnings and residency. We could link these data sets using individual identifiers and include information on residence, educational attainment and earnings. Parental identifiers enabled us to link parents and their children. The FHI maintains the medical birth registry and the death registry, which can be linked to other registry data. The medical birth registry provides information about the timing of infant deaths, birth weight, gestational age, and some indicators of the risky health behaviour of mothers (e.g. smoking). This enabled us, for example, to connect a child's birth outcomes to its parents' socioeconomic status. In addition, we made use of the detailed information on year of death and the municipality of residence at death, both of which are available after 1951.

Our income inequality measures were based on earnings data from 1967 to 2018. This earnings measure included discounted labour earnings, taxable sickness benefits, parental leave benefits, pensions, and unemployment benefits. For the death counts, we combined information from the death and medical birth registries.[3]

In Section 1.3, where we provide internationally comparable results on the mortality gradient in Norway, we aggregated the individual-level data to the municipality level to measure the link between poverty and mortality across local areas. We consider mu-

[3]Not all infant deaths that occur in the first few hours are included in the death registry. Therefore, we added additional information from the medical birth registry.

nicipalities to be a relevant unit of analysis as they constitute the administrative level responsible for the delivery of health care services, such as general practitioners, emergency rooms and infant health check-ups, in Norway. We used earnings to create poverty shares for all municipalities in each year.[4] We defined the poverty share as the proportion of working age individuals (20 to 60 years of age) within a municipality earning less than 50% of national median earnings in the relevant year (OECD, 2020). Using annual population data for one-year age groups together with death counts, we then computed the age- and gender-specific mortality rate at the municipality level for each year.

In Section 1.4, we focus on the relationship between income inequality and infant mortality at the individual level. Infant deaths are all deaths of children that occur between delivery until the end of the first year of life of the child. We calculated cohort-specific infant death rates per 1,000 live births for different deciles of the parental lifetime earnings distribution. To compute the parental income distribution, we ranked children by parental lifetime earnings. The lifetime earnings of mothers and fathers were calculated by taking the average incomes of individuals between ages 34 and 40 years. Income in mid-life has been shown to provide a good approximation for the lifetime earnings of individuals and reflects socioeconomic status much better than the earnings of parents at the birth of the child (Nybom and Stuhler, 2016; Bhuller, Mogstad and Salvanes, 2017; Attanasio and Nielsen, 2020). This analysis allows us to shed some light on the relationship between children's

[4]Municipalities in Norway are subject to continuous aggregation and deaggregation. Therefore, we harmonised municipalities in each year to match the municipality structure prevailing in 2019, consisting of 422 municipalities in mainland Norway. We excluded the Norwegian archipelagos in the Arctic Ocean (Svalbard and Jan Mayen) from the analysis.

socioeconomic status and infant mortality.

Table 1.1 provides some descriptive statistics for the sample used in Section 1.4. A birth cohort consisted of approximately 55,000 individuals. Overall, our sample included approximately 2.9 million live births between 1967 and 2018. Mothers were on average about three years younger than fathers and their average age at childbirth increased over the observational period. Education for parents also increased considerably, with 57% of women in 2015 obtaining some type of post-secondary degree. Likewise, earnings at birth increased substantially between 1970 and 2015. Note that we measured parental lifetime earnings between ages 34 and 40 years. For children of the youngest cohorts in our sample, this earnings measure was progressively based on earnings observations at younger parental ages and therefore did not increase in the same way as earnings at birth. As this measurement issue could influence our findings, we applied a sensitivity analysis whereby we grouped parents based on education instead, and employed three different education groups as an alternative to the lifetime earnings deciles. We classified parents into three broad educational groups based on their highest educational degree: individuals who had finished any post-secondary education (including short and long tertiary education), individuals who had only finished high school (secondary schooling), and those who had never finished high school (primary schooling).[5] For most of the analysis, the education of the mother was the primary measure, and only in cases where maternal education was not applicable did we use paternal education as our measure.

[5]This classification follows SSB using NUS-2000 codes and categories as follows: post-secondary (6,7,8), secondary (4,5), and primary (0,1,2,3)

Table 1.1: Summary Statistics: Infant Mortality Sample

	1970	**1985**	**1995**	**2005**	**2015**
Share LBW ($\leq$ 2500 gram)	0.044	0.043	0.045	0.048	0.044
Share PT ($\leq$ 37 gram)	0.054	0.057	0.067	0.069	0.061
IM Rate (per 1,000 LB)	12.28	8.14	3.90	2.46	2.08
Mother's Age at Birth	26.19	27.35	28.84	30.25	30.68
Father's Age at Birth	29.26	30.24	31.71	33.26	33.63
Mother Post-Sec. Educ.	0.15	0.32	0.39	0.53	0.57
Mother Sec. Education	0.10	0.25	0.33	0.29	0.25
Parental LT Earnings	438.70	554.90	694.40	869.90	538.10
Parental Earnings Birth	235.40	297.60	319.40	414.90	498.60

Note: Income in 1,000 Norwegian Kroner adjusted to 2015 level using CPI provided by SSB.

1.3 Trends in Mortality Inequality Across Age Groups — Part A

This section aims to provide an internationally comparable overview of the development of mortality rates at the geographic level for different age, gender and income groups in Norway following Currie and Schwandt (2016*b*). We first ranked the municipalities using their poverty shares from richest to poorest, and grouped them into bins (ventiles), each representing approximately 5% of the Norwegian population in 1991, 2000, 2010 and 2017.[6] In addition, we aggregated the age- and gender-specific mortality rates for six distinct age categories, namely 0–4, 5–19, 20–49, 50–64, 65–79 and 80 or more years. Combining the mortality measures and the poverty ventiles allowed us to compute the one-year death rates for each year, gender, age and ventile cell. We age-adjusted the age

[6] Owing to the large population size of the Oslo municipality, we divided Oslo into six sub-areas based on basic statistical units (SSB, 2020).

groups to the year 2017.

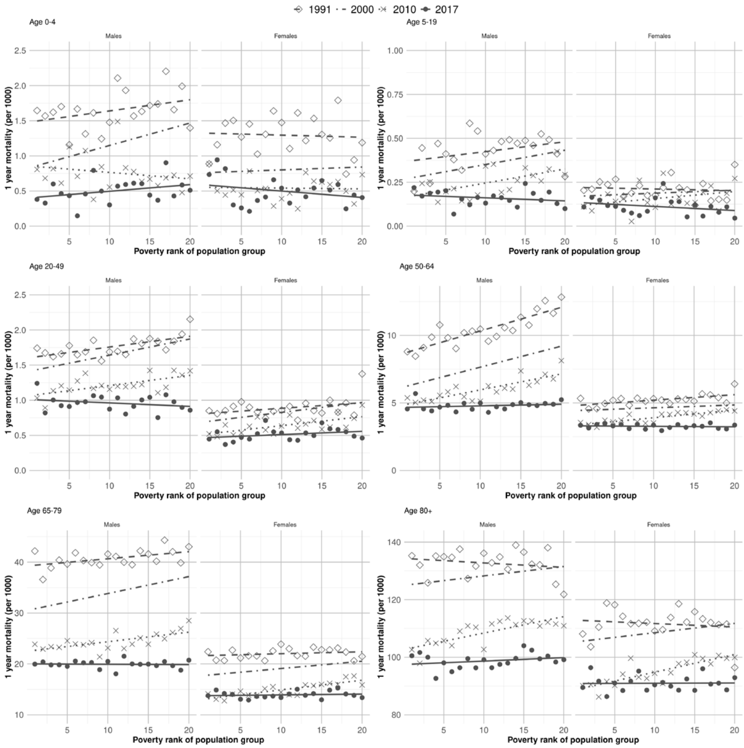


Figure 1.1: One-Year Mortality Rates by Municipality-Level Poverty Rates.

Note: Average one-year mortality rates are plotted across poverty rate percentiles. Each bin represents a group of municipalities with about 5% of the overall population in the respective year. Straight lines provide linear fits.

Figure 1.1 depicts the results for the internationally comparable analysis separately for each age group and gender in 1991, 2000, 2010 and 2017. The municipality-level poverty rank is on the x-axis, where municipalities with the lowest poverty share have the lowest poverty rank and vice versa, and the one-year age-adjusted mortality rates per 1,000 individuals is on the y-axis.[7] In Table 1.2, we report the age-specific mortality in the least and most deprived areas and the change in inequality for each group and time period. Moreover, we test whether the changes in slopes over time for each group are statistically significant.

Several patterns apply to all age groups. First, over the whole observation period, absolute mortality decreased in all age groups. Second, men of every age were more likely to die compared to women in the same age group. Even though the gap between female and male mortality narrowed over time, the gender disparities did not disappear for all age groups by 2017. Third, the income gradients in mortality were a predominantly male phenomenon at the municipality level. For women, there were no significant poverty gradients in morality. For men, however, several age categories initially depicted substantial differences in mortality between the more and less deprived areas. The fourth universal trend concerns the reductions in absolute mortality: even though the relative declines in mortality for men and women were similar in some age groups, the absolute declines in mortality were mostly driven by the large decrease in the mortality rates of men in the oldest age groups. Last, the poverty gradients in mortality declined over time, such that by the end of 2017 we did not observe any income gradients at the geographic level in any of the six age groups.

For men and women in the youngest age group (0 to 4 years), Figure 1.1 illustrates that their mortality decreased from approximately 1.7 to 0.5 per 1,000 and from 1.4 to 0.5 per 1,000, respec-

[7] Note that the results do not change substantially if we use municipality-level median and mean earnings for the ranking of municipalities instead of the poverty share.

tively. While there was no income–mortality gradient for girls in any of the years, the small gradient in mortality for boys disappeared between 2000 and 2010. The mortality in the age group 5 to 19 years was lower than for younger children. Nevertheless, the trends appear similar: a small decline in both gradient and mortality from 1991 to 2017 together with a very small mortality decline for men and no mortality inequality in any of these years for women. For the age group 20–49 years, we observed only small declines in mortality for both men and women and a very small decrease in the poverty gradient for men.

The largest gradient and decrease in gradient in mortality among men was in the age group 50 to 64 years. In 1991, men of the most advantaged ventile had a 50% lower mortality rate compared to men in the most disadvantaged ventile. By 2017, the differences across areas in mortality had evened out for both men and women in this age group. Both men and women in the age category 65 to 79 years experienced marked declines in absolute mortality of approximately 50% and 35%, respectively. For the age group 80 years and older, we observed the largest absolute declines in mortality. For men, deaths dropped by 37 per 1,000 and for women by approximately 35 per 1,000. Although there were weak gradients in mortality in the oldest age group before 2017, there was no longer inequality in mortality across municipalities in 2017.[8]

Note that the finding that mortality is equally distributed across

[8]Note that we compared all our results to national averages of the Human Mortality Database and they matched in all age categories except for the 80 years and older category, where we missed approximately 20% of deaths using our data. The difference likely arose for two main reasons. First, we used different data sources than the Human Mortality Database. Second, we only included individuals that were residents in a municipality at the time of death and therefore excluded Norwegian citizens with a place of residence outside Norway. Nonetheless, the discrepancy should not pose a problem in interpreting the gradients and trends in our mortality data for this age group.

Table 1.2: Age-Specific Mortality in Least and Most Deprived Areas and Change in Inequality

	Lowest Poverty Rate				Highest Poverty Rate				Slope Coefficient				P-Value		
	1991	2000	2010	2017	1991	2000	2010	2017	1991	2000	2010	2017	Δ^{1991}_{2000}	Δ^{2000}_{2010}	Δ^{2010}_{2017}
Men															
0-4	1.645	0.933	0.807	0.381	1.398	0.299	0.712	0.511	0.016	0.032	-0.009	0.009	0.422	0.042	0.112
									(0.010)	(0.017)	(0.010)	(0.006)			
5-19	0.199	0.361	0.209	0.219	0.283	0.500	0.293	0.100	0.006	0.008	0.008	-0.002	0.606	0.967	0.003
									(0.004)	(0.003)	(0.003)	(0.002)			
20-49	1.741	1.389	1.031	1.241	2.153	1.664	1.419	0.858	0.015	0.023	0.015	-0.005	0.280	0.238	0.003
									(0.004)	(0.006)	(0.004)	(0.004)			
50-64	8.795	5.728	5.376	4.559	12.840	9.482	8.133	5.239	0.175	0.156	0.118	0.012	0.560	0.214	0.000
									(0.024)	(0.023)	(0.019)	(0.013)			
65-79	42.180	32.520	23.870	19.970	43.030	38.880	28.510	20.740	0.140	0.333	0.188	-0.007	0.032	0.078	0.002
									(0.059)	(0.063)	(0.049)	(0.030)			
80+	135.300	127.100	102.600	100.500	121.900	115.500	110.900	99.170	-0.155	0.327	0.572	0.119	0.099	0.330	0.005
									(0.179)	(0.222)	(0.111)	(0.101)			
Women															
0-4	0.887	0.740	0.888	0.734	1.188	0.560	0.732	0.407	-0.003	0.004	-0.001	-0.009	0.616	0.643	0.430
									(0.011)	(0.009)	(0.007)	(0.008)			
5-19	0.204	0.100	0.111	0.109	0.351	0.336	0.272	0.046	-0.001	0.002	0.004	-0.002	0.467	0.553	0.049
									(0.002)	(0.003)	(0.003)	(0.002)			
20-49	0.849	0.769	0.520	0.447	1.375	1.117	0.930	0.463	0.008	0.014	0.013	0.005	0.341	0.805	0.106
									(0.005)	(0.004)	(0.004)	(0.003)			
50-64	5.335	4.541	3.553	3.343	6.410	4.299	4.406	3.358	0.041	0.021	0.057	-0.004	0.294	0.025	0.000
									(0.013)	(0.013)	(0.008)	(0.006)			
65-79	22.390	20.010	14.120	13.770	21.470	22.160	15.870	13.390	0.039	0.144	0.191	0.015	0.057	0.336	0.000
									(0.038)	(0.038)	(0.030)	(0.027)			
80+	108.100	102.300	90.100	89.510	96.430	106.000	99.930	92.930	-0.120	0.323	0.614	0.009	0.079	0.095	0.000
									(0.201)	(0.140)	(0.097)	(0.106)			

Note: Columns (1)-(8) report the means of (smoothed) 1-year mortality rates for each gender and age group in 1991, 2000, 2010 and 2017, in the bin of municipalities with lowest and highest poverty rate, respectively. Columns (9)-(12) report the coefficient of the fitted regression line in each year. Column (13)-(15) report p-values for the null hypothesis that the slopes in the respective years (see column header) are equal. Standard errors for regression coefficients are reported in parentheses.

municipality-level poverty rates does not mean that mortality in Norway is no longer income dependent. It is important to point out that the analysis presented in Figure 1.1 shows that the aggregate mortality difference between poor and rich areas evened out in 2017. In Section 1.4, we therefore focus on a municipality- and individual-level approach to describe how the gradients in infant mortality changed over time in terms of local and individual inequality.

1.4 Mortality Inequality Among Infants 1951–2018 — Part B

The great decline in infant mortality during the 20th century and its causes and consequences has been of great interest to researchers and policy makers alike. Like other developed countries, Norway experienced a marked decline in infant mortality between 1900 and 1955 from 80 deaths per 1,000 live births in 1900 to 23 deaths per 1,000 live births in 1955 (Backer, 1961). This period is often referred to as the 'Mortality Revolution' (Easterlin, 1995) and is characterised by a decrease in deaths from infectious disease. In particular, advances in public health such as water purification (Cutler and Miller, 2005; Beach et al., 2016; Alsan and Goldin, 2019) and infant health care services (Moehling and Thomasson, 2014; Hjort, Sølvsten and Wüst, 2017; Bhalotra, Karlsson and Nilsson, 2017; Bütikofer, Løken and Salvanes, 2019) and medical developments such as the introduction of antibiotics (Bhalotra and Venkataramani, 2015) contributed to the large decline. Although the decline slowed after 1955, further advances in medical technology, pharmaceutical innovations, and public health measures led to substantial decreases in infant mortality from 15 deaths per 1,000 live births in 1967 to 2.5 deaths per 1,000 live births in 2018 (see, e.g., Daltveit et al., 1997; Bharadwaj, Løken and Neilson, 2013). In this section, we investigate the development of the mortality inequality among infants at the municipality and individual levels to discuss, for example, how the expansion of the Norwegian welfare state and new medical technologies altered the gradient between income and infant mortality.

1.4.1 Methodology

Our main analysis concerns the comparison of infant mortality rates—deaths during the postnatal period until the end of the first year of life—from 1967 to 2018. We computed gender-specific mortality rates (deaths per 1,000 live births) for all infants as well as for four distinct age groups: 0–24 hours after birth, 2–7 days after birth, 8–28 days after birth and 29–365 days after birth.

To compare income gradients in infant mortality rates in Norway over time, we ranked infants according to their parental lifetime income and calculated the three-year mortality rates by decile of parental income. Note that we used three-year mortality rates as the annual rates of deaths per thousand live births for specific gender, age, decile and year groups can be volatile in a small population such as Norway. For consistent scaling, we adjusted the definition of the mortality rate in the following way:

$$\text{MR}_{d,g,t} = \frac{\sum_{t=t-1}^{1} \text{deaths}_{d,g,t}}{\text{births}_{d,g,t} \cdot 3},$$

where $\text{MR}_{d,g,t}$ is the three-year mortality rate divided by three for a specific decile d, gender of the child g, and year t. Specifically, we scaled the three-year infant mortality rate to an annual rate that enabled us to interpret it in a similar fashion as the definition of deaths per 1,000 live births. For the years 1951 and 1954 where we studied infant mortality by municipality income level, we used one-year mortality rates.

1.4.2 Infant Mortality by Gender and Age 1968–2018

As discussed, infant mortality in Norway has declined from approximately 15 deaths per 1,000 live births in 1967 to 2.5 deaths per 1,000 live births in 2018. Figure 1.2 depicts the development of infant death rates separately for boys and girls. The figure illustrates two distinct features. First, male infants were subject to an infant mortality rate that was about 40 percent higher than that of female infants in 1967. This gender mortality gap continuously declined and had almost disappeared by the beginning of the 21st century. Second, as also shown in Figure 1.2, the decrease in infant mortality for both men and women was steeper before the mid-1990s than afterwards.

Different medical innovations could affect infant death rates at different times after delivery. To investigate these differences, we plot the deaths per 1,000 live births in Figure 1.3 for four distinct age categories by gender. The first category captures infant deaths within the first 24 hours after delivery, the second category deaths occurring between 2 and 7 days after birth, the third deaths between 7 and 28 days after birth, and the last category shows all infant deaths that occurred 28 days or more after birth.

Figure 1.3 displays several distinct trends. While the decrease in the mortality rate from 1967 to 2018 was about 90% among the first two categories and about 80% among the second two cate-

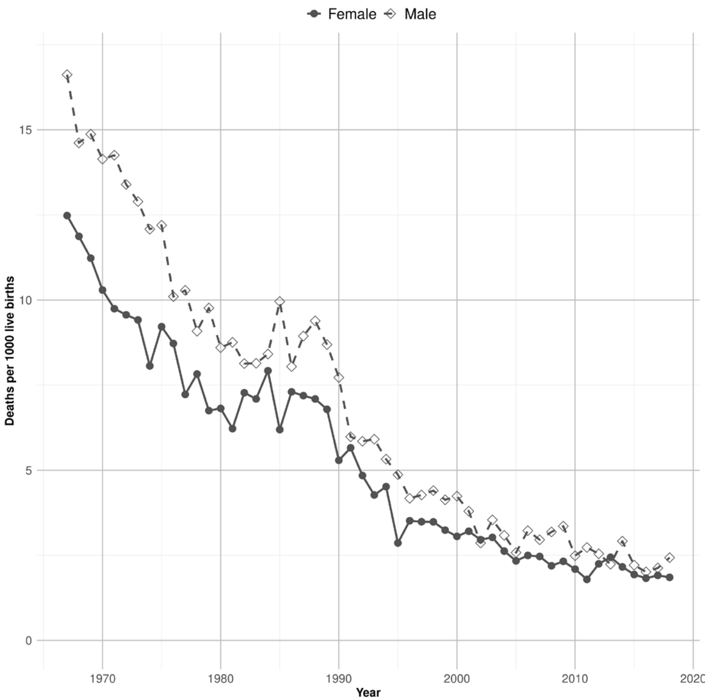


Figure 1.2: Infant Mortality 1967–2018 by Gender.
Note: Cohort-specific infant mortality by gender between 1967 and 2018. Each dot represents the cohort-specific deaths in the first year after birth per 1,000 live births.

gories, the decrease was more modest for the category containing deaths between 8 and 28 days after birth. Importantly, the sharpest declines in mortality occurred during different decades for each of these age categories. Deaths 0–24 hours after birth and deaths 2–7 days after birth fell most strongly between 1967 and 1980, likely because of technological advancements in neonatal medicine implemented during the 1970s and 1980s. In particular, treatments for infants with respiratory disorders became increasingly effective (see, e.g., Bharadwaj, Løken and Neilson, 2013). Jorgensen (2010), for example, suggests that the use of a surfactant as a treatment for respiratory distress syndrome reduced mortality from respiratory distress among infants by 40 percent in the US. In addition, the increasing availability of incubators, better respiratory management, ventilators, and overall improvements in neonatal care improved the chances of survival for low birth weights and preterm births in the critical post-delivery period (Lee et al., 1999).

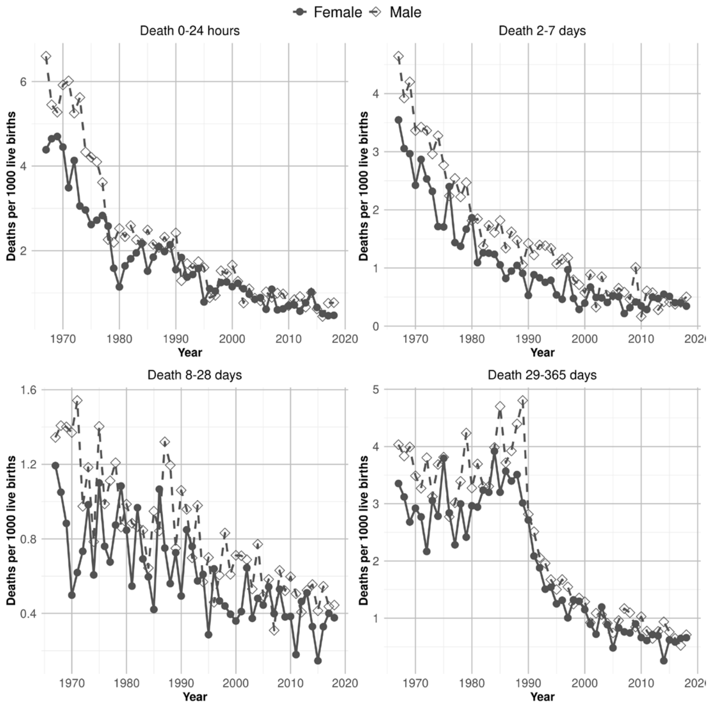


Figure 1.3: Infant Deaths 1967–2018 by Gender and Age Category.
Note: Cohort-specific infant mortality by gender and age category between 1967 and 2018. Each dot represents cohort-specific deaths in one of the four age categories (0–24 hours, 2–7 days, 7–28 days, and 29–365 days after birth) per 1,000 live births.

Although there is a steady decrease in infant mortality among the ages 8 to 28 days, the reduction does not display a clear pattern and is thus more difficult to link to a specific technological progress or public health campaign. Nevertheless, the decrease in mortality likely reflects general improvements in health care, health care access and the health behaviour of mothers.

The gender gap in the mortality rate for infants between 29 and 365 days is much smaller than for the other categories. The most visible decline in infant mortality for this category started in the late 1980s and early 1990s, where we see a drop in male infant mortality of almost 70% within just five years. This phenomenon coincides with an influential article published in one of Norway's largest newspapers discussing the association between the sleeping position of the infant and occurrences of sudden infant death syndrome (SIDS). In turn, the article was based on the results from a larger Norwegian surveillance system introduced to limit SIDS and marks the starting point of a major public health campaign encouraging parents to have their infants sleep on their backs (Daltveit et al., 1997). Later research corroborated that sleeping in a prone position influences infant respiration and their ability to move their body in life-threatening situations, both of which can have lethal consequences.

Overall, different technological developments and new medical insights reduced infant death rates at different ages. The remainder of the paper discusses whether the reductions in infant mortality observed for different age groups had differential consequences along the parental income distribution.

1.4.3 Inequality and Infant Mortality 1951–2018

Infant mortality decreased over many decades. However, the question remains of whether public health investment and widespread availability of medical technology can close the mortality gap between rich and poor children. In a first step, we extended the inequality analysis in Section 1.3 to include all years between 1967 and 2018. Although there was little gradient in mortality among the youngest group (ages 0–4 years) at the municipality level between 1990 and 2017 (see Figure 1.1), the expansion of the Norwegian welfare state and the advent of new industries in Norway such as oil and gas production in the 1960s and 1970s could have evened out the income disparities across municipalities and thereby lowered the inequality in infant mortality between 1967 and today. The results for this extension are shown in Figure 1.4; the left panel presents the results for males and the right panel for females. Both panels plot infant deaths per 1,000 live births on the y-axis and birth year on the x-axis. Municipalities are ranked by the municipality-level poverty rate. The triangles and dots identify the infant mortality for the poorest and richest 10 percent of municipalities, respectively.

Figure 1.4 displays the strong decline in infant mortality from 1967 to 2018 discussed earlier (see Figure 1.2). Although mortality is a little higher in 1967 and the decline in mortality for the poorest municipalities is slightly steeper than that for the richest municipalities, the differences between rich and poor municipalities are not significant. This indicates that geographical differences in infant mortality were already small in Norway in the late 1960s.[9] As a variety of factors affect infant deaths, including health care

[9] Note that infant mortality is an extreme health outcome and our findings do not necessarily suggest that newborns are equally healthy across all municipalities.

access, environmental influences, genetic endowment and health behaviour of the parents, particularly mothers' behaviour, there are several reasons that can explain these small geographic differences.

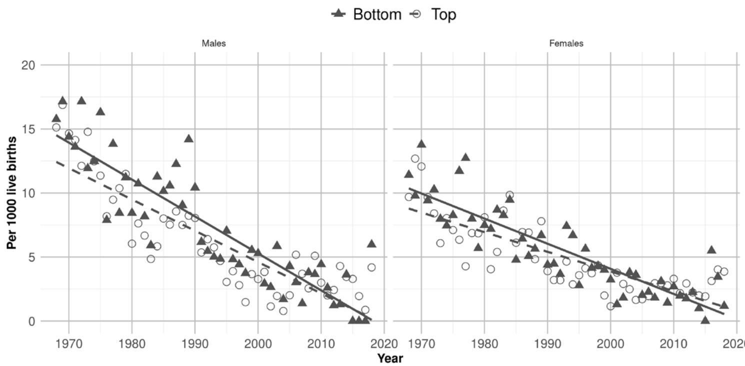


Figure 1.4: Municipality-Level Inequality of Infant Deaths 1967–2018.
Note: Cohort-specific infant mortality by gender in municipalities with the 10% highest and 10% lowest poverty rates from 1967 to 2018. Each triangle/dot represents infant deaths per 1,000 live births. Straight lines provide linear fits.

Access to infant health care was rolled out in Norway in the first half of the 20th century with the goal to reach out to everyone and to establish a unified and free primary health care system for infants (Bütikofer, Løken and Salvanes, 2019). Almost all municipalities had a centre providing infant health care in the late 1960s and in 1972, it became mandatory for all municipalities to operate state-funded health care centres for infants and mothers. The Health Directorate, through official guidelines and handbooks, regulated the services provided by the centres, with the centres being responsible for the national vaccines programme and information campaigns targeting new parents. This suggests that universal health care investments before the late 1960s may have evened out any health inequalities by geography to a large degree.

To investigate the importance of public health policies and the diffusion of medical innovations such as antibiotics before 1967, we documented inequality in mortality in the 1950s by combining infant mortality data from the Death Register in 1951 and 1954 with aggregate data on municipal-level average taxable income in the same years.[10] Figure 1.5 depicts the differences in infant deaths per 1,000 live births in 1951, 1954 and 1968 in the 10% richest and 10% poorest municipalities. The figure shows that even though geographical differences in infant mortality across municipalities were small in the late 1960s, there was a clear gradient in the early 1950s. In particular, the infant mortality rate in municipalities in the bottom decile in 1951 and 1954 was twice as high as the infant mortality rate in the top decile of the municipality-level income distribution. While the mortality rate only decreased to a small degree between 1951 and 1968 in the richest municipalities, infant deaths were decreasing dramatically between 1954 and 1968 in the poorest areas.

In addition to infant health care, there are several other reasons why the inequalities in infant mortality fell sharply in the 1950s and 1960s. Similar to the other Nordic countries, Norway introduced comprehensive social welfare institutions, which enhanced access to public services and social insurance with important implications for inequality and social mobility during the 1950s and 1960s (Pekkarinen, Salvanes and Sarvimäki, 2017). Moreover, economic recovery programmes in the aftermath of World War II brought changes to the pre-war economic structure and income

[10]The aggregate income data was compiled from the Norwegian municipality database (Kommunedatabase) provided by the Norwegian Centre for Research Data (NSD).

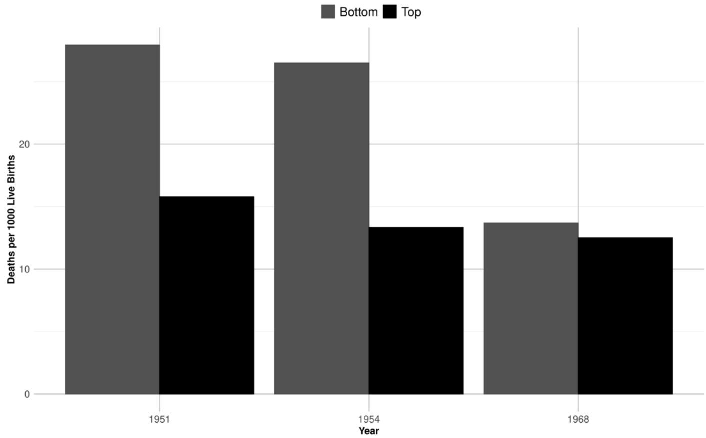


Figure 1.5: Municipality-Level Inequality of Infant Deaths 1951, 1954 and 1968.
Note: Cohort-specific infant mortality in the municipalities with the 10% highest and 10% lowest average tax incomes in 1951, 1954 and 1968. Each bar represents infant deaths per 1,000 live births.

distribution. In addition, the availability of effective antibiotics and advancements in disease detection techniques led to extensive disease control campaigns, for example against tuberculosis, which lowered the disease burden in poor Norwegian municipalities (Bütikofer and Salvanes, 2020). Moreover, 1953 saw a new law for child immunisation against diphtheria introduced, followed by a vaccine against polio in 1956 (FHI, 2017). Overall, these interventions lowered poverty or at least the consequences of being poor and had important equalising effects on the infant mortality rate across municipalities.

Even though the differences between rich and poor municipalities in infant mortality were small in the late 1960s in Norway, there may still have been large income disparities in infant mortality within municipalities. Because family socioeconomic status could affect parental behaviour or the possibilities of families shielding themselves from negative environmental influences, we additionally studied inequality at a less aggregated level. In particular, we analysed the differences in infant deaths between children born into the top and bottom deciles of the parental income distribution between 1967 and 2018. Moreover, we documented mortality inequality for different age groups to further our understanding of what policies are most central in closing the socioeconomic mortality gap.

Figures 1.6 and 1.7 depict the male and female infant death rates for the different age categories of infant death. For all age categories, male mortality rates are higher. The time patterns, however, are similar for both genders, indicating that men and women both benefited from improvements in health technologies and information at a similar point in time.

In addition to infant health care, there are several other reasons why the inequalities in infant mortality fell sharply in the 1950s and 1960s. Similar to the other Nordic countries, Norway introduced comprehensive social welfare institutions, which enhanced access to public services and social insurance with important implications for inequality and social mobility during the 1950s and 1960s (Pekkarinen, Salvanes and Sarvimäki, 2017). Moreover, economic recovery programmes in the aftermath of World War II brought changes to the pre-war economic structure and income distribution. In addition, the availability of effective antibiotics and advancements in disease detection techniques led to extensive disease control campaigns, for example against tuberculosis, which lowered the disease burden in poor Norwegian municipalities (Bütikofer and Salvanes, 2020). Moreover, 1953 saw a new law

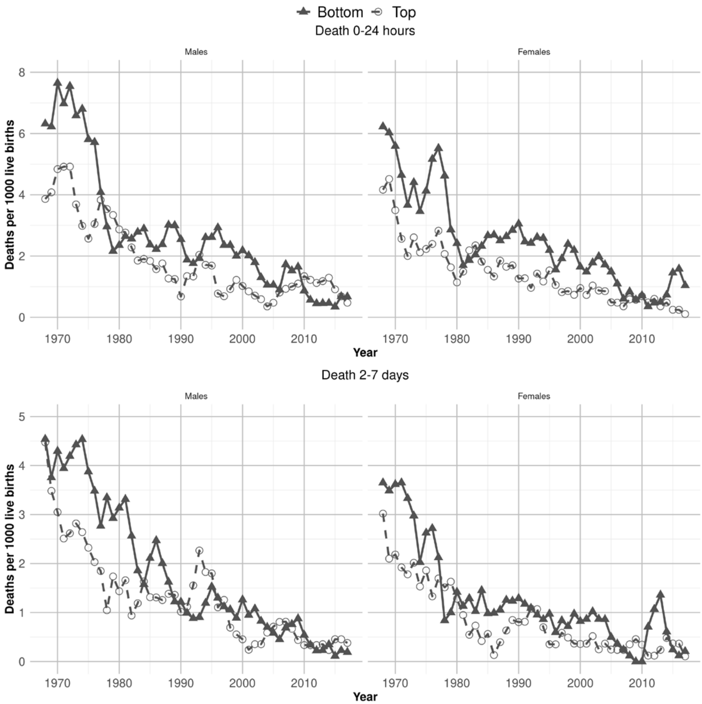


Figure 1.6: Individual-Level Inequality of Infant Deaths 0–24 Hours and 2–7 Days after Birth.

Note: Infant mortality in the top and bottom deciles of the parental income distribution by gender from 1968 to 2017. Each dot/triangle represents cohort-specific infant deaths per 1,000 live births 0–24 hours or 2–7 days after birth. Parental lifetime income is calculated by taking the average income of individuals between age 34 and 40.

for child immunisation against diphtheria introduced, followed by a vaccine against polio in 1956 (FHI, 2017). Overall, these interventions lowered poverty or at least the consequences of being poor and had important equalising effects on the infant mortality rate across municipalities.

Even though the differences between rich and poor municipalities in infant mortality were small in the late 1960s in Norway, there may still have been large income disparities in infant mortality within municipalities. Because family socioeconomic status could affect parental behaviour or the possibilities of families shielding themselves from negative environmental influences, we additionally studied inequality at a less aggregated level. In particular, we analysed the differences in infant deaths between children born into the top and bottom deciles of the parental income distribution between 1967 and 2018. Moreover, we documented mortality inequality for different age groups to further our understanding of what policies are most central in closing the socioeconomic mortality gap.

Figures 1.6 and 1.7 depict the male and female infant death rates for the different age categories of infant death. For all age categories, male mortality rates are higher. The time patterns, however, are similar for both genders, indicating that men and women both benefited from improvements in health technologies and information at a similar point in time.

The mortality gap between the poorest and richest families was quite substantial in the 1960s and 1970s among males and females in the first two categories (see Figure 1.6). As discussed in Section 1.4.2, deaths within the first 24 hours after birth declined sharply in the 1970s. Interestingly, this became evident about five years earlier for the richest individuals compared to the poorest individuals, suggesting that affluent families were faster adaptors of medical technology or acted more quickly to public awareness campaigns about the negative consequences of maternal smoking during pregnancy (see, e.g., Aizer and Stroud, 2010). Overall, the decline in infant mortality was larger for the bottom decile than the top decile. In particular, the mid-1970s brought tremendous

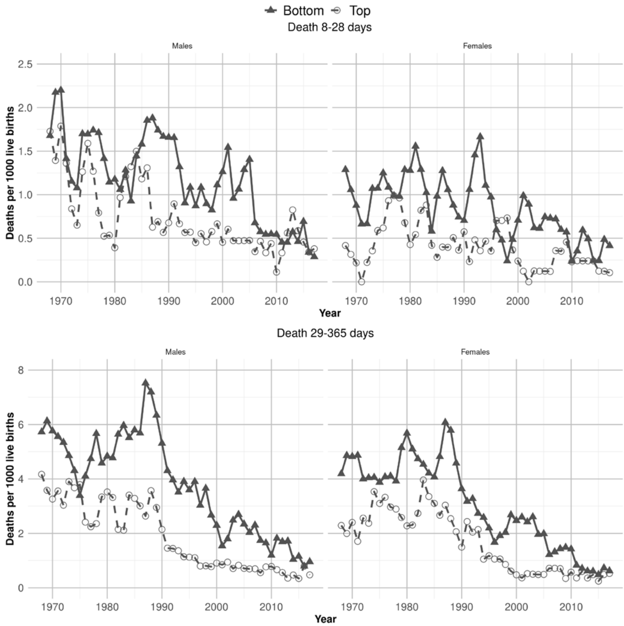


Figure 1.7: Individual-Level Inequality of Infant Deaths 8–28 and 29–365 Days after Birth.

Note: Infant mortality in the top and bottom decile of the parental income distribution by gender from 1968 to 2017. Each dot/triangle represents cohort-specific infant deaths per 1,000 live births 8–28 or 29–365 days after birth. Parental lifetime income is calculated by taking the average income of individuals between ages 34 and 40 years.

improvements regarding infant mortality for the poorest income groups. This suggests that the medical advances in neonatal care in the 1970s were likely to have reduced inequality and thereby benefited the more vulnerable disproportionately to higher socioeconomic status families.[11] The income gradient in infant mortal-

[11]Note that there is a negative correlation between socioeconomic status and the likelihood of low birth weight and preterm births. Hence, improvements in neonatal care might be more beneficial to newborns from poor families. In Section 1.4.4, we discuss this relationship in detail.

ity in the first 24 hours only disappeared between 2005 and 2010, while the mortality gap between children born into the poorest and the richest families closed in 2018. The pattern for deaths between 2–7 days after birth resembled that for the earliest timing category. The gap between rich and poor after the mid-1980s was, however, much smaller than for deaths within the first 24 hours.

The pattern for deaths between 8–28 days and 28 days after birth differed from those for the younger age categories. Deaths between 8–28 days after birth were slightly higher for infants from a lower socioeconomic background than those from a higher socioeconomic background during most of the analysis period. Due to the small number of deaths occurring for this timing category, the year-to-year variation in infant mortality in this category was very high. Nevertheless, Figure 1.7 reveals that both the mortality rate and the income differences declined over time and the rates for infants from less and more deprived backgrounds converge around 2010. Infant deaths between 29 and 365 days after birth did not decline significantly until 1990, and the gradient between the top and bottom deciles was substantial and persistent until the end of the 1980s. In 1987, the average infant mortality rate was approximately 50% higher in the bottom decile of the parental income distribution compared to the top decile. However, a substantial decrease in the mortality rate for this age category in the 1990s closed the income gap substantially. By the mid-1990s, the average infant mortality rate was 20% higher among the poorest than the richest families. As discussed, information campaigns encouraging parents to have their infants sleep on their backs in the 1990s coincide with this sharp decrease in death rates among this age category. The mortality inequality further decreased for death between 29 and 365 days after birth and disappeared around 2015.[12]

Overall, children born into both the top and the bottom part of the parental income distribution benefited from large declines in infant mortality between 1967 and 2018. The periods of decline in mortality rates mostly started among the most advantaged families first. However, the decrease was disproportionately larger among children from the lowest socioeconomic background and the income gap in infant mortality for all age groups closed by 2015. These differences in the timing of the closing of the infant mortality gap at the local and individual level are striking. Moreover, they may well explain why our results presented in Section 1.3 differ from the findings of Kinge et al. (2019), which concludes that the differences in life expectancy at 40 years of age by individual income increased from 2005 to 2015 in Norway.

1.4.4 Low Birth Weight and Preterm Deliveries

Children born with low birth weight (birth weights less than 2,500 grams) and preterm deliveries (born before the end of the 37th week of gestation) account for a large share of infant deaths. A question remains whether socioeconomic differences in low birth weight and preterm births could explain part of the link between income inequality and infant mortality and the disappearance of the income–mortality gradient in the 2000s. Low birth weight is frequently associated with higher rates of subnormal growth, illness and neuro-developmental problems (Hack, Klein and Taylor, 1995). Birth weight is determined by a variety of factors during the

[12] Note that a sensitivity analysis studying mortality differences based on parental education instead of parental lifetime income confirms that the infant mortality gap between high and low socioeconomic status at the individual level closed by 2018 for all four age categories. The results are in Appendix Figure A1.

in utero period such as nutritional intake, stress, illness, pollution and maternal health behaviour (e.g. smoking) (see, e.g., Almond, Hoynes and Schanzenbach, 2011; Aizer, Stroud and Buka, 2016; Almond, 2006; Currie and Walker, 2011; Lien and Evans, 2005) and birth weight is shown to affect later-in-life outcomes (Black, Devereux and Salvanes, 2007). Similarly, preterm births are related to poorer lung development, which also affects cognitive ability and labour market outcomes (Bharadwaj, Løken and Neilson, 2013). We leveraged detailed information in the medical birth registry to investigate how low birth weight and preterm birth relate to the decline in infant mortality and the potential underlying mechanisms.

Figure 1.8 illustrates the male and female rates of children born with low birth weight and preterm births over time. These rates are the average number of low birth weight or preterm births per 1,000 live births by decile of parental lifetime income. The figure shows a substantial socioeconomic gap in the frequency of low birth weight and preterm births prior to 2010, and the fact that between 1967 and 2010, both low birth weight and preterm births increased for the top and the bottom income deciles. There are two potential reasons for the latter. First, low birth weights and preterm birth are positively correlated with the mother's age at birth. As the average age of mothers at birth increased by four years from 1967 to 2018, these age differences might have resulted in a slight increase in low birth weight and preterm births over the analysis period. Second, the definitions of infant mortality near the threshold of viability differ between countries, and reporting standards can change over time (Chen, Oster and Williams, 2016).

Most importantly, Figure 1.8 shows that socioeconomic differences in low birth weight and preterm births decreased after about

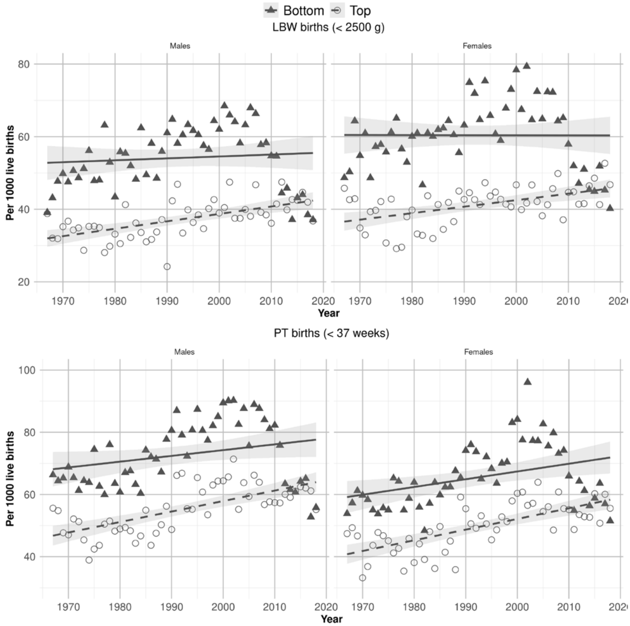


Figure 1.8: Individual-Level Inequality in Low Birth Weight and Preterm Births 1968–2018

Note: Cohort-specific numbers of children born with low birth weight (<2500 gram) and preterm deliveries (before the end of the 37th week of gestation) per 1,000 live births in the top and bottom deciles of the parental income distribution by gender from 1967 to 2018. Parental lifetime income was calculated by taking the average income of individuals between ages 34 and 40 years.

2005. After 2010, the gap in birth outcomes between the 10% richest and 10% poorest families closed. This convergence coincides with the convergence in infant mortality presented in Figures 1.6 and 1.7.[13]

As maternal smoking constrains foetal growth and increases

[13]Note that a sensitivity analysis studying differences based on parental education instead of parental lifetime income confirmed these findings. The results are presented in Appendix Figure A2.

the likelihood of preterm birth or low birth weight (see, e.g., Kramer, 1987, for an overview), socioeconomic differences in mothers' smoking behaviour could be a driver of the patterns observed above. In particular, Aizer and Stroud (2010) suggest that the information on the negative effects of smoking on health had an immediate impact on educated mothers and the health of their newborns. Less-educated mothers changed their smoking behaviour much later. Convergence in the smoking behaviour of rich and poor mothers could therefore be a driver of the closing of the socioeconomic gap in low birth weight, preterm births and the infant mortality rate. Since 2000, the Medical Birth Register has included several questions on maternal smoking behaviour. We plot the likelihood of smoking at the end of the pregnancy by children's birth cohort in Figure 1.9. While more than 20% of the most disadvantaged mothers reported smoking at the end of the pregnancy, only 5% of the most advantaged mothers smoked during pregnancy in 2000. The figure shows a rapid decline in the percentage of mothers smoking at the end of the pregnancy in the bottom decile of the income distribution. This decline is present for both male and female infants and is almost identical in magnitude. By 2018, almost no mothers in the top decile of the income distribution reported smoking while pregnant and only 1% of the mothers in the bottom decile still smoked during pregnancy.[14]

Changes in smoking policies and tobacco tax hikes are shown to influence maternal smoking behaviour and the exposure of infants to second-hand smoke (Lien and Evans, 2005; Adda and Cornaglia, 2010). Although Norway's Tobacco Act of 1988 forbids smoking in public premises and vehicles, a 2004 change in

[14]Note that a sensitivity analysis studying differences based on parental education instead of parental lifetime income confirmed these findings. The results are presented in Appendix Figure A3.

the law extended the smoking ban to bars and restaurants. This policy decreased maternal smoking among mothers working in bars and restaurants (Bharadwaj, Johnsen and Løken, 2014), accompanied by major smoking prevention campaigns. Hence, the decrease in the socioeconomic gap in smoking could have accelerated with the reform in 2004. While smaller in utero exposure to maternal smoking mostly affects low birth weight and preterm incidents and thereby infant mortality during the first week of life, less exposure to second-hand smoke during the first year of life strongly relates to lower rates of SIDS (Carpenter et al., 2013; CDC, 2020) and thus the decline in infant deaths 29–365 days after birth. Hence, the changes in the smoking behaviour of mothers are likely to have constituted a contributing factor to the closing of the socioeconomic gap in infant mortality after 2000.

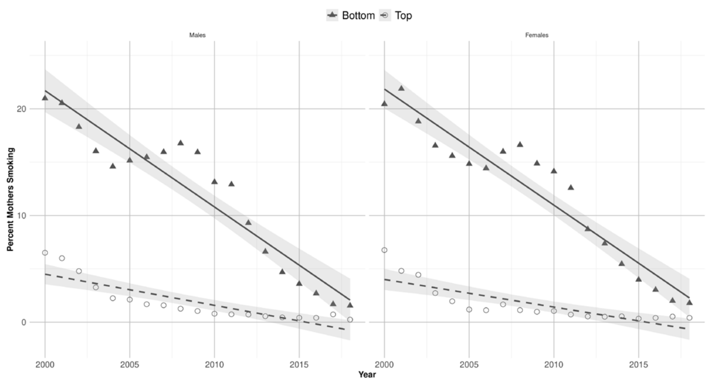


Figure 1.9: Individual-Level Inequality in Low Birth Weight and Preterm Births 1968–2018.

Note: The figure shows the percent of mothers smoking at the end of pregnancy in the top and bottom deciles of the parental income distribution by gender from 2000 to 2018. Parental lifetime income was calculated by taking the average income of individuals between ages 34 and 40 years.

1.5 Conclusion

Technological and medical advancements have prolonged human life during the past several decades. Whether the decrease in mortality is evenly distributed, or whether it compensates or reinforces mortality inequality is less clear and context dependent. In this regard, we used both municipality- and individual-level income and mortality data to analyse mortality gradients over the past 70 years in Norway.

We find that mortality, and in particular infant mortality in Norway, has greatly declined for individuals of all ages. Focusing on mortality at a geographic level (municipalities) for the years 1991 to 2017 revealed that mortality in absolute terms has fallen for all age groups. While there are income gradients at older ages in the early 1990s, a municipality's poverty level seems to have played a very small role in determining the risk of death at any given age in 2018. Nevertheless, while these findings document equality in mortality across Norwegian municipalities by 2018, they do not imply that there are no income gradients in mortality within Norwegian society. In particular, there may still exist a strong association between income and mortality at the individual level.

For infants, the decline of the income gap in mortality at the geographic level had already closed by the late 1960s. At the individual level, however, the inequalities in mortality persisted much longer and the risk of dying within the first year of life only converged in 2010 for children born into the richest and poorest Norwegian families. Our results suggest that the decline in infant mortality since 1951, as well as the income gradient in infant mortality, were strongly tied to advances in medical technology, the scientific discovery of the link between SIDS and sleeping positions, and the dangers of maternal smoking.

Overall, the results suggest that improvements in infant mortality over the past 70 years in Norway are an important example of how access to health care, the transmission of scientific knowledge, as well as societal level changes in health behaviour can affect children's lives—particularly among disadvantaged groups. Although the levelling out of the playing field among the rich and the poor in terms of infant mortality is an important achievement for a comprehensive welfare state such as Norway, it is important to note that this does not suggest that mortality equality at older ages holds for children born today. Occupational hazards and different health behaviours as teenagers by various socioeconomic groups could still contribute to mortality inequality at later ages for these individuals. This leaves room for further research to investigate how the closing of the early-life mortality gap between the rich and the poor translates into individual-level mortality differences by socioeconomic background later in life.

Appendix

1.A Additional Figures

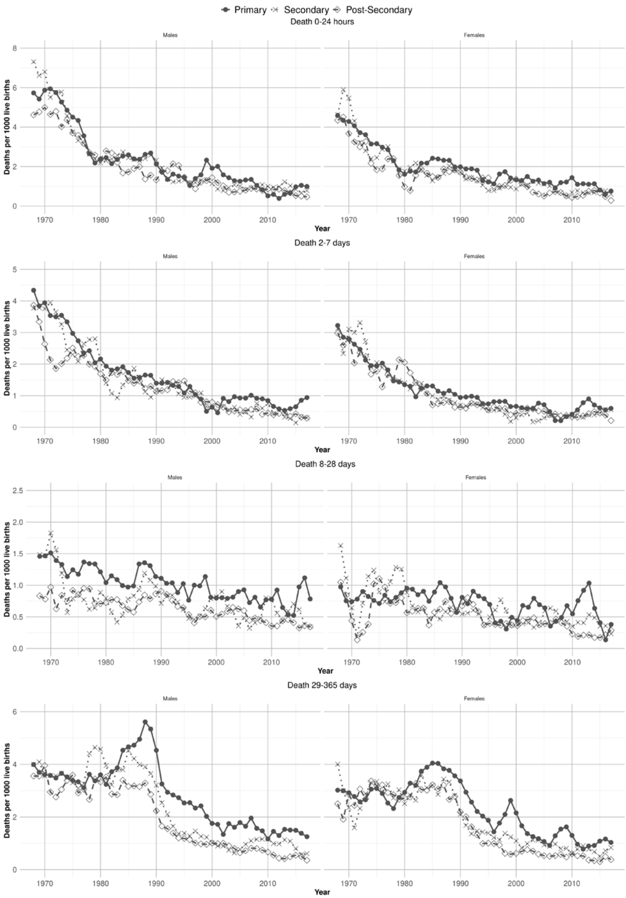


Figure A1: Individual-Level Inequality by Education and Age of Death.

Note: Infant mortality in three different education groups by gender from 1968 to 2017. Each dot/rectangle/triangle represents cohort-specific infant deaths per 1,000 live births (smoothed over three years) for the four different age categories. Education is classified based on NUS-2000 codes.

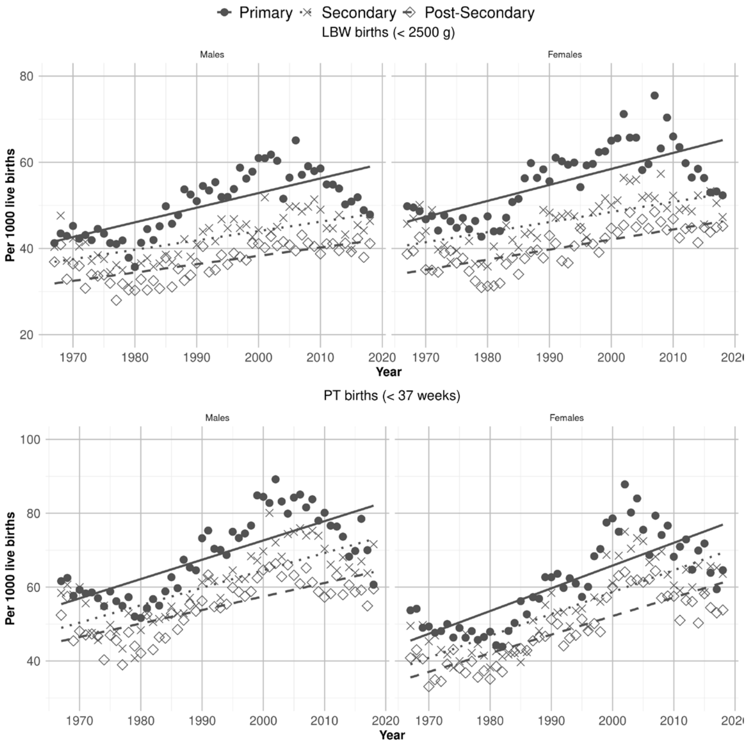


Figure A2: Individual-Level Inequality in Low Birth Weight and Preterm Births 1967-2018.

Note: Cohort-specific numbers of children born with low birth weight (<2,500 grams) and preterm deliveries (before the end of the 37th week of gestation) per 1,000 live births in the three educational groups by gender from 1967 to 2018. Education is classified based on NUS-2000 codes.

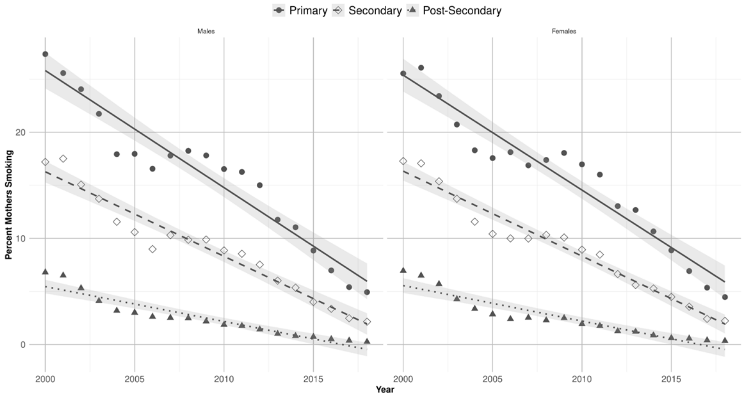


Figure A3: Individual-Level Inequality in Maternal Smoking by Education 2000-2018.

Note: Percentage of mothers smoking at the end of pregnancy in the three educational groups by gender from 2000 to 2018. Education is classified based on NUS-2000 codes.

Chapter 2

Intergenerational Mobility Trends and the Changing Role of Female Labor

Abstract: Using harmonized administrative data from Scandinavia, we find that intergenerational rank associations in income have increased uniformly across Sweden, Denmark, and Norway for cohorts born between 1951 and 1979. Splitting these trends by gender, we find that father-son mobility has been stable, while family correlations for mothers and daughters trend upwards. Similar patterns appear in US survey data, albeit with slightly different timing. Finally, based on evidence from records on occupations and educational attainments, we argue that the observed decline in intergenerational mobility is consistent with female skills becoming increasingly valued in the labor market.

This chapter is written together with Ulrika Ahrsjö and Joachim Kahr Rasmussen. We are grateful for advice and comments from Adrian Adermon, Eirik Berger, Aline Bütikofer, Claus Thustrup Kreiner, Søren Leth-Petersen, Mårten Palme, Kjell G. Salvanes, David Seim, Jósef Sigurdsson, Alexander Willén and participants at seminars in Bergen, Copenhagen, Stockholm and Uppsala. Additional thanks to Kevin Lang and two anonymous referees for constructive feedback on an earlier version of this paper. Ahrsjö gratefully acknowledges funding from Handelsbankens Forskningsstiftelser (Jan Wallanders och Tom Hedelius Stiftelse). Activities at NHH are funded by the Research Council of Norway through project No. 275800 and through its Centres of Excellence Scheme, FAIR project No. 262675 (Karadakic). The activities at CEBI are financed by the Danish National Research Foundation, Grant DNRF134 (Rasmussen).

2.1 Introduction

A central question in the social sciences is how the childhood family environment shapes economic fortune in adulthood. If the family environment plays an important role in determining socioeconomic outcomes, a common interpretation is that children are not born with equal opportunities in life. Empirical studies of the influence of family environment often estimate the relationship between income of parents and their children. Early work by for instance Becker and Tomes (1979) and Solon (1999) highlights that in such estimations, it is essential to account for the role of idiosyncratic labor market conditions. Accordingly, variation in labor market conditions may shape estimates of intergenerational mobility across space (Solon, 2002; Chetty et al., 2014*a*; Bratberg et al., 2017) and potentially also time (Corak, 2013). While substantial time-variation in mobility has been documented (see e.g. Lee and Solon (2009), Olivetti and Paserman (2015), Chetty et al. (2014*b*)), little is known about how changes in labor market conditions shape these patterns (see Song et al. (2020) for a notable exception).

In this paper, we ask what implications the "grand convergence" between men and women in labor market conditions (Goldin, 2014) has had for intergenerational income mobility. Over the past 50 years, women in all Western economies have become more likely to participate in market work (Olivetti and Petrongolo, 2016) and occupational segregation of men and women has decreased (Blau, Brummund and Liu, 2013). While it is widely acknowledged that economy-wide changes in female labor supply may change the precision with which female income indicates economic status (Chadwick and Solon, 2002), the implications of this change for the persistence of income between parents and children is *a priori* unclear due to two opposing forces. On the one hand, when female labor supply increases, the relative position of a woman in the female income distribution arguably reflects her underlying skills better. All else equal, this puts an upwards pressure on the intergenerational persistence of income. On the other hand, maternal income represents a larger share of joint parental income. If maternal income is initially less informative than paternal income about their children's income potential, this puts downwards pressure on measures of intergenerational income persistence. Because female income has arguably been an unreliable indicator of social status historically (Chadwick and Solon, 2002; Björklund, Jäntti and Lindquist, 2009; Blanden et al., 2004), the extent to which the secular trend in female labor supply affects time variation in measures of intergenerational mobility is largely unexplored.

We address this question by turning to the three Scandinavian countries. The high quality of Scandinavian administrative data allows us to follow how the changing patterns in female labor supply affect earnings at the individual level. Scandinavia provides an ideal setting for understanding how the changing role of women at the labor market can affect intergenerational mobility, as the development toward gender equality precedes that in other countries (Kleven, Landais and Søgaard, 2019). First, we document trends in intergenerational mobility in Sweden, Denmark, and Norway for cohorts of children born between 1951 and 1979 leveraging administrative income data from 1968 up until 2017. By applying a unified approach to long panels of full-population administrative data for three different countries, we can investigate the extent to which intergenerational mobility follows similar trends across countries, ensuring that any differences in findings are not related to country-specific developments, the choice of data period or income definition.

Our results reveal a substantial decline in intergenerational income mobility in Scandinavia that remains robust across a large set of common empirical specifications. In particular, we show that the results are largely unchanged when studying intergenerational correlations in log earnings rather than within-cohort earnings ranks, and when considering intergenerational correlations in gross or net-of-tax income rather than earnings. This suggests that the observed mobility trends were not driven by simultaneous rank-distorting changes in taxes or transfers across Scandinavia.

Second, we turn our attention towards understanding how changes in female labor market conditions and access to education have affected estimates of intergenerational mobility over time. When breaking mobility trends down by the gender of parents and children, it is evident that earnings of children have become increasingly correlated with maternal earnings over time, while the correlation with paternal earnings has remained close to constant. In the earliest cohorts in our analysis, child earnings — in particular earnings of sons — were virtually uncorrelated with earnings of mothers, while exhibiting a clear and economically significant correlation with earnings of fathers. Over time, these parent-specific mobility estimates between children and their mothers and fathers, respectively, have all converged to similar levels. Our results show that this is not solely an effect of extensive margin labor supply increasing among women, but also stems from women entering more skilled occupations. Conducting a similar analysis on Panel Study of Income Dynamics (PSID) data from the US, we find comparable patterns, albeit with a slightly lagged timing. This suggests that the observed patterns are not solely a Scandinavian phenomenon.

Third, we build a simple model of gender-specific mobility and latent productivity and calibrate it to the Scandinavian data. Inspired by some of the building blocks in the model by Becker and Tomes (1979), we assume that income consists of two components: one inheritable part, say skills or productivity, and one non-inheritable, idiosyncratic determinant. We decompose the trend in intergenerational mobility into parts that reflect assortative mating (correlations in parental skills), gender-neutral skills transmission, gender-specific skills transmission and gender-specific return on skill. Calibrating our model to country-specific aggregate data, we show that the observational downwards trend in intergenerational mobility is largely compatible with increasing return on inheritable skills among women, relative to men. This phenomenon explains an increase in the intergenerational rank association of five to six rank points in all three countries for cohorts of children born from 1962 to 1979. To build intuition for this rise in gender-specific return on skills and the associated implications for mobility, we can think of an early period where a woman with a significant cognitive endowment is more likely to become a secretary, compared to an equally skilled man with similar preferences who sorts into becoming a lawyer. In this case, the woman's skills are arguably less reflected by her earnings, which effectively attenuates the association between her earnings and that of her children. If this occupational and educational segregation becomes smaller over time, the correlation between maternal earnings and child earnings will increase. The decomposition thus suggests that the observed trends in income mobility is simply an artefact of changes in how women participate in the labor market.

In the final part of the paper, we corroborate this decomposition exercise empirically, by showing that gender-specific intergenerational correlations in *economic status* — measured by combining own income, years of education, and occupation using the proxy variable method developed by Lubotsky and Wittenberg (2006) — remain constant over time, or are only weakly increasing. Mobility also remains at a constant level when correlating earnings of sons with that of their maternal uncles, as another way to proxy for maternal skills. Our evidence thus suggests that the observed trends in intergenerational income mobility can be explained as the result of income rank correlations between children and parents gradually becoming less attenuated by frictions caused by gender segregation in the labor market. In other words, our results suggest that intergenerational mobility in income did in fact decline consistently in Scandinavia across cohorts born between 1951 and 1979, but that this was driven by female earnings becoming more reflective of their actual skills. The return on latent productivity of women has converged towards that of men. The observed development in intergenerational mobility can therefore potentially be seen as a natural implication of a socially desirable development, rather than a sign of actually declining equality of opportunity.

With this paper, we make three main contributions to the understanding of time variation in intergenerational mobility. The first contribution is related to a series of recent empirical studies from Western economies, which vary in their conclusions on mobility trends. Connolly, Haeck and Laliberté (2020), Harding and Munk (2020) and Markussen and Røed (2020) all find that intergenerational mobility has declined rapidly for cohorts of children born between 1960 and 1980 in the Canada, Denmark and Norway, respectively. On the other hand, Pekkarinen, Salvanes and Sarvimäki (2017), Song et al. (2020) and Brandén and Nybom (2019) only detect weakly declining — or even stable — trends in a similar set of countries. Davis and Mazumder (2022) find declining mobility in the US for children born between 1950 and 1960, while Chetty et al. (2014*b*) find no change in rank associations between children born in 1971 and later cohorts. A recent paper by Jácome, Kuziemko and Naidu (2021) shows evidence of stable or slightly increasing trends for birth cohorts from the 1940s to the 1970s. In this paper, we provide clear evidence of a uniform decline in intergenerational income mobility across Scandinavia for cohorts born between 1951 and 1979. In addition, we show that this trend is not simply a result of certain empirical specifications or country-specific policies. We also provide evidence of a similar pattern in the US from panels of linked survey data. To our knowledge, ours is the first paper to estimate and compare trends in relative mobility across multiple countries, thereby providing suggestive evidence of a general phenomenon in Western economies.

The second contribution lies in explicitly documenting substantial gender-variation in mobility trends and showing that gender-specific mobility trends are surprisingly similar across a range of Western economies. A noteworthy strand in the mobility literature has previously highlighted that cross-sectional estimates of intergenerational mobility may differ substantially by gender due to different opportunities for men and women in the labor market (Corak, 2013; Lee and Solon, 2009). With this paper, we are able to correlate income of children to *individual* labor earnings of their fathers and mothers across three decades of birth cohorts. We show that mobility has remained stable for father-son relations, while it has been declining considerably whenever female earnings are taken into account.[1] These findings suggest that not only do mo-

[1] This pattern has previously been documented — though not extensively discussed — in Engzell and Mood (2021); Brandén and Nybom (2019) and Jácome, Kuziemko and Naidu (2021).

bility *levels* vary by gender, but secular changes in gender-specific earnings determinants have also caused *trends* to differ substantially, in turn causing levels to converge. These patterns are present across all countries in our analysis, suggesting that one explanation for why the recent literature on mobility trends has reached different conclusions is choices regarding how to deal with female earnings.

The third contribution of this paper is that we provide an explanation for the observed pattern of declining overall mobility, which is compatible with the gender-specific mobility trends that we observe in Denmark, Sweden, and Norway. In recent studies, several explanations for downward trends in mobility have been proposed, none of which are consolidated across countries and specifications. One dominant explanation put forward by Davis and Mazumder (2022) is that the return to education has increased. Given that education and human capital are significant channels for the transmission of income across generations, this has led to a decline in mobility. A similar explanation put forward by Connolly, Haeck and Laliberté (2020) is that the degree to which women obtain secondary education has increased. Observing that conditional on parental income, income in the child generation is 'boosted' by a higher level of education among parents, the authors conclude that this upward trend in mothers' level of education must have led to a decline in social mobility. However, the underlying mechanism of this relationship remains unclear. Finally, Harding and Munk (2020) suggest other explanations, such as changes in family structure including marital status, assortative mating, and childbearing among women. Our paper is the first to explicitly show a connection to female participation in the labor market and valuation of female skills. In other words, we show that changes in female labor market conditions have caused parental earnings to be substantially better reflected in child earnings — hence inducing a *real* downwards shift in intergenerational mobility — in spite of the between-generation correlation in latent skills being relatively constant.

The remainder of the paper is structured as follows. Section 2.2 provides a brief overview of the key features of the Scandinavian welfare states, and Section 2.3 describes our data sources. In Section 2.4 we describe the common methodology used to estimate intergenerational income mobility and present our main results. Section 2.5 builds and calibrates a model for the connection between intergenerational rank associations and increasing female labor force attachment. In Section 2.6, we show how the estimated trends change when we use a measure of maternal economic status that better captures female earnings potential. Section 2.7 concludes.

2.2 Institutional Context

The Scandinavian countries share similar traits in terms of economic development, political culture, and institutions. The welfare states are of universal character, with access to social security benefits, health care, subsidized childcare, and tuition-free higher education (Baldacchinoel and Wivel, 2020). In order to finance the provision of these public goods, marginal tax rates at the top of the income distribution, as well as the average tax burden, are substantially higher in Scandinavia than in other developed countries (Kleven, 2014). Employees are to a large degree organized in unions and wages are often collectively bargained (Pareliussen et al., 2018). Historically, all three countries have been characterized by low levels of inequality and high levels of income mobility, in comparison to other Western countries (Søgaard, 2018; Bratberg et al., 2017).

During the second half of the 20th century, the role of women in society, and particularly at the labor market, experienced a "grand convergence" towards the position of men (Goldin, 2014). Contributing to this development were the individualization of the tax system (Selin, 2014), the introduction and expansion of paid paternity leave (Ruhm, 1998), and the expansion of compulsory and higher education (Meghir and Palme, 2005; Black, Devereux and Salvanes, 2005). As a result, female labor force participation increased from the early 1950s and is currently higher in Scandinavia than in most other Western economies.[2] Figure B1

In Figure 2.1 we provide some descriptive evidence on the development of female labor in the countries under study, for the parent and child generations separately. Panels A and B show how labor force participation among women converged to the male level.[3] Participation rates of mothers with children born in the 1950's were less than half the rate of fathers, but this gap had closed almost entirely for mothers of children born in the 1970s. It is even less pronounced when we compare sons and daughters of a given birth year. Panels C and D show the development of occupational segregation, i.e. the extent to which men and women work in the same occupations. The segregation index is calculated as the difference in the share of all women and men in the labor force who work in a given occupation, summed over all observed

[2]See Appendix Figure B1 for a comparison of labor force participation rates across Scandinavia and the United States.

[3]The labor force participation rate for men and women is based on the income definitions we use in our later analysis and always relates to the birth year of the child. A person is considered in the labor force in a given year if they have annual earnings exceeding the equivalent of 10,000 USD (2017).

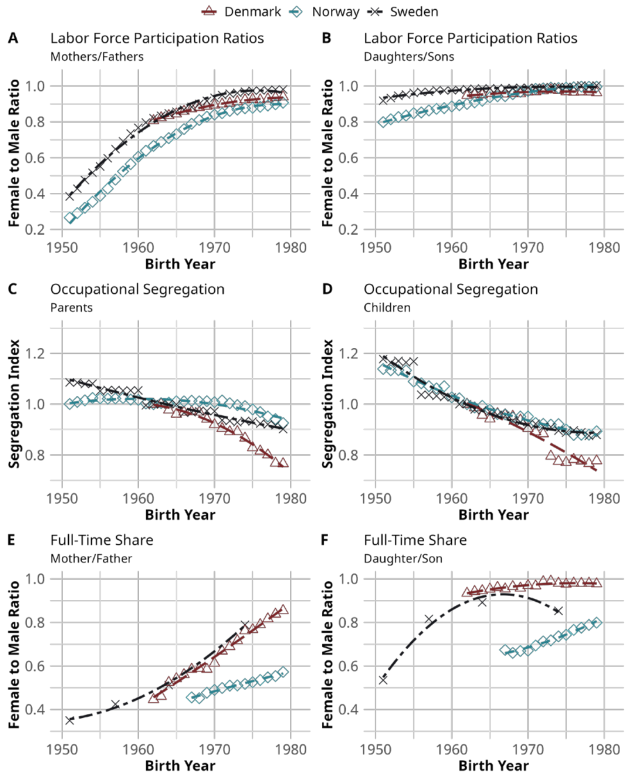


Figure 2.1: Labor Market Developments.

Note: Panel A and B depict female-to-male ratios of labor force participation in our main samples; for parents in Panel A and for children in Panel B. Panels C and D depict an index for labor market segregation, for parents and children respectively. The index is normalized to the base year 1962. In some years, Danish occupational codes have been imputed from other variables — therefore, the Danish trend in occupational segregation should be interpreted with caution (see more in Appendix 2.A). Panels E and F provide female-to-male ratios of full-time work. Full-time work is defined as working at least 27 hours in Denmark and Sweden, and at least 31 hours in Norway. Data for full-time shares is obtained from linked employer-employee data in Norway and Denmark, and nationally representative surveys in Sweden. In Denmark, there is a significant data break in the full time definition which only affects the child cohort — we attempt to adjust for this appropriately with a simple correction procedure (see more in Appendix 2.A). "Birth Year" refers to birth year of the **child** in each parent-child pair.

occupations (Duncan and Duncan, 1955). To make comparisons of trends easier, we normalize the index with 1962 as the base year, allowing for an interpretation of occupational segregation relative to the 1962 level.[4] Evidently, occupational segregation has declined persistently over time, similar to the development in the United States, as documented by Blau, Brummund and Liu (2013) and Blau and Kahn (2017). In contrast to the development of female labor force participation, the decline in occupational segregation is to a larger extent present in the child generation, rather than the parent generation.

In addition, the intensive margin labor supply of women increased during the time period under study. Panels E and F provide female-to-male ratios of the share of individuals working full-time, by birth year of the child. For Denmark and Sweden, full-time work is defined as at least 27 working hours per week, while full-time in the Norwegian data is defined as at least 31 weekly hours. In Panel E, we present this for mothers relative to fathers. Similar to the development in labor force participation, mothers have continuously caught up to the rate at which fathers work full time, although a sizable difference remains toward the end of our sample period.[5] Especially in Norway, our data reveals a significant remaining gender gap in hours worked. Notably, the Swedish female-to-male full-time ratio increases at a higher rate from around 1960 and on. Turning to the full-time ratio of daugh-

[4]The occupational segregation index is defined by three-digit occupation codes for Norway and Sweden and one-digit codes for Denmark due to data limitations. Therefore, the cross-country difference in trends should not be interpreted as hard evidence of deviating patterns of occupational segregation.

[5]The level differences between Sweden/Denmark and Norway stem from different full-time definitions in the data. Moreover, the convergence in intensive margin labor supply in all countries is almost entirely driven by increases in female full-time shares; male full-time shares are almost constant over the entire time-period under study.

ter compared to sons, in Panel F, this shows two conflicting observations. On the one hand, looking at Sweden and Denmark, the female full-time rate in the 1960s cohort was already quite high, with only small changes after that. On the other hand, the Norwegian series, using a stricter definition, suggests that a significant gender gap in working hours persists also among the child generation of our sample.

Overall, the three labor market measures presented in Figure 2.1 show a substantial gender convergence in labor market participation. Convergence in extensive margin labor force participation of mothers happened faster before the 1960s birth cohort, and was almost entirely equal to the fathers' level for the 1979 cohort. Changes in occupational choice and intensive margin labor supply of mothers were, however, more predominant in the second part of our sample, after the 1960 cohort, compared to those born before 1960. We will later argue that both the extensive and intensive margin developments in labor market participation have key implications for our measures of trends in intergenerational mobility.

2.3 Data

For our main analysis, we use register data from Denmark, Norway, and Sweden that cover the whole population of each country. This is available from 1968 to 2017 for Norway and Sweden and from 1980 to 2017 for Denmark. The data consist of linked administrative records that provide a variety of information, including birth year, educational attainment, earnings and other income measures, family status, and various demographic variables. Individuals can be linked to their parents, which allows us to create data sets containing all child-parent pairs in a given time frame, with relevant individual income measures. For more details about the registers used, see Appendix 2.A.

Our Scandinavian estimation sample consists of all children born between 1951 (1962 for Denmark) and 1979, who (i) have a valid personal identifier, and (ii) have at least one parent with a valid identifier. As this means that we remove a significant share of immigrants from our samples — in particular in early years — we remove all foreign-born individuals and all children with foreign-born parents. Sample sizes per birth year are approximately 70,000 child-parent pairs in Denmark, 60,000 pairs in Norway, and 100,000 pairs in Sweden, with variation over time.

The results involving US data are based on the Panel Study of Income Dynamics (PSID). The PSID is a nationally representative survey that covers information on employment, income, occupation, education, and family links, starting from 1968. The PSID follows families and individuals across time and has a relatively low attrition rate. With this data, we create a sample of child-parent pairs for the US in a comparable, yet more limited, fashion than our analysis on the main Scandinavian samples. In total, the US sample contains about 5,000 child-parent pairs. See e.g. Lee and Solon (2009) and Vosters (2018) for previous applications of the PSID to intergenerational mobility estimation.

The main income specifications are chosen for easy comparisons with much of the recent literature.[6] Income for the child generation is defined as three-year averages of annual labor income.[7] See Appendix Table A2 for an overview of the earnings components and how these compare across countries. This is measured at ages 35-37, which balances the needs for a measure of per-

[6]See e.g. Chetty et al. (2014*a*) and Lee and Solon (2009).

[7]Averages are calculated including zeroes. Individuals with one or more missing observations in the years averaged over are dropped from the sample.

manent income rank with the needs for measuring child incomes relatively early in order to maximize the number of cohorts that can be included in the analysis (Nybom and Stuhler, 2016; Bhuller, Mogstad and Salvanes, 2017).

Parental income is defined as the average of maternal and paternal individual income, measured as three-year averages of annual labor earings around age 18 of the child. In general, this means measuring the parents' income at age 40 or later, which is considered a meaningful proxy for lifetime income in the literature (Nybom and Stuhler, 2016). In our Appendix, we provide robustness checks to different income definitions for child and parent income variables, such as estimating trends in total factor (gross) income or net-of-tax income, and evaluating the sensitivity to the exact age at which we measure child income. Finally, due to the fact that we measure parent income at age 18 of the child, parent age may vary substantially in our main specification. In particular, parents who get children at a younger age mechanically have their income measured at a younger age as well. Ranking parent income within both child birth year *and* parental birth year jointly, we are able to verify that the observed mobility trends are not driven by this measurement issue.

2.4 Trends in Intergenerational Mobility

In this section, we first describe the empirical method we apply for measuring child-parent rank associations, and present the trend for Scandinavia. We then analyze rank associations when we split the sample into sons, daughter, mothers and fathers, and compare our Scandinavian results to suggestive US estimates. Finally, we discuss to what extent this trend can be attributed to changes in intensive- or extensive margin labor supply of women.

2.4.1 Empirical Method

In order to measure intergenerational income persistence, we transform observed income into cohort-specific ranks, as in Dahl and DeLeire (2008) and Chetty et al. (2014*a*). Using ranks, rather than levels or logs, offers certain advantages in this context. First, estimated rank correlations have proven to be less prone to life-cycle bias than other measures (Nybom and Stuhler, 2017), and in addition, the use of ranks enables the inclusion of zero incomes. However, in order to ensure that our results are not driven by the rank transformation, we also present mobility trends in intergenerational income elasticities (IGE) in the Appendix.

Rank correlations are estimated with the following regression, separately by birth cohort and country:

$$\text{Rank}^{\text{C}}_{it} = \alpha_t + \beta_t \text{Rank}^{\text{P}}_{it} + \varepsilon_{it}, \qquad \forall \quad t \in [1951, 1979] \tag{2.1}$$

where $\text{Rank}^{\text{C}}_{it}$ is the percentile rank of child i's average income at age 35-37 within the distribution of all children born in year t. When we analyze sons and daughters separately, we calculate their income rank separately by gender. $\text{Rank}^{\text{P}}_{it}$ is the percentile rank of the same child's parents' income within the distribution of all parents with children in birth cohort t, averaged over ages 17-19 of the child. The coefficient β_t captures the average cohort-specific parent-child correlation in income ranks, sometimes referred to as the intergenerational rank association (IRA). Lower values of β_t are interpreted as lower rank associations in income, and thus higher levels of intergenerational mobility.

Intuitively, one can think of the IRA as the correlation in inheritable skills and values that are transmitted across generations. These are attenuated by earnings determinants that cannot be passed on to children, which reduce the signal value of parental income. Such "noise" may stem from individual-specific idiosyncratic shocks to the earnings process or time-specific characteristics of the labor market. In particular, changes in the IRA over time are not necessarily driven by transmissible factors, but rather by the importance of earnings determinants that cannot be passed on to children. In the context of analyzing how changing female labor market participation may have affected the intergenerational association in income, this is a relevant consideration.

2.4.2 Estimated Trends

In Figure 2.2, we present estimates for country-specific trends in intergenerational rank associations in individual labor income. Each point in the graph represents a slope parameter for a cohort-specific regression of equation (2.1) with linear trends estimated separately for 1951-1961 and 1962-1979. We provide fitted lines separately to facilitate comparisons between Denmark, Norway, and Sweden for the cohorts where all countries have available data.[8]

From Figure 2.2, it appears that intergenerational mobility, measured using the IRA, has declined in all three countries, with the fastest rate of decline in Denmark. There, the rank association in income increased by 7.5 rank points (39 %) from 1962 to 1979 — equivalent to an average annual increase of 0.44 rank points. While smaller than in Denmark, the trends in Norway and Sweden are by no means negligible. From 1962 to 1979, the rank association

[8] In addition to providing graphical illustrations of the trends in the IRA, Appendix Table B4 provides an overview of IRA coefficients for different specifications and tests whether trends are statistically different across countries.

in income increased by 6.4 and 4.4 rank points (38 % vs. 25 %) in Norway and Sweden, respectively, yielding annual increases of 0.38 and 0.25. From 1951 to 1979, the total change in IRA for Norway is 7.8 rank points (50 %) and 5.8 rank points for Sweden (34 %).

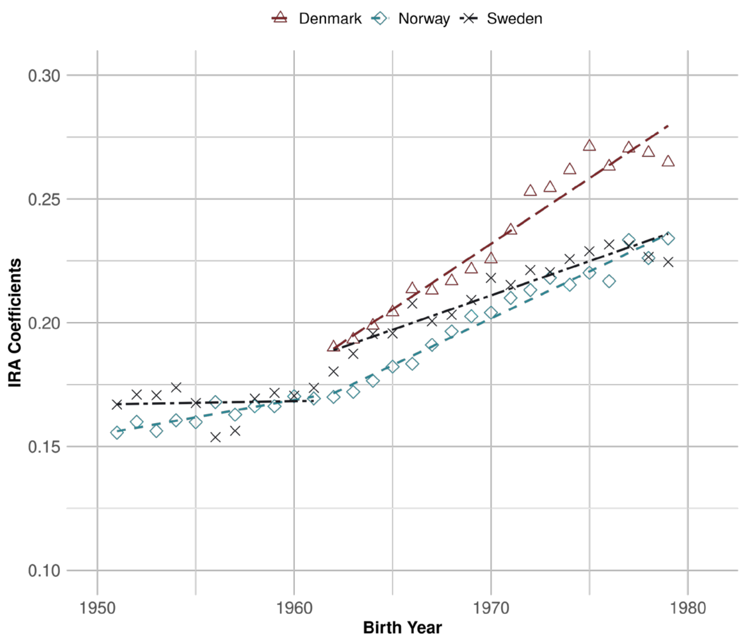


Figure 2.2: Trends in Intergenerational Mobility in Individual Labor Income.
Note: The figure plots the coefficients for the intergenerational rank association (eq. 2.1) in individual labor income for Sweden, Denmark and Norway for birth cohorts 1951 (1962) to 1979. "Birth Year" refers to birth year of the **child** in each parent-child pair. Each panel shows fitted trend lines separately for the period 1951 to 1962 and 1962 to 1979.

One may wonder what it actually means, in economic terms, that the rank association in income increased by up to 0.44 rank points per year in Scandinavia. Abstracting from non-linearities in the relationship between parent and child income ranks, a straightforward interpretation is the following: for two children born by parents in the bottom versus the top percentile, the difference in the conditional expectation of their income ranks as adults increased by 0.44 each year — amounting to as much as 4.4 rank points over a decade. Taking the Norwegian results as an example, another interpretation of the observed trends is that in the earliest observed birth cohort, a ten rank points difference in parental income corresponded to an average difference in income ranks of 1.6 between their children. In contrast, the same difference was 2.3 rank points for children born in the latest cohort. While still indicating relatively high levels of mobility by international standards, such changes over relatively short periods of time are by all means economically substantial.

In order to ensure that the trends are robust and reflect structural changes in the economy (as opposed to being something that purely exists within a narrow set of specifications), we document similar trends for a large set of different specifications in Appendix 2.B. Most importantly, we show that the trends remain largely similar when measured in net-of-tax- and gross income (Figure B3), and when measuring child income at various ages (Figure B4).[9]

In Figure B5, we restrict the sample to parent-child pairs with labor income surpassing 10,000 USD (2017). In other words, we calculate rank associations for the subset of the population that is fully active in the labor market. In general, the mobility trends persist and are similar in magnitude in this specification. However, some cross-country differences are also revealed. Rank asso-

[9] We also tested a specification where we rank a measure of parental income within both child cohort *and* their own cohort. The trends remain stable, but the results are not presented in the current version of the paper.

ciations in Denmark and Norway are lower when excluding non-participating workers from our samples, indicating that *intergenerational correlations in labor market participation* contribute greatly to intergenerational persistence in income — or at least that children of non-participating parents do disproportionately bad in the labor market themselves. In Sweden, on the other hand, the level of mobility largely remains the same after excluding non-participating parents from the estimation sample (Panel B), and even increases slightly when excluding both non-participating parents and children (Panel C).

2.4.3 Trends by Gender of the Child and Parent

Figure 2.3 presents estimates of country-specific IRA coefficients for pairs consisting of, in turn, sons and fathers (Panel A), sons and mothers (Panel B), daughters and fathers (Panel C), and daughters and mothers (Panel D). Each coefficient is again obtained by estimating equation (2.1) year by year, for the respective combination of child and parent and with individual mother or father earnings instead of the parental average. In Appendix Table B4, we test several hypotheses regarding the trends and also report slope coefficients for different specifications.

The four sets of graphs make clear that — at least starting with the 1962 cohort — the trends in IRA for all combinations of child and parent are similar in Sweden, Denmark, and Norway. Estimates for birth cohorts 1951-1979 are strikingly similar in Norway and Sweden: the trends are statistically indistinguishable for all combinations and years except for the trend in the mother-daughter IRAs after 1961. Across the panels, however, there are several distinct differences. Most importantly, we see that the

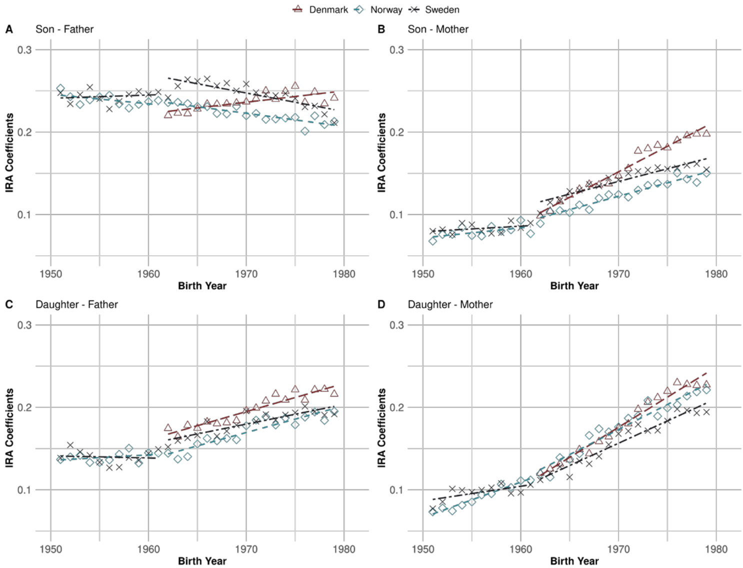


Figure 2.3: Trends in Intergenerational Mobility by Gender of the Parent and Child.
Note: The four panels plot coefficients for the intergenerational rank association (eq. 2.1) in individual labor income for Denmark, Sweden and Norway, by child year of birth. Each panel provides estimates separately by gender of the parent and child. Each marker indicates the coefficient of a separate regression and each line indicates fitted trend lines separately for birth cohorts 1951 to 1962 and 1962 to 1979. "Birth Year" refers to birth year of the **child** in each parent-child pair.

rank association between fathers and sons is generally *decreasing* (Sweden, Norway) or displays a relatively flatter trend over time (Denmark). The strongest trends in IRAs are found for mother-daughter correlations, closely followed by mother-son correlations. Father-daughter correlations depict slightly weaker trends.

Do these observed mobility patterns describe a phenomenon unique to Scandinavia? In order to understand this, we compute comparable mobility estimates for the US for birth cohorts from

1947 to 1983.[10] Results from this exercise are presented in Table 2.1.[11] US mobility trends are steepest for pairs involving women, and in particular daughters, while father-son rank associations appear to be relatively constant in the US, suggesting a comparable development to that observed in Scandinavia.[12] However, the US trends in mother-son correlations are not statistically significant.

Another feature of Figure 2.3 and Table 2.1 is that earnings are more strongly related for parent-child pairs within gender (i.e., son-father and daughter-mother) than across gender (i.e., son-mother and daughter-father). In fact, while the association in earnings ranks is generally higher among sons and fathers than among any other combination of genders, the daughter-mother correlation reaches almost the same level towards the end of the considered period in Scandinavia. For the US, we only provide a pooled IRA coefficient due to the small sample. Nevertheless, the pattern that within-gender correlations are stronger than cross-gender correlations is also found in the US data.[13]

To the extent that father-son correlations, which are stable over time, credibly measure equality of opportunity, it is hard to argue that an actual decline in opportunity has taken place over

[10]Due to the small sample sizes, trends have been estimated directly on the underlying micro data by regressing cohort-specific child ranks on cohort-specific parent ranks interacted with a linear time trend.

[11]In Appendix Table B1, we provide similar estimates with alternative sample specifications and weighting procedures. In Table B3, we document the cohort-specific number of parent-child pairs used to compute these trends.

[12]Recent evidence by Song et al. (2020) for the US supports relatively stable father-son trends for the relevant cohorts in our samples.

[13]This finding could have several reasons, such as intergenerational occupational mobility being lower within- than across gender, and the general tendency of men and women to sort into different occupations (see e.g., Blau and Kahn (2017) for a review on this latter point). Altonji and Dunn (2000) also find within-gender correlations in work hour preferences between parents and children and a recent working paper by Galassi, Koll and Mayr (2021) highlights how employment correlates between mothers and their children, especially so for daughters.

Table 2.1: IRA Coefficients and Trends (United States).

	Parents	**Father**		**Mother**	
	Child	Son	Daughter	Son	Daughter
Pooled IRA	0.317***	0.336***	0.195***	0.097***	0.137***
	(0.017)	(.022)	(0.031)	(0.025)	(0.029)
Trend × 100	0.603***	-0.240	0.980***	0.136	1.047***
	(0.149)	(0.205)	(0.277)	(0.253)	(0.292)
N	5,392	2,272	1,637	2,477	2,205

Note: The table presents estimates of the IRA and linear trends in the IRA separately for different child-parent combinations. Due to the small sample sizes, trends have been estimated directly on the underlying micro data by regressing cohort-specific child ranks on cohort-specific parent ranks interacted with a linear time trend. The trend coefficients and corresponding standard errors have been multiplied by 100 in order to avoid too many digits after the separator. Estimates are based on the full sample of individuals in the PSID born between 1947 and 1983 using PSID sample weights. Standard errors are in parentheses. P-values indicated by * < 0.1, ** < 0.05, *** < 0.01.

time in either Scandinavia or the US. Thinking of transmission of skills and values as something passive, this suggests that determinants of male income ranks, as well as the labor market valuation of skills that are passed on across generations, are unchanged over time. Instead, since all combinations of parent-child correlations that do yield upwards trends in IRAs (Panels B-D) involve women,[14] a close-at-hand explanation lies in that women's increasing integration into the labor force has changed the way that incomes are correlated across generations.[15] The difference in mater-

[14]Notably, father-son correlations in Denmark display a weakly increasing pattern in 1962-1975. This deviant pattern compared to Sweden and Norway is found also for trends in *absolute* mobility in Manduca et al. (2019), but their discussion of its sources from pre- vs. post tax income does not match our findings.

[15]The weak link between maternal income and skills for the earliest birth cohorts is also suggested by patterns of assortative mating on individual income. Appendix figure C1 provides evidence that maternal "skills" and earnings are virtually unrelated in the early period of our sample. In 1962 — and even more so in

nal trends between the US and Scandinavia would also be in line with such an explanation, as developments in female labor force participation started later in the United States and therefore likely impacted mothers only for later-born cohorts, while having a potentially larger impact through changing labor market equality for daughters.[16]

2.4.4 The Importance of Female Labor Market Developments

We began this article by arguing that the changes in female labor supply seen in the past half-century could affect mobility trends in several different ways. This makes the sum of the different effects unknown *a priori*. In fact, the evidence that has been provided this far shows that the estimated IRAs stayed relatively constant across birth cohorts 1951-1962, despite a great increase in maternal participation rates. Mechanically, whenever earnings are informative about heritable skills, we would expect mother-child earnings ranks to correlate more strongly at higher participation rates. The fact that this is not what we find speaks to a development where expansions on the extensive margin of employment happen in occupations where women's skills are not well reflected.

Figure 2.4 provides suggestive evidence to this fact. It shows

1979 — however, maternal income rises almost monotonically in paternal income. Assuming a time-invariant pattern of assortative mating, this is evidence favoring our hypothesis that mothers' income becomes more predictive of their true social status over time. An alternative explanation for the pattern in Figure C1 would involve rapid and strong changes in underlying mating patterns, which appear to be unlikely given recent research by e.g. Bratsberg et al. (2018).

[16]The validity of this explanation is confirmed in Table B2. Here, we estimate child incomes around age 30 rather than 36, allowing us to compute gender-specific rank-correlations for cohorts of children born in 1953 to 1989 rather than 1947 to 1983. Looking at this set of children born slightly later, we find that rank-correlations that include mothers exhibit a clear and significant upwards trend.

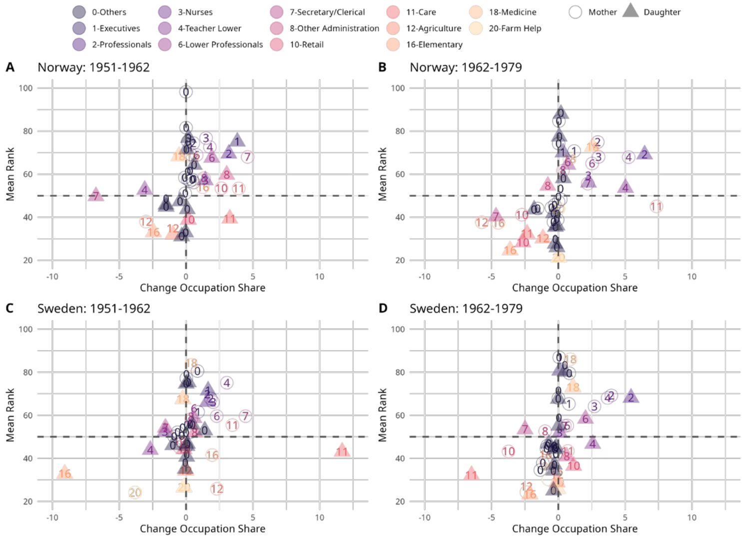


Figure 2.4: Changes in the Occupational Composition of Mothers and Daughters.

Note: The four panels depict percentage point changes of shares in each occupational group, for mothers and daughters respectively (on the x-axis), in relation to the average rank of mothers and daughters in the end of the respective period and country (on the y-axis). Panels A and C depict the change between birth cohorts 1951/1952 and 1961/1962 for Norway and Sweden respectively. Panel B and D depict the change between birth cohorts 1961/1962 and 1978/1979. Years refer to birth years of the children. The y-axis shows mean earnings ranks by occupational group for mothers and daughters, measured at the end of the respective period (61/62 and 78/79). Triangular shapes represent daughters, while circled shapes represent mothers. Occupations that never change more than two percentage points across periods are combined in the "Other" (0) group. Occupations are formed into groups based on (the Swedish and Norwegian national versions of) 3-digit ISCO codes, see the Online Data Appendix for a full list.

the relation between the change in the share of women (mothers and daughters, respectively) in an occupation, and the mean income rank of that same occupation. Panels A and C show this change in the occupational composition of the female labor force for Norway and Sweden, between cohorts born in the early 50's and in the early 60's.[17] Expanding occupations — especially among mothers — are health care workers, secretaries, personal services workers and pre-school teachers, i.e. relatively low-skilled occupations. In fact, in both Sweden and Norway, personal services and secretaries together increase by almost 10 percentage points, which corresponds to about two thirds of the increase in extensive margin employment among mothers. Additionally, the average rank of an occupation was evidently not very informative about the skill level required for that occupation; executive managers and secretaries were at almost the same average rank. Hence, one can argue that the earnings of mothers remained uninformative of transmissible earnings potential, and mother-child earnings correlations did not increase.

Conversely, our estimated gender-specific IRA-trends for the period from 1962 to 1979 show that mother-son and daughter-parent earnings ranks were converging rapidly in this period, while female employment increased at a slower pace than previously. From Figure B5 in the appendix (which shows IRA-trends exclusively for individuals active at the labor market), it is also evident that extensive margin entry can not explain the increasing rank correlations in this time periods, since correlations limited to only participants still display a trend toward lower mobility. Meanwhile, our descriptive statistics in Figure 2.1 (section 2.2) showed rapidly declining occupational segregation and increasing rates of full-time employment.

Panels B and D of Figure 2.4 depict changes in the composition of occupations among mothers and daughters during this time period. First, an increasing number of women entered high-skilled

[17] Denmark is excluded from this analysis for the sake of comparability, as the occupations data is different from that in Sweden and Norway.

occupations, such as the "professional" category, while low-skilled occupations such as cashiers and domestic help were declining. The average income rank of women in a particular occupation also became more reflective of their skill level, primarily through declining mean ranks of low-skilled occupations. Together, these pieces of evidence indicate an increased level of sorting along skill levels in the economy. Thus, we may describe these years as a period where female workers became better able over time to "earn their potential".

2.5 Decomposition by Earnings Determinants

In the previous section, we documented that the intergenerational income rank association has increased rapidly in Scandinavia, but that this phenomenon is found almost exclusively for parent-child pairs involving mothers or daughters. However, we cannot *a priori* distinguish a trend in the extent to which skills are transmitted across generations (for example if having a mother working *per se* generates higher child-mother correlations) from a trend in the extent to which female incomes reflect their inherent earnings potential ("skills"). In addition, our analysis might be influenced by any potential changes in assortative mating among parents.[18] We can nevertheless use the gender-specific variation in mobility trends, along with correlations in parental earnings, to quantify the importance of these two potential mechanisms for our observed trends. In this section, we build a simple model that allows us to quantify the importance of these channels through a decomposition exer-

[18]On the other hand, the influence of changes in assortative mating on intergenerational income associations is found to be small in Holmlund (2022), studying the case of Swedes born in 1945-1965.

cise.

2.5.1 Model Setup and Calibration

In our framework, individual gender-specific earnings at time t, y_{it}^{k}, are determined by two factors; inheritable skills, x_{it}^{k}, and a non-inheritable determinant ε_{it}^{k}. This generalizes to all fathers, mothers, sons, and daughters, i.e. all $k \in \{F, M, S, D\}$. Interpreting the setup in the context of a highly simplified version of the frameworks formulated by Becker and Tomes (1979) and Solon (2004), we can think of x_{it}^{k} as representing an aggregate measure of earnings determinants that can be transmitted across generations such as skills, values, and connections, while ε_{it}^{k} represents the value of all other income determinants that are uncorrelated to skills that can be transmitted across generations (it may be instructive — yet slightly naïve — to think of this as luck).

We assume that inheritable skills in the parental generation follow a bivariate Gaussian distribution. In particular, we assume that

$$x_{it}^{F} \sim \mathcal{N}(0,1), \quad x_{it}^{M} = \left(\psi_t x_{it}^{F} + (1-\psi_t)\, u_{it}^{0}\right) / \Gamma_t^{0}$$

where $u_{it}^{0} \sim \mathcal{N}(0,1)$. Standardizing the variance of maternal skills to one using Γ_t^{0}, ψ_t reflects cohort-specific correlations in parental skills, thus measuring assortative mating in the model.[19]

We assume that skills are transmitted passively from the parental generation to the child generation on the following form:

$$x_{it}^{k} = \begin{cases} \left(\kappa_t \left[\alpha_t x_{it}^{F} + (1-\alpha_t)\, x_{it}^{M}\right] + (1-\kappa_t)\, u_{it}^{1}\right) / \Gamma_t^{1}, & \text{for } k = S \\ \left(\kappa_t \left[\alpha_t x_{it}^{M} + (1-\alpha_t)\, x_{it}^{F}\right] + (1-\kappa_t)\, u_{it}^{1}\right) / \Gamma_t^{1}, & \text{for } k = D \end{cases}$$

[19]This is a trivial scaling coefficient that ensures that the distribution of maternal skills is standard normal.

Here, κ_t is a measure of correlation in inheritable skills — or the rate at which skills are transmitted — across generations within a given cohort of children, and α_t is a coefficient that allows the transmission of skills within gender to be stronger than skills across gender. Once again, Γ_t^1 is a trivial scaling coefficient that ensures that the distribution of skills is standard normal.

Individual income is a monotone transformation of a linear index composed of inheritable and non-inheritable determinants:

$$y_{it}^k = \hat{F}_t^k \left(\tilde{\phi}_t^k x_{it}^k + \left(1 - \tilde{\phi}_t^k\right) \varepsilon_{it}^k \right), \quad \text{for } k \in \{F, M, S, D\}$$

Here $\tilde{\phi}_t^k = \phi_t^k / \max\left(\phi_t^F, \phi_t^M\right)$ for $k \in \{M, F\}$ in the parental generation and $\tilde{\phi}_t^k = \phi_t^k / \max\left(\phi_t^S, \phi_t^D\right)$ for $k \in \{S, D\}$ in the child generation, respectively. Hence, ϕ_t^F, ϕ_t^M, ϕ_t^S and ϕ_t^D reflect the relative importance of inheritable skills in the income process for fathers, mothers, sons and daughters, respectively. Making the simple assumption that the distribution of non-inheritable determinants can be summarized by a standard normal distribution, $\varepsilon_{it}^m \sim \mathcal{N}(0, 1)$, the individual earnings index is standard normal[20].

When measuring gender-specific intergenerational mobility in income ranks, the functional form of the monotone transformation function, $\hat{F}_t^k(\cdot)$, is essentially unimportant; as long as it is monotone in the earnings index, any rank transformation of the earnings index will yield the same result as a rank transformation of earnings. However, in order to find both a pooled measure of child income ranks and a measure of joint parental earnings, such functional form can no longer be disregarded without also dis-

[20]Through simulations, it can be verified that composing the individual income index of two sets of Gaussian components, one inheritable and one non-inheritable, replicates the aggregate functional relationship between parental and child income ranks remarkably well.

regarding potentially non-negligible differences in gender-specific earnings distributions. Fortunately, we can obtain the functional forms directly from the data. Exploiting the assumed monotone relationship between the earnings index and earnings, we match index ranks to the earnings distribution observed in the data. This allows us to compute pooled earnings ranks in the child generation as well as a measure of joint parental earnings, y_{it}^{P}, that takes the true earnings distribution into account:

$$y_{it}^{P} = \hat{F}_t^F \left(\tilde{\phi}_t^F x_{it}^F + \left(1 - \tilde{\phi}_t^F\right) \varepsilon_{it}^F \right) + \hat{F}_t^M \left(\tilde{\phi}_t^M x_{it}^M + \left(1 - \tilde{\phi}_t^M\right) \varepsilon_{it}^M \right)$$

Here, $\hat{F}_t^F(\cdot)$ and $\hat{F}_t^M(\cdot)$ are year-specific estimates of the functions that map the earnings index to the earnings distribution observed in the data.

For each country and cohort, we are currently calibrating a vector of seven decomposition parameters from only five equations $[\ \psi_t \quad \kappa_t \quad \alpha_t \quad \phi_t^F \quad \phi_t^M \quad \phi_t^S \quad \phi_t^D\ ]'$. In order to avoid underidentification, we make two adjustments. First, we set $\phi_t^F = \phi_t^S$ such that the skill importance in earnings for mothers and daughters, ϕ_t^M and ϕ_t^D, must be interpreted relative to that of fathers and sons respectively — i.e. a generation-specific gender bias in the importance of skills for determination of earnings. Secondly, we set both ϕ_t^F and ϕ_t^S equal to one, thereby effectively pinning down the level around which κ_t trends over time[21]. Finally, the vector of decomposition parameters that are now left for us to calibrate across countries and years is given by: $[\ \psi_t \quad \kappa_t \quad \alpha_t \quad 1 \quad \phi_t^M \quad 1 \quad \phi_t^D\ ]'$.

[21] The more skills are reflected in earnings, the less skills need to be transmitted across generations in order to obtain a given correlation in earnings over time. Fixing the importance of skills for earnings among males therefore effectively pins down the skill transmission rate across time for a given intergenerational correlation in earnings.

The calibration procedure is explained in Appendix section 2.C, where we also document the quality of the calibration exercise for each set of country-year combinations of parameters.

2.5.2 Decomposition

By calibrating the model, we are eventually interested in understanding how country-specific changes in intergenerational mobility can be decomposed into changes in the rate at which inheritable skills manifest themselves as labor earnings among mothers and daughters relative to fathers and sons, and changes in assortative mating on skills among parents. Before doing so, we first investigate how the parameters associated with these channels have changed over time in our calibration exercise. Parameters for selected years are displayed in Table 2.2.[22]

Table 2.2: Decomposition Parameters.

	1951			1962			1979		
	SE	DK	NO	SE	DK	NO	SE	DK	NO
ψ_t	0.131	-	0.147	0.289	0.189	0.171	0.249	0.186	0.174
κ_t	0.301	-	0.300	0.257	0.267	0.274	0.261	0.290	0.286
α_t	0.580	-	0.603	0.632	0.582	0.626	0.561	0.560	0.564
ϕ_t^M	0.286	-	0.260	0.368	0.398	0.371	0.594	0.701	0.622
ϕ_t^D	0.511	-	0.501	0.591	0.721	0.619	0.935	1.011	0.951

Note: The table presents calibrated decomposition parameters for Sweden, Denmark, and Norway in three selected years. The coefficients have been obtained by matching a simulated version of the aforementioned model to empirical gender-specific IRA-coefficients as well as the relation between father and mother income.

Several noteworthy features of our calibration exercise stand out. First, the decomposition parameters generally evolve very

[22]The full set of parameters is available upon request.

similar across countries. This observation adds credibility to the decomposition approach. In particular, the parameters associated with skill-importance in earnings for mothers and daughters, ϕ_t^M and ϕ_t^D, have increased at a somewhat similar pace across all three countries. This, in turn, suggests that female earnings may have become more reflective of inheritable skills in both the parent and child generations. Second, the parameter associated with assortative mating is relatively constant over time in all three countries (at least from the early 1960s and onward) in spite of strongly increasing associations in maternal and paternal earnings over time. This may be an implication of the fact that maternal earnings have become more reflective of maternal inheritable skills, thereby mechanically increasing the observational correlation in father and mother earnings for a given correlation in skills. Third, within-gender correlations in skill do in fact seem to be stronger than cross-gender correlations in skills — α_t is approximately 0.6 across all countries but slowly declining from the early 1960s and onward. Finally, the coefficient associated with non-gendered skill-transmission is slowly downwards trending in both Sweden and Norway, while exhibiting a weak but robust upwards trend in Denmark.

While the trends in decomposition parameters are generally similar across countries, the direction and extent to which their changes may affect the intergenerational rank association in earnings between parents and children is unknown. In order to decompose changes in this main parameter into effects associated with changes in the modelling parameters, we compute "counterfactual" income associations holding one parameter fixed over time, while allowing the aggregate gender-specific income distributions that we obtained from the data to vary over time.

We do this by first defining $\tilde{\beta}_t$ as the rank association between joint parental and child earnings obtained from the calibrated set of parameters in the model stated above subject to a simulated set of data such that $\tilde{\beta}_t \equiv \beta\left(\psi_t, \kappa_t, \alpha_t \phi_t^M, \phi_t^D\right)$. Then we define $\tilde{\beta}_{t,\underline{t}}^b$ in a similar fashion, but we fix parameter $b_t \in \left(\psi_t, \kappa_t, \alpha_t, \phi_t^M, \phi_t^D\right)$ to the calibrated value in period $\underline{t}$ such that for instance $\tilde{\beta}_t^{\psi_{\underline{t}}} \equiv \beta\left(\psi_{\underline{t}}, \kappa_t, \alpha_t, \phi_t^M, \phi_t^D\right)$. Finally, the part of the trend in $\tilde{\beta}_t$ that can be attributed to parameter b is simply the difference in trend between $\tilde{\beta}_t$ and $\tilde{\beta}_t^{b_{\underline{t}}}$, while the part of the actual trend in β_t that can jointly be attributed other factors than decomposition parameters and changes in the aggregate gender-specific income distributions is the difference in trend between β_t and $\tilde{\beta}_t$. The results from this exercise are documented in Table 2.3.

Table 2.3: Decomposition Results.

	1952-1961			1962-1979		
	SE	DK	NO	SE	DK	NO
Trend in β_t	0.013	-	0.140	0.277	0.530	0.379
Trend in $\tilde{\beta}_t$	0.068	-	0.138	0.240	0.527	0.327
Due to ψ_t	0.189	-	0.000	-0.056	-0.001	0.001
Due to κ_t	-0.343	-	-0.164	-0.041	0.242	-0.035
Due to α_t	0.009	-	0.007	0.002	0.000	0.001
Due to ϕ_t^M	0.054	-	0.117	0.138	0.220	0.161
Due to ϕ_t^D	0.020	-	0.130	0.158	0.069	0.119

Note: The table presents trends in observational IRA coefficients, β_t, in the three countries as well as trends in IRA coefficients obtained from the calibrated models in the three countries, $\tilde{\beta}_t$. The contribution from each parameter is computed as the difference in $\tilde{\beta}_t$ that is obtained from holding one calibrated parameter fixed at a time. The sum of contributions from each parameter need not sum to the trend in $\tilde{\beta}_t$ as part of the trend will be driven by changes in the scale of gender-specific income distributions which is not modelled.

As the rank associations in earnings did not exhibit any clear upwards trend for cohorts born between 1952 and 1961 in Sweden and Norway, there is not much to be explained by the decomposition parameters. However, there *are* certain noteworthy patterns in this period. In particular, the parameter associated with non-gendered skills transmission, κ_t, contributes negatively to the IRA over time, while the opposite is the case for the parameters associated with the extent to which female earnings are reflective of parental skills, ϕ_t^M, and ϕ_t^D. This could possibly suggest that skills transmission may in fact have declined over time, thereby pushing mobility up, but that this effect was mitigated by the increasing extent to which women's individual income reflects their earnings potential, i.e. inheritable skills. However, one should be careful with drawing too strong conclusions based on this evidence.

For birth cohorts 1962 to 1979, the simple decomposition model captures the fact that IRAs are increasing uniformly across Scandinavia well. While both ψ_t and α_t generally contribute little to mobility trends in this period, our results suggest a bigger role for κ_t — at least in Denmark, where this component explains almost half of the observed trend in mobility. In both Sweden and Norway, however, the contribution of κ_t is negative and the importance is somewhat negligible. Finally, changes in the extent to which female earnings, and particularly maternal earnings, are reflective of inheritable skills are found to be important drivers of downwards trends in mobility across Denmark, Norway, and Sweden. These effects jointly contribute to a yearly increase in the earnings IRA of between 0.28 and 0.30 rank points in all three countries, amounting to a total increase in the IRA of between 5 and 6 rank points over the period. In the next section, we show that it is plausible to interpret this phenomenon as female earnings becoming more reflective of inheritable skills over time.

2.6 Intergenerational Correlations in Latent Economic Status

In this section, we provide estimates of trends in intergenerational mobility that are less affected by changing labor market conditions. While we speak here of *socioeconomic status* rather than, as before, inheritable skill-based earnings potential, we argue that for our application to intergenerational correlations in female individual labor earnings, socioeconomic status is more or less equivalent to potential earnings.

2.6.1 Estimation

Because female labor earnings have not constituted a good measure of their earnings potential during most of our studied time frame, estimates of the model in equation 2.1 for maternal income will not capture the intergenerational relationship between maternal and child labor market skills. To fix ideas, denote the underlying relationship of interest as:

$$x_{it}^{*C} = \alpha_t + x_{it}^{*P} + \varepsilon_{it},$$

where x_{it}^{*} is a person's true economic status, unobserved by the researcher. In our setting, it is reasonable to assume that lifetime earnings alone are a good measure of economic status among sons and fathers, but less so for mothers and daughters. We follow recent work by Vosters and Nybom (2017); Vosters (2018) and Adermon, Lindahl and Palme (2021) and apply the method from Lubotsky and Wittenberg (2006) (from now on "LW") in the intergenerational mobility context. In essence, this method amounts to using a set of proxy variables that together represent a single latent variable, economic status, and weighting these together optimally, given some outcome variable (in our case, child income). These optimal weights have been shown to result in an estimator which minimizes attenuation bias among its class of estimators (Lubotsky and Wittenberg (2006), p.552).[23] The intuition here is that observable variables constitute imperfect measures about a person's underlying, or "latent", socioeconomic status, but that a less attenuated measure of economic status can be constructed from a weighted average of several proxy variables.

We use income, years of education and occupation as proxy variables for mother and father economic status. These are denoted $x_j, \quad j \in 1, .., k$. The LW estimator is constructed as follows:

$$\beta_{LW} = \sum_{j=1}^{k} \rho_j b_j, \tag{2.2}$$

where $\rho_j = \frac{cov(Rank_{it}^C, x_{jit})}{cov(Rank_{it}^C, Rank_{it}^P)}$, and the $\mathbf{b_j}$'s are OLS coefficients from a multiple regression of child income on the set of proxy variables.[24]

In order to compare the estimates to IRAs, we want to correlate parent *ranks* in economic status to child *income ranks*. We make use of the explicit index construction mentioned in Lubotsky and Wittenberg (2006) (p.554):

[23] The procedure requires the theoretical assumption that each proxy measure affects the left-hand side variable — child economic status — only through latent economic status, but it does not assume independence between the proxy variables.

[24] This method has previously been used to estimate mother-child intergenerational income elasticities for Swedish birth cohorts 1951-1961 in Vosters and Nybom (2017). Our application uses the same set of proxy variables and the same methodology, with two exceptions. First, we calculate year-specific LW estimates, in order to study the time trend in latent economic status mobility. We also extend the analysis to later-born cohorts, which necessitates measuring parental income somewhat earlier in life than in Vosters and Nybom (2017).

$$x_{it}^{\rho,P} = \frac{1}{\beta_{LW}} \sum_{j=1}^{k} x_j b_j, \tag{2.3}$$

and calculate LW index values for each mother-son and father-son pair using the logarithm of child and parental labor income. In order to keep the interpretation as close as possible to that in our main analysis, we assign individuals with zero labor income a token low level of log earnings.[25] We then transform these into percentile ranks. Finally, we regress the child income ranks on these parental index ranks, for mothers and fathers separately.

The method described so far addresses the problem of unrepresentative maternal earnings. If trends in intergenerational rank correlations in latent economic status between mothers and sons resemble those found between fathers and sons, it stands to reason that the upwards trend in mother-son earnings correlations are attributable to increased economic opportunities of women, and subsequently less attenuation bias in rank correlations. In order to understand whether daughter-father correlations are subject to the same issue (and bias in estimation), we repeat the above procedure for daughters, and approximate their economic status with income, education and occupations. Since the LW method deals with measurement error in the right-hand-side (independent) variable, this requires "flipping" the intergenerational model (eq. 2.1), and estimating rank associations between fathers and their daughters:

$$\text{Rank}_{it}^{\text{P}} = \alpha_t + \beta_t \text{Rank}_{it}^{\text{C}} + \varepsilon_{it}. \tag{2.4}$$

This has only minor impacts on the year-specific IRA estimates and does not alter the trend. Apart from this first step, the analysis

[25] Sensitivity checks show that the exact level of earnings assigned does not alter the conclusions from this analysis. Results are available on request.

proceeds in an identical manner as for son-parent estimates.

2.6.2 Results

Figure 2.5 plots the trend in IRA and LW estimates for birth cohorts 1962 to 1979, separately by country.[26] In Appendix Table B5, we also report the difference between the trend estimates, and test whether trends in intergenerational mobility are statistically distinguishable between the IRA and LW approaches.

First, the son-father trends obtained from the LW method are almost identical to the son-father IRA trends, which validates that the LW method captures the rank-rank association in economic status. For Norway and Sweden, IRA and LW trends are negative, indicating a development towards *increased* mobility, while Denmark's decline in mobility is supported by both the IRA and LW methods. The middle column of graphs show son-mother estimates. Compared to the IRA trend, our LW trends are noticeably weaker, and in the case of Norway even negative. This suggests a development similar to that of father-son estimates. The difference between the trends in the IRA and LW coefficients is statistically meaningful and different from zero, and also similar in magnitude across all three countries. The LW method thus mitigates attenuation bias to approximately the same extent across countries. Evidently, when using mothers' years of education and occupations - rather than just labor earnings - to proxy for their latent economic status, income persistence between male children and their mothers as well as their fathers has remained relatively constant over time. In the last column, we present comparisons between trends in LW and IRA coefficients for daughters and fathers. Again, the

[26]We focus here on the birth years 1962-79 because this is the part of the sample for whom we observe the largest increase in rank correlations.

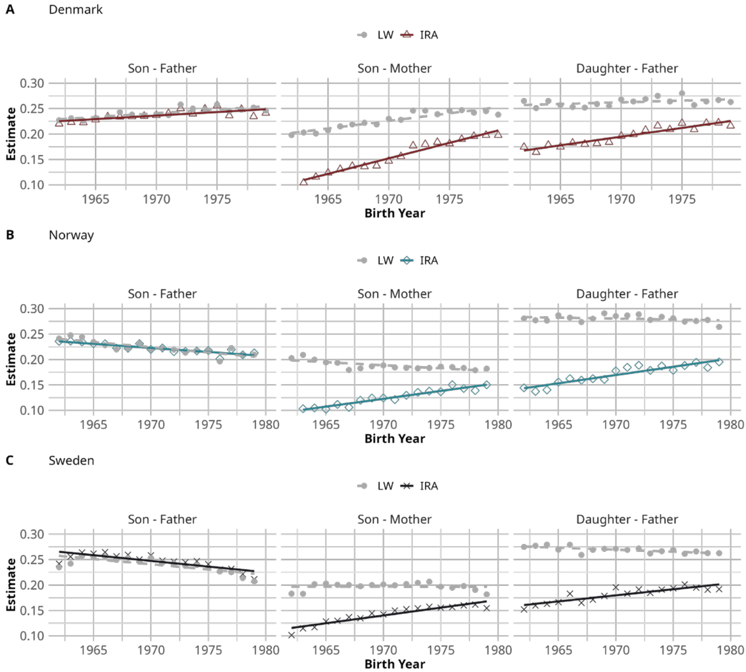


Figure 2.5: Trends in Intergenerational Mobility in Latent Economic Status.

Note: The three panels plot comparisons of intergenerational rank associations and rank associations in latent economic status (calculated using the proxy variable method from Lubotsky and Wittenberg (2006)) for Denmark, Sweden and Norway. Each panel shows, in turns, son-father, son-mother and daughter-father correlations. Each marker indicates the coefficient of a separate regression (eq. 2.2 for LW; eq. 2.1 for IRA) and each line indicates fitted trend lines for the period 1962 to 1979. "Birth Year" refers to the birth year of the child in each parent-child pair.

LW trends are significantly less steep than the IRA trends. The differences between them are again almost identical across countries. For Denmark, the adjusted trend still indicates that over time, mobility in economic status decreases, albeit at a lower rate. In Norway the relationship is stable, while daughters in Sweden experience a small increase in mobility over time.

In summary, these results provide three important takeaways. First, son-father LW trends are similar to IRA trends, indicating that the IRA captures the development of intergenerational mobility in latent economic status. Second, trends in the son-mother and daughter-father IRA appear to overestimate declines in mobility and, third, differences in trends between the IRA and LW method are comparable across countries. In addition to the comparison of trends, the levels of the son-father, son-mother, and daughter-father LW coefficients are more similar to the IRA coefficients of son-father pairs, as would be expected when accounting for attenuation in the coefficients.[27] Estimating rank associations in latent economic status by birth cohort shows that over time, father-daughter correlations have remained roughly constant at a level just below 0.3. The transmission of economic potential between parents and their female children, as well as their male children, has thus seen little change across birth cohorts from 1962 to 1979. That girls are not over time increasingly "invested in" by their parents might reflect the particular setting, with relatively equal schooling opportunities among boys and girls already for individuals born in the 1950s. On the other hand, the fact that father-daughter correlations are as high as the father-son ones suggests that whatever skills relevant to economic success are transmitted between parents and their children, these are gender-neutral.

By estimating correlations in "latent economic status" rather than observed income, our goal is a measure that better approximates the transmission of income-generating skills between parents and their children. One could argue, however, that occupational and educational choices suffer from the same low correla-

[27] This result is also supported by findings in Vosters and Nybom (2017).

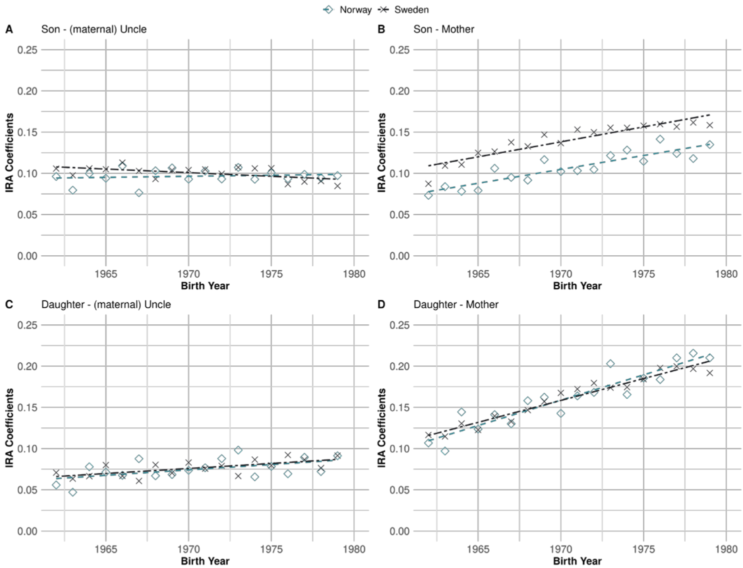


Figure 2.6: IRA Estimates Between Children and Their Maternal Uncles.

Note: The four panels depict IRA coefficients (eq. 2.1) for the correlation between children and their mothers' brothers, i.e. maternal uncles; Sons in Panel A and daughters in Panel C. Panels B and D show the estimated IRA between sons and daughters and their mothers for the sample where information on maternal uncles is available. "Birth Year" refers to the birth year of the child in each parent-child or uncle-child pair. Denmark is left out of this analysis due to issues with data availability.

tion with underlying skills as does income. To corroborate the LW results, we also estimate the intergenerational rank association in labor income between sons and their maternal uncles. Given a constant level of brother-sister correlation in earnings potential, this estimated trend captures changes in the importance of parental earnings potential for child outcomes.[28] Using observed skills of

[28]Using Swedish data, Björklund, Jäntti and Lindquist (2009) show that brother correlations in income remain similar for cohorts born between 1953 and 1968.

maternal uncles to proxy for unobserved female values is a strategy previously used by e.g. Grönqvist, Öckert and Vlachos (2017). Because the data needed for generating parental sibling links is partly unavailable, the sample size used to estimate these correlations is smaller, particularly for the earliest birth cohorts; and Denmark is left out of the analysis. Figure 2.6, Panel A, presents the results, which reveals a constant level of rank associations over time. Panel B shows the original mother-son associations for comparison, and in Panels C-D, the same results are shown for daughters and maternal uncles. Daughter-uncle trends are substantially flatter than daughter-mother trends, indicating that a certain part of the mother-daughter trends is driven by the mothers. However, the remaining IRA trend shows that increased labor force attachment by daughters over time also contributes to the observed mobility trend.[29]

2.7 Conclusion

In this paper, we document trends in intergenerational income mobility in Denmark, Norway, and Sweden, for children born between 1951 and 1979. Harmonizing data and definitions, we show that the intergenerational rank association between parents and children in individual labor earnings has increased significantly in all three countries. These trends are robust to using different types of income measures, as well as to restricting the analysis to labor market active individuals. Splitting trends by gender of par-

[29] An additional validation exercise, carried out but not featured in this paper, is to impute female incomes based on observed average incomes among men with the same level of education and occupation, within a given birth cohort. The results show stable trends between mother-son and daughter-father pairs in imputed income ranks, and are available upon request.

ents and children, son-father correlations exhibit the weakest trend in all three countries, whereas all correlations involving mothers and daughters increase over time. The strongest trend is found between mothers and daughters. We also show that a similar, but delayed, pattern of changes in mobility can be found for US parent-child pairs from the Panel Study of Income Dynamics.

Our results suggest that increasing female labor supply at the extensive and intensive margin results in higher child-parent rank associations. We show that this is the result of better manifestation of skills in income for women, such that the intergenerational correlation in "potential income", or *latent economic status* is revealed. In other words, the fact that maternal economic status was poorly reflected in maternal income among early cohorts of our sample caused rank associations between child income and joint parental income to be an attenuated measure of mobility in economic status, or "opportunity". Over time, as female labor force participation has increased and women are better able to choose occupations that match their inherent skills, this attenuation has declined.

These results highlight the importance of accounting for changes in female economic status when estimating trends in intergenerational mobility. The interpretation that higher income rank associations between children and parents reflect a lower degree of social mobility or equality of opportunity is not always easily applicable when labor market conditions change substantially. Our findings suggest that over time, female income becomes increasingly determined by their earnings potential, meaning that the traits and norms that women inherit from their parents are also better reflected in their income. While such a development must be seen as a necessary side-effect of increased gender equality in the labor market, it is not clear whether it should be seen as a reduction or advancement in equality of opportunity.

Appendices

2.A Data Registers and Variable Definitions

Denmark

The Danish income registries start in 1980 and contain detailed information on the individual income composition of Danish adults. The registries are based on information from the Danish tax authorities, supplemented with information from other Danish authorities, including unemployment insurance funds and local authorities.

The measure of labor income that is being used in this paper consists of wage payments (incl. non-wage benefits, non-taxable wage payments, stock options, and more) and any net surplus from own, private company. Gross income is equal to labor income, transfers, property income, and any other non-classifiable income that the individual may have received throughout the year. Net-of-tax income is finally equivalent to gross income net of all taxes that have been paid to either the government, municipalities, or other public authorities. Individuals with no parents in the sample (generally people who moved to Denmark, whose parents have moved abroad, or whose parents are deceased) are naturally dropped from the sample. In order to ensure comparability between income definitions across countries, negative observations of income have been set to income.

When constructing household income measures, individuals are being linked to their spouses. In the Danish sample, a spouse is generally defined by marriage, registered partnership or registered as a cohabiting couple. Matching individuals to spouses as well as parents is done using the population registries of Denmark.

Occupation data are obtained from the Danish employment classification module, AKM. Occupations are characterized by the Danish ISCO classification system (DISCO)[30] and only observable from 1992 and onward.[31] In order to impute data from 1980 to 1991, we use a random forest-algorithm in order to map a range of occupation-related variables to occupation codes in 1992.[32] We then use this mapping in order to characterize individual occupations from 1980 and up until 1990 where the occupation codes are missing. Starting from 1991, there is a range of data-breaks in the sense that certain types of occupations are re-classified, split up or gathered into one group. In order to make occupations as comparable as possible over time, we re-classify occupations across years so that classifications are approximately constant over time. For certain individuals, the occupational status is either unknown or missing. These are instead assigned to the occupational status that was observed most recently within a window of +- 3 years. Due to these imputation and re-classification procedures, any results in this paper that rely on occupation codes from Denmark should be interpreted with some caution.

Information on individuals working part time is obtained from a matched employer-employee data module, IDAP. For a series of individuals (people who are not self-employed and do not have multiple occupations), full time work is assessed in a relatively straightforward way. In cases of doubt (individuals are e.g. self-employed or have multiple part-time occupations), we assign full employment status to a individual whose labor income amounts

[30] In this paper, we focus on the 1-digit classification codes.

[31] There is a series of data breaks occurring from 1992 and onward — we attempt to handle these appropriately.

[32] Code is available upon request

to at least the *mean* labor income of full-time employed individuals (specifically, full-time employed individuals whose earnings fall within the range of the 15th and 25th earnings percentile as observed within the same age-category). Up until 2007, full-time work is classified as at least 27 weekly hours. From 2008 and onward, the equivalent number is 32. This data break does not show up in the parental generation in our sample (as these are observed up until 1998). In the child generation, gender-specific full time rates are re-scaled in the following way in order to account for the data-break when computing aggregate cohort- and gender-specific statistics:

$$\widetilde{\text{ftr}}_t^g = \begin{cases} \text{ftr}_t^g, & \text{if } t < 1971 \\ \widehat{\text{ftr}}_t^g, & \text{if } t = 1971 \\ 1 - \left(1 - \text{ftr}_t^g\right) \frac{\left(1-\widehat{\text{ftr}}_{1971}^g\right)}{\left(1-\text{ftr}_{1971}^g\right)}, & \text{if } t > 1971 \end{cases}$$

where $\widetilde{\text{ftr}}_t^g$ are the full time rates that we report in Figure 2.1, ftr_t^g are full time rates as observed in the data (hence, without accounting for the data break) and $\widehat{\text{ftr}}_t^g$ is the fitted gender-specific full time rate in 1971 that is obtained from a linear regression of full time rates, ftr_t^g, on t, using full time rates from 1966 up until 1970.

An individual's level of education reflects the highest obtained level of education. In some years, the duration of the education is not available in the data. In these cases, the duration of the education is either imputed from the duration of that same education in an earlier year, or it is imputed from educations that are characterized as being similar.

Norway

Our Norwegian data set combines information from the central population registry with information about income and earnings from the tax registry. Income data in Norway is available from 1967 to 2018. Labor income includes payments related to employment, including overtime pay, taxable sickness benefits, parental leave pay, short-term disability pay, and rehabilitation benefits. This is top-coded for a few years in the 1970s at the maximum amount for contributions to the national social security scheme (folketrygden). Gross income is the sum of labor income and taxable and non-taxable transfers and income from capital. Disposable income is defined as gross income minus taxes and is also sometimes referred to as net-of-tax income. The definitions change somewhat over time due to reforms of the benefit, insurance, and tax system. For the net-of-tax and gross income variable, the data series ends in 2014, which is why these income measures are constructed from more detailed income data only available from 1993. Spouses (married couples as well as couples in civil unions) are linked through their personal identifiers

The occupation data used for implementing the method proposed by Lubotsky and Wittenberg (2006) is pooled from matched employer-employee data (*Registerbasert sysselsettingsstatistikk*) available annually starting with the year 2000. In addition occupation data from the censuses 1960, 1970, and 1980 are added. To achieve a comparable classification of occupations we use the STYRK-98 one-digit code to group individuals into broad occupational groups (see Table A1). Individuals are assigned the occupation they have at age 36. In cases where this is not possible we use the closest applicable occupation we observe in the data. Due to the long break in occupational data between the 1980 census and the start of the employer-employee data, there might be some differences in the age at which we observe occupations for individuals that are also connected to the relevant birth year.

Data on intensive margin labor supply is obtained from an earlier version of matched-employer employee data covering the years 1986 to 2010. These data are available annually and include information on individuals' employment characteristics in connection with Norwegian firms. In order to construct the full-time share in Figure 2.1 we use a variable categorizing employment relationships into three categories depending on the contracted weekly working hours. For Norway all individuals working at least 31 hours a week are counted as full-time workers. In order to account for year-to-year fluctuations in intensive margin labor supply, we take the mode of this variable over three years to capture individual hours worked. In detail, this means that parents hours are constructed as the mode of hours worked during the three years around the time their child turns 18. For children hours are constructed as the mode during the three years around age 30. Due to data limitations the age at which child hours are captured is slightly lower than the age used for our main life-time income measure.

The educational data for the LW method is also pooled from different registries. Most individuals we observe are included in the national education database available from 1970. These data include variables for the highest achieved education of all individuals which we can link via personal identifiers. For individuals who are not included in the national education database, we obtain information about their educational attainment via census data from 1960, 1970, and 1980.

Sweden

The Swedish Income and Taxation registry starts in 1968 and holds official records of income for all individuals with any recorded income. In general, it contains all earned income from employment or businesses, capital income, taxable (mostly social insurances), and non-taxable transfers (social welfare, educational grants, child benefits, etc.). Identifiers for biological or adoptive parents are linked to the child identifier through the multi-generational register. Households are constructed by linking individuals (children, mothers, and fathers) to their spouses. This is available only for married couples (and those in registered partnerships) and thus excludes households formed by cohabiting partners.

Data on occupations are taken from two sources. First, the population censuses (Folk- och bostadsräkningarna) contain occupational codes corresponding to the ISCO-58 classification system, called NYK. This information is available in 1960, and then every five years between 1970 and 1990 for the whole adult population. Individuals who are in the labor force, but whose occupation is unknown are dropped from the sample, while individuals not in the labor force are coded as "no occupation". The census data are used to infer occupations for all parents in our Swedish sample, and we assign each parent an occupational code from the census closest in time to when the child is 18 years old (for example, a mother with a child born in 1951 will primarily be assigned an occupational code from the 1970 census, and occupations for fathers with children born in 1975 will be taken from the 1990 census). If no occupations is observed in this focal year, we search iteratively through the second and third closest waves, and so on. Parents who are missing an occupational code after this procedure, and who are at least 18 years old in 1960, are assigned occupations from that year's census. This mainly serves to capture occupations of women who are out of the labor force continuously after the birth of their first child; about 6.5 percent of the mother sample (3 percent of the fathers).

Occupational codes for the child generation are primarily taken from the 1985 and 1990 censuses for individuals born in the years 1951-1955, and primarily from population register data for birth cohorts 1956 to 1979. Since these sources use different classification systems, we have created a mapping between the NYK and the SSYK occupational codes at the 3-digit level, available upon request. The population occupation register, contained in the *Integrated database for labour market research (LISA)* uses an adapted version of the ISCO-08 classifications, called SSYK-2012, and is available in our data for the years 2012-2017[33]. As a result, the age at which occupations are observed among the child sample varies between 35 and 56, which might induce noise in between-birth cohort comparisons. On the other hand, this age span corresponds to prime working age, and occupational choice is relatively constant, especially given the broad classes we use in our analysis.

Information on parent- and children-specific rates of full time work are calculated using the Swedish Level of Living Survey, which is a nationally representative survey on about 0.1 % of the adult population from years 1968, 1974, 1981, 1991, 2000 and 2010. Respondents are asked about their contracted number of weekly working hours (1974 and on), alternatively about their number of hours worked last week (1968). Only respondents with non-zero survey weights are kept in the sample. *Parents* are defined as in-

[33] Years 2012 and 2013 are re-coded at the 4-digit level from the earlier SSYK-96 occupation codes, using the official translation key from Statistics Sweden (https://www.scb.se/dokumentation/klassifikationer-och-standarder/standard-for-svensk-yrkesklassificering-ssyk/ 2022-09-08.)

dividual between age 45 and 55 with at least one child in survey waves 1968-1991. We set the "birth year of the child" to 1951, 1957, 1964 and 1974, respectively, i.e. 18 years before the observation. *Children* are defined as individuals aged between 30 and 36 in years 1981 to 2010.

The highest attained level of education is observed in the 1970 census, and in the annual population registers that start in 1990. Each person is assigned the level of education that he or she displays in the year closest in time to when income is observed (age 36 for children; age 18 of the child for the parents). Years of education is then inferred from these categorical data (e.g. completing a three-year secondary education program is coded as twelve years of education, or eleven years if the person completed primary school when it was still only seven years in duration).

Table A1: Occupation Classification by Country.

Code	Definition
Norway	
0	Armed forces and unspecified
1	Managers
2	Professionals
3	Technicians and associate professionals
4	Clerical support workers
5	Service and sales workers
6	Skilled agricultural, forestry and fishery workers
7	Craft and related trades workers
8	Plant and machine operators and assemblers
9	Elementary occupations
Sweden	
1	Professional work (arts and sciences)
2	Managerial work
3	Clerical Work
4	Wholesale, retail and commerce
5	Agriculture, forestry, hunting, and fishing
6	Mining and quarrying
7	Transportation and communication
8	Manufacturing
9	Services
10	Military/Armed Forces
Denmark	
0	Military work
1	Management work
2	Work that requires knowledge at the highest level in the area in question
3	Work that requires knowledge at intermediate level
4	Ordinary office and customer service work
5	Service and service work
6	Work in agriculture, forestry and fisheries
7	Craft and related trades workers
8	Operator and assembly work, transport work
9	Elementary occupations

Note: Occupational categories for Norway are assigned using the STYRK-08 classificaiton provided by SSB. For Sweden the classification follows SSYK-2012 similar to Vosters and Nybom (2017). For Denmark, we use the first integer from the Danish ISCO classication (link). In the Danish case, note that this variable is not available for all years in the data. For this reason, we generate it from a set of other available occupation related variables. Code is available upon request.

Table A2: Overview Income Definitions by Country.

		Denmark	Norway	Sweden
1	Salary	taxable salary incl. fringe benefits, tax-free salary, anniversary and severance pay and value of stock options	all payments related to employment including overtime pay	all payments from employment
2	Net Profit	net profit from self-employment incl. profit of foreign company and net income as employed spouse	net income from self-employment and income from other businesses	net profit from self-employment, income from other businesses
3	Transfers	cash benefits, unemployment insurance benefits, sickness benefits, unemployment benefits, pensions, child allowance, and more	taxable sickness benefits, parental leave benefits, unemployment benefits, short-term disability payments, rehabilitation benefits	sickness benefit from employer (sjuklö), value of e.g. car, travel expenses (förmånsvärden)
		Earnings/Labor Income = Combination of 1+2+3		
4	Transfers	cash benefits, unemployment insurance benefits, sickness benefits, unemployment benefits, pensions, child allowance, and more	taxable transfers: benefits from the national insurance scheme (disability insurance, pensions, etc.) non-taxable: child benefits, housing allowance, scholarships, parental leave benefits, social assistance payments	taxable: social insurances (unemployment, parental leave etc), private pension income, stipends etc. non-taxable: pensions and annuities housing support child support social welfare alimony conscript support grants and loans for students
5	Property Income	capital and wealth income excl. calculated rental value of real estate	gross interest income, dividend income, return on life insurance, net realised capital gains (e.g. shares, house, land), other capital income (taxable rental income)	capital income (gross pre-2004, net post-2004) and after-tax rental income
		Gross Income = Combination of Earnings +4+5		
6	Other Income	other non-classifiable income		
7	Taxes	taxes on earnings, wealth taxes, property value tax, tax on share dividends/gains and more	taxes, maintenance paid, mandatory insurance premia	all taxes
8	Negative Transfers			repayments of study loans, paid alimony
		Disposable Income = Combination of Gross Income - 7+10		

Table A3: Occupational Groups.

Code	Label	Examples
00	missing	Undefined or missing occupation
01	executives	Politicians, business executives, managers
02	professionals	Professions requiring advanced college degrees
03	nurses	Nurses, midwives, physical therapy
04	teacher_lower	Pre-school and elementary school teachers
05	law	Law professionals
06	lower_professionals	Professions requiring shorter college education
07	secretary_clerical	Secretary, bank clerks, administrators
08	other_administrative	Sales persons, customer service agents, postal workers, local politicians
09	services	Waiters, beauticians, security personnel, public transport workers
10	retail	Shop assistants, cashiers, phone marketing
11	care	Assistant nurse, personal assistant, nursery staff
12	agriculture	Farming, forestry, fishery
13	construction_craft	Carpenters, welders, printers, food processing, tailors
14	machine_operators	Mining workers, steelworkers, fitters, industry machine operators
15	transportation	Truck drivers, sailors, bus drivers, train drivers
16	cleaning	Cleaning and domestic services
17	military	Military personnel, officers
18	medecine	Medical doctors, veterinaries, dentists, psychologists
19	teaching_professionals	University teachers, high school teachers, vocational teachers
20	farm_help	Planters, croppers
21	manual	Dock workers, factory workers
22	other_services	Garbage collectors, market vendors, fast food workers, janitors

Notes: The mapping between these codes and ISCO 3-digit codes is done by the authors, with the overall aim of separating female-dominated occupations from other occupations with the same 2-digit codes in the official ISCO system. These are used in the descriptive exercise in Section **??**. The exact mapping between 3-digit codes and these groups can be obtained from the authors upon request.

2.B Additional Figures and Tables

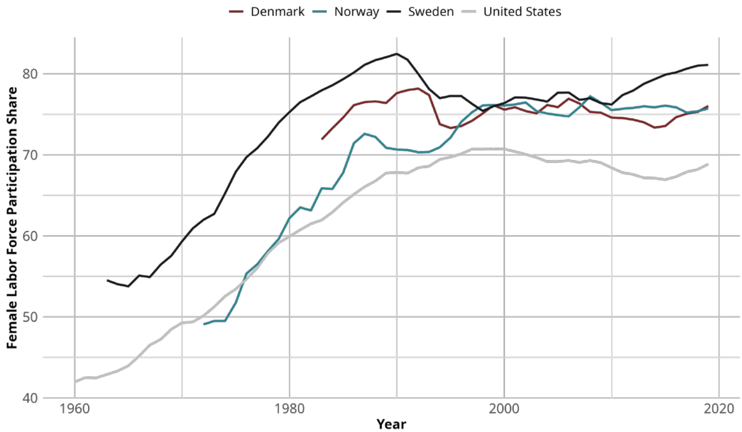


Figure B1: Female Labor Force Participation Rate.

Note: The figure depicts the labor force participation rates of women aged 15 to 64 for Denmark, Norway, Sweden and the United States. The data was obtained from the OECD (2021) and covers all years available for the respective countries.

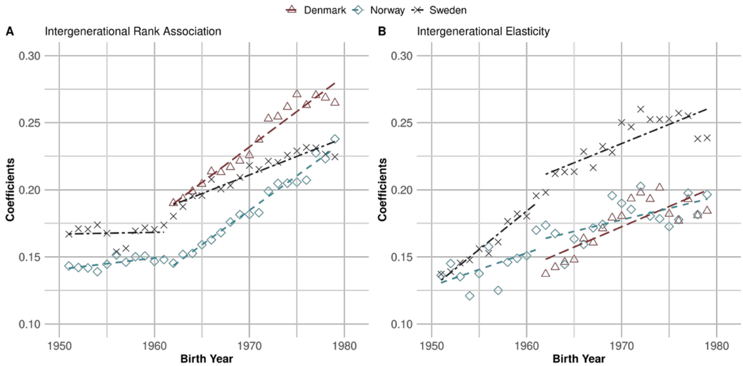


Figure B2: Estimates of IRA and IGE in Labor Income.

Note: Panel A depicts intergenerational rank associations (eq. 2.1) between parents and children for Sweden, Norway and Denmark and by birth year of the child. Panel B shows intergenerational income elasticities, i.e. correlations in log income between parent and child pairs (with zero incomes excluded from analysis). Parental income averaged over child ages 17-19, and child income averaged over ages 35-37 in all estimates. "Birth Year" refers to birth year of the **child** in each parent-child pair.

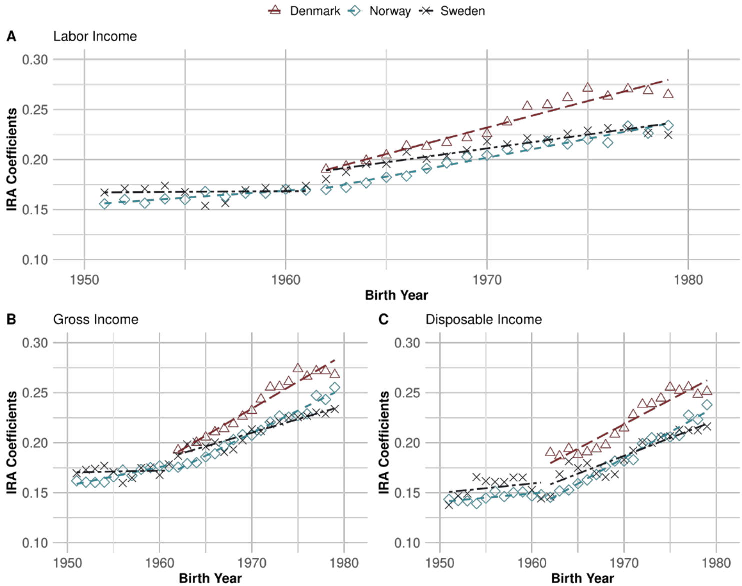


Figure B3: Estimates of IRA in Labor, Net-of-tax and Gross Income.

Note: Each panel depicts intergenerational rank associations (eq. 2.1) between parents and children, for each country and by birth year of the child. Panel A shows estimates of the main specification: labor earnings. In panel B, total factor (gross) income is used, and panel C depicts net-of-tax (disposable) income. Parental income averaged over child ages 17-19, and child income averaged over ages 35-37 in all estimates. "Birth Year" refers to birth year of the **child** in each parent-child pair.

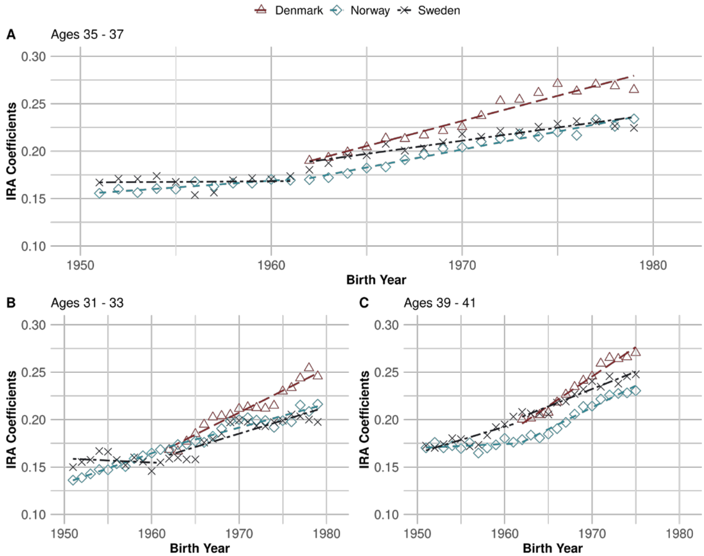


Figure B4: Estimates of IRA at Different Ages of the Child (Labor Income).

Note: Each panel depicts intergenerational rank associations (eq. 2.1) between parents and children, for each country and by birth year of the child. Panel A shows estimates of the main specification: Average income at child ages 35-37. In panel B, child income is measured at ages 31-33, and in panel C, it is measured at ages 39-41. Parental income averaged over child ages 17-19 in all estimations. "Birth Year" refers to birth year of the **child** in each parent-child pair.

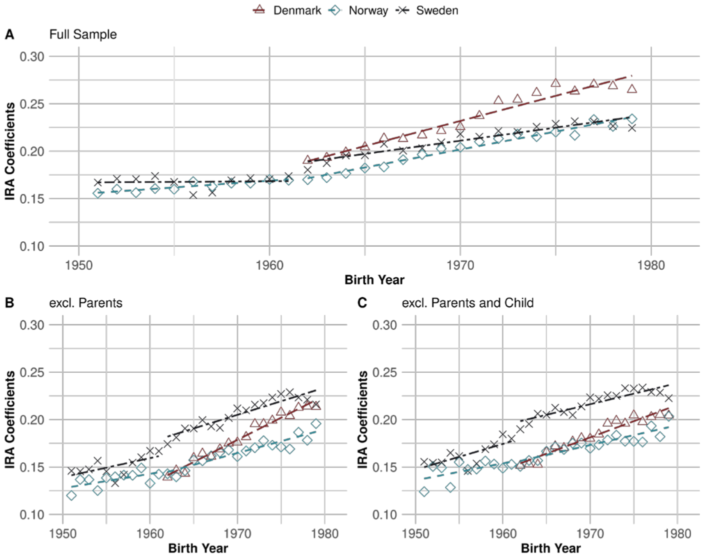


Figure B5: Estimates of IRA, Labor Force Participants Only (Labor Income).

Note: This figure shows IRA coefficients by birth year of the child, when removing non-participants from the sample. Panel A shows the baseline estimates on the full sample; Panel B removes non-participants of the child generation; Panel C removes non-participants of the child and parent generations. "Participation" is determined by an indicator variable for having annual labor earnings exceeding 10,000 USD in a year. Parental income averaged over child ages 17-19 in all estimations; child income measured as the average over ages 35-37. "Birth Year" refers to birth year of the **child** in each parent-child pair.

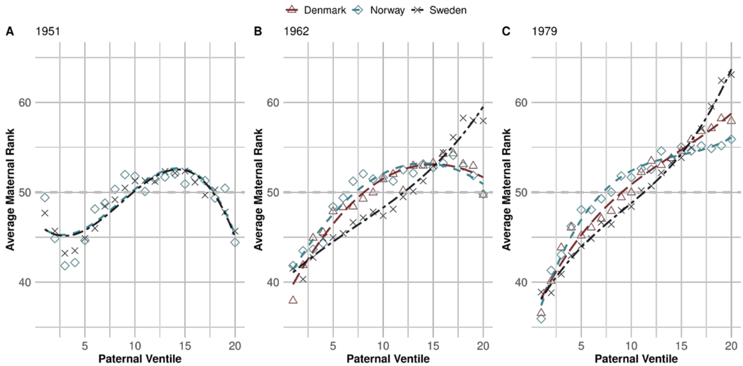


Figure B6: Average Maternal Income Percentile Rank by Paternal Income Ventile.

Note: The three panels show the average maternal income rank of mothers with children in the same birth cohort, by paternal (within parental pairs) income ventile. Each panel depicts these measures separately by country; for birth cohorts 1951 in Panel A; birth cohort 1962 in Panel B, and birth cohort 1979 in Panel C. The fitted lines are estimated with local polynomial (third order) regressions.

Table B1: IRA Coefficients and Trends - Details (United States).

	Parents	Father		Mother	
	Child	Son	Daughter	Son	Daughter
A: Weights					
Pooled IRA	0.317***	0.336***	0.195***	0.097***	0.137***
	(0.017)	(.022)	(0.031)	(0.025)	(0.029)
Trend × 100	0.603***	-0.240	0.980***	0.136	1.047***
	(0.149)	(0.205)	(0.277)	(0.253)	(0.292)
N	5,392	2,272	1,637	2,477	2,205
B: No weights					
Pooled IRA	0.335***	0.360***	0.237***	0.107***	.152***
	(0.013)	(0.020)	(0.025)	(0.021)	(0.022)
Trend × 100	0.449***	-0.263*	0.728***	0.268	0.917***
	(0.118)	(0.178)	(0.229)	(0.202)	(0.213)
N	5,392	2,272	1,637	2,477	2,205
C: SRC sample					
Pooled IRA	0.294***	0.327***	0.192***	0.098***	0.126***
	(0.018)	(0.023)	(0.0353)	(0.026)	(0.032)
Trend × 100	0.433**	-0.393	1.156***	0.180	0.727
	(0.162)	(0.218)	(0.305)	(0.266)	(0.327)
N	2,927	1,583	904	1,497	1,001

Note: The table presents estimates of the IRA and linear trends in the IRA separately for different child-parent combinations. Due to the small sample sizes, trends have been estimated directly on the underlying micro data by regressing cohort-specific child ranks on cohort-specific parent ranks interacted with a linear time trend. The trend coefficients and standard errors have been multiplied by 100 in order to avoid too many digits after the separator. Panel A contains estimates for the full PSID sample using provided sample weights, Panel B uses the full sample without weights and Panel C includes estimates on the nationally representative SRC sample. Standard errors are in parentheses. P-values indicated by * < 0.1, ** < 0.05, *** < 0.01.

Table B2: IRA Coefficients and Trends - Age 30 (United States).

	Parents	Father		Mother	
	Child	Son	Daughter	Son	Daughter
A: Weights					
Pooled IRA	0.327***	0.318***	0.222***	0.120***	0.151***
	(0.015)	(0.022)	(0.026)	(0.025)	(0.027)
Trend × 100	0.643***	0.133	0.661***	0.610**	0.571**
	(0.129)	(0.193)	(0.220)	(0.239)	(0.262)
N	6,652	2,664	2,109	2,685	2,611
B: No weights					
Pooled IRA	0.345***	0.341***	0.263***	0.148***	0.168***
	(0.012)	(0.018)	(0.021)	(0.019)	(0.020)
Trend × 100	0.457***	0.102	0.510***	0.429**	0.567***
	(0.101)	(0.59)	(0.183)	(0.176)	(0.181)
N	6,652	2,663	2,109	2,686	2,611
C: SRC sample					
Pooled IRA	0.303***	0.310***	0.225***	0.097***	0.133***
	(0.016)	(0.023)	(0.028)	(0.027)	(0.030)
Trend × 100	0.528***	0.020	0.586**	0.661**	0.352
	(0.146)	(0.210)	(0.245)	(0.261)	(0.307)
N	3,451	1,757	1,161	1,460	1,142

Note: The table presents estimates of the IRA and linear trends in the IRA separately for different child-parent combinations Children's income is measure at age 30. Due to the small sample sizes, trends have been estimated directly on the underlying micro data by regressing cohort-specific child ranks on cohort-specific parent ranks interacted with a linear time trend. The trend coefficients and standard errors have been multiplied by 100 in order to avoid too many digits after the separator. Panel A contains estimates for the full PSID sample using provided sample weights, Panel B uses the full sample without weights and Panel C includes estimates on the nationally representative SRC sample. Standard errors are in parentheses. P-values indicated by * < 0.1, ** < 0.05, *** < 0.01.

Table B3: PSID Sample Size with Parent-Child Links by Birth Cohort (United States).

	Parents	**Father**		**Mother**	
Child birth year	Child	Son	Daughter	Son	Daughter
1947	76	27	26	35	39
1948	107	38	39	48	56
1949	143	52	44	73	65
1950	171	57	72	68	99
1951	218	76	81	101	108
1952	193	70	78	88	102
1953	239	87	98	116	117
1954	235	78	96	101	128
1955	267	103	98	122	136
1956	263	86	106	107	147
1957	247	95	85	126	115
1958	220	75	95	95	119
1959	159	79	37	107	46
1960	177	96	34	121	54
1961	105	54	27	62	38
1962	117	50	37	63	53
1963	125	49	41	64	56
1964	100	47	30	56	43
1965	91	42	20	49	40
1966	88	39	17	52	34
1967	95	49	21	60	33
1968	66	32	17	39	26
1969	99	49	32	55	42
1970	87	40	20	50	35
1971	92	40	21	60	32
1972	111	49	22	63	43
1973	107	50	20	65	36
1974	117	51	25	64	48
1975	128	55	36	70	48
1976	132	58	37	63	53
1977	130	66	23	41	21
1978	138	66	34	27	30
1979	179	93	43	29	41
1980	142	65	36	33	26
1981	148	77	30	40	24
1982	120	58	27	26	30
1983	160	74	32	38	42
Total	5,392	2,272	1,637	2,477	2,205

Note: The table presents the number of cohort-specific parent-child links that were used to produce the main results from the PSID survey data.

Table B4: IRA Coefficients, Trends and Differences Across Countries and Time.

	1951		1962			1979			Trend 1962-1979			Δ P-value		
IRA Spec.	NO	SE	DK	NO	SE	DK	NO	SE	DK	NO	SE	DK-NO	DK-SE	NO-SE
All	0.156	0.167	0.190	0.170	0.180	0.265	0.234	0.225	0.530	0.379	0.277	0.065	0.004	0.176
									(0.035)	(0.018)	(0.033)			
Son-Parent	0.242	0.245	0.225	0.222	0.233	0.280	0.241	0.235	0.360	0.085	0.014	0.000	0.000	0.736
									(0.036)	(0.024)	(0.061)			
Daughter-Parent	0.146	0.158	0.197	0.173	0.169	0.276	0.262	0.240	0.592	0.552	0.428	0.363	0.067	0.020
									(0.035)	(0.037)	(0.038)			
Son-Father	0.253	0.248	0.220	0.236	0.242	0.241	0.213	0.211	0.139	-0.160	-0.224	0.000	0.000	0.378
									(0.035)	(0.022)	(0.060)			
Son-Mother	0.068	0.080	0.098	0.089	0.101	0.198	0.150	0.155	0.619	0.324	0.307	0.000	0.007	0.723
									(0.026)	(0.026)	(0.041)			
Daughter-Father	0.137	0.139	0.175	0.144	0.152	0.216	0.195	0.192	0.342	0.325	0.239	0.778	0.570	0.421
									(0.031)	(0.032)	(0.034)			
Daughter-Mother	0.073	0.077	0.120	0.119	0.112	0.227	0.221	0.194	0.731	0.607	0.541	0.439	0.181	0.107
									(0.036)	(0.037)	(0.034)			

Note: Columns (1)-(7) report the IRA coefficients of separated regressions in the years 1951, 1962 and 1979 separately for Denmark, Norway and Sweden. Columns (8)-(10) report the coefficient of the fitted regression lines of country specific regressions of the IRA coefficient on a linear trend for the years 1962 to 1979. The trend coefficients and corresponding standard errors have been multiplied by 100 in order to avoid too many digits after the separator. Columns (11)-(13) report rounded p-values for the null hypothesis that the slopes for the respective countries (see column header) are equal. Robust standard errors are reported in parentheses.

Table B5: Comparison of Trends 1962 - 1979.

	Denmark	Norway	Sweden
Panel A: Son - Father			
IRA	0.1385	-0.1598	-0.2243
	(0.0349)	(0.0222)	(0.0605)
LW	0.1504	-0.2062	-0.1898
	(0.0239)	(0.0350)	(0.0613)
Difference	-0.0118	0.0464	-0.0346
	(0.0423)	(0.0414)	(0.0861)
Panel B: Son - Mother			
IRA	0.6186	0.3244	0.3069
	(0.0256)	(0.0262)	(0.0408)
LW	0.2994	-0.1200	0.0175
	(0.0353)	(0.0273)	(0.0495)
Difference	0.3192	0.4444	0.2894
	(0.0436)	(0.0379)	(0.0642)
Panel C: Daughter - Father			
IRA	0.3416	0.3247	0.2388
	(0.0309)	(0.0318)	(0.0339)
LW	0.0658	-0.0385	-0.0897
	(0.0324)	(0.0314)	(0.0201)
Difference	0.2758	0.3632	0.3285
	(0.0447)	(0.0447)	(0.0394)

Note: IRA indicates linear trends estimated through all coefficients of the intergenerational rank association. LW specifies linear trends estimated through all coefficients obtained from applying the Lubotsky-Wittenberg method. The trend coefficients and corresponding standard errors have been multiplied by 100 in order to avoid too many digits after the separator. Difference indicates differences between LW and IRA trends and tests the null-hypothesis of equality in trends between the IRA and LW coefficients. Heteroskedasticity robust standard errors are in parentheses.

2.C Calibrating Parameters in Model

Each set of country-year model parameters for trend decomposition are — loosely described — calibrated in the following steps:

1. If the year is the first year of observation for a given country, draw a random set of parameters. If the year is not the first year of observation, initialize the algorithm with the optimal set of parameters from the last year associated with the same country. These become the 'search parameters' until they are replaced.

2. Draw 100,000 parent-child pairs (the same in each year), and repeat the following procedure until there is a sufficiently close match between empirical rank associations and modeled rank associations[34]:

   (a) Compute skills and incomes for all individuals (father, mother, son and daughter) using the set of 'search parameters' along with randomly drawn values for x_{it}^{k} and ε_{it}^{k}.

   (b) Compute associations in income ranks between (i) fathers and sons, (ii) fathers and daughters, and (iii) mothers and sons, and (iv) mothers and daughters, while (v) matching the relationship between mother and father income ranks.

   (c) If the convergence criterion is not met, adjust the parameters using a customized variation of gradient descent. These now become the 'search parameters'.

[34]Or stop the algorithm early if it stops converging, i.e. it seems that a much better match cannot be achieved.

In the following set of graphs, we illustrate how the implied empirical association between the two types of income in the (calibrated) simulated data compares to the empirical association between the same two incomes as observed in the data. These figures validate the quality of the calibration exercise.

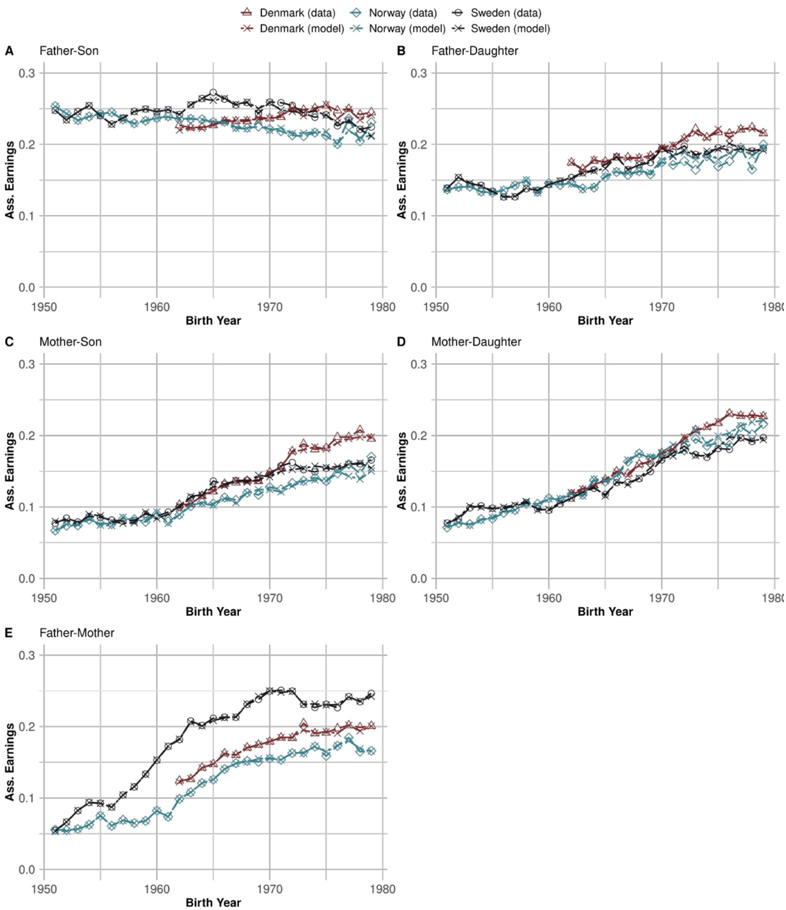


Figure C1: Validation of Calibration Exercise.

Note: Each panel displays the empirical association between two incomes as observed in the data as well as the implied empirical association between the same two types of income in the simulated data as calibrated in the decomposition model. "Birth Year" refers to birth year of the **child** in each parent-child pair.

Chapter 3

Consequences of Rapid Structural Transformation - Evidence from Hydropwer Expansions

Abstract: The establishment of hydroelectric power plants resulted in a rapid structural transformation of Norwegian municipalities around the beginning of the 20th century. Using a novel dataset linking individuals born between 1890 and 1910 to historic death data, I find that experiencing childhood in rapidly transforming local areas leads to an increase of ten months in age at death for men. This effect is entirely driven by individuals born into higher socioeconomic status households. I find that incomes, manufacturing, immigration and economic inequality in local areas in the short/medium-term increase after the introduction of hydropower, while public health deteriorates at the same time. This suggests that, in the long term, economic development through structural transformation outweighs the negative consequences of a deteriorating public health environment and thereby increases the lifespan of individuals.

I am thankful for advice and comments from Eirik Berger, Prashant Bharadwaj, Aline Bütikofer, Andreas Haller, Sam Hirshman, Katrine V. Løken, Kjell G. Salvanes, Alexander Willén and participants at seminars at NHH, UCSD and the FAIR Midway Conference. The activities at NHH are funded by the Research Council of Norway through project No 275800 and through its Centres of Excellence Scheme, FAIR project No 262675.

3.1 Introduction

New technologies and large infrastructure investments can have significant transformative effects on local economies. In the United Kingdom, the adoption of the spinning jenny started the industrial revolution (Allen, 2009). The early adoption of electricity in the United States in the 20th century led to increased economic activity up to seven decades later (Lewis and Severnini, 2020), while railway access in India resulted in significant income gains, as well as increases in trade volumes around the turn of the 20th century (Donaldson, 2018). Technology adoption and infrastructure investments can have important long-term impacts through structural change accompanied by industrialisation, economic development and economic growth. Conversely, these changes may come at a cost. Externalities of increased economic activity might create pollution and result in rising mortality rates (Beach and Hanlon, 2018), while increased economic activity and urbanisation can lead to crowded housing and poor sanitation (Cain and Hong, 2009; Boustan, 2018), and improvements in transportation networks potentially facilitate the spread of disease (Adda, 2016; Oster, 2012). Moreover, the distributional consequences of positive and negative effects of structural change are *a priori* ambiguous. These examples raise important questions about the long-term effects of structural transformation and the channels through which they operate.

Understanding the channels through which structural transformation impacts local areas and their population, in the short, medium, and long term is challenging for various reasons. Policies leading to structural change are usually implemented after careful consideration or with certain objectives in mind, and it might therefore prove difficult to identify causal effects due to correlated unobservables and selection into areas which differ on various dimensions. An additional complication concerns data requirements when investigating the long-term consequences of structural transformation and its implications. For example, to investigate the individual-level impact of policies as far back as the 19th century, a researcher needs detailed historical accounts linked to contemporary individual-level data in order to investigate the long-term implications of said policies.

Consequently, evidence of the long-term implications of structural change and its short- and medium-term channels is hard to obtain. Most of the previous literature on the implications of structural transformation has a narrow focus on aspects connected to or coinciding with industrialisation, such as pollution, electrification and public health infrastructure. My setting allows me to focus more broadly on the implications of local comparative advantages in electricity production and consumption, and how this shaped local areas, on the one hand, and individuals in those areas, on the other. I overcome the identification challenges by leveraging the staggered adoption of hydropower as a shock to local economies, inducing structural change in the context of rural Norwegian municipalities during the period between 1890 and 1920. I introduce novel historical death data starting in 1928 linked to full population census data from 1910 to create a dataset spanning almost 120 years. This thereby links the structural transformation process at the beginning of the 20th century to outcomes up until today. I supplement these data with newly collected information on public health and economic indicators in local areas obtained from historical documents.

There are two main advantages of focusing on the Norwegian context. First, around the turn of the century, some Norwegian municipalities underwent rapid structural change mainly driven by their topography, which was favourable for the establishment of hydropower plants. These plants significantly increased the availability of electricity in predominantly rural areas and made it possible for large industrial factories to operate in previously very agriculture-based communities (Venneslan, 2009). At this time, no electric grid existed, which made hydroelectric power a comparative local advantage. I use the staggered establishment of these hydropower plants as my main identification strategy. I will also provide evidence of certain outcomes using a topographic instrument inspired by Duflo and Pande (2007), and further tailored to the Norwegian context by Borge, Parmer and Torvik (2015), to provide causal estimates of the short-term implications of hydropower-induced industrialisation between 1890 and 1920. The second advantage comes from the availability of detailed historical health and economic records in Norwegian districts and municipalities, which I am able to combine with full population census data from 1900 and 1910. Furthermore, I can link individuals in the 1910 census to individual-level death records starting in the year 1928 by applying a data-linking algorithm set out in Abramitzky et al. (2021). In addition, matched individuals dying after 1960 can be linked to modern Norwegian administrative data, enabling me to assess the effect of childhood exposure to hydropower-induced structural change on lifetime education and residential mobility. In combination, the rich set of historical and individual data enables me to investigate the long-term implications of structural change decades after childhood exposure, while also addressing the short- and medium-term economic and health consequences at the local level that could provide important channels for these long-term effects.

The main results I present in this paper concern the long-term implications of childhood exposure to a structurally transforming environment. Using linked historical census data connected to individual-level death data starting in 1928, I am able to provide evidence that men born in hydropower municipalities experience an increase of approximately ten months in age at death. Using event-study methodology to estimate the long-term implications of staggered hydropower adoption for age at death suggests that hydropower adoption has a particularly strong effect on age at death if it occurs at least five years prior to birth. This thereby suggests that, in order for structural transformation to have beneficial long-run implications on longevity, a certain maturation or adoption period is required. Breaking down samples by occupation-based socioeconomic status of the household head, I find that the benefits of hydropower establishment in terms of age at death are entirely driven by individuals from high socioeconomic status families. Hence, hydropower establishment probably contributed to widening the gap in longevity among men and could explain some of the differences in mortality we observe among old-aged individuals today.

After documenting the longevity consequences of hydropower adoption, I will turn to the short and medium-term developments in local areas that provide possible channels for these longevity effects. Building on and extending previous research by Leknes and Modalsli (2020) on the industry and occupational changes resulting from hydropower adoption in Norway, I show that economic opportunities in terms of income, wealth and manufacturing jobs increased rapidly in response to hydropower adoption, whereas the share of individuals working in agriculture declined. This suggests that municipalities adopting hydropower were subject to rapid structural transformation and offered greater economic opportunities than comparable municipalities without access to hydropower.

These improved economic prospects attracted internal migrants from other counties. Using a linked sample of individuals from both the 1900 and 1910 full population censuses, I show that individuals residing in hydropower municipalities in 1910 are significantly more likely to have moved there from a different municipality, most often from within the same county.[1] Moreover, movers were positively selected in terms of an occupation-based socioeconomic status measure, suggesting that individuals with marketable skills, as measured by their occupation, were able to relocate to areas with potentially greater economic prospects. This attraction of individuals from other Norwegian municipalities combined with a generally growing population led to a significant population increase compared to non-hydropower areas. Rural municipalities gaining access to hydropower and industrialisation had a 3.5 times higher population growth rate during the period 1890 to 1920 than areas that had not established a hydropower plant by 1920.[2] The combination of rising incomes and the influx of skilled migrants also translated into rising inequality among the local population. Gini coefficients of an occupation-based stratification measure increased by approximately 11 per cent in hydropower areas between 1900 and 1910, while only increasing by approximately 6 per cent in non-hydropower municipalities.[3]

[1] In 1900, Norway was divided into 20 counties (*fylker* in Norwegian) with an average population of approximately 120,000 inhabitants and an average area of 16,000 sq. kilometres.

[2] Rural hydropower municipalities grew by 42 per cent from 1890 to 1920, while other rural municipalities only saw a population increase of approximately 12 per cent.

[3] Since occupation-based measures of inequality only capture differences across

After documenting the way in which hydropower municipalities structurally transform and industrialise, I provide novel evidence of the impact of hydropower adoption on public health in Norwegian health districts.[4] Reports from district doctors during this period suggest that population growth that went hand in hand with the industrialisation process outpaced infrastructure development in terms of guaranteeing hygienic standards. Estimating event-study specifications examining the number of cases of common infectious diseases confirms that areas adopting hydropower saw up to a 25 per cent increase in annual infectious disease cases per 100,000 inhabitants. At the same time, the overall supply of health professionals (e.g. doctors, midwives, pharmacists) relative to the population size remains unchanged. Taken together, these developments created two main counteracting forces. On the one hand, hydropower areas experienced increased economic prosperity, which likely increased existing inequalities, while, on the other hand, the prospect of economic opportunity led to population growth and a worsening of the public health environment in the short- to medium-term.

In combination, these results suggest that, despite a deterioration of the public health environment, the positive implications of structural change, such as the increase in per capita income, had a stronger effect on the longevity of individuals. Multiple channels could potentially be at work simultaneously. Income could generate better nutritional standards, improve education and also change the occupational sorting of individuals into less hazardous

occupations, this could be seen as a type of job diversification measure, rather than as pure socioeconomic inequality. However, assuming that income differences do exist within occupations, this measure could be seen as a lower bound to socioeconomic inequality.

[4]Health districts are aggregations of municipalities and there were approximately 150 health districts during the period 1880 to 1920.

occupations. My results shed light on the importance of family background in connection with these longevity effects, but do not inform the specific mechanism due to the multidimensionality of structural transformation.

With this paper, I make three main contributions to the understanding of the economic and health implications of structural transformation. The first and main contribution concerns the long-term implications for longevity of structural transformation of local areas induced by hydropower adoption during childhood. Importantly, longevity in this context should be seen as more than a simple long-term measure of health, but rather as a complex accumulation of advantages and disadvantages over the life cycle. By thinking in terms of this much broader definition of longevity, these long-term effects of structural transformation go beyond direct health effects. Longevity can then be influenced through channels such as occupational choice, geographic mobility, and marital sorting, but also by the direct implications of the public health environment (Case, Lubotsky and Paxson, 2002).

The structural change in hydropower areas in this context has two main aspects I consider relevant to longevity. First, the structurally changing environment is connected to population growth and increased economic activity and potentially leads to a worse public health environment during childhood. Urbanisation and economic activity have previously been associated with higher levels of infant mortality, the spread of infectious disease and higher levels of pollution (Boustan, 2018; Beach and Hanlon, 2018; Alsan and Goldin, 2019). This is also the focus of a large body of literature summarised by Almond, Currie and Duque (2018), which finds significant and lasting impacts of early-life exposure to pollutants and pathogens on human and health capital in adulthood.

Ferrie, Rolf and Troesken (2015) provide evidence of the negative impact of lead exposure during childhood on later-life cognition. Exposure to pathogens at different stages of childhood has also been causally linked to negative health and cognitive outcomes in later life (Almond, 2006; Daysal et al., 2021). Even though air pollution was not a direct externality from hydroelectric power generation, the associated industrial production created pollution in previously pristine environments. Air pollution's negative short and long-term impacts on human capital and health have also been documented in a plethora of articles (Beach and Hanlon, 2018; Bailey, Hatton and Inwood, 2018). An additional counteracting force, positively associated with longevity, would be the increased availability of resources and economic opportunities during childhood, setting the stage for economic advantage over the entire life cycle. Chetty, Hendren and Katz (2016) show that neighbourhood resources during childhood matter significantly and that longer exposure to lower-poverty neighbourhoods has significant benefits in terms of college attendance and income later in life. In their overview article, Deryugina and Molitor (2021) show that places affect health through a variety of channels, including socioeconomic composition, pollution, health care provision and nutrition. In the Norwegian context, such place-based policies have been shown to improve infant and later-life outcomes through the provision of health care at local infant health clinics (Bütikofer, Løken and Salvanes, 2019). These particular clinics were introduced at a much later point in time, however. Even though household electrification was simply a side effect of this early hydropower expansion it is possible that improved indoor lighting and connected improvements in indoor air quality could have long-run benefits for children as well.[5] The literature identifies a variety of channels through which places undergoing structural transformation induced by hydropower can potentially affect the longevity of individuals. Due to the counteracting force of some of these channels, however, it is ambiguous what the long-term effect on longevity will be in such cases.

The second contribution lies in providing causal evidence of the impact of hydropower adoption on public health and public health infrastructure in areas undergoing hydropower-induced structural change. Previous literature predominantly focused on certain aspects of public health connected to industrialisation and urbanisation. A large body of literature on the urban mortality penalty emphasises the importance of waste, water and sewage disposal in urban areas, which is strongly connected to the increased spread of infectious diseases. Alsan and Goldin (2019) show that the combination of clean water and effective sewage systems reduced urban child mortality in the Boston area by approximately 33 per cent between 1880 and 1920. In their seminal paper, Cutler and Miller (2005) point to the importance of water filtration and chlorination technologies to mortality in US cities at the beginning of the 20th century, and show that large parts of the decline in overall, child and infant mortality can be attributed to clean water technologies.[6] Beach and Hanlon (2018) focus on the negative consequences of air

[5]The main use of electricity in households at the time was lighting, while heating was still mostly wood-based. Improvements in indoor air quality therefore mainly came from lamps, which likely contributed little to overall indoor air pollution.

[6]In a recent article, Anderson, Charles and Rees (2018) find no effect on overall and infant mortality due to clean-water technologies. They attribute this discrepancy to transcription errors in the article by Cutler and Miller (2005). Nevertheless, multiple other papers (see, e.g., Alsan and Goldin (2019); Beach (2021); Beach et al. (2016); Ferrie and Troesken (2008); Troesken (2001)) have provided credible evidence of the importance of clean water and clean-water technologies in terms of reducing mortality, particularly in urban areas in the US.

pollution resulting from coal power plants in British cities at the end of the 19th century. They find that, during industrialisation, coal-induced air pollution can explain approximately one-third of the urban mortality penalty. My context differs from the aforementioned articles in that it does not consider cities, but rather focuses on predominantly rural areas that experience rapid population growth due to the economic opportunities they offer. Moreover, the negative health development in response to hydropower adoption mainly comes from the pace at which the population grows, rather than from urbanisation and pollution through, e.g., coal. I will show that this population growth, outpacing infrastructure development, contributes to the spread of infectious diseases, while simultaneously not being addressed through a larger supply of medical professionals. I argue that the increased availability of resources in hydropower areas likely had a mitigating effect on the consequences of the spread of infectious diseases.

The final contribution is connected to a large body of literature on the implications of structural transformation for local areas. First, I contribute to the literature investigating the transformative effects of infrastructure investments on local areas. Kline and Moretti (2014), for example, investigate the implications of large federal transfers in the US in the 1930s and show that, in the short term, both manufacturing and agricultural employment increased significantly. Lewis and Severnini (2020) show that rural electrification only impacted agricultural productivity in the US, and had very limited short-term effects on employment in other sectors. In the long term, early access to electricity increased economic growth even after the US was fully electrified. This path dependency is also supported by evidence from US portage sites (Bleakley and Lin, 2012). Leknes and Modalsli (2020) show, in the context of

Norway, that areas adopting hydropower experienced rapid structural transformation and higher occupational mobility than areas without early hydropower adoption. I add to this understanding by providing event-study estimates of population changes in local areas, and I also provide novel event-study evidence of income and wealth development in response to hydropower adoption using newly transcribed data on municipality-level income, wealth and taxpayer statistics. Moreover, I show that internal migrants were positively selected from other municipalities. This positive selection into booming local economies is consistent with predictions of the adoption of the Roy model (Roy, 1951) in Borjas (1991), which shows that high returns to skill in the receiving location led to a disproportional selection of immigrants from the top end of the source location's income distribution.

The remainder of this paper is structured as follows. Section 3.2 provides an overview of the data and background. In section 3.3, I present the event-study research design as well as the instrumental variable approach used for some specifications. The results are included in section 3.4, and are subdivided into sections discussing the i) implications of childhood exposure to hydropower-induced structural change for longevity, ii) short/medium-term local structural transformation, and iii) the short/medium-term public health effects. Finally, section 3.5 concludes the paper.

3.2 Data, Background and Definitions

I will first provide a brief overview of the historical context with the focus on industrial development and Norway's adoption of hydroelectric power. I will then describe the data sources in more detail in separate sections. I will also discuss the historical census data, the historical death register and the linking process used to combine these data sources. I then provide a short overview of the aggregate economic indicators for Norwegian municipalities between 1890 and 1920, obtained from various sources. In the final part of this section, I will discuss the newly transcribed public health records and provide a brief overview of public health in Norway at the beginning of the 20th century. In Appendix 3.C, I provide more detailed information about sources and how to access the data used in this project.

3.2.1 Hydropower Establishment

The idea of using the power of rivers to drive machines has been around for several centuries. Water was used in mills to grind grain as far back as the 13th century (Skansen, 1958). Mechanical use of hydropower such as sawmills and other large machinery was also introduced later. Hydroelectric power production started already during the 1880s, but initially with very low capacity. This energy production was mainly used for street lighting in Norwegian towns and would not have been sufficient to power large industrial enterprises. Starting in 1890, larger hydroelectric power plants were built, which could finally produce energy in quantities sufficient for industrial use. These hydropower plants were ideally suited to Norway due to two main factors: Norway's abundance of water bodies such as rivers, waterfalls and streams, and the country's mountainous topography, which opens for the exploitation of altitude and the potential energy of water. The period around the beginning of the 20th century was characterised by a strong movement towards modernisation, both in society in general and economically. Norway had already developed light industry in the first phase of industrialisation during the 19th century, but, due to a lack of financial institutions, it was slightly slower in industrialising than, for example, its neighbouring country Sweden, with which it was in a union until 1905.[7] In contrast to Sweden, Norway had not developed banks that served the whole country but mostly relied on small, local mercantile banks (Sejersted, 2021). However, these banks were not able to finance large infrastructure projects. Additionally, there was no organised technological higher education that could form the basis for science-driven economic innovations. The latter problem was addressed by the establishment of the Norwegian Technical Institute (*Norges tekniske høgskole*) in 1910. The initial lack of know-how and financial institutions made it necessary to rely on foreign capitalists in the beginning. They were interested in Norway's water as a natural resource, and Norwegian legislation made investments extremely attractive. At the turn of the century, waterways were privately owned in contrast to other countries. This meant that waterways could be bought and used freely and would potentially not be available for the public benefit. Increased foreign ownership of waterways and of newly constructed hydropower plants led to heated debate about how to protect Norwegian natural resources from foreign ownership. This ultimately resulted in the adoption of the concession laws, also often called the 'panic laws' (Sejersted, 2021). The first such laws enacted in 1906 limited the duration of ownership and acquisition of waterways in Norway (Faugli, 2012). Later extensions of the scope of these laws were adopted in 1909 and 1917, with a particular focus on the speed of new developments. They also stipulated that, after a period of 60 to 80 years, any private developments

[7] The union with Sweden mainly involved foreign policy and the sharing of a common monarch. Almost all other affairs of state were organised separately in the two states.

would pass back into public ownership without compensation.

The analysis in this paper is centred around data on hydropower adoption obtained from (NVE, 1946). This report includes information about all hydropower plants established before 1920, including relevant information about their characteristics. I transcribed information on the municipality, the establishment year and the ownership of all installations with a capacity larger than 500 kW, in a similar way to Leknes and Modalsli (2020).[8] Smaller hydropower plants were probably not suited to supplying electricity to large factories and were used for mechanical power transmission, for example for mills. By excluding smaller hydropower plants, I also shifted the focus to areas that had local electricity production that was capable of meeting the energy demand of industries, thereby excluding electricity use in households. The ownership status of hydropower plants can be broken down into private, municipal and state-owned hydropower plants. In Figure 3.1, the cumulative development of hydropower plants is plotted. Between 1890 and 1920, around 150 hydropower plants were established all over Norway. It is clearly shown that, up until 1906, most hydropower plants were privately owned. It was only after the first concession laws were passed that public hydropower developments increased in number.

The ownership structure was the result of foreign investors' interest in hydropower production to power local industry, and it made it possible to prioritise industrial over public electricity consumption. Even though around 40 per cent of hydropower plants were already owned by municipalities by 1920, most of the elec-

[8]Municipalities in Norway changed quite frequently. From the end of the 19th century to the 1930s, the number of municipalities grew from around 600 to over 700. I harmonised the data and applied the municipality structure of 1900, when there were 594 municipalities.

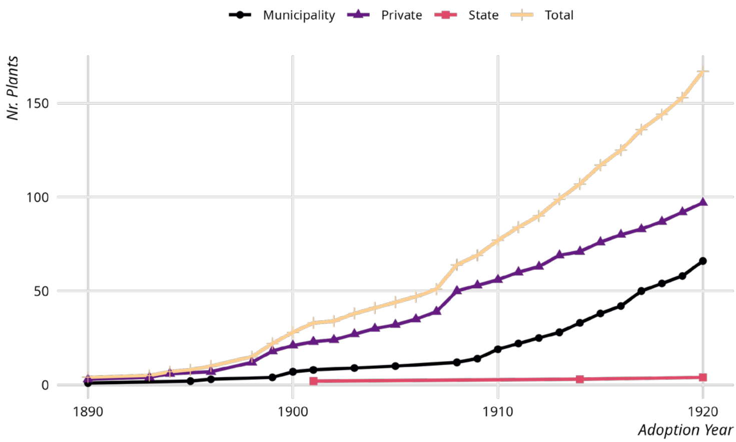


Figure 3.1: Cumulative Number of Hydrowpower Plants by Ownership.
Note: The figure depicts the cumulative number of hydropower plants by ownership. The data were obtained from NVE (1946) and include all hydropower establishments with a capacity of at least 500 kW.

tricity was used for industrial purposes (NVE, 2016). Norway also had a relatively high electrification rate early on. By 1934, approximately 74 per cent of households had access to electricity, which was mainly used for indoor lighting. Importantly, hydroelectric power had previously been used before for public electricity consumption such as street lighting, but, overall, this only accounted for a small share of total electricity consumption. In 1932, electricity consumption by industry accounted for 80 per cent of total energy consumption. This share was on a declining trend and had likely been higher in earlier years. Another important feature of this early electrification period was the hyperlocal use of electricity due to limitations on electricity transmission through an electric grid. The limitation on transporting electricity is an important factor because it allowed previously rural and remote areas to suddenly acquire a comparative advantage in relation to industrial production.

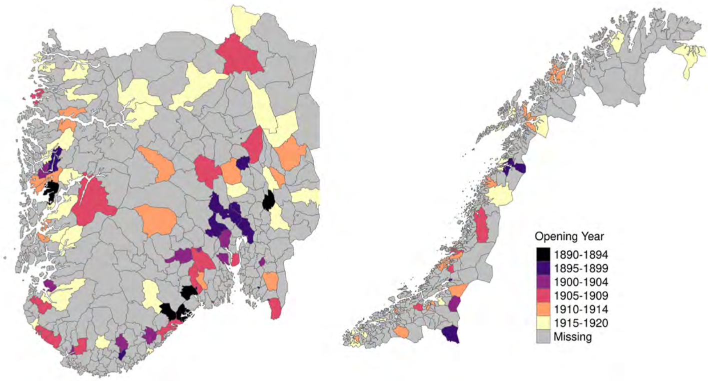


Figure 3.2: Geographical Distribution of Hydropower Plants by Establishment Year.
Note: The figure shows the establishment year and geographical location of hydropower plants in Norway during the period 1890 to 1920 for the entire country. The left side of the figure plots the southern part of Norway, while the northern part is shown on the right. The data were obtained from NVE (1946) and include all hydropower plants with a capacity of at least 500 kW.

The expansion of hydropower started relatively slowly in the late 19th century and then really gained pace from the beginning of the 1900s. Figure 3.2 illustrates the geographical and temporal dispersion of hydropower, indicating a nationwide expansion of hydropower over time. The use of hydroelectric power by industry was initially very small. Venneslan (2009) argues that by 1896 horsepower produced by electricity made up only 1.2 per cent of the total horsepower, while this share had already increased to 44.6 and 79.8 per cent by 1910 and 1920, respectively. This exemplifies how important hydroelectric power became in Norway's industrialisation process. The development of these plants led to significant structural change in local areas and had important implications for multiple margins, which I will elaborate on later in this paper using novel data linkages and newly digitised data from Norwegian municipalities and health districts.

3.2.2 Historical Census Data

The individual-level data used in this project mainly come from the full population censuses of 1900 and 1910. The data were obtained from IPUMS International (Minnesota Population Center, 2020*a*,*b*). The censuses were carried out by Statistics Norway (SSB) at the beginning of December in the respective years. They have been harmonised and contain information about the first name, last name, birth year, birth municipality, sex, occupation, municipality of residence, occupations and information about household members.[9] These censuses have multiple use cases. Firstly, they provide a detailed record of changes in demographics and background characteristics in hydropower areas and non-hydropower areas between 1900 and 1910. The detailed record of individuals in Norway during this period enables me to compute aggregate statistics about demographics in local areas. In addition, new linking methods, summarised and developed by Abramitzky et al. (2021), enable the construction of panel data by linking individuals using fixed characteristics such as birthdate, birthplace and name. This procedure makes it possible to link the two censuses over time and obtain panel data for a subset of linked individuals. The most novel use case for the historical census data is the linkage to individual-level death records for the periods 1928 to 1945 and 1951 to 2014. This connection allows for death data to be

[9]Detailed information on available variables and definitions are provided on the IPUMS website and in Appendix 3.C.

supplemented by detailed information about individuals' place of birth, residence and household structure around the time of their birth. I will use these data to present novel evidence of the long-term implications for the age of death of structural change during childhood.

The historical census data also include important information about the relationship between an individual and the household head. In most cases, this is either the father, the husband or the person himself/herself.[10] I use this information to assign occupations of the household head to individuals. These occupations are reported in the Historical International Classification of Occupation (HISCO) format and are harmonised across censuses. Moreover, they can be used to assign a social stratification value to an individual's occupation using the Historical International Standard Classification of Occupations (HISCAM) (Lambert et al., 2013). HISCAM provides a measure of socioeconomic status for occupations and makes it possible to construct basic measures of inequality. Ideally, these would also include information about income and education, but due to a lack of such records, I rely on the HISCAM-based measure of inequality.

The linking procedure applied in this paper was first presented in Abramitzky, Boustan and Eriksson (2012, 2014). It has been slightly adapted to the Norwegian-specific context and data. For the sake of simplicity, the linking procedure will be named ABE (Abramitzky, Boustan, Eriksson) matching from now on. ABE matching enables otherwise standalone cross-sectional datasets to gain a longitudinal dimension. For this paper, it enables me to link samples from the 1900 and 1910 censuses with data from the his-

[10] In some cases when individuals were interviewed for the census while they were somewhere else than their usual residence, this would be the household head in the temporary residence.

toric death register. The ABE matching algorithm uses characteristics that are fixed for an individual across different datasets (e.g. sex, birthplace, birthdate, names) and creates possible matches. These possible matches are then assigned a string distance value based on the first and last names of individuals. Through a few decision rules, matches meeting a string distance threshold and various other criteria are then selected as successful matches. The overall goal of ABE matching is to obtain a linked dataset that includes a significant number of correct matches. The decision rules are put in place to balance the number of overall matches against a high number of false positive matches. The exact procedure applied in this paper is described in Appendix 3.D together with details of the linking rates. In total, I can match 34 per cent of individuals in the 1900 census to individuals in the 1910 census, with significantly more men being successfully matched due to patrilineal naming conventions.[11]

3.2.3 Historic Death Register

The main death records consist of two separate datasets that were both obtained from the Digital Archives (*Digitalarkivet*) of Norway and have previously been maintained by SSB. The first dataset covers the period from 1928 to 1945, while the second dataset covers all deaths from 1951 to 2014. Both datasets include information on the first name, last name, birthdate, death date, municipality of death and sex of the deceased individual. All these variables enable me to link the death data to the census data from 1910, using the previously described ABE matching procedure.

Death data include the cause of death, and registration dates

[11] The exact matching procedure, summary statistics and matching rates can be found in Appendix 3.D.

go back to the 17th century, when priests were required to keep track of basic demographic changes such as marriage, births and deaths. From 1859, death statistics were published in annual reports and were based on doctors' reports obtained from the health district doctors. From 1928 onwards, SSB took over responsibility for the central processing of death reports from doctors, priests and other local sources (e.g. police officers). Harmonisation and central processing led to significant improvements in the completion and quality of the reports of deaths (Gjertsen, 2002). Death records from 1928 to 1945 at a detailed individual level have not been available in formats that are useful for research and have just recently been finalised by the Digital Archives. I obtained these data upon request and they should be available to the general public in future.

The raw data obtained from the Digital Archives required some light processing to obtain a consistent sample and sufficient coverage of the municipality of residence at the time of death. In the raw data, residence locations were reported using various different sub-areas such as counties, municipalities, parishes and villages. We used the municipality structure of 1930 obtained from the Norwegian Centre for Research Data (NSD) for ABE matching of municipality names. For all non-matched residence locations in the data, we manually assigned a municipality from the 1930 structure. This procedure yielded a municipality coverage of more than 99 per cent.

The death data for the period 1951 to 2014 are also publicly available via the Digital Archives. They cover all deaths in the relevant period in Norway. These data have been available since 2019 and correspond to the cause of death register of Norway, which is available for registered researchers via the website of the Norwegian Public Health Institute (FHI). The publicly available death data do not include the cause of death since that is exempt from the public domain and considered to be sensitive private information. Even though Norwegian administrative data cover deaths from 1961 to 2018 including the cause of death, they do not include the names of individuals. The first and last name are crucial, however, if the ABE matching algorithm is to be applied and death data are to be linked to the census data from 1910. The information from the 1910 census is in this case the missing piece needed to assess exposure to structural transformation, a key component of this paper.

3.2.4 Municipality-Level Data

The structural change that occurred in municipalities across Norway has to some extent been documented in previous research (Leknes and Modalsli, 2020). I have obtained similar data on industry employment and the labour force, but I have also collected new data on income, wealth and the number of taxpayers in Norwegian municipalities. The combination of these data allows for a more detailed picture to be created of the structural and economic changes in Norwegian municipalities around the time of the adoption of hydropower.

The first part of the data concerning Norwegian municipalities' population, employment and industry composition is provided through the Norwegian Municipality Database (*Kommunedatabase*). These data have mainly been collected through the full population censuses of 1891, 1900, 1910 and 1920, which means that most variables will only be available for those years. The information on employment and industry affiliation is provided in varying detail across years. These aggregate variables are more detailed for the subpopulations (e.g. age groups and sex) they cover. I aggregate these sector employment variables into five groups, agriculture, manufacturing, services, shipping and others, to obtain a comparable measure of sector employment across years.[12] Using these data, I am able to quantify the hydropower-induced structural change as well as labour force development.

In addition to the data available from the Municipality Database, I obtained data on taxpayers, income and wealth in Norwegian municipalities for the time period 1894 to 1920 through our research centre. These data were digitised from PDF documents provided by SSB. Importantly, taxpayers in these documents are defined as both physical and non-physical entities and therefore also include firms. Historically, these data were collected by SSB and the central government to assess the public finances of Norwegian municipalities. The information does not initially correspond to the municipality structure of 1900, because these data are reported at the tax district level, which corresponds relatively closely but not perfectly to the structure of municipalities. I match them to create a balanced panel of municipalities with income, wealth and taxpayer information.

In addition, I add data on pre-existing infrastructure such as railway stations and steamship stops in a municipality from a dataset named 'Norwegian Ecological Data', which has been made publicly available (Aarebrot and Kuhnle, 1984). These data contain information about infrastructure for the year 1880 for all Norwegian municipalities. Summary statistics for the municipality-level data are provided in Table A2.

[12] A detailed overview of the industry and employment classification can be found in Table A9.

3.2.5 Health District Data

In order to examine the public health externalities of hydropower-induced structural change, I collected data on health district-level health personnel, infectious disease cases and deaths, as well as infant deaths and live births. During the late 19th century and the early 20th century, Norway was divided into approx. 160 health districts. These districts were overseen by district doctors who were responsible for collecting information about and reporting on the hygienic, living and public health conditions in their respective health districts. Health districts are aggregations of municipalities and, especially in the northern and western parts of Norway, they cover large and relatively inaccessible areas.[13] Due to Norway's size and terrain, it was not possible to perfectly observe the exact conditions in each part of a health district. The data I collected are published in statistical yearbooks about public health in Norway. The data on health personnel include the number of doctors, midwives, dentists, pharmacists and vaccinators per health district. Midwives, in particular, have been shown to have a bearing on maternal mortality, but not on infant mortality, in rural Norway during this period (Kotsadam, Lind and Modalsli, 2022). Even though treatment for most communicable diseases was limited during the early 20th century, health personnel were likely one of the few sources of reliable information about the transmission of infectious diseases and mitigation strategies. From detailed reports, it is clear that the district doctors had a clear understanding of what causes disease and how to prevent

[13] The number of health districts expanded from the 1880s onwards. In order to obtain a consistent geographical classification of health districts, I harmonised all information to the 1880 health district structure.

infections.[14] Overall variables for health personnel are intended to capture the public health resources assigned to health districts and can give us an understanding of the potential disparities, which are much more difficult to measure (e.g. sewage systems, hygienic standards), particularly in historical contexts.

Infectious diseases, and their connection to urbanisation and population growth, have been addressed in a large body of literature on urban mortality. For this purpose, I collected data on the number of cases of and deaths from four communicable diseases that are relevant to this period. Two diseases commonly transmitted via the faecal-oral route, diarrhoea (incl. cholera) and typhoid fever, and two respiratory diseases, pneumonia and diphtheria. Waterborne diseases which are mostly transmitted via the faecal-oral route are argued to have mainly been caused by contaminated food and drinking water and they declined to a large extent with the development of waste and water infrastructure (Cutler and Miller, 2005; Alsan and Goldin, 2019).[15] Food standards such as milk inspections in the context of early-20th century US cities have shown insignificant effects on infant mortality (Anderson et al., 2022). The disease counts are available for all years and are used to compute population standardised cases per 100,000 inhabitants and deaths per 100,000 inhabitants. Summary statistics of the health district data are available in Table A3.

Moreover, the annual reports contain information on birth outcomes, such as the number of live births, infant deaths within the first year after birth and the number of stillbirths. The number of infant deaths per 1,000 live births in Norway was signifi-

[14]One district doctor clearly addresses the issue of sewage and waste disposal infrastructure as a public health and hygiene concern.

[15]Appendix Table A8 provides an overview of important characteristics of the different infectious diseases.

cantly lower in rural areas than in urban areas such as Oslo and Bergen, but rapidly decreased from the beginning of the 20th century (Backer, 1961). The health district data indicate that, between 1890 and 1920, infant mortality in health districts including a city municipality declined by 43 per cent, while in purely rural health districts, the decline was significantly smaller, only amounting to 34 per cent. Areas with higher population density were generally more affected by infectious diseases. This positive correlation between population density and cases per 100,000 population is also confirmed by the health district data. Pneumonia, diphtheria and diarrhoea cases per population are significantly and positively correlated with population density measured as the number of people per square kilometre. As infectious diseases decline, infant mortality from infectious diseases also declines, and disproportionately favours a mortality decline in densely populated, high infectious disease areas.[16]

The data on health district-level health personnel, infectious disease deaths and cases in combination with municipality-level data on hydropower adoption make it possible to look in detail at the development of public health in response to hydropower establishment, and thereby to discuss changes in longevity in the context of public health developments as well as economic developments.

[16] Appendix Figure A1 shows the massive decline in infant mortality, which was mainly driven by a decline in infectious disease mortality. From 1899 to 1940, infant mortality due to infectious diseases fell by approximately 80 per cent, while most other causes only saw minor declines.

3.3 Research Design

I will utilise empirical strategies that are based on the staggered establishment of hydropower plants to obtain causal estimates of the impact of hydropower-induced structural change on various outcomes. For simplicity's sake, I will start by describing the empirical approach I use to evaluate the impact of hydropower establishment on local areas. I will then explain the specification I estimate to disentangle the effect of childhood exposure to a structurally changing environment on the longevity of individuals. The longevity effect must be viewed as a two-stage process in this setting. First, hydropower establishment impacts local areas and, through the transformation of local areas, impacts the longevity of individuals.

3.3.1 Event-Study Approach

The main empirical approach to identifying the impact of hydropower establishment on outcomes is an event-study specification utilising the staggered adoption of hydropower plants across Norwegian municipalities. I restrict the samples to five years before and ten years after the hydropower establishment. A hydropower plant is considered as established in the year it starts operating, i.e. providing electricity to the local area. Note that not all municipalities in the sample are treated and that hydropower adoption is an absorbing state, thereby ruling out the possibility that, once a municipality has adopted hydropower, it can potentially revert to no hydropower access.[17] Moreover, I follow the methodology set out in Sun and Abraham (2020) to estimate the specification presented

[17] I have not found any evidence of municipalities gaining access to electricity from hydropower and later losing this access during the period between 1890 and 1920. In addition, I do not have any data on the potential downtime of hydropower plants.

in Equation 3.1. Using this relatively new estimator addresses issues such as the negative weighting of positive treatment effects, and should provide sensible estimates of dynamic causal effects under heterogeneity across cohorts (Roth et al., 2022).

$$Y_{mt} = \sum_{k=-5}^{-2} \beta_k \times \mathbb{1}\left[t - t_m^* = k\right]_{mk} + \sum_{k=0}^{10} \beta_k \times \mathbb{1}\left[t - t_m^* = k\right]_{mk} + \mu_m + \tau_t + \epsilon_{mt} \tag{3.1}$$

Here Y_{mt} is the outcome of interest in area m in year t. The expression t_m^* indicates the year when the first hydropower plant was established in area m.[18] In addition, I include fixed effects for local areas μ_m, which are intended to capture time-invariant differences between local areas across years and time-fixed effects τ_t, capturing differences over time that do not vary across local areas. The estimates β_k thus capture the change in outcomes in the years prior to and ten years after the establishment of a local hydropower plant relative to the omitted period $k = -1$.

In order to estimate causal effects of hydropower establishment on the relevant outcomes using this event-study approach, non-hydropower areas must serve as appropriate counterfactuals for hydropower areas. This assumption means that, conditional on time and area-fixed effects, outcomes in hydropower and non-hydropower areas would have developed in a parallel fashion. Moreover, I implicitly assume that there are no unobservable changes in determinants of the outcome variables that affect hydropower and non-hydropower places differentially and coincide with the establishment of a hydropower plant.

[18] Note that area in this paper can either be the municipality or the health district, since information on infectious diseases and other public health outcomes is only available at the more aggregate health district level.

Like the event-study approach for variables measured at the local area level, I am able to use the individual-level linked census and death data to estimate event-study regressions on individual outcomes such as the age at death. Equation 3.2 represents the individual-level equivalent to Equation 3.1, which compares the outcomes of individuals born in different years relative to the establishment of a local hydropower plant, conditional on the municipality of birth and birth year-fixed effects. The estimate β_k then captures the impact of hydropower establishment on the outcome of individual i, born in municipality m in birth year t relative to individuals born in non-hydropower municipalities.

$$Y_{imt} = \sum_{k=-5}^{-2} \beta_k \times \mathbb{1}\left[t - t_m^* = k\right]_{mk} + \sum_{k=0}^{14} \beta_k \times \mathbb{1}\left[t - t_m^* = k\right]_{mk} + \mu_m + \tau_t + \epsilon_{imt} \tag{3.2}$$

The main identifying assumption here also assumes that non-hydropower municipalities serve as appropriate counterfactuals for hydropower areas. This identifying assumption can in turn be separated into two parts. First, the age at death of individuals in hydropower and non-hydropower areas would have followed parallel trends in the absence of hydropower adoption. Second, determinants of outcomes that occur simultaneously with hydropower establishment do not affect individuals born into hydropower and non-hydropower areas differently. There are several potential reasons why violations of the second identifying assumption can be envisaged.

I. Selection of Areas into Hydropower Adoption:
   Pre-determined characteristics of local areas could potentially

impact the selection of areas into hydropower adoption, either through private investors selectively choosing areas with suitable economic/infrastructure conditions or other characteristics, for hydropower adoption conditional on the prevailing economic environment or infrastructure availability. In this case, selection into treatment might be correlated with potential outcomes post-hydropower introduction.

II. Alternative Policies:
Other local policies, such as the construction of new transport infrastructure, could potentially be correlated with the introduction of hydropower plants and have independent effects on relevant outcomes in addition to the impact of hydropower.

III. Selection of Individuals into Hydropower Areas:
For the analysis of individual-level outcomes, the selection process for migrants may create a pool of parents who are positively selected on margins relevant to longevity, which could be transmitted across generations and lead to improved longevity among their offspring. Effects of hydropower establishment could then simply reflect a beneficial selection of individuals into areas, rather than improved longevity resulting from structural transformation.

I address these threats to identification in several ways. As regards point I, I argue that characteristics of hydropower and non-hydropower municipalities in terms of, e.g., pre-existing infrastructure and economic conditions, are to some degree correlated. However, I show that there are no pre-trends in most of the economic and health outcomes when estimating Equations 3.1 and 3.2. This strongly suggests that any changes in outcomes observed after hydropower adoption are driven by the actual adoption of hydropower rather than selection into treatment. I argue that, without hydropower, potential areas would not have been chosen for infrastructure development and that hydropower plants were a necessary condition for the further development of infrastructure. Sejersted (2021) mentions that domestic know-how and financial constraints would not have sufficed for the development of these plants and that strategic interest from foreign investors was therefore necessary in relation to both financing and the adoption of the new technology. These investors mainly decided the location of plants with respect to the suitability for hydroelectric power production of an area rather than other pre-determined characteristics. I will later show that pre-existing railway and shipping infrastructure does not seem to have influenced economic development in these areas, which is suggestive evidence supporting the idea that hydropower adoption was driving structural change rather than pre-existing comparative advantages in infrastructure.[19]

The second point argues that other local policies might have coincided with the introduction of hydropower and that subsequent effects on individuals and local economies were the result of alternative policies rather than of the introduction of hydropower. This potential threat to identification is very difficult to address. Given the strong and positive results as regards short/medium-term economic responses to hydropower establishment, alternative policies would need to have been recorded or reported by economic historians. These alternative policies would also need to coincide with hydropower establishment all across the country, which is unlikely and not supported by the evidence found by using the instrumen-

[19]In Appendix Table A4, I provide information on differences between hydropower and non-hydropower municipalities in terms of certain pre-determined characteristics.

tal variable approach.

Point III raises the potential issue of selection into treatment. In order to interpret any effect of hydropower establishment on the age at death of individuals in a causal fashion, it is necessary to ask whether hydropower establishment and structural change caused gains in longevity or whether this was driven by a different selection of individuals into the treatment areas. It could be argued that, given that hydropower leads to attraction of migrants, part of the hydropower effect on longevity works through selection. I think this is a partly valid interpretation. However, what we are ultimately interested in is whether the structural change in local areas causes longevity effects. In order to answer this question, I apply several sample restrictions and show that the results are in general robust to excluding individuals from families who selectively immigrated to hydropower areas. I will discuss these robustness checks after presenting the main results in Section 3.4.

3.3.2 Two-Way Fixed Effects

There are two main disadvantages of the event-study approach presented in Equation 3.1 and 3.2, which mainly concern the requirement of sufficiently long panel data. First, not all outcome variables are applicable over the whole period between 1890 and 1920, but are only available for a few years, when Norway conducted full population censuses. The second potential downside of the event-study approach concerns the ability to perform heterogeneity analyses for the long-term analysis, due to the fact that some groups used for the estimation of treatment effects become relatively small when applying the dynamic difference-in-difference framework. Due to these data limitations, I provide evidence of the implications of hydropower establishment for some outcomes and heterogeneity tests in a simple two-way fixed effects framework.

$$Y_{mt} = \alpha + \beta_1 \times \text{HPP}_{mt} + \mu_m + \tau_t + \varepsilon_{mt} \tag{3.3}$$

Equation 3.3 implements this two-way fixed effects strategy for local areas, where Y_{mt} is again the outcome of interest in area m and year t, HPP_{mt} is an indicator variable equal to one if in area m in year t there was an operating hydropower plant present, and zero otherwise. The fixed effect for an area μ_m and time τ_t again capture unobservables that are time-invariant or invariant across different areas but do not account for changes in unobservables across both time and areas simultaneously.

Since I am not able to provide a heterogeneity analysis using Equation 3.2 due to a small sample size at some event times, I will supplement my main results on longevity using the individual-level pendant of Equation 3.3.

$$Y_{imt} = \alpha + \beta_1 \times \text{HPP}_{imt} + \mu_m + \tau_t + \varepsilon_{imt} \tag{3.4}$$

Here, Y_{imt} is the outcome of individual i born in municipality m in birth year t. Birth municipality fixed effects are represented by μ_m and capture time-invariant differences across birth municipalities, while τ_t captures variation across time that does not change within the birth municipality. HPP_{mt} is a dummy equal to one if individual i was born in a municipality with an operating hydropower plant in the same year. β_1 will then again capture the average difference in outcomes for individuals who were born into hydropower municipalities, compared to those who were not. A major concern with this specification is that the two-way fixed effects estimator assumes no heterogeneity in treatment across time and units. In cases where this identification strategy is used, I will provide additional information about the weights and size of the different two-by-two estimates, as suggested by Goodman-Bacon (2021). This is mainly necessary in order to avoid negative weighting issues and to understand the coefficient weights resulting from 'forbidden comparisons' where earlier-treated units are compared to later-treated units and *vice versa* (Roth et al., 2022).

Alternatively, I will also present estimates from an instrumental variable strategy following Borge, Parmer and Torvik (2015), where I instrument HPP_{mt} from Equation 3.3 using a topographic variable that captures the length of rivers within an area that falls into terrain with a gradient larger or equal to four degrees. The main idea behind this instrument is that hydropower adoption happened in areas that were most suitable and efficient for the production of electricity. Since water and potential energy from elevation differences in the terrain are the main ingredients of the production of hydroelectric power, they should strongly determine the adoption of hydropower plants in a local area.

$$\text{HydroPotential}_m = z_m = \frac{\sum_{w=10}^{w=750} \text{river}_{mw} * w}{\text{area}_m}$$

Hydropower potential z_m is defined as the product of river length within a municipality that falls into areas with a slope larger or equal to four degrees within the water flow category w times the water flow strength.[20] In addition, I scale this product to the total area of the municipality, so that large municipalities do not disproportionately benefit simply due to their larger land area. In the

[20]Water flow strength is constructed from data obtained from the HydroRivers database and includes the following categories: 1-10, 10-50, 50-100, 100-150, 150-200, 200-250, 250-300, 300-400, 400-600, and 600-750 m^3/s.

instrumental variable regressions, I interact the instrument with a year dummy, so that the influence of hydropower potential is allowed to differ annually. This step is necessary due to the time-invariant nature of the topographic variables entered into the instrument. I will mainly use the instrument later on to conduct a robustness check of the simple two-way fixed effects estimates provided by estimating Equation 3.3. For most of the analysis where I am able to provide event-study estimates using Equation 3.1 or 3.2, I will not present results using the instrumental variable approach.

3.4 Results

In this section, I will present the main results, showing that in the long-term men born into hydropower areas outlive comparable men by ten months on average. I will discuss the short and medium-term changes in the local economy, which are related to the improvements in longevity I am observing. This includes the changes in local industry composition, income and inequality, as well as the selection and movement of migrants towards hydropower municipalities. I will then discuss potential negative externalities in terms of public health of hydropower adoption in the short as well as the medium term, and argue that, despite the worsening of the public health environment, improved economic prospects have a stronger impact on longevity.

3.4.1 Long-term Implications of Hydropower Adoption

Figure 3.3 presents the main result of this paper obtained from estimating Equation 3.2 using age at death as the outcome variable. Each point represents the respective estimate of the impact of a hydropower establishment relative to the birth year on the age at death of men. The estimate at event-time five can be interpreted as the relative increase in the age at death of individuals born five years after a hydropower plant was established in their birth municipality relative to individuals who were born into non-hydropower municipalities. The first main finding from the event-study results suggests that hydropower establishment mainly impacts individuals' age at death if it has been established at least five years prior to birth. For individuals born between five and six years after hydropower establishment, the hydropower-induced structural change has progressed sufficiently to impact longevity by approximately ten months. Estimates from event time ten onwards suggest an even larger impact on longevity of up to 20 additional life months relative to control individuals.

The second important feature suggested by Figure 3.3 is that hydropower plants that opened five years before and four years after birth do not seem to significantly increase the age at death. One interpretation consistent with this finding is that, to significantly benefit from hydropower-associated changes in local areas, a certain maturation or development process needs to have set in. This would mean that only after a hydropower plant opened would necessary changes to local economies occur that significantly benefit the longevity of individuals. Given that hydroelectric power production does not create relevant pollution or other direct negative health externalities, the main impact on individuals must take place through changes to the amenities/environment in the local area. It also suggests that the benefits from structural change are reaped to a larger extent early in life. This can be inferred from the fact that those born before or soon after hydropower adoption

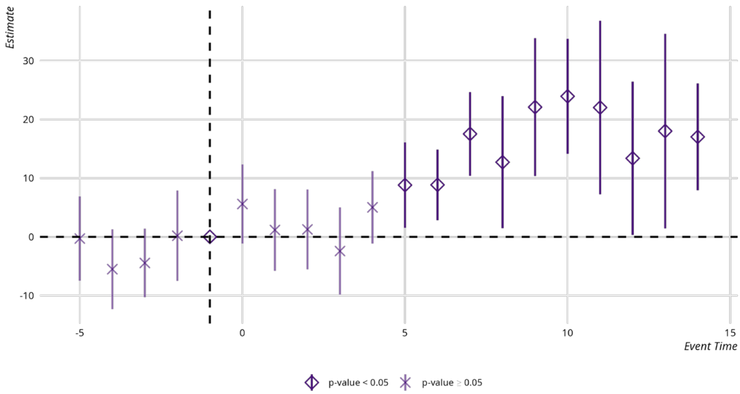


Figure 3.3: Impact of Hydropower Openings on Age of Death of Men.
Note: The figure shows event-study estimates of the effect of hydropower establishment on the age at death for men born between 1890 and 1910. The data used for the estimation come from the linked historical death register and the full population census of 1910. Event-study estimates are obtained using the methodology presented in Sun and Abraham (2020). Estimates can be interpreted as changes relative to event time $t = -1$ indicated by the vertical dashed line. They were obtained by estimating Equation 3.2. Standard errors used for the calculation of the 95-% confidence intervals are clustered at the municipality level.

do not experience longevity gains, despite their exposure later in life. This is similar to findings from Chetty, Hendren and Katz (2016), who find that longer exposure of children to higher income neighbourhoods has larger effects on adult economic outcomes. Even though the confidence intervals become larger from event time eight onwards, this effect appears to be persistent and does not disappear, which also suggests that the benefits of hydropower adoption are permanent and potentially benefit individuals born after 1910 in a similar way.

One candidate mechanism in line with the findings relating to occupational upward mobility resulting from hydropower could be that improvements in longevity mainly operated through improved living conditions and higher incomes for individuals whose parents gained economically as a result of hydropower. Upward mobility in this historic context has been documented to be a result of hydropower establishment in Norway (Leknes and Modalsli, 2020). This mechanism would also imply that hydropower adoption increased inequality in longevity through the occupational choices of individuals and their parents. In Table 3.1, I provide some evidence of this mechanism by estimating Equation 3.3 on the linked historical death register.

Column (1) shows the two-way fixed effect estimate of hydropower adoption on the age at death of men. In line with the results from Figure 3.3, individuals born after hydropower adoption live on average approximately ten months longer than non-hydropower-born individuals. In columns (2) to (4), I then separate the estimation sample. I do this by dividing individuals within the same birth year and born into the same municipality into three quantiles depending on the HISCAM value of their household heads' occupation. In column (2), I exclude all individuals born into hydropower municipalities that do not fall into the lowest quantile. For column (3), I proceed identically but exclude everybody born into a hydropower municipality who does not fall into the second quantile, and in column (4) everybody who does not fall into the top quantile. This ensures that the comparison group is the same for both columns, but the individuals in the treated group are from different socioeconomic backgrounds.[21]

Turning to the estimated coefficients, we see that the impact of hydropower establishment on the age at death of individuals of lower socioeconomic status is not distinguishable from zero. On

[21] The difference in the number of observations is the result of not everybody in a certain municipality being treated, but the position in the SES distribution is determined by all individuals within a given municipality.

Table 3.1: Hydropower Impact on Age of Death by Socio-Economic Status.

	All (1)	Low SES (2)	Medium SES (3)	High SES (4)
HPP	10.2***	-2.05	13.7***	22.2***
	(2.85)	(3.02)	(1.99)	(5.88)
R^2	0.021	0.023	0.022	0.022
Observations	146,634	134,322	134,160	133,975
Mean D.V.	891.8	892.5	893.2	893.4
Birth Municipality FE	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓

Note: Standard errors are clustered at the municipality level. Data comprise all men born between 1890 to 1910 linked to the historical death register. Estimates are obtained by estimating Equation 3.4. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$.

the other hand, the estimates in columns (3) and (4) are significant and larger than the overall estimate, suggesting that individuals from high-SES backgrounds born into areas with an already operating hydropower plant live 14 and 22 months longer than comparable individuals born into non-hydropower areas. This strongly suggests that the overall effect is driven by changes benefiting men born to families where the household head is of higher socioeconomic status. Since overall longevity is increasing and the impact of hydropower establishment on children from low socioeconomic status backgrounds are indistinguishable from zero, it appears that these hydropower establishments and the associated changes to local areas have been Pareto optimal in terms of longevity.

To put the average effect of an increase in the age at death of ten months into perspective, we can compare this estimate to the effects of education on longevity. Lleras-Muney, Price and Yue (2022) estimate that an additional year of schooling for cohorts born between 1906 and 1915 in the US is associated with an increase of approximately five months in age at death. The magnitude of this result is similar to findings in Halpern-Manners et al. (2020). This suggests that the longevity effect of structural change induced by hydropower establishment is on average twice as large as the effect of an additional year of education in the US context.

Since, as discussed in Section 3.3, selection into hydropower areas is a potential threat to the causal interpretation of these estimates, I will provide some evidence that the selection effect of hydropower areas does not drive the main results presented in Figure 3.3. There are two main reasons why selection could be problematic. First, it changes the composition of individuals in treatment municipalities and could therefore affect longevity. This could be the result of immigration by parents who are positively selected in terms of longevity. Their offspring born into hydropower municipalities could then simply have a longer lifespan due to the intergenerational transmission of longevity from their parents. I test this hypothesis directly by limiting the sample to individuals who have a family member who was born in the same municipality before hydropower adoption started in Norway in 1890. This assures that I capture families that did not move into hydropower areas due to the adoption of hydropower. The second problem with selection could arise if people emigrate from hydropower municipalities after birth and thereby are not truly exposed to the structural transformation process. This can be tested by conditioning the sample on individuals who still lived in their birth municipality during the 1910 census. Results for both these sample selections show very similar point estimates compared to the full sample, albeit with slightly larger confidence intervals. The results for this exercise are presented in Appendix Figure A2 and suggest that the main effect is not simply driven by migration but also by individuals from the local stayer population.[22]

Another concern about the interpretation of these results is the potential for selection bias resulting from the fact that the death data only include individuals who died between 1928 and 1945, and 1951 and 2014. If hydropower establishment is somehow correlated with population developments prior to 1928, the sample of individuals dying after 1928 might be different in hydropower and non-hydropower municipalities. Three main developments are relevant to consider: i) emigration responses are different for hydropower and non-hydropower municipalities, ii) population changes (e.g. mortality and fertility rates) are different in hydropower and non-hydropower areas, and iii) the linking algorithm matches individuals from hydropower and non-hydropower municipalities with differing success. Table A6 provides evidence of the emigration responses to hydropower establishment using data obtained from the municipality database. The two columns estimate the effect of hydropower establishment on emigration in the years following a hydropower establishment, once using the

[22] An additional concern is that hydropower does not just affect the treatment group composition but could also affect the composition of the control groups through selective emigration. For example, through the emigration of parents who would have children with higher longevity. This would be a SUTVA violation since treatment has spillovers to non-treated individuals. I used a sample of individuals who were not born into hydropower municipalities to test whether increasing distance to a hydropower establishment has a differential effect on longevity in those areas. The idea would be that greater proximity should lead to more emigration and a more negatively selected local population if spillovers are relevant. I could not find any statistically meaningful indications that this is the case. Moreover, I also split samples into high and low-emigration municipalities and tested whether the hydropower establishment in the geographically closest municipality affected longevity. I cannot find evidence for a longevity effect in this case either.

simple two-way fixed effect setup from Equation 3.3 in column (1) and using the instrumental variable approach in column (2). Emigration in this context is defined as emigration outside of Europe, which at this point was mainly the United States.[23] Hydropower establishments did not significantly influence emigration responses towards the United States according to Table A6, since both coefficients are statistically indistinguishable from zero. This means that areas with and without hydropower adoption did not have significantly different numbers of emigrants leaving for the US.

The concern about potentially different population changes is relevant, but using aggregate data for deaths and births from the municipality database, I cannot confirm that mortality and fertility rates significantly differ in response to hydropower establishment. I will discuss these findings in more detail later on when discussing public health development in the health districts. Finally, a big concern is that the linking of men between census and death data is correlated with treatment. Since this could potentially influence the composition of the comparison groups, I have provided evidence of the impact of hydropower establishment at birth on the probability of being linked to the historical death register. In Table A7, I test whether hydropower establishment impacts the probability of being matched to the death data. Columns (1) to (3) present simple regressions of an indicator variable on the probability of being matched to the historic death data, with different sets of fixed effects. Overall, there is a correlation between hydropower establishment and the probability of a successful match. This effect entirely disappears, however, once I control for birth

[23] Semmingsen (1960) documents that the mass migration starting in 1865 and lasting until 1915 resulted in approximately 600,000 Norwegians leaving Norway, mainly to start new lives in the United States.

year and birth municipality fixed effects, suggesting that Equation 3.3 should not suffer from bias due to differential matching of individuals in hydropower and non-hydropower areas.

3.4.2 Structural Transformation and Industrialisation

So far, I have described results suggesting that hydropower establishment prior to birth within the birth municipality positively impacted the longevity of individuals and that this effect appears to be largely driven by individuals from families of higher socioeconomic status. In this section, I will provide evidence of the short- and medium-term changes to the local economy that explain part of this longevity effect. Table 3.2 starts by presenting the first set of main results for industrialisation, employment composition and population changes. In Panel A, I provide results from estimating Equation 3.3 on various outcomes. Columns (1) to (4) show the effect of hydropower establishment on the employment share of manufacturing, agriculture, service and shipping, respectively. In column (5), the impact on the total labour force is estimated, while column (6) uses the labour force to population ratio as the outcome variable.

Like Leknes and Modalsli (2020), I find that the labour force in general and the share of the labour force working in manufacturing increase significantly in response to hydropower establishment. This is the main evidence of structural change in hydropower areas. Employment increases by 13.2 per cent while the manufacturing share rises by approximately 25 per cent in response to hydropower establishment. For a municipality of average size with approximately 3,740 inhabitants in 1900, this would roughly translate into 150 additional manufacturing jobs.[24] The share of the labour force working in agriculture declines in response to hydropower establishment by approximately ten per cent (relative to the mean). The service and shipping shares remain constant. Importantly, even though the agricultural share declines in response to hydropower, overall employment in agriculture remains constant. This suggests that manufacturing did not crowd out jobs in agriculture, but instead created new jobs in addition to those that already existed in the municipality. Employment in services increases, while the employment share in column (4) remains stable. This indicates that there is some complementarity between manufacturing and service employment in this setting. Moreover, a constant service share also confirms that the provision of services for the local population does not decline in response to hydropower establishment. Moreover, the labour force to population ratios in column (6) remain fairly constant over time. Hence, it is likely that the overall increase in employment is driven by external increases in population size, such as immigration, rather than through individuals already residing in local areas joining the labour force (e.g. women, children).[25] The results presented in Panel A are in general compatible with results obtained from instrumental variable regressions presented in Panel B, where the hydropower indicator variable is instrumented by the interaction of hydropower potential and year dummies. [26]

[24]To provide a contemporary comparison, this roughly corresponds to the opening of Tesla's large new factory in Brandenburg (Germany), creating an estimated 10,000 jobs for a larger area of approximately 200,000 inhabitants.

[25]I tested this hypothesis, and between 1900 and 1910 there is no notable change in female employment in hydropower municipalities.

[26]In Appendix Figure A3, I provide estimates from Bacon decompositions following Goodman-Bacon (2021), which show that the effects presented in Table 3.2 are mainly driven by comparisons between untreated and treated municipalities. In addition, the decomposition shows that no negative weighting occurs.

Table 3.2: Employment and Population Responses to Hydropower Openings.

	Employment Share					
	Manu. (1)	Agri. (2)	Serv. (3)	Ship. (4)	Ln(LF) (5)	LF/Pop (6)
Panel A: FE						
HPP	0.054***	-0.062***	0.006	0.002	0.132***	-0.454
	(0.009)	(0.010)	(0.004)	(0.002)	(0.025)	(0.339)
R^2	0.912	0.956	0.923	0.871	0.967	0.500
Observations	2,376	2,376	2,376	2,376	2,376	2,376
Mean D.V.	0.225	0.665	0.078	0.031	6.75	32.8
Panel B: IV						
HPP	0.079*	-0.048	-0.030	0.000	0.690**	-1.27
	(0.045)	(0.060)	(0.047)	(0.020)	(0.315)	(2.38)
R^2	0.912	0.956	0.917	0.871	0.947	0.499
Observations	2,376	2,376	2,376	2,376	2,376	2,376
Mean D.V.	0.225	0.665	0.078	0.031	6.75	32.8
F-test (1st stage)	46.4	46.4	46.4	46.4	46.4	46.4
Municipality FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓

Note: Standard errors are clustered at the municipality level. Estimated using aggregated census data for the years 1891, 1900, 1910 and 1920 following Leknes and Modalsli (2020). Estimates were obtained by estimating Equation 3.3. Intrumental variable is hydropower potential interacted with year dummies. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$

Since the labour force-to-population ratio remains constant after hydropower adoption, I assume that the main reason for the increase in the labour force is permanent family reallocation rather than simple worker migration. Electrification has been shown to impact female participation in the labour market in, e.g., the United States, mainly through a reduction in time spent on household work and increased demand for skilled labour in manufacturing Vidart et al. (2021). However, since electricity was at this time mainly used as an energy source for manufacturing and time savings due to electric household appliances were not relevant, reductions in household work have likely not contributed to increased female labour market participation. Moreover, using municipality-level aggregates from the full population census data, I can construct female and male employment shares, which allows me to test the hypothesis concerning whether hydropower adoption between 1900 and 1910 impacted female and male employment shares. According to these data, there is no difference in female employment changes between hydropower and non-hydropower areas.[27] Returning to population changes and immigration into hydropower municipalities, in Figure A4, I confirm the hypothesis that the share of individuals who resided in a different municipality in 1900 than in 1910 is significantly higher in hydropower than in non-hydropower areas and that the probability of meeting an internal migrant is at least six percentage points higher in municipalities that have adopted hydropower by 1910.[28]

In addition to changing industry composition, employment and population, hydropower also had important implications for public finances and economic prospects in municipalities. Through historical records, information is available about income, wealth and the number of taxpayers. Taxpayers in this setting include both individuals and impersonal entities and do not include individu-

[27] Individuals are defined as employed if they have a recorded occupation in the census. According to source variable documentation, occupations should have been recorded for women if they were working in any industry.

[28] In Figure A5, I provide some descriptive statistics on mover selection, where I show that individuals are generally positively selected in terms of an occupation-based socioeconomic status measure and are generally younger and more likely to be single.

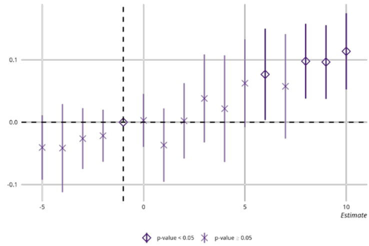
als under a certain income threshold (Berger and Vagle, 2017).[29] In Figure 3.4, I present estimates of Equation 3.1 for three different outcomes. Panel 3.4c confirms that the overall number of taxpayers in hydropower municipalities rises significantly in the years after hydropower adoption. Compared to non-hydropower municipalities, hydropower municipalities grow by approximately ten percentage points. Since taxpayers include both individuals and impersonal entities, this could be due to both an increase in the number of firms and a population increase.

Panel 3.4a provides evidence of the growth of income per taxpayer in the years after hydropower establishment. On average, the structural change in hydropower municipalities significantly increases income per taxpayer, but only within approximately five years of hydropower adoption. This rise in income relative to non-hydropower adopters also continues in the years thereafter, suggesting a significant long-term impact on income per taxpayer in treated municipalities. Panel 3.4b presents a similar set of results for wealth per taxpayer, which slightly increases in the first few years after hydropower adoption and stagnates thereafter. It is not possible to distinguish between the wealth accumulation of private individuals and firms in this case either. However, the rapid increase in wealth suggests that asset acquisition by firms connected to the construction of manufacturing and infrastructure might be driving these results.[30]

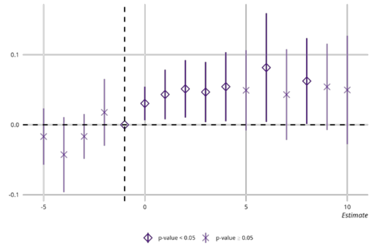
As a result of improving economic conditions and increased labour demand from newly constructed manufacturing sites, hy-

[29] Gerdrup (1998) argues that, on average, about 50 to 60 per cent of the total income within a municipality was taxable and would therefore be included in these data.

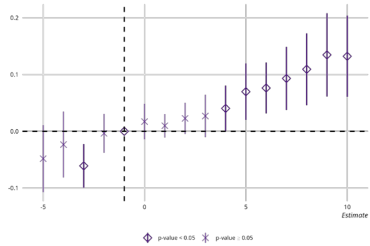
[30] Results in Figure 3.4 are corroborated by Table A10 providing estimates of Equation 3.3 for the income, wealth and taxpayer variables.



(a) Income per Taxpayer



(b) Wealth per Taxpayer



(c) No. of Taxpayers

Figure 3.4: Income, Wealth and Taxpayer Change in Response to Hydropower Establishment

Note: The figure shows event-study estimates of the impact of hydropower establishment on income per taxpayer, wealth per taxpayer and the total number of taxpayers in Panels a, b and c, respectively. Event-study estimates follow methodology proposed in Sun and Abraham (2020). All outcomes are in logarithmic form to allow for comparable interpretations relative to the reference period $t = -1$. Estimates are obtained by estimating Equation 3.1. Standard errors are clustered at the municipality level. Shapes and transparency indicate significance at the 5 % level and error bars represent 95 % confidence intervals.

dropower municipalities see a larger influx of immigrants compared to areas without hydropower adoption. Figure A4 presents point estimates, which can be interpreted as mover shares, by the hydropower status of the municipality. It shows that the share of individuals who have moved from a municipality into a municipality that has acquired a hydropower plant by 1910 is approximately 22 per cent, while the mover share is three percentage points lower in areas without hydropower by 1910. Importantly, these differences are statistically different from zero and confirm that hydropower municipalities disproportionately attracted migrants. Overall, the migrants I can identify using the linked censuses move from the broader general region (*fylke*) rather than from areas far away. Nevertheless, compared to movers who do not move to hydropower municipalities, they are more likely to move from different regions to a hydropower municipality, as can be seen in Figure A6. This increased immigration towards hydropower areas also attracts positively selected individuals in terms of occupation-based socioeconomic status.

One final feature that can be derived from a combination of the full (non-ABE matched) censuses of 1900 and 1910 are changes in inequality derived from occupation-based (HISCAM) measures of socioeconomic status. I first combine the full population censuses and subset the data to only include men aged 15 and older. I then construct Gini coefficients for both years in each municipality, and calculate the percentage change in Gini coefficients across the two census years for each municipality. In the final step, I regress hydropower status on the percentage change in Gini coefficients. The results of this exercise are presented in Figure 3.5. They suggest that, even based on this rudimentary inequality measure, inequality in hydropower areas rises at almost twice the rate it does in non-hydropower areas. Given that this measure only captures inequality due to differences between occupations of individuals it could also be interpreted as a job diversification measure. Given that, in addition to between-occupation differences in earnings one would expect further inequality to result from within-occupation

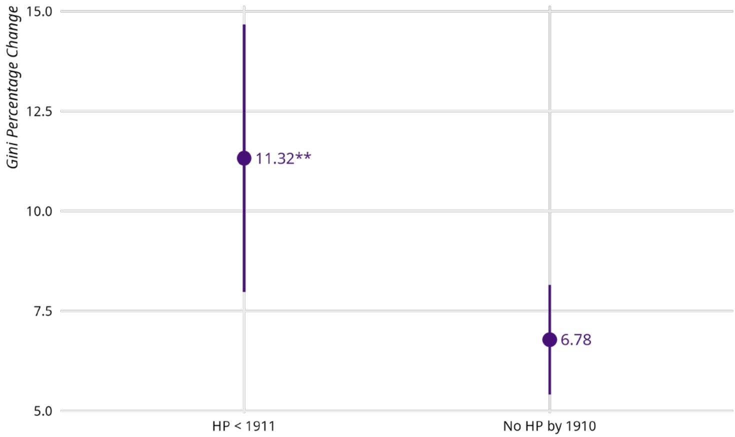


Figure 3.5: Change in Socioeconomic Inequality by Hydropower Status.

Note: The figure shows the change in HISCAM value-based Gini coefficients from 1900 and 1910 in municipalities by hydropower status. The data are from full population census data in 1900 and 1910 and include all men aged 15 and older. The 95% confidence intervals are calculated using heteroskedasticity-robust standard errors. Stars indicate the result of the hypothesis test of difference to No HP by 1920 group. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$.

differences in earnings, the occupation-based measure of inequality I present in this paper can be seen as a lower bound for actual changes in inequality.

In summary, the availability of hydroelectric power has been crucial to the economic development of local areas and started a process of structural change on multiple dimensions. A variety of economic indicators suggest that hydropower municipalities significantly improve the economic prospects of individuals. Moreover, this increase in economic activity increases the influx of positively selected migrants, which leads to population increases in local areas. Simultaneously, the diversification of the local economy appears to have an increasing effect on socioeconomic inequality. Overall, these economic developments seem to be in line with higher levels of longevity, simply through classical income effects. There are various channels through which one might envisage these economic improvements in local areas impacting longevity. Improved nutrition, education, but also overall improvements in life-cycle economic opportunities, have all been shown to have a positive effect on mortality and later-life outcomes. Given the event-study results on longevity from Figure 3.3, a mechanism operating partly through parental and local area resources seems plausible. Due to the limited availability of data, however, I am not able to test these hypotheses directly. In the next section, I will supplement the mostly positive economic developments with the development of public health in the health districts.

3.4.3 Public Health and Infectious Diseases

The previous section established that hydropower improved the economic situation in local areas in terms of incomes, resources and employment opportunities, while simultaneously leading to relatively rapid population growth. This population growth can be interpreted as an externality for the local area. There are several downsides of rapidly growing populations in this historical context, (e.g. housing affordability, food scarcity), I will focus, however, on the transmission of infectious diseases in these areas. Crowded living conditions, sewage and waste have previously been connected to the increased spread of infectious diseases, particularly in urban areas (Alsan and Goldin, 2019; Beach et al., 2016). In contrast, hydropower areas in Norway were predominantly rural and issues concerning the lack of hygienic infrastructure predominantly arose as a result of rapid population growth.[31] In Figure 3.6, I present event-study estimates of Equation 3.1 for the public health outcomes of Norwegian health districts, taken from historical documents.

In Panel 3.6a, I present results concerning the impact of hydropower establishment on the number of cases of four common infectious diseases.[32] Hydropower establishment, with a slight lag of up to four years, led to increased spread of disease. Hydropower plants by themselves do not have a direct influence on water contamination or the spread of disease more generally. However, as described in previous sections, hydropower establishment led to rapidly changing environments in local areas, which directly affected the transmission of infectious diseases. The two main arguments for the increased spread of infectious diseases in more populous areas are usually twofold. The first argument, which is relatively well documented and researched, concerns increased exposure to pathogens due to contaminated drinking water and poor waste and sewage disposal (Cutler and Miller, 2005; Ferrie and Troesken, 2008). There is qualitative evidence of this channel, from written reports of health district doctors at the time, in particular from hydropower areas.[33] This increased risk of infec-

[31] In his 1910 report about living conditions in Ullensvang health district (Western Norway), which had just acquired a hydropower plant, health district physician Garmann-Andreesen writes the following: '... the population is growing and conditions are developing so fast that it has not been possible to keep up with infrastructure in terms of hygiene. Sewage and waste disposal are waiting for a sensible solution.'

[32] The four diseases included in the analysis are diarrhoea (including cholera), pneumonia, diphtheria and typhoid fever. They were among the most common infectious diseases, and effective treatment and prevention were not available at the time. In Table A8, I provide a brief overview of some key characteristics of these infectious diseases.

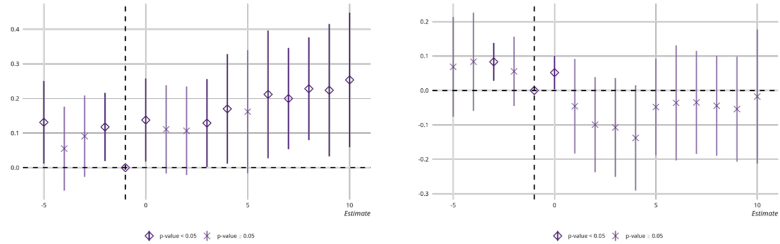
[33] In his annual report in 1910, the health district doctor in Ullensvang writes the

tion due to larger exposure to contaminated water and faeces is also supported by the increase in cases of diarrhoea as presented in Figure A7a, but it cannot be confirmed from the number of typhoid fever cases, which do not see a significant increase in response to hydropower establishment. The second main argument usually made about densely populated areas and their connection to declining public health are crowding and poor housing conditions and their potential link to the increased spread of airborne diseases. This hypothesis is extremely hard to test, but the increased number of cases in Figure 3.6a is also largely driven by the increased spread of pneumonia and diphtheria cases, which are diseases mostly transmitted via aerosols and respiratory droplets. This could have been facilitated by more crowded housing conditions and a generally larger degree of indoor activities as a result of factory work compared to outdoor work in agriculture.

Panel b of Figure 3.6 shows the impact of hydropower establishment on the number of health professionals per 100,000 inhabitants in a health district. Health professionals are doctors, midwives, dentists, pharmacists and vaccinators, and can, taken together, be seen as health care supply during this historical period.[34] The results suggest that hydropower health districts do not necessarily benefit in terms of better health care supply compared to areas without such hydropower establishment since there is no consistent significant impact of establishment on the number of health personnel per population. Reports from this period more generally suggest that the medical profession was not always perceived

following: 'The population is growing and conditions are developing so fast that it has not been possible to keep up with infrastructure in terms of hygiene. Sewage and waste disposal are waiting for a sensible solution.'

[34]Especially dentists and pharmacies were extremely unevenly distributed and quite rare in more remote areas. Over 50 per cent of health districts in 1910 did not have any dentists or pharmacists.



(a) Cases of Common Infectious Diseases

(b) Health Professionals

Figure 3.6: Public Health Development in Response to Hydropower Establishment.

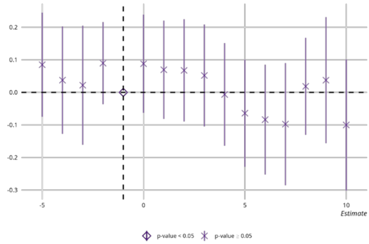
Note: The figure shows event-study estimates of the impact of hydropower establishment on the total number of cases of four common infectious diseases (diphtheria, typhoid fever, pneumonia and diarrhoea/cholera) and the number of health professionals (doctors, midwives, vaccinators, pharmacists and dentists) per 100,000 inhabitants in all health districts. Event-study estimates follow the methodology proposed in Sun and Abraham (2020). All outcomes are in logarithmic form to allow for comparable interpretations relative to the reference period $t = -1$. Estimates were obtained by estimating Equation 3.1. Standard errors are clustered at the municipality level. Shapes and transparency indicate significance at 5% level and error bars represent 95% confidence intervals.

by the local population as beneficial to society, but rather as a burden. It was only after treatment and medical practice improved that local communities changed their views of doctors in particular, which also led to more doctors relocating to rural areas of the country (Sandvik, 2000). Concerning the opening of hydropower areas, it is important to note that the healthcare supply in these areas did not disproportionately increase during the period of rapid structural change. If any change in health care provision per capita happened at all, it probably slightly declined according to Figure 3.6b. This general pattern is also corroborated by estimates of Equation 3.3 in Appendix Table A11 which indicates an average increase in common infectious disease cases of approximately ten percent as a response to hydropwer establishment, while finding no statistically significant effect on health care supply and deaths.

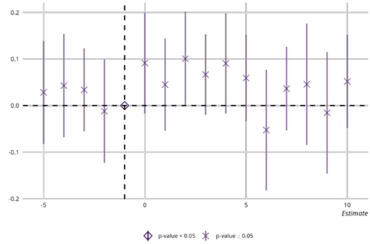
Even though hydropower municipalities appear to suffer from higher levels of morbidity due to infectious disease, this does not seem to translate into higher levels of infectious disease mortality, as can be seen in Panel a of Figure 3.7. Infectious disease deaths per 100,000 inhabitants do not seem to change significantly in response to hydropower establishment. A similar pattern is observable when looking at the change in infant mortality, which does not significantly increase in the ten years following a hydropower establishment. This development is important because it slightly alleviates selection concerns regarding longevity results. If mortality had increased immediately after hydropower establishment, survivors might be positively selected and longevity would therefore be mainly driven by differential survival up until the time data are available to measure the age of death. Since this does not seem to be the case, this potential selection channel is of minor concern.

There are several reasons why mortality, both from infectious disease and infant mortality, does not significantly respond to structural transformation induced by hydropower, despite the increasing transmission of infectious disease. The first reason is that the age structure in municipalities changes, such that the share of vulnerable individuals in proportion to the total population is smaller. Using the combined censuses from 1900 and 1910, I find evidence of this development. In addition, income and resources provided through higher paying jobs in manufacturing potentially increase the nutritional standards of individuals, which in previous research has been linked to higher resilience to infectious diseases (Schneider, 2022).

Taken together, public health development in the health districts does not explain the longevity results. If longevity was driven



(a) Common Infectious Disease Mortality



(b) Infant Mortality

Figure 3.7: Hydropower Establishment, Infectious Disease and Infant Mortality.

Note: The figure shows event-study estimates of the impact of hydropower establishment on deaths from common infectious diseases (diarrhoea, diphtheria, pneumonia, typhoid fever) per 100,000 on the left and infant mortality per 1,000 live births in the right panel. Event-study estimates follow methodology proposed in Sun and Abraham (2020). All outcomes are in logarithmic form to allow for comparable interpretations. Standard errors are clustered at the municipality level. Shapes and transparency indicate significance at the 5% level and error bars represent 95% confidence intervals.

by public health developments, we would expect morbidity to decrease. I find that hydropower-induced structural change has the opposite effect. Morbidity, as measured by infectious disease cases, increased in response to hydropower establishment. Infectious diseases could increase mortality in later life and would therefore attenuate the longevity effect in Figure 3.3, thereby suggesting that the economic benefits resulting from the structural change outweigh the negative developments. It could also be possible that the economic benefits and the negative externalities of structural change impact individuals from different backgrounds. Given the zero effects on longevity for individuals from a low SES background, it is possible that these groups suffer most from infectious disease morbidity. In summary, however, the public health developments do not suggest that longevity was driven by improved health and access to higher-quality care.

3.5 Conclusion

In this paper, I document that rapid structural change in rural Norwegian municipalities around the beginning of the 20th century increased the age at death of individuals experiencing childhood in such areas by ten months on average This impact on longevity is mainly driven by increases in the longevity of individuals from higher socioeconomic backgrounds as measured by the occupation of the household head. Moreover, this longevity effect only appears to be present if structural change has progressed significantly by the time of birth of individuals. This would be consistent with households gaining access to better-paying manufacturing jobs and therefore increasing investments in children, through education, nutrition and labour market choices. I then provide more detailed information about how local areas change in response to the adoption of hydropower. In line with previous findings by Leknes and Modalsli (2020), I find that hydropower adoption leads to significant increases in manufacturing employment, commonly interpreted as structural change. I also show that income per capita increases in response to this hydropower adoption, while inequality in socioeconomic status appears to be increasing, thereby providing a potential channel through which inequality in longevity might be affected. In the last step, I show that increases in longevity occur in areas despite the increased transmission of infectious diseases. Evidence of the spread of infectious disease in areas under rapid structural transformation increases, thereby suggesting that income gains heavily outweigh negative public health externalities.

Appendices

3.A Additional Tables

Table A1: Summary Statistics - Linked Census and Historical Death Register

Statistic	N	Mean	St. Dev.	Min	Max
Age at Death (months)	146,634	891.80	172.10	211.90	1,303.00
Birth Year	146,634	1,901.00	5.93	1,890	1,910
Death Year	146,634	1,974.00	14.91	1,928	2,014
HISCAM Value	131,480	54.97	9.44	37.34	99.00
SES Group	131,480	1.94	0.81	1	3
Birth Municipality	146,634	957.00	594.70	101	2,030

Note: The Table includes individuals linked from the 1910 census to the historical death register. It comprises men born between 1890 and 1910.

Table A2: Summary Statistics - Municipality Level Data

Statistic	N	Mean	St. Dev.	Min	Max
Panel A:					
Share Agriculture	2,376	0.67	0.25	0.00	0.98
Share Manufacturing	2,376	0.23	0.15	0.01	0.74
Share Service	2,376	0.08	0.09	0.00	0.55
Share Shipping	2,376	0.03	0.06	0.00	0.50
Labour Force	2,376	1,271.00	3,525.00	31	105,964
Panel B:					
Municipality	18,414	1,158.00	544.80	101	2,030
Year	18,414	1,905.00	8.94	1,890	1,920
Hydro Power Status	18,414	0.08	0.27	0	1
City Status	18,414	0.12	0.32	0	1
Hydro Power Potential	18,414	1.78	5.80	0.00	92.92
Area km^2	18,414	540.50	839.30	0.41	9,732.00
Population	18,414	3,863.00	10,055.00	162	259,364
Nr. Taxpayers	15,582	1,314.00	6,620.00	1	673,256
Income in 1,000 NOK	15,582	1,838.00	16,316.00	0	946,700
Wealth in 1,000 NOK	15,582	7,089.00	45,994.00	0	2,497,057
Nr. RWS	18,414	0.40	1.21	0	8
Nr. SSS	18,414	3.83	4.90	0	27

Note: The Table includes summary statistics for municipality level outcome variables obtained from the Municipality Database, Norwegian Ecological Data and transcribed from historical documents provided by SSB. For more detail on the data sources consult the data appendix. Panel A includes variables on sectoral employment for the years 1891, 1900, 1910 and 1920. Panel B includes data on other variables for the period 1890 to 1920, with the exception for income, wealth and taxpayers which are availbel from 1894 only.

Table A3: Summary Statistics - Health District Data

Statistic	N	Mean	St. Dev.	Min	Max
Health District Nr.	3,771	343.00	529.70	11	1,817
Year	3,771	1,902.00	11.43	1,880	1,920
Time to Treatment	3,771	0.49	2.52	−5	10
Opening Year HPP	1,088	1,906.00	8.03	1,890	1,920
Health Personal per 100,000	3,689	137.10	84.23	0.00	1,045.00
Nr. Doctors	3,771	5.16	18.88	0	289
Nr. Midwives	3,771	6.40	9.35	0	130
Nr. Vaccinators	3,771	3.93	3.04	0	59
Nr Pharmacists	3,771	0.71	1.85	0	25
Nr. Dentists	3,771	1.31	7.99	0	151
Cases per 100,000	3,689	1,450.00	1,246.00	0.00	19,896.00
Diarrhoea Cases	3,710	124.90	234.90	0	3,135
Diphtheria Cases	3,710	25.24	57.99	0	1,310
Pneumonia Cases	3,710	59.09	86.78	0	1,849
Typhoid Fever Cases	3,710	6.31	14.48	0	348
Deaths per 100,000	3,689	126.70	123.80	0.00	2,035.00
Diarrhoea Deaths	3,710	3.14	6.89	0	116
Diphtheria Deaths	3,710	3.41	8.35	0	211
Pneumonia Deaths	3,710	8.85	12.63	0	272
Typhoid Fever Deaths	3,710	0.72	1.95	0	60
Live Births	3,771	393.60	537.70	36	8,286
Infant Mortality per 1,000 LB	3,771	78.59	39.65	0.00	549.20

Note: The Table provides summary statistics for health district data collected from SSB's historical health statistic documents. The data were transcribed and harmonized to the 1890 health district level. More detail regarding data sources can be found in the data appendix.

Table A4: Balancing Table Hydropower versus Non-Hydropower Municipalities.

	No HP by 1920		HP by 1920			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. Means	P-Value
Hydro Potential	1.27	2.65	3.82	11.66	2.54	0.02
Steamship Stops 1880	3.94	4.93	3.39	4.79	-0.55	0.27
Railway Stops 1880	0.34	1.10	0.65	1.58	0.31	0.04
Mean Altitude	331.34	302.02	376.32	277.27	44.98	0.12
Area (km^2)	517.38	860.09	632.62	750.91	115.24	0.15
Rugedness	43.43	25.68	46.27	24.59	2.84	0.27
Distance to Coast (km)	20.50	33.07	24.80	37.65	4.30	0.26
City Status	0.12	0.32	0.12	0.32	0.00	0.96
Population Density	181.62	882.59	299.55	2621.80	117.93	0.63
Infant Deaths/LB	173.88	284.70	121.89	185.24	-51.99	0.02
Population	3016.70	3338.62	5053.50	13466.74	2036.80	0.10
Health Personal/Pop.	107.61	111.01	99.00	51.28	-8.61	0.22
Cases/Pop.	1788.70	1134.36	1914.79	1263.74	126.09	0.32
Deaths/Pop.	212.52	161.04	203.34	145.26	-9.18	0.55
Stillbirths/LB	27.15	10.98	27.82	12.23	0.67	0.59

Note: The table provides information on differences in 1890 characteristics between municipalities, which received hydropower by 1920 compared to those who did not. Health Personal, Cases and Deaths are rported per 100,000 inhabitants. Infant Deaths and Stillbirths are reported per 1,000 live births.

Table A5: Hydropower Openings and Infrastructure Dependence.

	Employment Share					
	Manu. (1)	Agri (2)	Serv (3)	Ship. (4)	Ln(LF) (5)	LF/Pop (6)
HPP	0.052***	-0.055***	0.002	0.001	0.101***	-0.733
	(0.012)	(0.013)	(0.005)	(0.003)	(0.032)	(0.458)
HPP × RWS	-0.010**	0.010**	0.0003	-0.0003	0.016	-0.045
	(0.004)	(0.005)	(0.003)	(0.0009)	(0.012)	(0.134)
HPP × SSS	0.002	-0.004*	0.001**	0.0002	0.006*	0.089
	(0.002)	(0.002)	(0.0006)	(0.0004)	(0.003)	(0.057)
R^2	0.913	0.956	0.923	0.871	0.967	0.501
Observations	2,376	2,376	2,376	2,376	2,376	2,376
Mean D.V.	0.225	0.665	0.078	0.031	6.75	32.8
Municipality FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓

Note: Standard errors are clustered at the municipality level. Estimated using aggregated census data for the years 1891, 1900, 1910 and 1920 following Leknes and Modalsli (2020). Railway stations and steamship stops taken from Norwegian Ecological Data covering the year 1880. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$

Table A6: Emigration Response to Hydropower Openings.

	FE	IV+FE
	(1)	(2)
HPP	-0.256	0.712
	(0.486)	(0.870)
R^2	0.303	0.300
Observations	3,564	3,564
Mean D.V.	0.833	0.833
F-test (1st stage)		97.4
Municipality FE	✓	✓
Year FE	✓	✓

Note: Estimated using aggregated emigration data from KDB for the years 1890-1920 in five year intervals. Instrument is hydropower potential interacted with year. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$. Standard errors are clustered at the municipality level.

Table A7: ABE Linking Probability and Hydropower Openings.

	(1)	(2)	(3)
Constant	0.262***		
	(0.009)		
HPP	0.073***	0.024**	0.003
	(0.019)	(0.011)	(0.004)
R^2	0.002	0.017	0.037
Observations	545,747	545,747	545,747
Mean D.V.	0.269	0.269	0.269
County FE		✓	
Birth Year FE		✓	✓
Birth Municipality FE			✓

Note: Estimated using full population census data from 1910 and historical death register. Sample consists of men born between 1890 and 1910 with a valid municipality of birth. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$. Standard errors are clustered at the municipality level.

Table A8: Communicable Disease Characteristics.

	Diarrhoea/Cholera	Diphtheria	Pneumonia	Typhoid Fever
Agent	bacterial	bacterial	bacterial/viral	bacterial
Transmission	fecal-oral	resp. droplets	resp. droplets	fecal-oral
Immunity Lifelong	no	no	no	no
Sub-Clinical Infections	yes	yes	yes	yes
Average CFR (1890-1920)	2.64%	13.95%	14.47%	11.29%
Treatment Option in 1920	Hydration	Antitoxin	None	Hydration

Note: Case fatality rates are own calculations using health average annual case fatality rate for the years 1890 to 1910 for all of Norway.

Table A9: Industry and Employment Variables.

Sector	Year	Variable Number
Agriculture	1891	67638, 67639, 67640, 67641, 67650
Manufacturing	1891	67642, 67643, 67644, 67645
Other	1891	67651, 67652, 67653, 67654
Services	1891	67646, 67647, 67648
Shipping	1891	67649
Agriculture	1900	67685, 67687
Manufacturing	1900	67688, 67689, 67690
Other	1900	67693, 67694, 67695, 67696, 67697, 67698
Services	1900	67691
Shipping	1900	67692
Agriculture	1910	67716, 67717, 67718, 67719, 67720, 67721, 67722, 67723, 67724, 67741, 67742, 67743
Manufacturing	1910	67725, 67726, 67727, 67744, 67745, 67746, 67747
Other	1910	67731, 67732, 67733, 67734, 67735, 67736, 67737, 67738, 67739, 67740, 67758, 67759, 67760, 67761, 67762, 67763, 67764, 67765, 67766
Services	1910	67728, 67729, 67748, 67749, 67757
Shipping	1910	67730, 67750, 67751, 67752, 67753, 67754, 67755, 67756
Agriculture	1920	67776, 67777, 67778, 67779, 67780, 67781, 67782, 67783, 67784, 67800, 67801, 67802, 67803, 67804, 67805, 67806, 67807, 67808, 67825, 67826, 67848, 67849
Manufacturing	1920	67785, 67786, 67787, 67788, 67809, 67810, 67811, 67827, 67828, 67829, 67830, 67831, 67850, 67851, 67852, 67853, 67854
Other	1920	67793, 67794, 67795, 67796, 67797, 67798, 67816, 67817, 67818, 67819, 67820, 67821, 67822, 67823, 67840, 67841, 67842, 67843, 67844, 67845, 67846, 67863, 67864, 67865, 67866, 67867, 67868, 67869, 67870, 67871
Services	1920	67789, 67790, 67791, 67812, 67813, 67814, 67832, 67833, 67834, 67835, 67836, 67837, 67838, 67855, 67856, 67857, 67858, 67859, 67860, 67861
Shipping	1920	67792, 67815, 67839, 67862

Note: The table shows information about the variables used for industry and employment data. The variable numbers in the third column correspond to variable numbers in the Municipality Database. More information on this data source can be found in Appendix 3.C.

Table A10: Income, Wealth and Taxpayer Responses to Hydropwer Establishment.

	Ln(Income/TP)	Ln(Wealth/TP)	Ln(TP)
	(1)	(2)	(3)
HPP	0.090***	0.075**	0.101***
	(0.026)	(0.031)	(0.032)
R^2	0.877	0.896	0.912
Observations	15,558	15,568	15,582
Mean D.V.	-0.413	1.10	6.73
Municipality FE	✓	✓	✓
Year FE	✓	✓	✓

Note: The table shows estimates of Equation 3.3 at the municipality level. All outcomes are in logarithmic form to allow for comparable interpretations relative to the reference period $t = -1$. Monetary values are reported in 1,000 NOK. Standard errors are clustered at the municipality level. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$.

Table A11: Hydropower Establishment, Infectious Disease and Health Professionals.

	Ln(Cases/Pop)	Ln(HP/Pop)	Ln(Deaths/Pop)
	(1)	(2)	(3)
HPP	0.102**	-0.084	-0.020
	(0.052)	(0.065)	(0.047)
R^2	0.556	0.519	0.350
Observations	3,666	3,504	3,526
Mean D.V.	6.98	4.78	4.59
Health District FE	✓	✓	✓
Year FE	✓	✓	✓

Note: The table shows estimates of Equation 3.3 for the health district level. All outcomes are in logarithmic form to allow for comparable interpretations relative to the reference period $t = -1$ and are measured per 100,000 inhabitants. Standard errors are clustered at the health district level. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$.

3.B Additional Figures

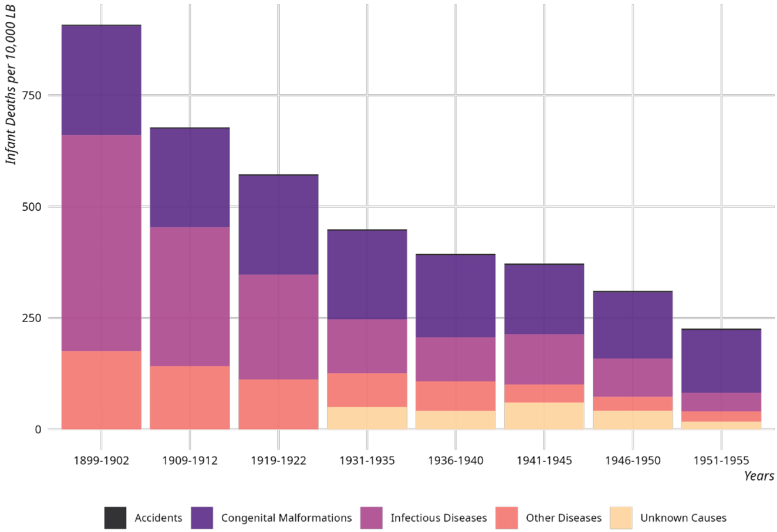


Figure A1: Infant Mortality by Cause of Death 1899 - 1955.

Note: The figure shows infant mortality per 1,000 live births in four year bins by cause of death. The infectious disease category includes deaths due to tuberculosis, pneumonia, diarrhea, influenza, scarlet fever, diphtheria, whooping cough and measles. The data are taken from Backer (1961).

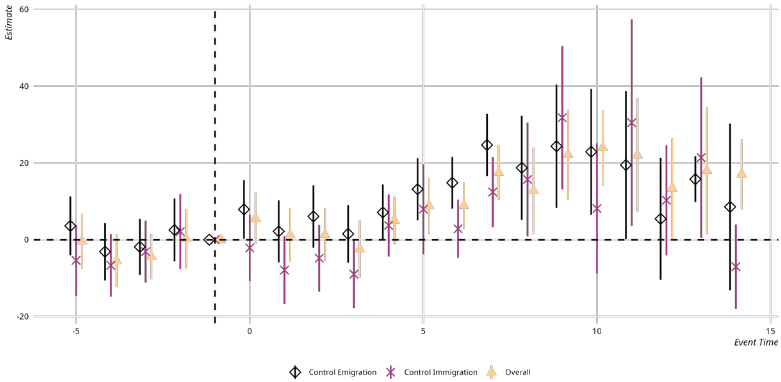


Figure A2: Longevity Effects and Robustness to Immigration and Emigration.

Note: The figure shows estimates of Equation 3.2 for three different sample selections estimated using Sun and Abraham (2020). Estimates can be interpreted as changes relative to event time $t = -1$. The different samples all include men born between 1890 and 1920. Triangle shapes represent the result for the full sample (n = 122,552), cross shapes exclude individuals who have no household member living in the birth municipality before 1890 (n = 82,743) and diamond shaped estimates are based on a sample of individuals who have the same residence municipality in 1910 as their birth municipality (n = 99,193). Standard errors used for the calculation of the 95-% confidence intervals are clustered at the municipality level.

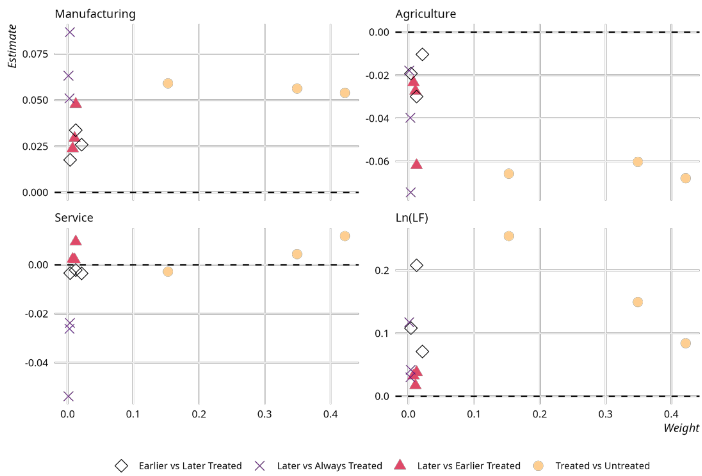


Figure A3: Bacon Decomposition for TWFE Models.

Note: The figure provides estimates and weights obtained from a decomposition exercise following Goodman-Bacon (2021). Each panel provides decomposition estimates and weights for a separate TWFE model presented earlier in Table 3.3.

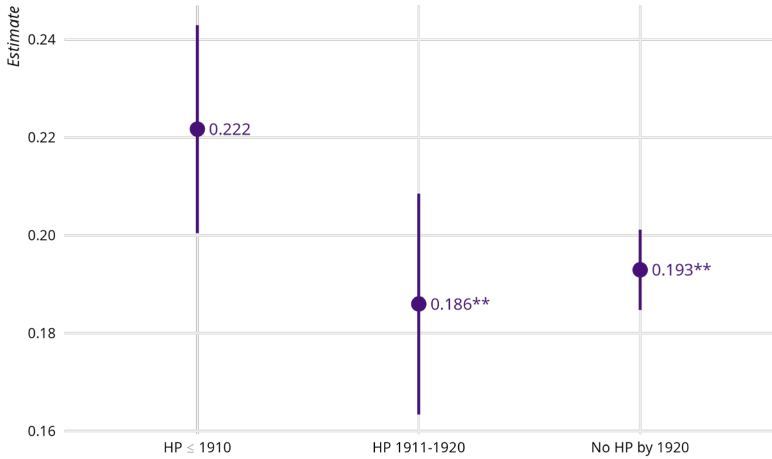


Figure A4: Mover Share by Hydropower Status of Municipality.

Note: The figure depicts the share of movers in municipalities in the 1910 census by hydropower status. Estimates were obtained from a simple OLS regression of hydropower status on a dummy indicating mover status of individuals. The data used for this excercise is the sample of linked individuals between the 1900 and 1910 census. Stars indicate result of hypothesis test of difference to HP $\leq$ 1910. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$.

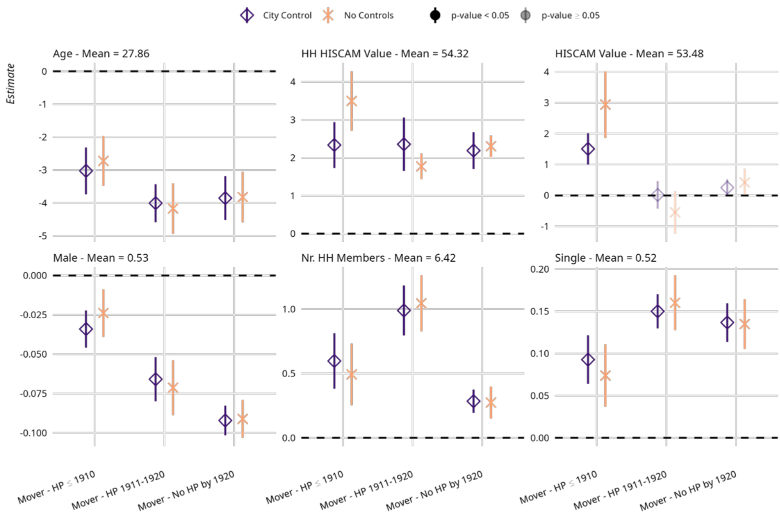


Figure A5: Characteristics of Movers in Comparison to Stayers.

Note: The figure shows estimates of simple OLS regressions of hydropower status on relevant outcomes. Each panel was obtained from a separate estimation. Dark diamond shaped estimates control additionally for a city dummy if the person has moved from a city . The data contain individuals linked across the 1900 and 1910 census. The estimates should be interpreted as differences in observables relative to the omitted stayer category. Standard errors used for the 95-% confidence intervals are clustered at the municipality level.

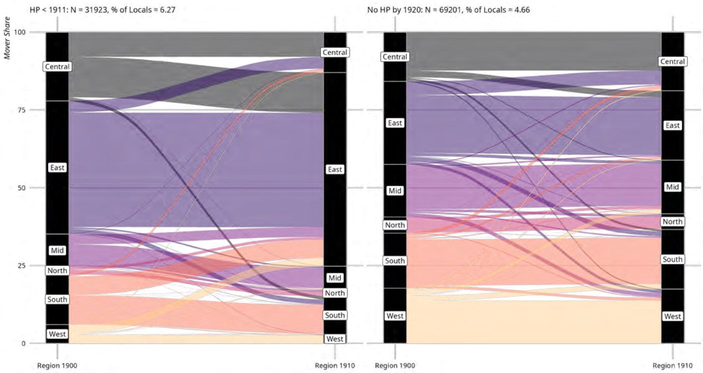
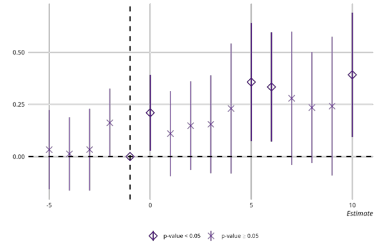
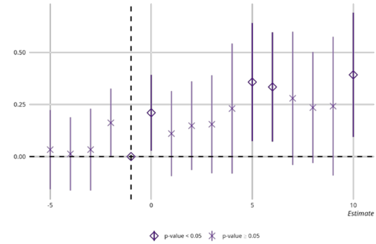


Figure A6: Mover Flows by Hydropower Status.

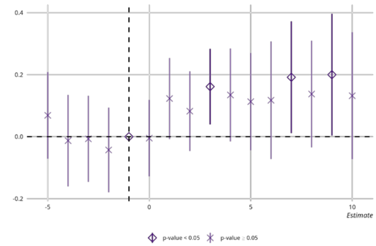
Note: The figure shows changes regional flows of migrants from the linked 1900 and 1910 census data. The graph shows the relative flows of movers. % of Locals indicates the overall number of movers in this category relative to the receiving municipalities popualtion in 1900. The left panel shows te mover flows towards hydropower municipalities, while the right one indicates mover flows to non-hydropower municipalities.



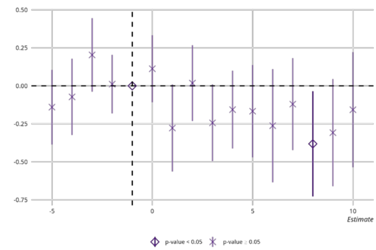
(a) Diarrhoea



(b) Diphtheria



(c) Pneumonia



(d) Typhoid Fever

Figure A7: Hydropower Openings and Impact on Common Infectious Disease Cases

Note: The figure shows event-study estimates of the impact of hydropower openings on diarrhoea, diphtheria, pneumonia and typhoid fever cases per 100,000 inhabitants. Event-study estimates follow methodology proposed in Sun and Abraham (2020). All outcomes are in logarithmic form to allow for comparable interpretations relative to the reference period $t = -1$ and were obtained by estimating Equation 3.1. Standard errors are clustered at the municipality level. Shapes and transparency indicates significance at 5-% level and error bars represent 95-% confidence intervals.

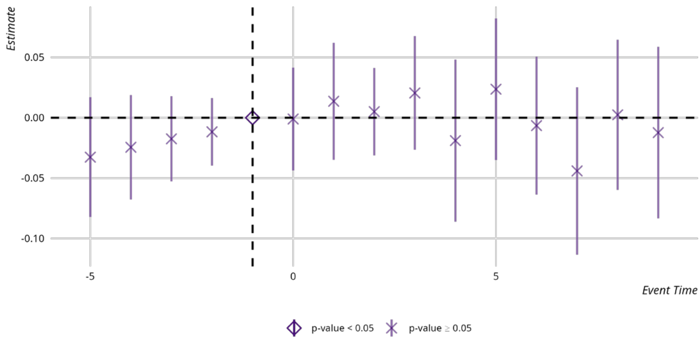


Figure A8: Impact of Hydropower Openings on Fertility.

Note: The figure shows event-study estimates of hydropower openings on the lead number of live births per 100 inhabitants. The outcome variable is transformed using the natural logarithm for easier interpretation. The data used for estimation was digitised from historical health records. Event-study estimates are obtained using the methodology presented in Sun and Abraham (2020). Standard errors used for the calculation of the 95-% confidence intervals are clustered at the health district level.

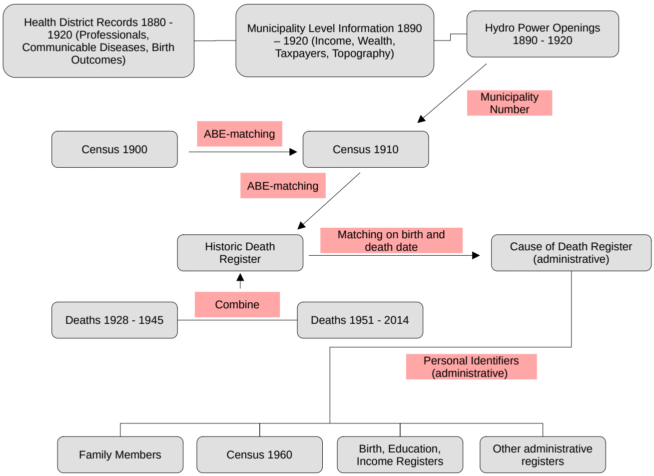


Figure A9: Connection of Historical and Modern Administrative Data.

Note: The figure shows a simplified version of the data connection and linking process. Information included here contains data on various levels of aggregation.

3.C Data Sources and Definitions

IPUMS Census Data

Individual-level census data for the years 1900 and 1910 come from IPUMS International, and access is restricted. Access to these data can be obtained by following this link. Information on sample characteristics can also be found via the IPUMS website.

Death Data 1928 - 1946

Individual-level death records for the entire Norwegian population for the period 1928 to 1945 have been obtained through the Digital Archives (Digitalarkivet) of Norway. The data are in general publicly available, excluding the cause of death. In order to receive the full dataset, it is necessary to get in touch with the Digital Archives, which were very generous in assisting with the provision of these data. A link to additional information on the population movement register for 1928 - 1946 can be found here. In order to use these data, some pre-processing and cleaning of municipalities, first and last names is necessary.

Death Data 1951 - 2014

Individual-level death records for the entire Norwegian population with death dates between 1951 and 2014 can be accessed via the Digital Archives. They provide a web tool enabling all deaths to be searched by first name, last name, sex, birthdate and death date using the following link. Since web scraping seemed to be very cumbersome for data that appear to be publicly available, the Digital Archives provided the full dataset for all deaths directly.

Geographic Information System Data

For the construction of instrumental variables, maps and geographic control variables, I have used data from various sources.

- **Elevation Data:**
  Elevation data for Norway were obtained from the Copernicus Land Monitoring Service and obtained in 25 m resolution. The data come in geotif format.

- **River data:**
  Data on rivers and river flow classes were obtained from the HydroRivers database, which is publicly available, including relevant documentation. These data include all rivers in Norway and the relevant flow strength for the construction of the instrumental variable.

- **Municipality Structure 1900:**
  A shapefile including all 594 Norwegian municipalities and their geographic extent was obtained from the Norwegian Centre for Research Data (NSD).

Norwegian Ecological Data

Data on pre-existing infrastructure in Norwegian municipalities for the year 1880, including steamship and railway access during winter and summer were obtained via ICPSR. The data were compiled by Frank H Aarebrot and Stein Kuhnle at the Department of Sociology at the University of Bergen. The data generally include four separate datasets and the variables for this analysis were obtained from dataset 4 on inter-municipal communication. The data are publicly accessible via the following link. The data include information about 461 municipalities which are matched to accord with the municipality structure in the year 1900.

Norwegian Municipality Database (Kommunedatabase)

The Norwegian Municipality Database (Kommunedatabase) contains information on various topics concerning Norwegian municipalities (kommuner), from the 19th century up until today. Historical data concerning the period 1890 to 1920 are available for some variables. During this project, the Norwegian Centre for Research Data restructured the database. There should not have been any change to the source material, but no documentation of this change is available. Below I provide a short description of the relevant variables and datasets from the Municipality Database used in this project.

- **Population Movement Data:**
  I obtained data on infant deaths, stillbirths, live births, emigration and population for the years 1890 to 1920. Some of these variables are only available in five-year aggregates.

- **Occupation and Industry Statistics:**
  Data on the number of individuals working in specific sectors and industries, as well as labour force counts per municipality for the years 1891, 1900, 1910 and 1920 were also obtained here and aggregated into four categories (manufacturing, agriculture, services, shipping)

Public Health Data

Data concerning the developmentof public health in Norwegian health districts during the period 1880 to 1920 can be found as PDF files through SSB's historical statistics page. All health data at the health district level is organised under the subsection 'Population. Health' (Befolkning. Helseforhold). The health statistics can then be found in annual publications under the subsection 'Health Statistics' (Helsestatistikk). The documents sometimes change from year to year, but, in general, they are fairly consistent in their content. Since the data are only available as PDF documents, I transcribed the information and organised it to match the health district structure in 1880. If you wish to look at an example of such a document, please see health statistics report 1890. These data are available annually for all health districts. Below, I will provide more detailed information about which data were obtained from this source.

- **Infant Deaths:**
  Number of infant deaths during first 24 hours after birth, infant deaths first year after birth.

- **Live Births:**
  Number of live births in all health districts.

- **Health Personnel:**
  Number of doctors, dentists, midwives, pharmacists and vaccinators per health district.

- **Cases of Infectious Diseases:**
  Number of reported cases of diarrhoea (incl. cholera), diphtheria, pneumonia and typhoid fever in the health district.

- **Deaths from Infectious Diseases:**
  Number of reported deaths from diarrhoea (incl. cholera), diphtheria, pneumonia and typhoid fever in the health district.

Income, Wealth and Taxpayer Data:

The information about the number of taxpayers, and income and wealth by municipality also comes from historical documents from SSB, which were published under the title 'Monthly Statistical Booklets' (Statiskk Månedshefte). These data were collected for our research centre and are only available from 1894 to 1920. The data include information on local tax collection, incomes and wealth in local communities. An example booklet for the year 1895 can be found under the following link. Note that the taxpayer definition is very broad and includes both physical and non-physical entities.

Historical International Standard Classification of Occupations

The censuses for the years 1900 and 1910 obtained via IPUMS include detailed records of individuals' occupations. These occupations are used to assign a value to socioeconomic status, referred to as the Historical International Standard Classification of Occupations (HISCAM). HISCAM provides a measure of social stratification that is based on historical records from the Netherlands, Germany, France, Sweden, the UK, Canada and Belgium. Lambert et al. (2013) describe the construction of this measure and how the HISCAM measure can be used to explain social stratification and inequality in a historic context. Norway is not one of the countries on which HISCAM is based and the most closely related country, Sweden, only contains a relatively small sample. For this reason, the analysis in this paper is based on the universal scale for the later period 1890 to 1930. The data can be accessed via this link. They are described in detail in Lambert et al. (2013).

3.D Linked Historical Data

This paper uses novel data linkages between census data from 1900 and 1910, as well as linkages to death data starting in 1928. The most important parts of the data-linking process have already been described in the main paper. The main linking procedure follows the algorithm first outlined in Abramitzky, Boustan and Eriksson

(2012, 2014) and recently summarised in Abramitzky et al. (2021). In the following, it will be referred to as ABE (Abramitzky, Boustan and Eriksson) matching. This procedure endeavours to solve one main problem, which is particularly prominent when working with individual-level historical data. It mainly concerns the problem of observing individuals (or firms and other entities) across time. Since identifiers were not used historically, censuses simply counted individuals in a historical context. It is not possible, however, to follow the same person across different censuses over time. ABE matching tries to solve this lack of a longitudinal dimension, but it could also be applied to similar non-longitudinal problems. The general idea of this matching procedure is that individuals in two different datasets can be matched on characteristics that do not change over time, and should therefore be identical in the two different datasets (e.g. birth year, birthplace, first name).

ABE matching thus means applying an automated function that goes through the following steps. In the first step, first and last names are cleaned so that they do not contain any symbols and specific non-English characters. In Norwegian, there are special characters such as the letters Æ, Ø, and Å in both capitalised and non-capitalised forms. These characters are substituted by a, o and a, for consistency. These characters are also often prone to transcription errors. Multiple first and last names are subset to only contain the first word of each name. For example, the name Lars Ivar Hansen will become Lars Hansen. This is necessary because of individuals over time changing how they report their full names. Moreover, all names are decapitalised for the sake of consistency. In the next step, I then create blocks of potential matches, by creating Cartesian products of individuals sharing the same sex, place of birth, birth year and the first letter of both the standardised first name and last name. The standardisation is performed using the New York State Identification and Intelligence System (NYSIIS) algorithm and is only used for the creation of standardised first letters of first and last names. The actual string distance is computed on the cleaned and non-standardised names. The blocking is mainly used to increase the potential match rate and to increase computing time by creating smaller Cartesian products of potential matches. For all matches within blocks, the string distance is computed using the Jaro-Winkler (JW) method. This measure provides a numeric value for how much a string would need to be edited in order to be identical to a comparison string. For example the JW string distance between *hans kristiansen* and *hans kristensen* is 0.0996 using the stringdist package in R (van der Loo, 2014). I then keep all individuals that are unique matches in the 1900 census and the 1910 census if the JW score is smaller or equal to 0.1. All other matches are only selected as a quality match if the match is the best match in both the 1900 and 1910 censuses, the JW score is smaller or equal to 0.1 and the next best match of this name has a JW score higher than 0.1. This is done to avoid arbitrary choices of matches that are of almost identical quality.

Envisage matching *hans kristiansen* with *hans kristensen* or *hans kristianson*. The first pair has a JW score of 0.0996, while the second has a score of 0.0417. Conditional on sex, birthplace and birth year, the difference between those two names is very arbitrary and a successful match is therefore ruled out. After all matches have been selected using these criteria, I iteratively link all remaining individuals in the censuses who have not been matched by relaxing the blocking variables. Note that the blocking criteria do not lower the quality requirement for the linkages, but only increase the number of comparisons and therefore increase computing time.

3.D.1 Linking the 1900 and 1910 Censuses

In order to obtain a longitudinal sample of migrants between the years 1900 and 1910, I link the 1900 and 1910 censuses using the matching procedure outlined above. Overall, I am able to link 31 per cent of all individuals in the 1900 census to individuals in 1910. This overall linking rate is a weighted average of the 34 per cent of men and 28 percent of women I am able to link. Especially among women, the false positive rate is likely higher, mainly due to patrilineal surnames. I account for this by only matching women by first names if they are above a certain age, but this is still a fuzzy method with a higher false positive rate compared to the male matches. For the census data, I block on the municipality of birth, birth year, sex, and first letter of the last name and first name. For unmarried women, I do not block on the first letter of the last name.

Table A12: Characteristics of Linked and Non-Linked Individuals.

	Linked (N=556,218)		Not Linked (N=1,224,838)	
	Mean	Std. Dev.	Mean	Std. Dev.
Male	0.518	0.500	0.460	0.498
Birth Year	1872.834	19.811	1875.785	19.321
Single	0.449	0.497	0.538	0.499
Birth Mun. City	0.215	0.411	0.220	0.414
Three Largest Cities	0.075	0.263	0.090	0.287
Distance to Coast	21.437	35.828	22.155	35.661
Mun. Size (km^2)	483.036	720.596	495.654	732.380
Steamship Stops	5.339	7.033	5.526	7.357
Railway Stops	0.738	1.610	0.769	1.634

Note: The table presents summary statistics (mean and standard deviation) for individuals from the 1900 census. Summary statistics are presented separately for individuals linked to the 1910 census, and individuals who were not linked successfully. Variables indicating municipality characteristics refer to the municipality of birth.

In Table A12, I provide some summary statistics, comparing linked and non-linked individuals. This sample includes all individuals born in 1900 or earlier who have a valid birth municipality. On average, linked individuals are more male, older and less likely to be single. This has to do with the aforementioned patrilineal surnames and the fact that especially women who were already married in 1900 are more likely to be matched because their last name should not change across censuses due to a very low divorce rate. The same argument holds for the lower rate of unmarried/single individuals. Focusing on the municipality of birth characteristics, individuals are less likely to be matched if they were born in a city municipality.[35] This becomes even more distinct when looking at the biggest three cities (Oslo (Kristiania), Bergen and Trondheim). The reason why linking rates in cities are lower has to do with the ABE algorithm which, as a selection criterion, requires matches to be significantly better than the next best match. With a larger pool of potential matches (e.g. in cities) this probability decreases. Other characteristics such as steamship stops and railway stops are also smaller in the linked sample, which is probably due to a correlation between transport infrastructure and population density (e.g. city status).

3.D.2 Linking the 1910 Census and the Historic Death Register

In order to obtain a sample of individuals including their municipality of birth and the date of death, I link the male individuals

[35]Note that city status in Norway did not necessarily mean that these were metropolitan areas. The mean city population in 1900 was just about 10,000 (median 3,500), while the mean population in municipalities without city status was 3,000 (median 2,500).

Table A13: Characteristics of Linked and Non-Linked Individuals.

	Linked (N=148,306)		Not Linked (N=405,181)	
	Mean	Std. Dev.	Mean	Std. Dev.
Birth Year	1900.597	5.937	1900.253	5.937
HH. HISCAM Value	55.118	9.582	53.590	7.426
Nr. H. Members	7.100	4.869	7.069	4.551
Birth Mun. City	0.378	0.485	0.247	0.432
Distance to Coast	18.211	34.366	20.694	35.320
Mun. Size (km^2)	396.278	694.518	483.771	738.913
Steamship Stops	7.666	8.982	6.088	7.762
Railway Stops	1.179	1.964	0.834	1.712
Three Largest Cities	0.182	0.386	0.106	0.308

Note: The table presents summary statistics (mean and standard deviation) for men born between 1890 and 1910. Summary statistics are presented separately for individuals linked to the historical death register, and individuals who were not linked successfully. Variables indicating municipality characteristics refer to the municipality of birth.

born between 1890 and 1910 to individuals in the historical death register. Since the time difference between the occurrence of death and 1910 might span over 100 years, the linking rate is slightly smaller than for the census data. Overall, I can link approximately 27 per cent of men born between 1890 and 1910 to individuals whom I observe in the historical death register. I also use slightly different blocking variables for this linking process. I block on the first letter of the last name and first name, birth year, birth month and day of birth. I am not able to block on the birth municipality since I do not observe this variable in the historical death register. Note that, when linking the censuses, I could not rely on the birth month and day of birth, since this variable is not applicable in the 1900 census. Blocking on the birth month and day of birth significantly improves this linking process in the absence of the municipality of birth.

Besides the linking rate, some key characteristics of individuals are slightly different in the linked sample of individuals from the 1910 census than in the non-linked sample. We see that, overall, individuals are slightly positively selected in terms of the socioeconomic status of the household head. They are more likely to be born in areas that had city status and are more likely to be born in smaller municipalities.

Bibliography

Aaberge, Rolf, Anthony B Atkinson, and Jørgen Modalsli. 2020. "Estimating long-run income inequality from mixed tabular data: Empirical evidence from Norway, 1875–2017." *Journal of Public Economics*, 187: 104196.

Aarebrot, Frank H, and Stein Kuhnle. 1984. *Norwegian Ecological Data, 1868-1903.* Inter-university Consortium for Political and Social Research.

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, James Feigenbaum, and Santiago Pérez. 2021. "Automated linking of historical data." *Journal of Economic Literature*, 59(3): 865–918.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2012. "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." *American Economic Review*, 102(5): 1832–56.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration." *Journal of Political Economy*, 122(3): 467–506.

Adda, Jérôme. 2016. "Economic activity and the spread of viral diseases: Evidence from high frequency data." *The Quarterly Journal of Economics*, 131(2): 891–941.

Adda, Jérôme, and Francesca Cornaglia. 2010. "The Effect of Bans and Taxes on Passive Smoking." *American Economic Journal: Applied Economics*, 2(1): 1–32.

Adermon, Adrian, Mikael Lindahl, and Mårten Palme. 2021. "Dynastic human capital, inequality, and intergenerational mobility." *American Economic Review*, 111(5): 1523–48.

Aizer, Anna, and Laura Stroud. 2010. "Education, knowledge and the evolution of disparities in health." National Bureau of Economic Research.

Aizer, Anna, Laura Stroud, and Stephen Buka. 2016. "Maternal stress and child outcomes: Evidence from siblings." *Journal of Human Resources*, 51(3): 523–555.

Allen, Robert C. 2009. "The industrial revolution in miniature: The spinning jenny in Britain, France, and India." *The Journal of Economic History*, 69(4): 901–927.

Almond, Douglas. 2006. "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population." *Journal of political Economy*, 114(4): 672–712.

Almond, Douglas, Hilary W Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the war on poverty: The impact of food stamps on birth outcomes." *The review of economics and statistics*, 93(2): 387–403.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood circumstances and adult outcomes: Act II." *Journal of Economic Literature*, 56(4): 1360–1446.

Alsan, Marcella, and Claudia Goldin. 2019. "Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880–1920, mortality." *Journal of Political Economy*, 127(2): 586–638.

Altonji, Joseph, and Thomas A. Dunn. 2000. "An Intergenerational Model of Wages, Hours, and Earnings." *Journal of Human Resources*, 35(2): 221–258.

Anderson, D Mark, Kerwin Kofi Charles, and Daniel I Rees. 2018. "Public health efforts and the decline in urban mortality." National Bureau of Economic Research.

Anderson, D Mark, Kerwin Kofi Charles, Michael McKelligott, and Daniel I Rees. 2022. "Estimating the Effects of Milk Inspections on Infant and Child Mortality, 1880- 1910." Vol. 112, 188–92.

Attanasio, Orazio, and Torben Heien Nielsen. 2020. "Economic Resources, Mortality and Inequality."

Backer, Julie E. 1961. *Dødeligheten og dens årsaker i Norge: 1856-1955.* Statistisk Sentralbyrå.

Bailey, Roy E, Timothy J Hatton, and Kris Inwood. 2018. "Atmospheric pollution, health, and height in late nineteenth century Britain." *The Journal of Economic History*, 78(4): 1210–1247.

Baldacchinoel, Godfrey, and Anders Wivel. 2020. *Handbook on the Politics of Small States.* Cheltenham, Gloucestershire, United Kingdom:Edward Elgar Publishing Limited.

Beach, Brian. 2021. "Water Infrastructure and Health in U.S. Cities." National Bureau of Economic Research Working Paper 28563.

Beach, Brian, and W Walker Hanlon. 2018. "Coal smoke and mortality in an early industrial economy." *The Economic Journal*, 128(615): 2652–2675.

Beach, Brian, Joseph Ferrie, Martin Saavedra, and Werner Troesken. 2016. "Typhoid fever, water quality, and human capital formation, mortality." *The Journal of Economic History*, 76(1): 41–75.

Becker, Gary S, and Nigel Tomes. 1979. "An equilibrium theory of the distribution of income and intergenerational mobility." *Journal of political Economy*, 87(6): 1153–1189.

Berger, Eirik, and Håkon Block Vagle. 2017. "Equality before the welfare state: the Norwegian income distribution 1892-1929." Master's diss.

Bhalotra, Sonia R, and Atheendar Venkataramani. 2015. "Shadows of the captain of the men of death: Early life health interventions, human capital investments, and institutions, mortality." *Human Capital Investments, and Institutions (August 8, 2015).*

Bhalotra, Sonia R., Martin Karlsson, and Therese Nilsson. 2017. "Infant Health and Longevity: Evidence from a Historical Trial in Sweden." *Journal of European Economic Association*, 15(5): 1101–1157.

Bharadwaj, Prashant, Julian V Johnsen, and Katrine V Løken. 2014. "Smoking bans, maternal smoking and birth outcomes, mortality." *Journal of Public economics*, 115: 72–93.

Bharadwaj, Prashant, Katrine Vellesen Løken, and Christopher Neilson. 2013. "Early life health interventions and academic achievement, mortality." *American Economic Review*, 103(5): 1862–91.

Bhuller, Manudeep, Magne Mogstad, and Kjell G Salvanes. 2017. "Life-cycle earnings, education premiums, and internal rates of return, mortality." *Journal of Labor Economics*, 35(4): 993–1030.

Björklund, Anders, and Markus Jäntti. 1997. "Intergenerational income mobility in Sweden compared to the United States." *The American Economic Review*, 87(5): 1009–1018.

Björklund, Anders, Markus Jäntti, and Matthew J Lindquist. 2009. "Family background and income during the rise of the welfare state: brother correlations in income for Swedish men born 1932–1968." *Journal of Public Economics*, 93(5-6): 671–680.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2005. "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital." *American economic review*, 95(1): 437–449.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *The Quarterly Journal of Economics*, 122(1): 409–439.

Blanden, Jo, Alissa Goodman, Paul Gregg, and Stephen Machin. 2004. "Changes in intergenerational mobility in Britain." *Generational Income Mobility in North America and Europe*, , ed. MilesEditor Corak, 122–146. Cambridge University Press.

Blau, Francine D, and Lawrence M Kahn. 2017. "The gender wage gap: Extent, trends, and explanations." *Journal of Economic Literature*, 55(3): 789–865.

Blau, Francine D., Peter Brummund, and Albert Yung-Hsu Liu. 2013. "Trends in Occupational Segregation by Gender 1970-2009: Adjusting for the Impact of Changes in the Occupational Coding System." *Demography*, 50(2): 471–494.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and path dependence." *The quarterly journal of economics*, 127(2): 587–644.

Borge, Lars-Erik, Pernille Parmer, and Ragnar Torvik. 2015. "Local natural resource curse?" *Journal of Public Economics*, 131: 101–114.

Borjas, George J. 1991. "Immigration and self-selection." In *Immigration, trade, and the labor market*. 29–76. University of Chicago Press.

Boustan, Leah. 2018. "Urbanization in American Economic History, 1800–2000." In *The Oxford Handbook of American Economic History*. Oxford University Press.

Brandén, Gunnar, and Martin Nybom. 2019. "Utvecklingen av intergenerationell rörlighet i Sverige."

Bratberg, Espen, Jonathan Davis, Bhashkar Mazumder, Martin Nybom, Daniel D Schnitzlein, and Kjell Vaage. 2017. "A comparison of intergenerational mobility curves in Germany, Norway, Sweden, and the US." *The Scandinavian Journal of Economics*, 119(1): 72–101.

Bratsberg, Bernt, Simen Markussen, Oddbjorn Raaum, Knut Roed, and Ole Jorgen Røgeberg. 2018. "Trends in assortative mating and offspring outcomes."

Bryan, Jennifer. 2017. "gapminder: Data from Gapminder." R package version 0.3.0.

Bütikofer, Aline, and Kjell G Salvanes. 2020. "Disease control and inequality reduction: Evidence from a tuberculosis testing and vaccination campaign, radio." *The Review of Economic Studies*, 87(5): 2087–2125.

Bütikofer, Aline, Katrine V Løken, and Kjell G Salvanes. 2019. "Infant health care and long-term outcomes, mortality." *Review of Economics and Statistics*, 101(2): 341–354.

Cain, Louis, and Sok Chul Hong. 2009. "Survival in 19th century cities: The larger the city, the smaller your chances." *Explorations in Economic History*, 46(4): 450–463.

Carpenter, Robert, Cliona McGarvey, Edwin A Mitchell, David M Tappin, Mechtild M Vennemann, Melanie Smuk, and James R Carpenter. 2013. "Bed sharing when parents do not smoke: is there a risk of SIDS? An individual level analysis of five major case–control studies." *BMJ open*, 3(5).

Case, Anne, and Angus Deaton. 2015. "Rising Morbidity and Mortality in Midlife Among White non-Hispanic Americans in the 21s Century." *Proceedings of the National Academy of Sciences of the United States of America*, 112(49): 15078—-15083.

Case, Anne, and Angus Deaton. 2017. "Mortality and morbidity in the 21st century." *Brookings papers on economic activity*, 2017(1): 397–476.

Case, Anne, Darren Lubotsky, and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review*, 92(5): 1308–1334.

CDC. 2020. "Health Effects of Secondhand Smoke." *https://www.cdc.gov/tobacco/data_statistics/fact_sheets/secondhand_smoke/health_effects/index.htm#:~:text=There%20is%20no%20risk%2Dfree,infant%20death%20syndrome%20(SIDS).*, Accessed: 2020-10-30.

Chadwick, Laura, and Gary Solon. 2002. "Intergenerational income mobility among daughters." *American Economic Review*, 92(1): 335–344.

Chen, Alice, Emily Oster, and Heidi Williams. 2016. "Why is infant mortality higher in the United States than in Europe?" *American Economic Journal: Economic Policy*, 8(2): 89–124.

Chetty, Raj, and Nathaniel Hendren. 2018. "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects." *The Quarterly Journal of Economics*, 133(3): 1107–1162.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. "The association between income and life expectancy in the United States, 2001-2014." *Jama*, 315(16): 1750–1766.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment." *American Economic Review*, 106(4): 855–902.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014*a*. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014*b*. "Is the United States still a land of opportunity? Recent trends in intergenerational mobility." *American Economic Review*, 104(5): 141–47.

Connolly, Marie, Catherine Haeck, and Jean-William P Laliberté. 2020. "Parental Education Mitigates the Rising Transmission of Income between Generations." In *Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth*. University of Chicago Press.

Corak, Miles. 2013. “Income inequality, equality of opportunity, and intergenerational mobility.” *Journal of Economic Perspectives*, 27(3): 79–102.

Currie, Janet, and Hannes Schwandt. 2016*a*. “Inequality in mortality decreased among the young while increasing for older adults, 1990–2010.” *Science*, 352(6286): 708–712.

Currie, Janet, and Hannes Schwandt. 2016*b*. “Mortality inequality: the good news from a county-level approach.” *Journal of Economic Perspectives*, 30(2): 29–52.

Currie, Janet, and Reed Walker. 2011. “Traffic Congestion and Infant Health: Evidence from E-ZPass.” *American Economic Journal: Applied Economics*, 3(1): 65–90.

Currie, Janet, Hannes Schwandt, and Josselin Thuilliez. 2020. “Pauvreté, Egalité, Mortalité: mortality (in) equality in France and the United States.” *Journal of Population Economics*, 33(1): 197–231.

Cutler, David, and Grant Miller. 2005. “The role of public health improvements in health advances: the twentieth-century United States.” *Demography*, 42(1): 1–22.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney. 2006. “The determinants of mortality.” *Journal of economic perspectives*, 20(3): 97–120.

Dahl, Molly W, and Thomas DeLeire. 2008. “The association between children’s earnings and fathers’ lifetime earnings: estimates using administrative data.” University of Wisconsin-Madison, Institute for Research on Poverty.

Daltveit, Anne Kjersti, Nina Øyen, Rolv Skjærven, and Lorentz M Irgens. 1997. “The epidemic of SIDS in Norway 1967-93: changing effects of risk factors, infant, mortality.” *Archives of disease in childhood*, 77(1): 23–27.

Davis, Jonathan, and Bhashkar Mazumder. 2022. “The decline in intergenerational mobility after 1980.”

Daysal, N Meltem, Hui Ding, Maya Rossin-Slater, and Hannes Schwandt. 2021. "Germs in the Family: The Long-Term Consequences of Intra-Household Endemic Respiratory Disease Spread." National Bureau of Economic Research.

Deryugina, Tatyana, and David Molitor. 2021. "The causal effects of place on health and longevity." *Journal of Economic Perspectives*, 35(4): 147–70.

de Walque, Damien. 2010. "Education, Information and Smoking Decisions: Evidence from Smoking Histories in the United States, 1940–2000." *Journal of Human Resources*, 45(3): 682–717.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review*, 108(4-5): 899–934.

Doyle, Orla. 2020. "The first 2,000 days and child skills." *Journal of Political Economy*, 128(6): 2067–2122.

Duflo, Esther, and Rohini Pande. 2007. "Dams." *The Quarterly Journal of Economics*, 122(2): 601–646.

Duncan, Otis Dudley, and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indexes." *American Sociological Review*, 20(2): 210–217.

Easterlin, Richard A. 1995. "Industrial revolution and mortality revolution: Two of a kind?" *Journal of Evolutionary Economics*, 5(4): 393–408.

Engzell, Per, and Carina Mood. 2021. "How Robust Are Estimates of Intergenerational Income Mobility?"

Faugli, Per Einar. 2012. "Vann og energiforvaltning – glimt fra NVEs historie." Norges Vassdrags- og Energidirektorat NVE Rapport 26-2012.

Ferrie, Joseph P, and Werner Troesken. 2008. "Water and Chicago's mortality transition, 1850–1925." *Explorations in Economic History*, 45(1): 1–16.

Ferrie, Joseph P, Karen Rolf, and Werner Troesken. 2015. "Lead exposure and the perpetuation of low socioeconomic status." *cliometrics. org*.

FHI. 2017. "History of the Norwegian Public Health Institue." `https://www.fhi.no/en/about/this-is-the-norwegian-institute-of-public-health/history-of-the-norwegian-institute-/`, Accessed: 2020-10-12.

Galassi, Gabriela, David Koll, and Lukas Mayr. 2021. "The Intergenerational Correlation of Employment." University of Bonn and University of Mannheim, Germany CRC TR 224 Discussion Paper Series.

Gerdrup, Karsten R. 1998. *Skattesystem og skattestatistikk i et historisk perspektiv.* Statistisk sentralbyrå.

Gjertsen, Finn. 2002. "Dødsårsaksregisteret–en viktig datakilde for medisinsk forskning." *Tidsskrift for Den norske legeforening*.

Glied, Sherry, and Adriana Lleras-Muney. 2008. "Technological innovation and inequality in health." *Demography*, 45(3): 741–761.

Goldin, Claudia. 2014. "A grand gender convergence: Its last chapter." *American Economic Review*, 104(4): 1091–1119.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*.

Grönqvist, Erik, Björn Öckert, and Jonas Vlachos. 2017. "The intergenerational transmission of cognitive and noncognitive abilities." *Journal of Human Resources*, 52(4): 887–918.

Hack, Maureen, Nancy K Klein, and H Gerry Taylor. 1995. "Long-term developmental outcomes of low birth weight infants, mortality." *The future of children*, 176–196.

Halpern-Manners, Andrew, Jonas Helgertz, John Robert Warren, and Evan Roberts. 2020. "The effects of education on mortality:

Evidence from linked US Census and administrative mortality data." *Demography*, 57(4): 1513–1541.

Harding, David J, and Martin D Munk. 2020. "The decline of intergenerational income mobility in Denmark: Returns to education, demographic change, and labor market experience." *Social Forces*, 98(4): 1436–1464.

Hjort, Jonas, Mikkel Sølvsten, and Miriam Wüst. 2017. "Universal Investment in Infants and Long-Run Health: Evidence from Denmark's 1937 Home Visiting Program." *American Economic Journal: Applied Economics*, 9(4): 78–104.

Holmlund, Helena. 2022. "How much does marital sorting contribute to intergenerational socioeconomic persistence?" *Journal of Human Resources*, 57(2): 372–399.

Jorgensen, Anne M. 2010. "Born in the USA–the history of neonatology in the United States: a century of caring, mortality." *NICU Currents*, 1(1): 8–11.

Jácome, Elisa, Ilyana Kuziemko, and Suresh Naidu. 2021. "Mobility for All: Representative Intergenerational Mobility Estimates over the 20th Century." National Bureau of Economic Research Working Paper 29289.

Kinge, Jonas Minet, Jørgen Heibø Modalsli, Simon Øverland, Håkon Kristian Gjessing, Mette Christophersen Tollånes, Ann Kristin Knudsen, Vegard Skirbekk, Bjørn Heine Strand, Siri Eldevik Håberg, and Stein Emil Vollset. 2019. "Association of household income with life expectancy and cause-specific mortality in Norway, 2005-2015." *Jama*, 321(19): 1916–1925.

Kjellsson, Gustav, Ulf-G Gerdtham, and Carl Hampus Lyttkens. 2011. "Breaking bad habits by education – smoking dynamics among Swedish women." *Health Economics*, 20(7): 876–881.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and gender inequality: Evidence from Denmark." *American Economic Journal: Applied Economics*, 11(4): 181–209.

Kleven, Henrik Jacobsen. 2014. "How can Scandinavians tax so much?" *Journal of Economic Perspectives*, 28(4): 77–98.

Kline, Patrick, and Enrico Moretti. 2014. "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority." *The Quarterly journal of economics*, 129(1): 275–331.

Kotsadam, Andreas, Jo Thori Lind, and Jørgen Modalsli. 2022. "Call the midwife. health personnel and mortality in norway 1887–1920." *Cliometrica*, 16(2): 243–276.

Kramer, Michael S. 1987. "Determinants of low birth weight: methodological assessment and meta-analysis." *Bulletin of the world health organization*, 65(5): 663.

Lambert, Paul S, Richard L Zijdeman, Marco HD Van Leeuwen, Ineke Maas, and Kenneth Prandy. 2013. "The construction of HISCAM: A stratification scale based on social interactions for historical comparative research." *Historical methods: a journal of quantitative and interdisciplinary history*, 46(2): 77–89.

Lee, Chul-In, and Gary Solon. 2009. "Trends in intergenerational income mobility." *The Review of Economics and Statistics*, 91(4): 766–772.

Lee, Kwang-sun, Babak Khoshnood, Stephen N Wall, Young-pyo Chang, Hui-Lung Hsieh, and Jaideep K Singh. 1999. "Trend in mortality from respiratory distress syndrome in the United States, 1970-1995." *The Journal of pediatrics*, 134(4): 434–440.

Leknes, Stefan, and Jørgen Modalsli. 2020. "Who benefited from industrialization? The local effects of hydropower technology adoption in Norway." *The Journal of Economic History*, 80(1): 207–245.

Lewis, Joshua, and Edson Severnini. 2020. "Short-and long-run impacts of rural electrification: evidence from the historical roll-out of the US power grid." *Journal of Development Economics*, 143: 102412.

Lien, Diana S, and William N Evans. 2005. "Estimating the impact of large cigarette tax hikes the case of maternal smoking and infant birth weight." *Journal of Human resources*, 40(2): 373–392.

Lindahl, Mikael. 2005. "Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income." *Journal of Human resources*, 40(1): 144–168.

Lleras-Muney, Adriana, Joseph Price, and Dahai Yue. 2022. "The association between educational attainment and longevity using individual-level data from the 1940 census." *Journal of Health Economics*, 84: 102649.

Lubotsky, Darren, and Martin Wittenberg. 2006. "Interpretation of regressions with multiple proxies." *The Review of Economics and Statistics*, 88(3): 549–562.

Manduca, Robert, Maximilian Hell, Adrian Adermon, Jo Blanden, Espen Bratberg, Keun Bok Lee, Stephen Machin, Martin D Munk, Martin Nybom, Yuri Ostrovsky, et al. 2019. "Trends in Absolute Income Mobility in North America and Europe." Mimeo.

Markussen, Simen, and Knut Røed. 2020. "Economic mobility under pressure." *Journal of the European Economic Association*, 18(4): 1844–1885.

Meghir, Costas, and Mårten Palme. 2005. "Educational reform, ability, and family background." *American Economic Review*, 95(1): 414–424.

Minnesota Population Center. 2020*a*. "Integrated Public Use Microdata Series, International: Version 7.3." *Norway Census 1900*, Minneapolis. `https://doi.org/10.18128/D020.V7.3`.

Minnesota Population Center. 2020*b*. "Integrated Public Use Microdata Series, International: Version 7.3." *Norway Census 1910*, Minneapolis. `https://doi.org/10.18128/D020.V7.3`.

Moehling, Carolyn, and Melissa Thomasson. 2014. "Saving Babies: The Impact of Public Health Education Programs on Infant Mortality." *Demography*, 51(2): 367–386.

NVE. 1946. "Utbygd Vannkraft i Norge." Norges Vassdrags- og Elektrisitetsvesen. https://publikasjoner.nve.no/diverse/1946/utbygdvannkraftinorge1946.pdf.

NVE. 2016. "Overview of Norway's Electricity History." Norges Vassdrags- og Elektrisitetsvesen. https://publikasjoner.nve.no/rapport/2017/rapport2017_15.pdf.

Nybom, Martin, and Jan Stuhler. 2016. "Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation, mortality." *Journal of Human Resources*, 51(1): 239–268.

Nybom, Martin, and Jan Stuhler. 2017. "Biases in Standard Measures of Intergenerational Income Dependence." *The Journal of Human Resources*, 52: 800 – 825.

OECD. 2020. "Poverty Rate Definition." *hhttps://data.oecd.org/inequality/poverty-rate.htm*, Accessed: 2020-10-16.

OECD. 2021. "LFS by Sex and Age." *https://stats.oecd.org/Index.aspx?DataSetCode=lfs_sexage_i_r*, Accessed: 2021-02-25.

Olivetti, Claudia, and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual Review of Economics*, 8(1): 405–434.

Olivetti, Claudia, and M Daniele Paserman. 2015. "In the name of the son (and the daughter): Intergenerational mobility in the United States, 1850-1940." *American Economic Review*, 105(8): 2695–2724.

Oster, Emily. 2012. "Routes of infection: exports and HIV incidence in Sub-Saharan Africa." *Journal of the European Economic Association*, 10(5): 1025–1058.

Pareliussen, Jon Kristian, Mikkel Hermansen, Christophe André, and Orsetta Causa. 2018. "Income Inequality in the Nordics from an OECD Perspective." *Nordic Economic Policy Review*, 17–57.

Pekkarinen, Tuomas, Kjell G Salvanes, and Matti Sarvimäki. 2017. "The evolution of social mobility: Norway during the twentieth century." *The Scandinavian Journal of Economics*, 119(1): 5–33.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2022. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *arXiv preprint arXiv:2201.01194*.

Roy, Andrew Donald. 1951. "Some thoughts on the distribution of earnings." *Oxford economic papers*, 3(2): 135–146.

Ruhm, Christopher J. 1998. "The economic consequences of parental leave mandates: Lessons from Europe." *The quarterly journal of economics*, 113(1): 285–317.

Sandvik, Hogne. 2000. "Distriktslegen (1836–1984)." *Tidsskrift for Den norske legeforening*.

Schneider, Eric B. 2022. "The effect of nutritional status on historical infectious disease morbidity: evidence from the London Foundling Hospital, 1892-1919." *The History of the Family*, 1–31.

Schwandt, Hannes, Janet Currie, Marlies Bär, James Banks, Paola Bertoli, Aline Bütikofer, Sarah Cattan, Beatrice Zong-Ying Chao, Claudia Costa, Libertad González, et al. 2021. "Inequality in mortality between Black and White Americans by age, place, and cause and in comparison to Europe, 1990 to 2018." *Proceedings of the National Academy of Sciences*, 118(40).

Sejersted, Francis. 2021. "The age of social democracy." In *The Age of Social Democracy*. Princeton University Press.

Selin, Håkan. 2014. "The Rise in Female Employment and the Role of Tax Incentives. An Empirical Analysis of the Swedish Individual Tax Reform of 1971." *International Tax and Public Finance*, 21(5): 894–922.

Semmingsen, Ingrid. 1960. "Norwegian emigration in the nineteenth century." *Scandinavian Economic History Review*, 8(2): 150–160.

Skansen, Jobs. 1958. *Bygdemøllene i Norge.* Mariendals Boktrykkeri AS.

Søgaard, Jakob Egholt. 2018. "Top Incomes in Scandinavia – Recent Developments and the Role of Capital Income." *Nordic Economic Policy Review 2018: Increasing Income Inequality in the Nordics*, 66–94.

Solon, Gary. 1999. "Intergenerational mobility in the labor market." In *Handbook of labor economics*. Vol. 3, 1761–1800. Elsevier.

Solon, Gary. 2002. "Cross-Country Differences in Intergenerational Earnings Mobility." *Journal of Economic Perspectives*, 16(3): 59–66.

Solon, Gary. 2004. "A Model of Intergenerational Mobility Variation Over Time and Place." *Generational Income Mobility in North America and Europe*, , ed. MilesEditor Corak, 38–47. Cambridge University Press.

Song, Xi, Catherine G Massey, Karen A Rolf, Joseph P Ferrie, Jonathan L Rothbaum, and Yu Xie. 2020. "Long-term decline in intergenerational mobility in the United States since the 1850s." *Proceedings of the National Academy of Sciences*, 117(1): 251–258.

SSB. 2020. "Metadata: Basic Statistical Units." `https://www.ssb.no/a/metadata/conceptvariable/vardok/135/en`, Accessed: 2020-10-16.

Sun, Liyang, and Sarah Abraham. 2020. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*.

Troesken, Werner. 2001. "Race, disease, and the provision of water in American cities, 1889–1921." *The Journal of Economic History*, 61(3): 750–776.

van der Loo, M.P.J. 2014. "The stringdist package for approximate string matching." *The R Journal*, 6: 111–122.

Venneslan, Christian. 2009. "Electrification and industrialisation: An assessment of the industrial breakthrough in Norway." *Scandinavian Economic History Review*, 57(2): 124–155.

Vidart, Daniela, et al. 2021. "Human capital, female employment, and electricity: Evidence from the early 20th century united states."

Vosters, Kelly. 2018. "Is the Simple Law of Mobility Really a Law? Testing Clark's Hypothesis." *The Economic Journal*, 128(612): F404–F421.

Vosters, Kelly, and Martin Nybom. 2017. "Intergenerational persistence in latent socioeconomic status: evidence from Sweden and the United States." *Journal of Labor Economics*, 35(3): 869–901.

World Inequality Database. 2020. "Pre-Tax Disposable Income Gini Index." data retrieved from World Inequality database on 2020-01-11, `https://wid.world/`.