Consequences of Rapid Structural Change - Evidence from Hydropower Expansions*

René Karadakic
Harvard University

August 24, 2023

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.

JEL Classification: I15, I14, O14
Keywords: Longevity, Structural Transformation, Public Health

*Karadakic: rkaradakic@hsph.harvard.edu, Harvard T. H. Chan School of Public Health, Department of Health Policy and Management. 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. Funding by the Research Council of Norway through project No 275800 and through its Centres of Excellence Scheme, FAIR project No 262675, is greatfully acknowledged.

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

[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.

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]

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. 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

[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 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.

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 (**?**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 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 simul-

[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.

taneously 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 2 provides an overview of the data and background. In section 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 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 5 concludes the paper.

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 C, I provide more detailed information about sources and how to access the data used in this project.

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

[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.

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 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 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 electricity was used for indus-

[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.

trial 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.

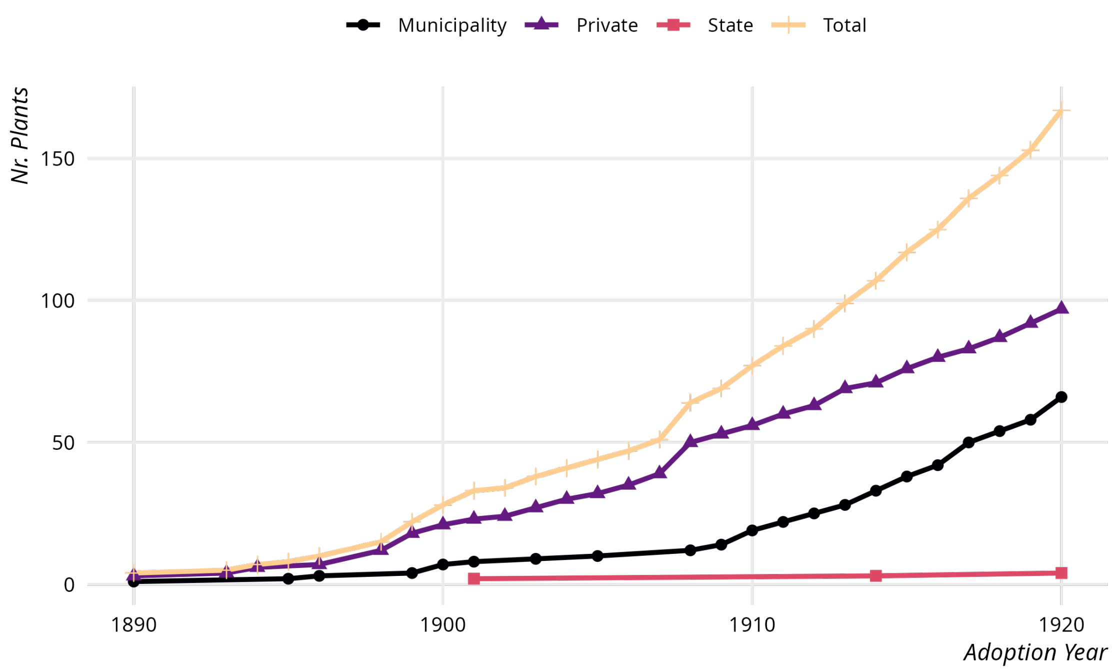


Figure 1: Cumulative Number of Hydrowpower Plants by Ownership.

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.

The expansion of hydropower started relatively slowly in the late 19th century and then really gained pace from the beginning of the 1900s. Figure 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.

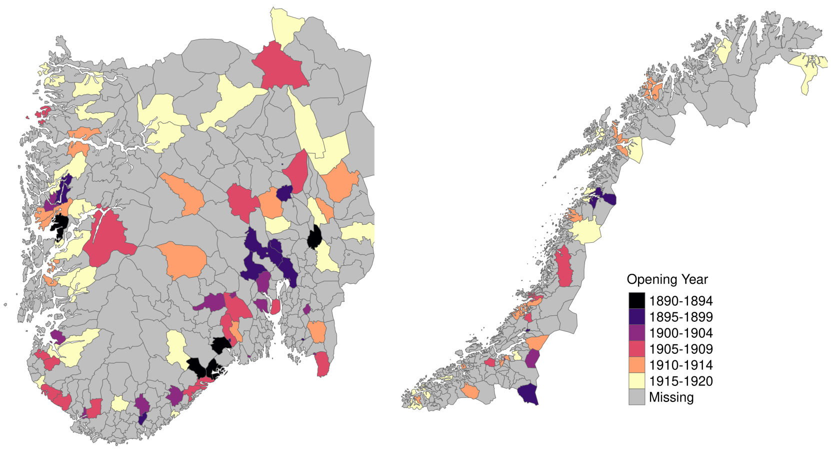


Figure 2: Geographical Distribution of Hydropower Plants by Establishment Year.

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.

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

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

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 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 historic 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.

[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.

The exact procedure applied in this paper is described in Appendix 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]

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 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

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

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.

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

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

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.

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

[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.

causes disease and how to prevent 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 significantly 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

[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.

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.

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.

3.1 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 in Equation 1. Using this relatively new estimator addresses issues such as the

[16] Appendix Figure B1 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.

[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.

negative weighting of positive treatment effects, and should provide sensible estimates of dynamic causal effects under heterogeneity across cohorts (Roth et al., 2022).

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

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.

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 2 represents the individual-level equivalent to Equation 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.

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

[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.

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 1 and 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 instrumental 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 4.

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

3.2 Two-Way Fixed Effects

There are two main disadvantages of the event-study approach presented in Equation 1 and 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.

$$Y_{mt} = \alpha + \beta_1 \times \text{HPP}_{mt} + \mu_m + \tau_t + \varepsilon_{mt} \tag{3}$$

Equation 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 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.

$$Y_{imt} = \alpha + \beta_1 \times \text{HPP}_{imt} + \mu_m + \tau_t + \varepsilon_{imt} \tag{4}$$

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 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 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. For most of the analysis where I am able to provide event-study estimates using Equation 1 or 2, I will not present results using the instrumental variable approach.

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

[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.

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.

4.1 Long-term Implications of Hydropower Adoption

Figure 3 presents the main result of this paper obtained from estimating Equation 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 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 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.

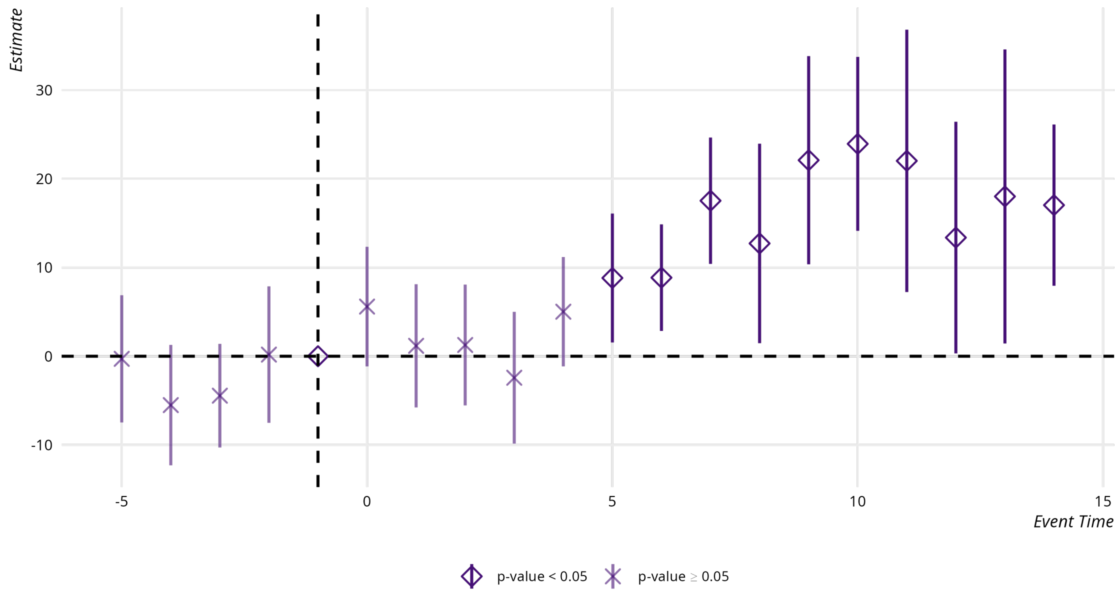


Figure 3: Impact of Hydropower Openings on Age of Death of Men.

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 2. Standard errors used for the calculation of the 95-% confidence intervals are clustered at the municipality level.

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 1, I provide some evidence of this mechanism by estimating Equation 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, 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]

Table 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 4. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$.

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 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

[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.

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, 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. 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 B2 and suggest that the main effect is not simply driven by migration but also by individuals from the local stayer population.[22]

[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

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 simple two-way fixed effect setup from Equation 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

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.

[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.

establishment and the probability of a successful match. This effect entirely disappears, however, once I control for birth year and birth municipality fixed effects, suggesting that Equation 3 should not suffer from bias due to differential matching of individuals in hydropower and non-hydropower areas.

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 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 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

[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.

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]

Table 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. Intrumental variable is hydropower potential interacted with year dummies. P-value thresholds $*** = p < 0.01$, $** = p < 0.05$ and $* = p < 0.1$

[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 B3, I provide estimates from Bacon decompositions following Goodman-Bacon (2021), which show that the effects presented in Table 2 are mainly driven by comparisons between untreated and treated municipalities. In addition, the decomposition shows that no negative weighting occurs.

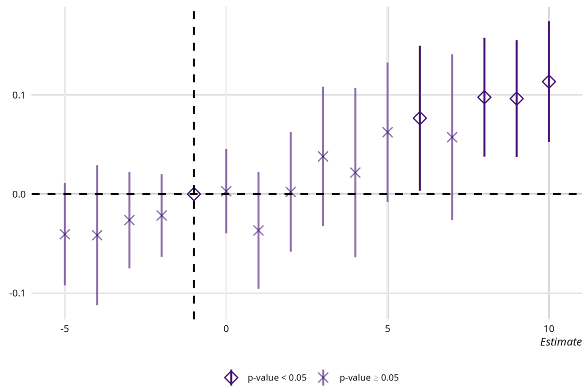
Since the labour force-to-population ratio remains constant after hydropower adoption, I assume that the main reason for the increase in the labour force is permanent family reallocation rather than simple worker migration. Electrification has been shown to impact female participation in the labour market 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 B4, 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 individuals under a certain income threshold (Berger and Vagle, 2017).[29] In Figure 4, I present estimates of Equation 1 for three different outcomes. Panel 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.

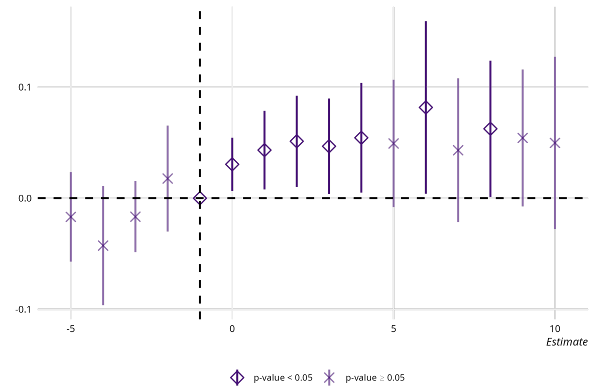
[27] Individuals are defined as employed if they have a recorded occupation in the census. According to source variable documentation, occupations should have been recorded for women if they were working in any industry.

[28] In Figure B5, 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.

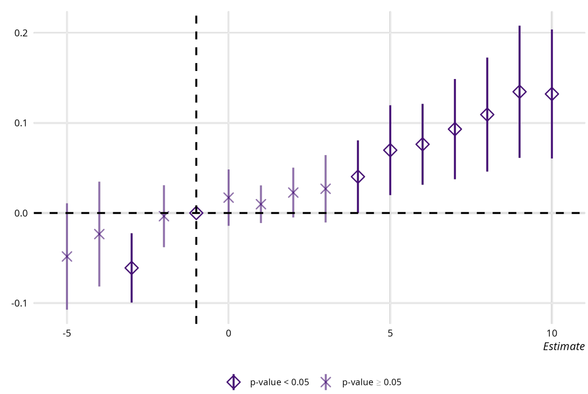
[29] Gerdrup (1998) argues that, on average, about 50 to 60 per cent of the total income within a municipality was taxable and would therefore be included in these data.



(a) Income per Taxpayer



(b) Wealth per Taxpayer



(c) No. of Taxpayers

Figure 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 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.

Panel 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 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, hydropower municipalities see a larger influx of immigrants compared to areas without hydropower adoption. Figure B4 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 B6. This increased immigration towards hydropower areas also attracts positively selected individuals in terms of occupation-based socioeconomic status.

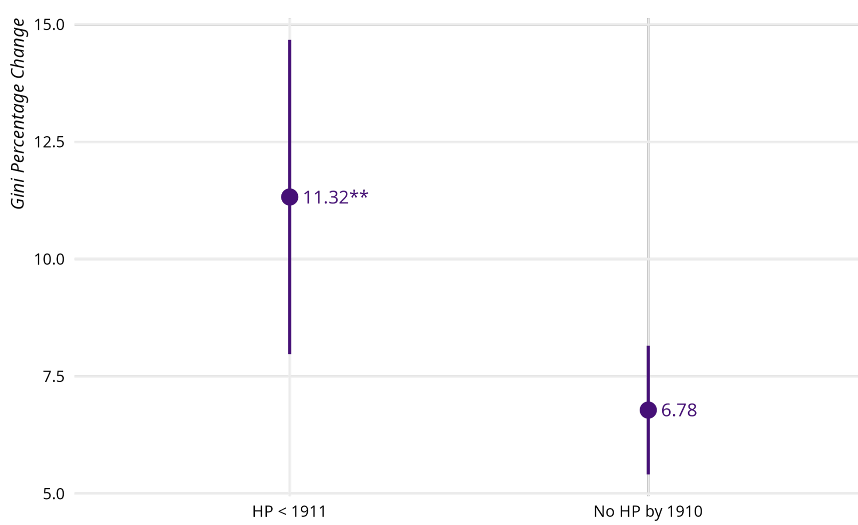


Figure 5: Change in Socioeconomic Inequality by Hydropower Status.
Note: The figure shows the change in HISCAM value-based Gini coefficients from 1900 and 1910 in municipalities by hydropower status. The data are from full population census data in 1900 and 1910 and include all men aged 15 and older. 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 $* = p < 0.1$.

[30] Results in Figure 4 are corroborated by Table A10 providing estimates of Equation 3 for the income, wealth and taxpayer variables.

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 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 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, 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.

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 6, I present event-study estimates of Equation 1 for the public health outcomes of Norwegian health districts, taken from historical documents.

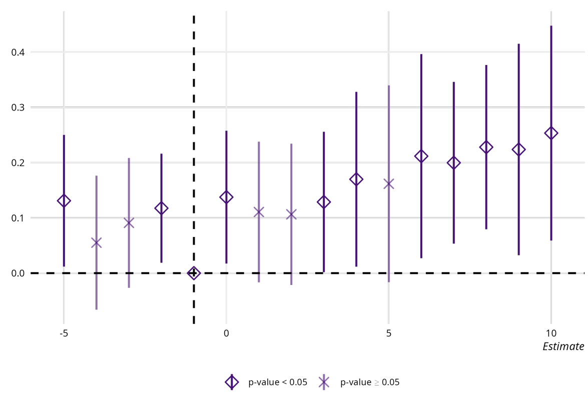
In Panel 6a, I present results concerning the impact of hydropower establishment on the number of cases of four common infectious diseases.[32] Hydropower establishment, with a slight lag of up to four years, led to increased spread of disease. Hydropower plants by themselves do not have a direct influence 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 infection due to larger exposure to contaminated water and faeces is also supported by the increase in cases of diarrhoea as

[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.'

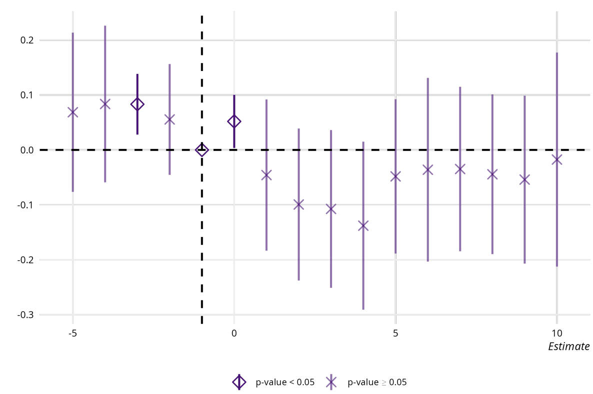
[32]The four diseases included in the analysis are diarrhoea (including cholera), pneumonia, diphtheria and typhoid fever. They were among the most common infectious diseases, and effective treatment and prevention were not available at the time. In Table A8, I provide a brief overview of some key characteristics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

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

presented in Figure B7a, 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 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.



(a) Cases of Common Infectious Diseases



(b) Health Professionals

Figure 6: Public Health Development in Response to Hydropower Establishment.

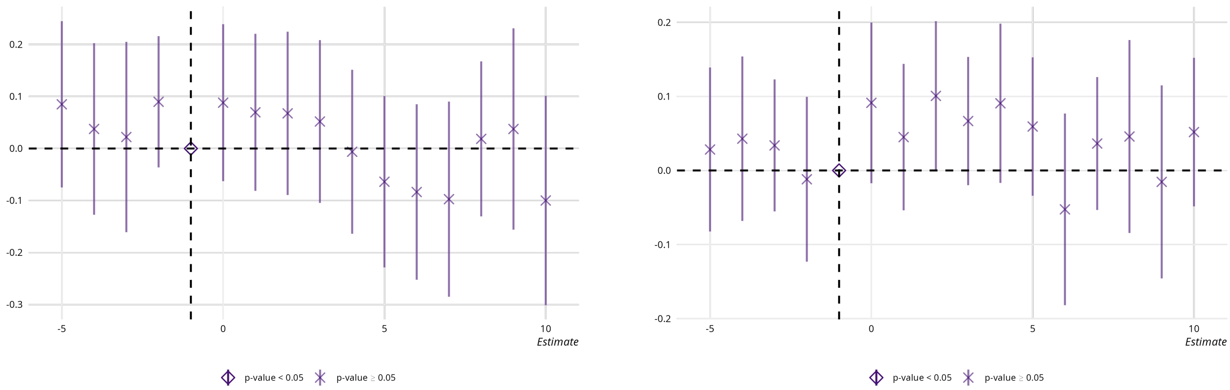
Note: The figure shows event-study estimates of the impact of hydropower establishment on the total number of cases of four common infectious diseases (diphtheria, typhoid fever, pneumonia and diarrhoea/cholera) and the number of health professionals (doctors, midwives, vaccinators, pharmacists and dentists) 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 1. Standard errors are clustered at the municipality level. Shapes and transparency indicate significance at 5% level and error bars represent 95% confidence intervals.

Panel b of Figure 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 al-

[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.

ways perceived 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 6b. This general pattern is also corroborated by estimates of Equation 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 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.



(a) Common Infectious Disease Mortality

(b) Infant Mortality

Figure 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.

There are several reasons why mortality, both from infectious disease and infant mortality, does not significantly respond to hydropower-induced structural transformation, 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 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, 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.

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.

References

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.

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, 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.

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.

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.

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.

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

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.

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.

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.

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.

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.

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.

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

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

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*.

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*.

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

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.

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.

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.

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 rollout of the US power grid." *Journal of Development Economics*, 143: 102412.

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.

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.

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.

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

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.

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.

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

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.

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."

Appendix

A 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. in 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 (1)	IV+FE (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

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

Note: Case fatality rates are own calculations using health average annual case fatility 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 C.

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

	Ln(Income/TP) (1)	Ln(Wealth/TP) (2)	Ln(TP) (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 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) (1)	Ln(HP/Pop) (2)	Ln(Deaths/Pop) (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 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 $* = p < 0.1$.

B Figures

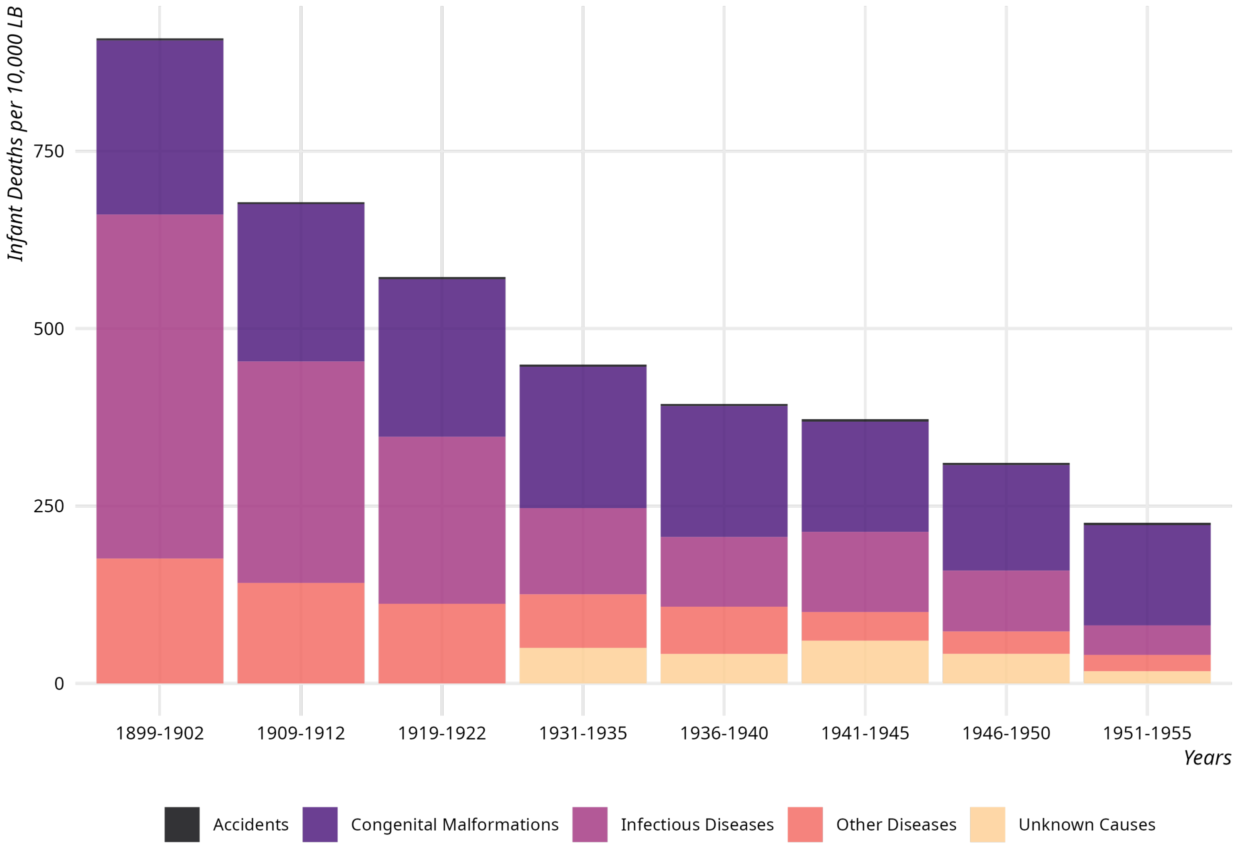


Figure B1: Infant Mortality by Cause of Death 1899 - 1955.

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).

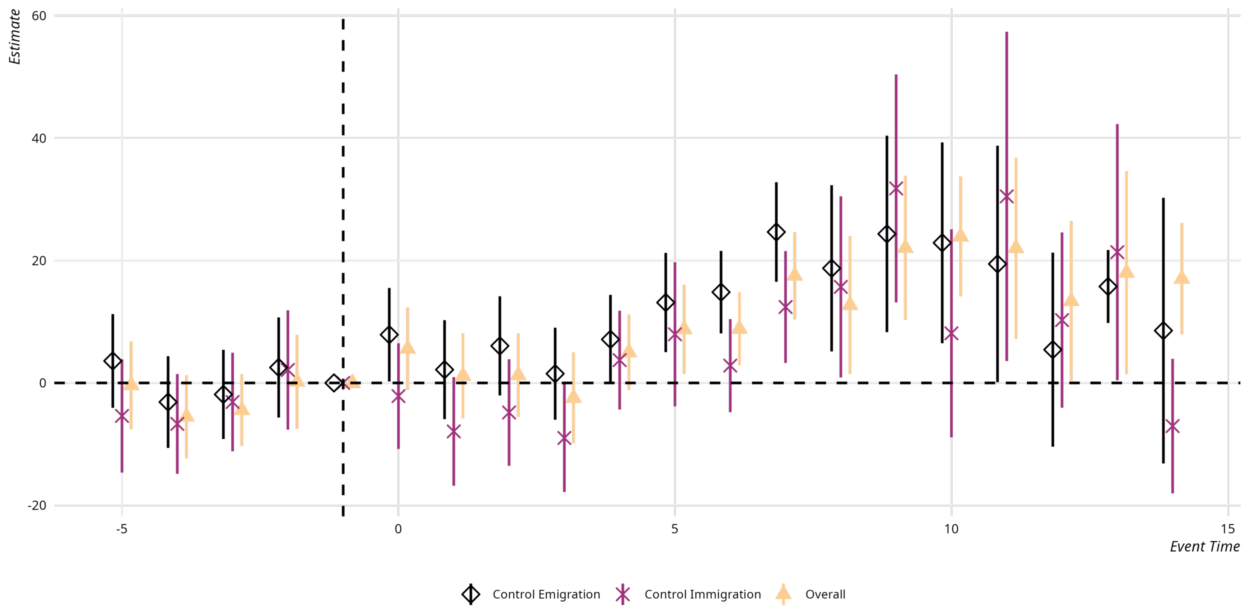


Figure B2: Longevity Effects and Robustness to Immigration and Emigration.

Note: The figure shows estimates of Equation 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.

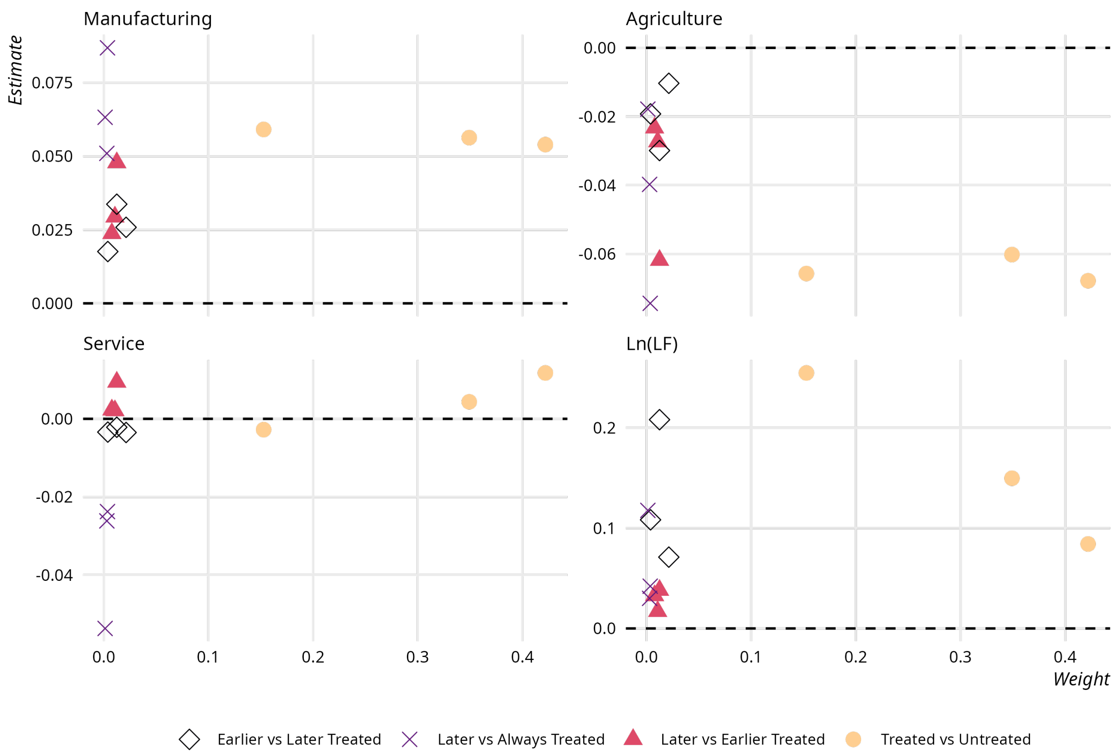


Figure B3: Bacon Decomposition for TWFE Models.

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.

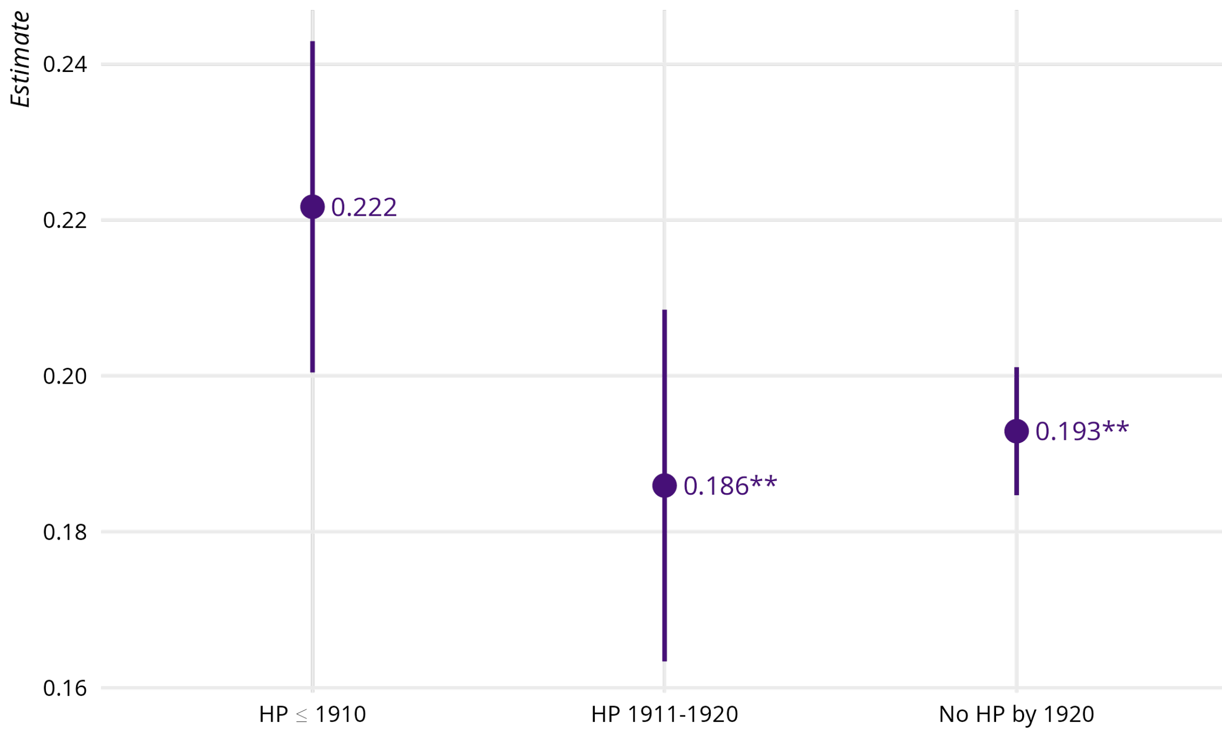


Figure B4: Mover Share by Hydropower Status of Municipality.

Note: The figure depicts the share of movers in municipalities in the 1910 census by hydropower status. Estimates were obtained froma 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 $* = p < 0.1$.

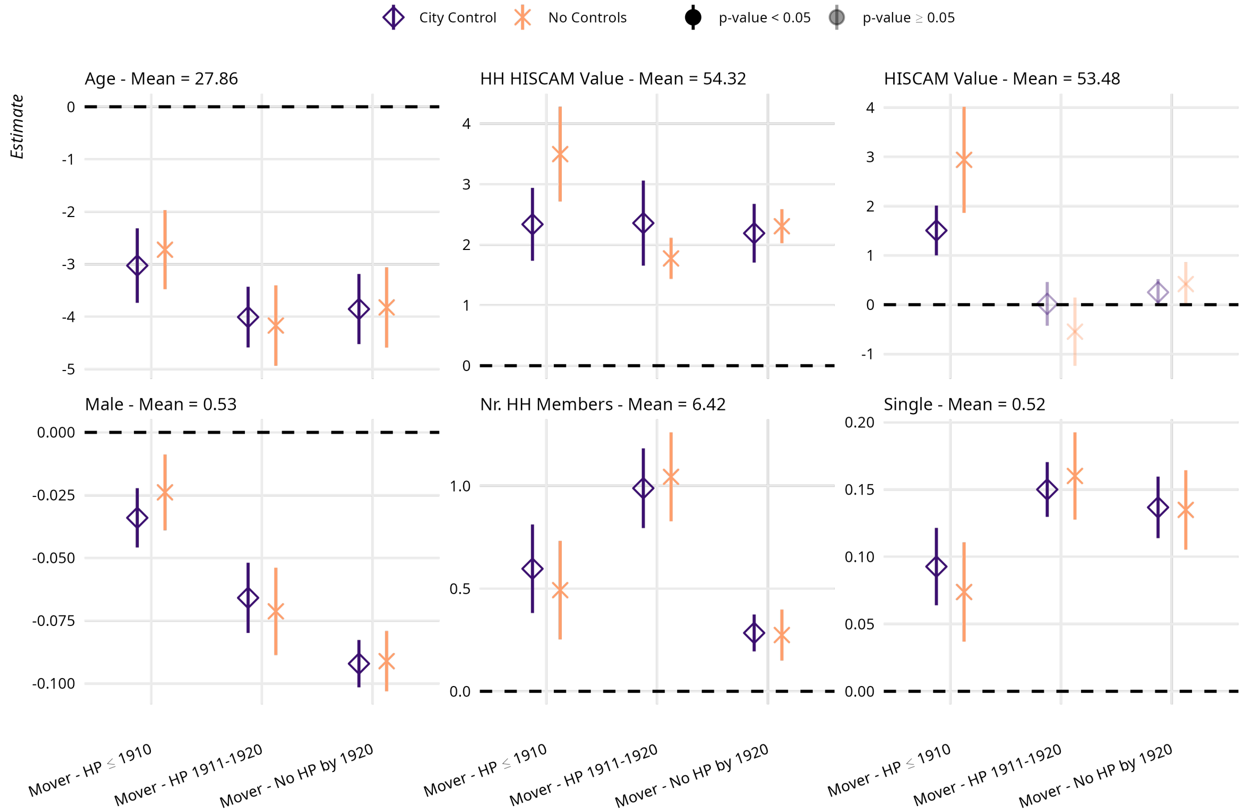


Figure B5: Characteristics of Movers in Comparison to Stayers.

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.

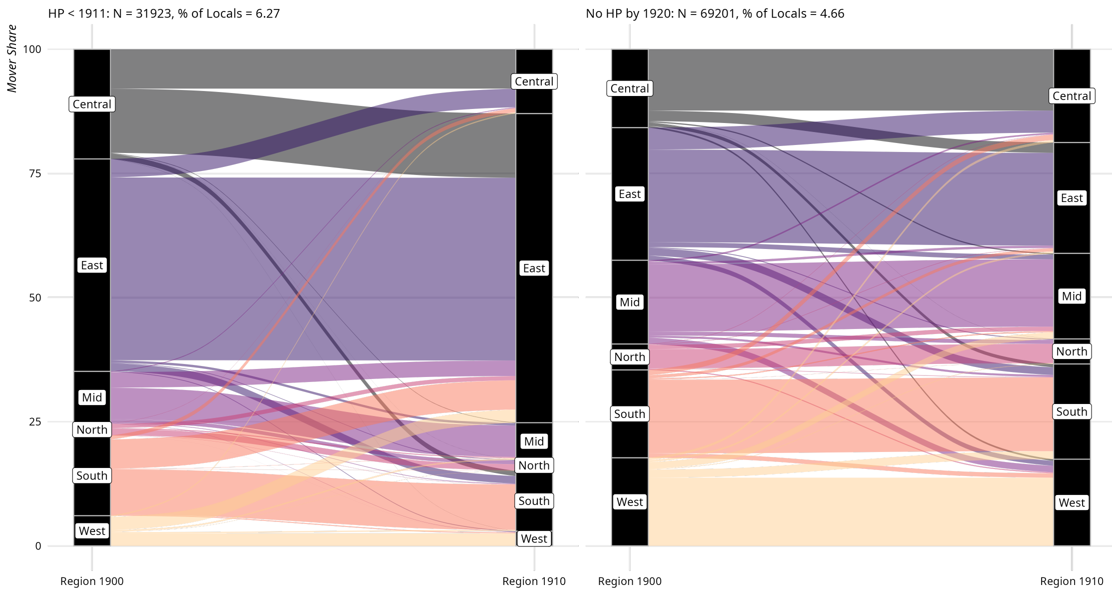
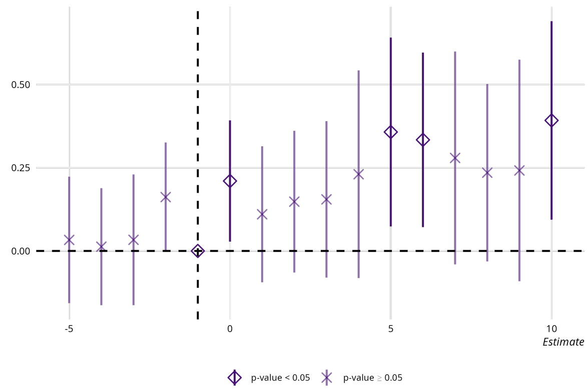
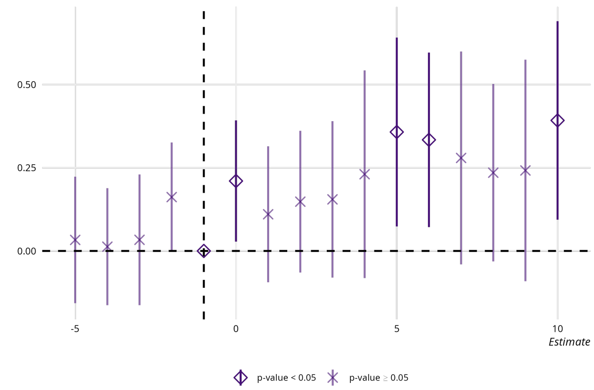


Figure B6: Mover Flows by Hydropower Status.

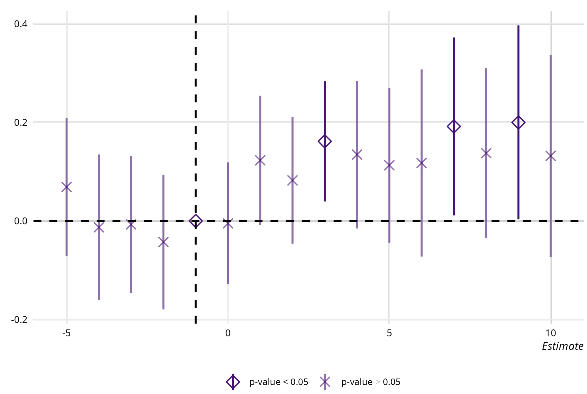
Note: The figure shows changes regional flows of migrants from the linked 1900 and 1910 census data. The graph shows the relative flows of movers. % of Locals indicates the overall number of movers in this category relative to the receiving municipalities popualtion in 1900. The left panel shows te mover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



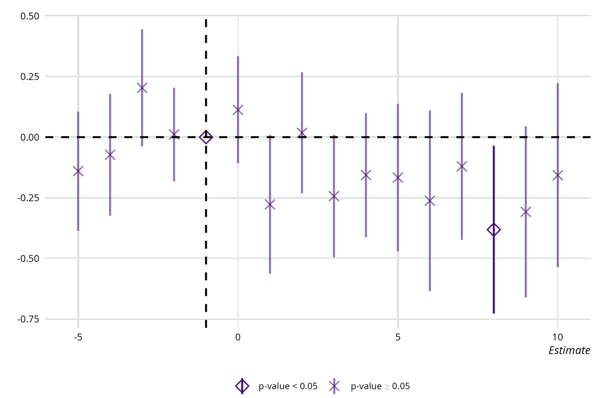
(a) Diarrhoea



(b) Diphtheria



(c) Pneumonia



(d) Typhoid Fever

Figure B7: 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 1. Standard errors are clustered at the municipality level. Shapes and transparency indicates significance at 5-% level and error bars represent 95-% confidence intervals.

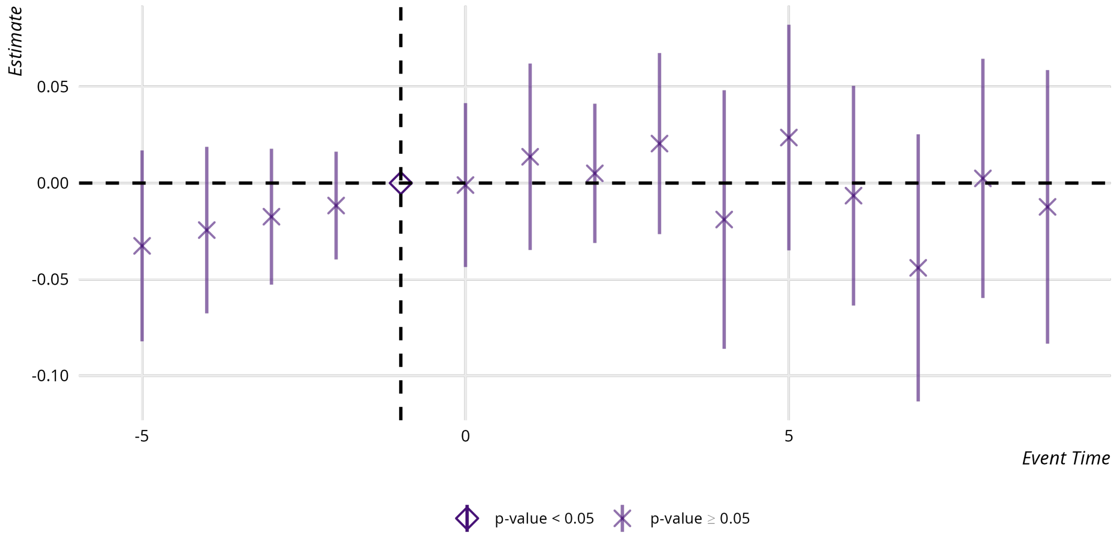


Figure B8: Impact of Hydropower Openings on Fertility.

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.

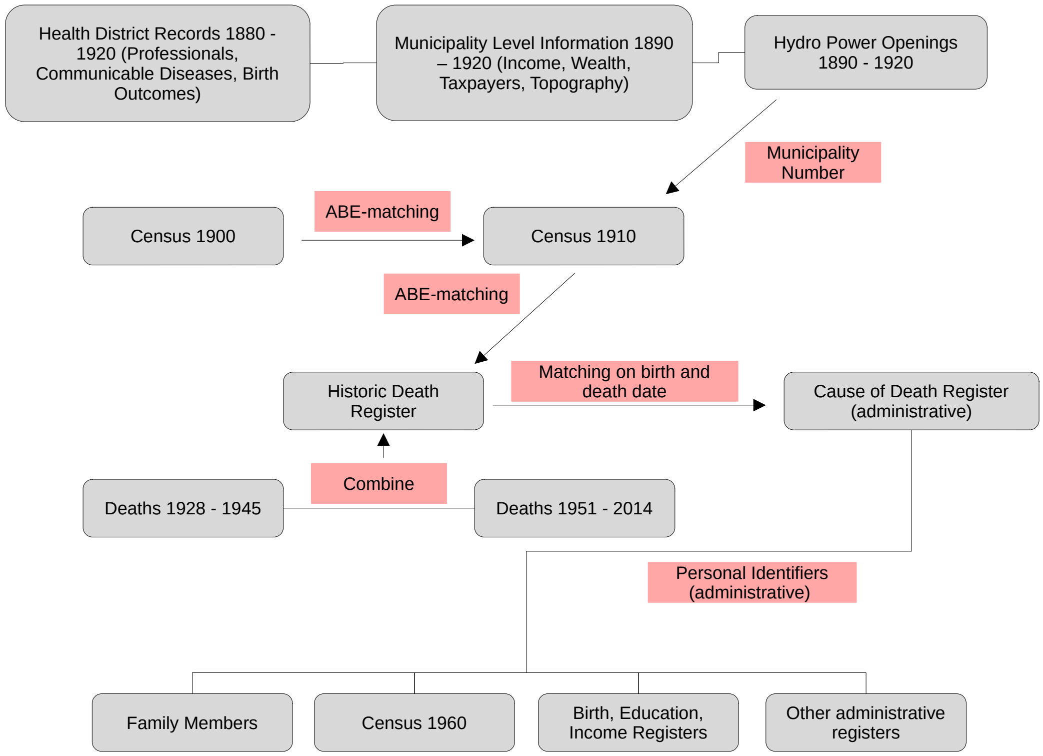


Figure B9: Connection of Historical and Modern Administrative Data.

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

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 intermunicipal 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).

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.

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.

In Table D12, 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).

Table D12: 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.

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 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

[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).

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.

Table D13: 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.