Norwegian School of Economics
Bergen, Spring 2016

An Empirical Analysis of Health Effects from Electrification:

Evidence from Norway 1900-1921

Ellen Balke Hveem, Lotte Leming Rognsås

Supervisor: Aline Bütikofer

Master thesis, MSc in Economic and Business Administration, Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

Working on this thesis for the past months has been very rewarding. We feel fortunate to be able to study a topic of our own choice, which also is meaningful in a societal context.

We would like to thank our supervisor Aline Bütikofer, for giving us advice on the choice of a feasible topic and for valuable input and feedback on our work. Throughout the process, her commitment to and knowledge of the topics studied has been highly motivating.

We also wish to thank Jonas Cederlöf and Kasper Kragh-Sørensen for helpful insights and proofreading.

Bergen, June 2016

Ellen Balke Hveem

Lotte Leming Rognsås

Abstract

Does access to electricity from hydroelectricity plants affect the health of a population? The aim of this thesis is to analyze the health effects from hydroelectricity at a time when electrification was more or less synonymous with gaining access to electric light from simple light bulbs. Exploiting data on health and the roll out of hydroelectricity plants in Norway in the period from 1900 to 1921, we assess the relationship between electrification and three health variables: infant mortality, tuberculosis and diarrhea.

Utilizing a staggered differences-in-differences approach, we find a positive effect of electrification on health. This implies that access to even simple electric appliances has positive consequences for health standards. The findings are interesting in two regards. Firstly, they contribute to the existing literature on the modernization and industrialization of Norway, as well as to the previous literature on health effects from electrification. Secondly, they are interesting in a contemporary policy perspective, as about 15 percent of the world population lack access to electricity today.

Our findings are consistent with the previous literature, although the estimated effect is smaller in magnitude. Partly, this can be explained by the simplicity of the appliances resulting from electrification at the time. Also, our estimated effects are intention-to-treat effects averaged across the entire population, whereas the actual effect in treated households is likely to be larger in magnitude.

Acronyms

DiD	Differences-in-Differences
GDP	Gross Domestic Product
ITT	Intention-to-treat
NSD	Norwegian Centre for Research Data
NVE	Norwegian Water Resources and Energy Directorate
OLS	Ordinary Least Squares
OVB	Omitted Variable Bias
SSB	Statistics Norway
TOT	Treatment on the treated
UNICEF	United Nations Children's Fund
WHO	World Health Organization

Table of contents

1 Introduction

1.1 Motivation and Purpose

In the early 1900s, Norway had low infant mortality rates as compared to other countries, despite having considerably lower GDP levels than other western nations. GDP levels are often used as an indicator of development, and are found to be correlated with health outcomes (see for instance Swift, 2011; Lago-Peñas et al., 2013; Bhargava et al., 2001). For instance, while Norway in 1920 had an income per person of $5,483[1], the US had an income per person of $9,181[2] (Gapminder, 2016). However, US infant mortality rates were 85.8[3] (Linder and Grove, 1947), while the comparable number for Norway was 53.6[4] (SSB, 1924). Norway was one of the earliest adaptors of hydroelectricity plants and household electrification, and this paper aims to investigate whether the early electrification could be a contributing factor to the relatively low infant mortality rate.

Infant mortality is traditionally among the main indicators for welfare in a society (Pedersen, 2003). Infant mortality rates respond to changes in standard of living, and the electrification of households could be such a change. In addition to infant mortality rates, we also analyze the effect on lung tuberculosis and diarrhea. These infectious diseases are chosen as variables because they were among the most fatal diseases in Norway, and children and infants were especially vulnerable to them (Pedersen, 2003). Together, the three health outcomes provide a good insight into the overall health status of infants.

The electrification of Norway gained momentum in the first decades of the 1900s through the rollout of hydroelectricity plants. The period was also characterized by a considerable decline in diseases and infant mortality. Historical sources attribute these changes to, among others, improved nutrition, sanitation and economic conditions (see for example Liestøl et al., 2007). In this context it is interesting to analyze whether the introduction of electricity can be a contributing factor to the positive trend in the health status of the population in the early 1900s. To our knowledge, this is the first paper of its sort where the health effects of access to hydroelectricity in Norway are analyzed. The results can therefore contribute to the existing literature on the modernization and industrialization of Norway. Also, to our knowledge, this

[1,2] GDP/Capital, PPP$ inflation adjusted

[3,4] Dead before age one per 1000 born

is the first paper utilizing European data to pinpoint these effects. Therefore, we expect to find results not directly comparable to previous studies conducted in other parts of the world.

Based on findings in previous literature, we link the effect on health from receiving a hydroelectricity plant through three main channels: the effect on outdoor pollution when there is a switch from one energy source to another (see for example Clay, Lewis, and Severnini, 2015; Severnini, 2014); the effect on the indoor environment when kerosene and wood is replaced by electric light bulbs (see for example Bruce et al., 2000; Apple et al., 2010); and the effect from alterations to the outdoor environment when a hydroelectricity plant is constructed (see European Commission, 1995). The first two effects are expected to be positive on health, while the last effect can be both positive and negative, meaning the overall effect from receiving a hydroelectricity plant is ambiguous.

Our estimation strategy is a staggered Differences-in-Differences (DiD) approach, utilizing the rollout of electricity in Norway from 1900-1921. We contrast the health status of municipalities that received a hydroelectricity plant at any given time with the health status of municipalities that had not yet received a plant. Our data set is constructed by linking data on hydroelectricity plants in Norwegian municipalities from the Norwegian Water Resources and Energy Directorate (NVE) with data on infant mortality rates, tuberculosis cases and diarrhea cases in Norwegian medical districts from Statistics Norway (SSB). At the time, electrification of households was more or less synonymous with receiving electric light, as technical appliances for household use were still unsophisticated. Thus, the resulting estimates will show the effect on health from access to electric light from a hydroelectricity plant. By controlling for time and year specific effects, as well as several other covariates, we separate the effect of hydroelectricity from other trends in health outcomes.

The link between hydroelectricity and health has implications for current policy, as more than 15 percent of the world population lack access to electricity (World Bank, 2012), and an estimated 500 million households still use kerosene, wood and other fuels for lighting (Lam et al., 2012 I). The World Health Organization (WHO) reports that 4.5 million infant deaths occurred in 2015 (WHO, 2016 I). Tuberculosis and diarrhea are still among the most serious public health problems globally (WHO, 2016 I and II), and especially so in developing nations. In 2014, 9.6 million people fell sick from tuberculosis (WHO, 2016 II), and diarrhea kills 760,000 children under five each year (WHO, 2016 III). These numbers further motivate why finding a significant effect of access to electric light on health is still relevant in today's context, and for policy making.

We find that access to electricity had positive effects on health. Specifically, it led to a statistically significant reduction in infant mortality of 4.74 percent and in tuberculosis rates of 8.84 percent. the qualitative results are robust, and the estimated coefficients vary only somewhat. The effect on infant mortality is persistent in the four years following the opening of a hydroelectricity plant. For diarrhea, we find no significant reduction in the year of the intervention; however, there seem to be significant effects in the years following the intervention.

1.2 Research Question

Based on the previous subsection, this thesis aims to investigate the following research question:

> *Did access to electricity from hydroelectricity plants affect health outcomes in Norway from 1900 to 1921?*

The remainder of the paper continues as follows. We present the historical background on the expansion of electricity and health conditions in Norway in Section 2. In Section 3 we provide a review of the previous literature. Section 4 describe the mechanisms through which hydroelectricity can affect health outcomes. We describe our data in Section 5, followed by an overview of our empirical strategy in Section 6. We present our results in Section 7, which also includes robustness checks. In Section 8 we discuss our findings and possible shortcomings of our analysis. Finally, Section 9 concludes.

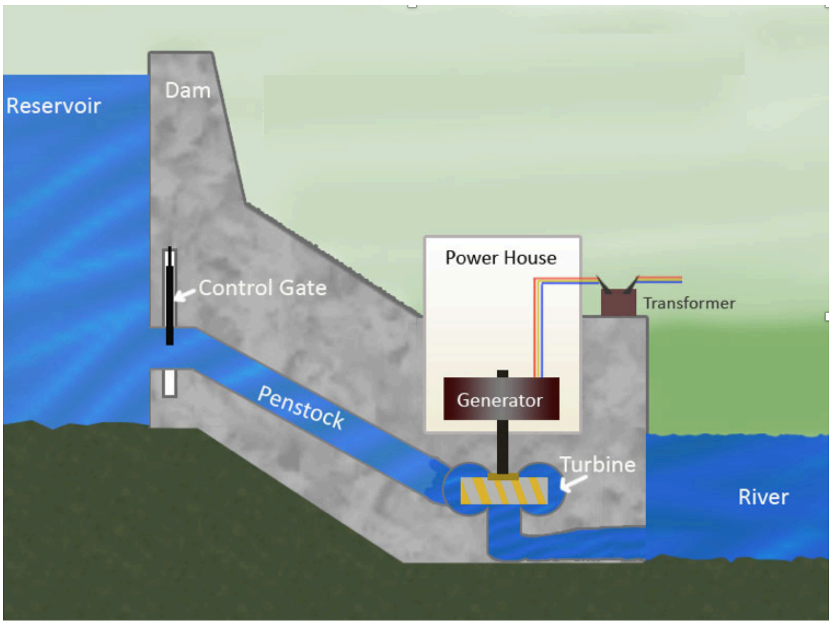
2 Historical Background

The rollout of electricity in Norway gained considerable momentum between 1900 and 1921. In the following subsections, we will look at the historic development of hydroelectricity, and the development of health, in Norway in the period of interest.

2.1 Hydroelectricity in Norway

Norway has excellent natural conditions for hydroelectricity generation, and was at the forefront of making use of the technology in the late 1800s and the early 1900s. The mountainous landscape, along with heavy precipitation, makes Norway well suited for hydroelectricity generation. The technology utilizes the energy from falling water to generate electricity (Rosvold, 2010), as illustrated in Figure 1.

Figure 1: Illustration of hydroelectricity plants



Notes: The figure illustrates the basics of how a hydroelectricity plant function. Water is stored in the reservoir, flows through the penstock and spins the turbine, causing electricity to be generated (Daware, 2016).

The new technology was rapidly adopted across the country. The first hydroelectricity plant opened in 1882, and in 1885 the first plant supplying electricity to private consumers opened. Hammerfest became the first town with street lamps supplied from hydroelectricity in 1890.

In 1901 there were a total of 614 power plants, out of which 215 were hydroelectricity plants (Hveding, 1992). The remaining were operated by steam, gas or petroleum engines. At the time, the production capacity of each hydroelectricity plant was low, and most supplied electric lights to one building only, usually factories (Sandberg, 1951).

The number of hydroelectricity plants rapidly increased during the two first decades of the 20^{th} century. Because water resources are spread around the country, suitable waterways could be found near virtually every town or village (Hveding, 1992). However, the roll out of hydroelectricity, like most other new technology, happened at spatial and temporal variation. The 1920 census showed that 64 percent of Norway's entire population lived in houses with installed electricity - the highest percentage in the world at the time (Hveding, 1992). By 1921 all cities had been electrified, and the number of rural power stations had reached 200.

As the number of hydroelectricity plants grew, the number of households connected to the electricity grid also increased. Simultaneously, production capacity of hydroelectricity stations increased rapidly, meeting the growing demand for electricity from households and industry. Access to cheap electricity from hydroelectricity gave energy intensive industry a comparative advantage in Norway, and in the period until 1920 an extensive industrial expansion occurred. Sites with abundant water resources became the new industrial centers, and small villages like Rjukan and Sauda quickly expanded (SSB, 2001).

Increasing demand for electricity, also at sites far from water resources, led to a gradual expansion of transmission lines and cables (Øberg and Bråten, 2014). At the beginning of the century, transmission of power was difficult, even over relatively short distances, because a large portion of the power was lost during transfer. This made transmission ineffective and expensive (Statnett.no, 2013). Therefore, most hydroelectricity plants supplied electricity only to the nearby area. In 1922, the first transmission line capable of transferring electricity over a great distance opened when Oslo was connected to the power station in Rjukan (Øberg and Bråten, 2014).

In the beginning of our period, waterfalls were often bought and developed by foreign investors. However, the so-called “panic laws” of 1906 increased the role of the Norwegian authorities in trade of development rights. The municipality authorities gradually gained a key ownership role in the local water resources. The municipalities made it their priority to supply electricity to private households, small companies and agriculture, rather than heavy industry (Faugli, 2012). The concession laws of 1917 established the municipalities' legal obligation to produce electricity to households (NVE, 2013).

The exact time to plan and build a hydroelectricity plant varied according to the size and complexity of the plant, as well as local resources (Vogt, 1971). In the four largest cities, the construction time was between two and three years from intention to opening (See Vasskrafta, 2016; Rosvold, 2012; NVE, 2015; and Trondheim Energiverk, 2001). In addition, preparing a construction proposal for city council takes time, so the real time from intention to opening was probably longer. It seems reasonable to assume that these were among the most sizable plants in the country, and that for smaller plants, the construction and planning time was no longer than two to three years.

In the early 1900s, general use of electricity was more or less synonymous with switching from kerosene lamps, candles and other polluting light sources to cleaner electric light bulbs (Øberg and Bråten, 2014). This was particularly true for private households, where electricity mainly was used for lighting until 1920 (Statisics Norway, 2001). The number of incandescent lamps increased rapidly, in line with the increasing number of hydroelectricity stations and the possibility to connect to the electricity grid. While the number of electric lamps was 805,000 in 1911, it had increased to more than 5 million by 1923 (Øberg and Bråten, 2014). The use of modern household equipment gained momentum only after 1920 (SSB, 2001).

2.2 Health in Norway

In the beginning of the 1900s Norway was one of the poorest countries in Europe (Eika, 2008). Poor nutrition and inadequate hygiene led to high infection rates and low life expectancy (Stene-Larsen, 2005). However, in the period leading up to 1920 there was rapid development in economic as well as social conditions, and living standard increased steadily.

At the end of the 1800s, the Norwegian healthcare system was quite well established and modern for the time (Larsen, 2003). By 1900, Norway was divided into 158 medical districts, each normally employing one to three doctors, and cities often had more. General health insurance that would cover the costs of medical care during and after pregnancy were introduced in 1911 (Pedersen, 2003). In 1914, the first mother and child health care center opened. The centers targeted poor families, and the main goal was to provide free medical check-ups for infants (Bütikofer et al., 2015). Furthermore, public health saw a big lift in the 1900s. The early years are referred to as the "hygiene period", where light, fresh air and cleanliness was considered important contributors to improved public health (Nordhagen et al., 2014).

In the beginning of the 20th century, tuberculosis was among the most common causes of death, and one in five deaths were caused by the disease (Kjeldstadli, 2015). However, between 1900 and 1920, the mortality from tuberculosis fell by about 30 percent (Backer, 1961). People of higher social standard had lower mortality from tuberculosis, even if the occurrence in number of cases was equal, and know as "class-less". As tuberculosis affected such a large part of the population, disease trends were closely monitored and reported (Liestøl et al., 2007). The second most common cause of death was diarrhea, and around the turn of the century, the disease caused every six to seven deaths (Kjeldstadli, 2015).

The early 1900s were characterized by large geographical variations in infant mortality rates between urban and rural areas, where the latter generally was better off (Pedersen, 2003). A large number of infants died from infectious diseases such as whooping cough, diphtheria, scarlet fever and tuberculosis, pneumonia and diarrhea. As house and working conditions improved, the standard of living increased in general, but the pace was higher in cities than in rural areas. Better health care and more doctors per capita was also an important contributing factor to increased standard of living. However, there were few new discoveries in medicine during the first two decades of the 1900s, implying that the observed decline in infant mortality could be attributed to other factors. For example, knowledge on how tuberculosis infected were known before 1900 (Liestøl et al., 2007), the only mandatory vaccine was against smallpox (Nøkleby and Feiring, 2006), and antibiotics were invented only in the late 1920s (Trueman, 2016).

3 Literature Review

The literature on the relationship between electricity and health has increased substantially after the turn of the millennium, and especially evolved after researchers convincingly documented the relationship between airborne particles and mortality in the 1990s (Clay, Lewis and Severnini, 2015). Subsection 3.1 review studies that analyze the effect on health from installing electricity or a change of energy source used in electricity generation. Subsection 3.2 provides an overview of studies that specifically analyze the effect on health when changing from a polluting to a non-polluting light source indoors. In Subsection 3.3, we make some concluding remarks and relate the previous literature to our thesis.

3.1 Health Effects from Switching Energy Source

Our paper is closest to Lewis (2014). He finds a negative causal relationship suggesting that household electricity can explain 25 – 30 percent of the decline in white infants' mortality in the US in the years from 1930 to 1960. However, in the period of analysis, access to electricity implied installation of modern household equipment like washing machines, vacuum cleaners and dishwashers. Hence, much of the effect is attributed to time saving and the increased hygiene conditions resulting from utilization of modern household equipment (also other papers find a positive effect on infant health from time spent on child care, see for example Miller and Urdinola, 2010).

Furthermore, a large number of studies analyze the health effects of changing pollution levels associated with changing energy sources on health. Severnini (2014) examines the health effects of increased electricity generation from coal-fired power plants, following the shut-down of nuclear power plants in the Tennessee Valley in the 1980´s. He shows that an increase in electricity from coal-fired power plants led to higher levels of air polluting particles in counties where the power plants were located. The estimated effect of increased air pollution in the most affected counties is a reduction of 10 grams in birth weight, and in gestational length by 0.39 days, when exposure to air pollution increases by 1μg/m3 during pregnancy.

Recent work by Clay, Lewis and Severnini (2015) analyze the health effects of coal-fired power generation by linking plant-level coal consumption with county-level air quality measures and infant mortality rates. They find a positive relationship between coal consumption in electricity generation and higher levels of polluting particles. They also associate increasing air pollution with increased infant mortality rates. The total rise in emissions are estimated to lead to an additional 12,270 infant deaths in the US in the period between 1938 and 1962. Further, a one percent increase in air pollution is associated with a 0.15 to 0.29 percent increase in infant mortality.

3.2 Effect on the Indoor Environment from Switching Light Source

Several articles link the use of kerosene lamps with indoor air pollution and the possibility of general health problems. The most important sources to harmful indoor air pollution from combustion are particles, carbon monoxide, nitrous oxides, sculpture oxides and formaldehyde, all of which can be emitted from kerosene lamps (Bruce et al., 2000). Lam et al. (2012 I) points out that hazards from kerosene, including poisonings, fires and explosions, are well-documented. Lam et al. (2012 II) finds that hazardous particles are a common result of burning kerosene. A study conducted by Savitha et al. (2007) points out that the use of kerosene lamps are among the most significant indoors environmental risk factors in India.

Apple et al. (2010) demonstrates that using a lamp with bad combustion could lead to an exposure to particles that are smaller in size and greater in numbers than ambient health guidelines, potentially increasing the risk for respiratory illnesses. Their study also finds that a change from polluting light sources to electric lights are associated with health benefits due to reduced particulate matter, even if the main reason for switching is enhanced lighting.

Bruce et al. (2000) links indoor air pollution to the risk of acute respiratory infections in children. They point out that it is not only the concentration of pollution that matters, but also the exposure; the time spent in the polluted area as well as the proximity to the source. Kerosene lamps are oftentimes located in close proximity to the users, according to Apple et al. (2010).

Furthermore, Joseph (2006) finds that artificial light impacts health by enabling the performance of visual tasks in the absence of natural light. In a properly lighted environment, people make fewer mistakes when undertaking a task. A better lightning source could make it easier to spot dirt and increase the number of hours available to conduct household chores, which again can improve the hygiene conditions. Lastly, access to a good light source at night increases the number of hours it is possible to read or work, and thereby increasing standard of living (Sovacool and Vera, 2014).

3.3 Implications for our Study

The literature on health effects from switching energy source originates from the US, and analyze the effects on infant health in the US. However, Arceo et al. (2016) finds that the effect of pollution on infant mortality differ when comparing data from the US and Mexico. As cultural, geographical and other conditions affecting health are likely to differ quite substantially in the US as compared to Norway, it is not certain that the effects are transferable to Norway. For example, it could be that the estimated effects in Norway will be smaller in magnitude since polluting energy sources were still used for indoor heating, and a cold climate increased the use of these energy sources.

Theliterature on the switch from kerosene lamps to electric light use more recent data than our period of interest. This could mean that the effects substantially differ from the effect found in Norway in the early 1900s.

The literature focuses on a shift between polluting and non-polluting energy production, and how this affect health outcomes. In our research, we look at the effect of a change from no access to electricity, or electricity supplied from a small polluting power plant fired by coal or gas, to a hydroelectricity source. As will be explained in Section 4, the transition from no power plant to a hydroelectricity plant has ambiguous effects, as the indoor environment might improve, while there are substantial alterations to the outdoor environment.

Our paper contributes to the existing literature in a number of areas. Firstly, we expect to find results not directly comparable to previous studies, as this, to our knowledge, is the first study of its sort using European data. Secondly, since Norway was an early implementer of hydroelectricity, our study provides an opportunity to analyze the health effects of hydroelectricity when electricity is mainly used for lightning. Based on Severnini (2014), it seems likely that our findings will deviate from the estimated results when electricity was synonymous with the installation of modern household equipment. Thirdly, only a small portion of the population switched from other energy sources to hydroelectricity, whilst the majority gained access to electricity for the first time when hydroelectricity stations were built. Hence, this paper mainly analyzes the effect of households gaining access to electricity for the first time, and not a shift in the source of power generation. Furthermore, our paper analyzes the effect of introduction of a "clean" energy source. Our results might give an indication of the expected effect on health when receiving a clean energy source for the first time.

4 Suggestive Mechanisms

Our analysis is built on the hypothesis that the opening of hydroelectricity plants affect health. While Section 3 looked at previous empirical studies analyzing the effect on health from access to electricity or switching energy source, this chapter will explore the direct health externalities from hydroelectricity plants, and seek to understand the mechanisms behind changes in health outcomes.

In the following section, WHO's three main determinants of health will be used in order to explain the mechanisms behind health effects from electrification. Further, the positive and negative impacts on health from alterations in the outdoor environment and opportunities will be discussed, before we discuss the direction in which we expect that the mechanisms will drive the results in the analysis.

4.1 Determinants of Health

WHO identifies three main determinants of general health, namely (WHO IV, 2016): The physical environment; The social and economic environment; The person’s individual characteristics and behaviors.

The physical environment, includes the natural and built environment in which people live. Researchers link changes in the energy mix offered in a municipality to changes in households' indoor and outdoor environment (see Section 3 for discussion). The links from hydroelectricity supply to changes in environment will be further explained in the next subsection.

The social and economic environment comprise the close physical surroundings, social relationships and cultural communities of a household (Barnett and Casper, 2001), as well as economic conditions. It seems likely that the economic environment might be affected by the opening of hydroelectricity plants. In this study, the municipality's income per taxpayer and number of poor people will be used as controls. Subsection 4.2 further discuss how hydroelectricity stations might affect economic conditions.

A person´s individual characteristics and behaviors, can be linked to genetics as well as personal choices. In our study, this is not possible to account for, since we do not have individual level data. The lack of individual level and education data, and the possible implications for this study, will be discussed further in Subsection 8.2.

4.2 Direct Effects from the Hydroelectricity Plant

In Section 3, we discussed how the literature associate a change in energy source to changes in indoor and outdoor pollution levels. In this subsection, we will explore how constructing a hydroelectricity plant alters the environment, and thereby opportunities and health. As the indoor environment is not significantly affected by the energy source fueling a power plant, the discussion will be limited to how hydroelectricity plants can change the outdoor environment.

In contrast to fossil fuels, where burdens are dominated by polluting emissions, the burdens from constructing a hydroelectricity plant are much more site and technology specific. Therefore, it is hard to generalize the overall effect (European Commission, 1995). However, building a hydroelectricity plant generally include constructing a dam, which will flood an area and change the opportunities for land use. The direct impacts will differ from site to site. For example, some sites would experience problems from the alteration of river flows, while others could have issues with the release of temperate water into fjords.

The European Commission's (1995) ExternE project, which looked at the combined externalities from various energy sources, identified a negative effect on the physical environment through a loss of private goods, like land, and thereby income from e.g. traditional agriculture, and forestry. Also, aquatic and wildlife might have a loss of habitat, which could lower the socioeconomic position for individuals in the municipality involved in forestry, farming, or hunting as a source of income. This could further have a negative impact on general health in the municipality.

Furthermore, there could be a loss of water supply to the general public. Sauda municipality is one of the main cases in the ExternE report (European Commission, 1995). When Sauda received a hydroelectricity plant, several households experienced that wells became dry, so that new wells had to be built. Access to clean drinking water is essential for keeping good health, and diarrhea caused by unclean water sources may lead to infectious diseases in infants. Further, some farms experienced reduced capacity of water intake for irrigation, and parts of the municipality needed a new source of drinking water. Loss of irrigation opportunities might cause crops to fail, which can harm both nutritional and socioeconomic conditions in the municipality, and thereby health.

A potential benefit could be that some switched from coal and gas driven power plants to the new hydroelectricity plants. There might be indications that this happened to some degree in

Norway, as coal consumption peaked in 1916 (SSB, 2001). This could mean lower emissions from coal burning, and contribute to better respiratory conditions in general.

Cheap access to electricity from hydroelectricity plants gave Norway a comparative advantage within energy intensive industry, as discussed in Subsection 2.1. Hence, when a hydroelectricity plant was build, it is also likely that the industrialization of the municipality would gain momentum. In terms of health, this can have ambiguous effects (European Commission, 1995). Firstly, the industrialization might lead to higher pollution levels, either directly from the industry, or from transport emissions associated with the industry. This could have had a negative impact on health, and lead to an increase in e.g. respiratory disease. Also, population density could increase, which might lead to poorer hygiene conditions and higher exposure to infectious disease. However, there could be an increase in job creation in the municipality (European Commission, 1995), leading to an increased average social and economic position for the population, which again might increase health status.

4.3 Expected Effects on Health

The underlying mechanisms from hydroelectricity to health are expected to drive our results in two opposing directions. Firstly, we expect the decreased levels of indoor pollution, as described in Subsection 3.2, to affect health positively. Electric light bulbs represented a substantial improvement, both in terms of particulate matter and in terms of the visual ability to carry through tasks effectively and without error. However, the direct effects from the hydroelectricity plant itself is more ambiguous. Negative effects include potential nutritional loss from the loss of forestry and farming opportunities, as well as a potential loss of water supply from altered waterways. Positive effects include possible lower pollution levels if the hydroelectric plant caused less activity in coal plants, as well as the possibility of economic gains from electricity intensive industry. Thus the expected sign of our estimates is ambiguous, as it is hard to say which underlying mechanism will affect our results the most.

5 Data

In order to investigate the effect of electrification of households on health, we have constructed a panel data set that links data on the rollout of hydroelectricity plants with health statistics from the beginning of the 20th century, as well as with a number of control variables.

Our data are drawn from four main sources: Data on the rollout of hydroelectricity plants are collected from the archives of NVE. We used health statistics reports from SSB to collect data on infant mortality, prevalence of tuberculosis and diarrhea, and the number of inhabitants and doctors. Data on poverty are collected from the Norwegian Centre for Research Data (NSD), and data on tax payments are from SSB`s archive of historical statistics. With the exception of the poverty data, all data have been digitized from original sources. Data is digitized from 1900 until 1921, as relevant information was not provided in the medical reports after 1921. In the following subsections, we describe our data in detail.

5.1 Health Data

The data on health status in the different municipalities is collected by digitizing SSB's historical yearly health statistics reports from 1900 to 1921.[5] The information is reported by the chief physician of each medical district. The data available in these reports is therefore the most accurate data covering health status at the time. From the reports, we have digitized the number of inhabitants, doctors, tuberculosis cases, diarrhea infection cases, newborns and infants who died during the first year of life. All data is reported at medical district level.

To obtain comparable figures, all variables are calculated as ratios. The infection rates are obtained by dividing the number of persons infected by tuberculosis and diarrhea by the number of inhabitants in each medical district. The infant mortality rate is obtained by dividing the number of children between 0 and 1 years of age who passed away in a given year by the number of children born in the same year. All the rates are then multiplied by 1000.

Medical data is reported detailed and consistently throughout the data set, with the exception of a few years in the beginning of the period. The number of tuberculosis cases are reported in particular detail and description. The exceptions are 1900, 1901, 1902 and 1918. These

[5] Sources: SSB, 1902; 1903; 1904; 1905; 1906 I; 1908; 1908; 1909; 1910; 1911 I; 1912; 1913; 1914; 1915; 1917; 1918; 1920; 1921 I; 1922; 1923; 1924; 1925

years are therefore excluded from the analysis when tuberculosis is the dependent variable. Diarrhea is reported by number of cases and deaths, where the number of cases has been digitized. The excpetions are 1900, 1901 and 1902. The number of births and deaths for children aged 0 to 1 are reported each year from 1900 to 1921. Some medical districts lack observations on one or more health variable for different years, and this is accounted for by dropping these observations in the analysis.

5.2 Hydroelectricity Data

Data on the rollout of hydroelectricity plants were obtained by digitizing a report from NVE containing all hydroelectricity plants in operation by 1922 (NVE, 1922). The plants are shown in Figure 2. The data available contain information on the year of construction and localization on municipality level. Without more information on the size of the turbines and weather conditions, it is not possible to produce accurate predictions of production volume for each plant. We therefore simplify by assuming that the effect of receiving a plant is equal across all municipalities, and that the supply of electricity begins in the construction year. We also assume that municipalities gains access to electricity when its first hydroelectrcity plant is built, and that a plant supplies only the municipality in which it is built.

An issue with this data is that we miss information on construction year for 32 percent of our observations of hydroelectricity plants, and NVE is not able to provide clarifying information. Municipalities where there are hydroelectricity plants with unknown construction year, are treated as municipalities without hydroelectricity plants until they receive a plant with known construction year, under the assumption that unrecorded plants are very small or otherwise insignificant enough to not be properly registered. However, the majority of the affected municipalities have other plants with known construction year, and are therefore part of the treatment group when the first hydroelectricity plant with known construction year is recorded.

Another issue is that the opening years of hydroelectricity plants in the four largest cities, Oslo, Bergen, Trondheim and Stavanger, were not reported in the official statistics. The hydroelectric plants supplying cities were not necessarily located within the municipality borders. Rather, the cities owned development rights of waterways in another municipality, and we therefore supplemented the official statistics with the opening years of the first hydroelectricity plants in the largest cities (see Lyse, 2016; NVE, 2015; Riksantikvaren, 2013; and Rosvold, 2012).

Figure 2: The rollout of hydroelectricity plants between 1900 and 1922

Notes: The opening years are grouped into four periods: municipalities with plants constructed before 1900; between 1900 and 1910; between 1911 and 1922; and municipalities without hydroelectricity plants by 1922 (NVE, 1922).

5.3 Geographical Units

Health statistics are reported on medical district level, while data on hydroelectricity plants are available on municipality level. In the beginning of the 20th century each medical district could stretch across several municipalities. This difference in geographical units cause problems for merging the data, and require attention.

Firstly, all municipalities within the same medical district are assigned the same health rates, as health data are available on this level only. At the same time, this might cause single municipalities' sensitivity to hydroelectricity plants to decrease. We cannot account for how much each municipality contributes to a medical district's total health statistic. Therefore, we report each municipality's medical data in the medical district it resides at any given time.

Secondly, a number of municipalities have been affected by changes, either by being split into several smaller municipalities or by border adjustments. Judkvam (1999) provides a historic overview of splits, merges and border adjustments in the municipalities, which has been used to deal with this. Split municipalities have been included as distinct municipalities in all years of the data set prior to the split, and assigned the value of the medical district in which they later will reside. In the case of border adjustments, the changes were oftentimes so small that they had no practical implications. For these observations, we kept the municipalities constant throughout the period. In the few cases where the municipality border adjustment led to considerable changes in population between medical districts, we moved a corresponding proportion of all variables to the other medical district. This is also done for all years prior to the adjustment in order to construct comparable units. There are almost no cases of municipalities being merged within the period of the data set. A small number of municipalities are dropped from the data set, as it was not possible to account for the new municipality names and in what medical districts they belong.

Furthermore, the composition of medical districts changes over time, where some municipalities either switch district, or a district is split into smaller units. Therefore, some of the variation in health outcomes within municipalities will be caused by this arbitrary change. This will be a problem if the change in medical districts coincides with an opening of a hydroelectriciyt plant, but not otherwise.

Finally, we have created variables to indicate each municipality's affiliation with counties and regions. We have identified 20 counties and five regions. The counties follow the official classification at county level in 1920. The regions are defined as follows: Eastern Norway, with 30.6 percent of the observations; Southern Norway, with 11.1 percent of the observations; Western Norway, with 29.3 percent of the observations; Trøndelag, with 7.2 percent of the observations; and Northern Norway, with 21.8 percent of the observations.

5.4 Control Variables

The number of doctors per 1000 inhabitant is collected from the previously described medical statistical reports published by SSB. Including this variable allows us to control for differences in health care services in the medical districts across the period.

Data on the number of poor people per 1000 inhabitant is collected from NSD.[6] For the period of interest, poverty data is available for 1900, 1905, 1910, 1915 and 1920. The data is reported at municipality level. The years in between the observations are assigned data from the closest observation.

Data on the number of taxpayers and total taxable income in each municipality is collected from SSB's yearly statistical reports.[7] Data is collected for 1900, 1905, 1910, 1915 and 1920, and shown in 1000s. As earlier, years in between observations are assigned data from the closest observation. The income data is adjusted to real 1998 Norwegian kroner (SSB, 2016), in order to make the numbers comparable over time. Including the two economic variables allow us to control for differences in economic conditions in the municipalities across the period.

For missing observations on the control variables, we include a dummy variable indicating that the variable is missing. This keep the sample constant across the specification with and without control variables.

5.5 Descriptive Statistics

Infant mortality rates display a huge improvement during our period of interest, as displayed in Figure A1 panel (a) in the Appendix. As the figure shows, there are clear spikes and troughs. However, the overall trend is a steep decrease, which can be attributed to improved economic conditions, increased focus on hygiene, better health care for infants at the end of the period (see Bütikofer et.al., 2015 for a further discussion) and improvements in nutrition (Liestøl et al., 2007).

[6] Sources: NSD, 1900; 1905; 1910; 1915; 1920
[7] Sources: Statistics Norway 1901; 1906 II; 1911 II; 1916; 1921 II

The number of tuberculosis incidents decreased during our period of interest, as shown in Figure A1 panel (b) in the Appendix. However, the decrease is more unsystematic than that of infant mortality. While infection rates were rapidly declining between 1903 to 1909, the period between 1910 and 1917 show major tuberculosis outbreaks. The infection rates were especially high in 1913 and 1917. Tuberculosis is highly contagious, which might explain why some years stand out from the statistics.

The development in diarrhea rates does not seem to have a declining trend during our period of interest, as shown in Figure A1 panel (c) in the Appendix. The graph shows several spikes and troughs, where 1918 clearly stands out with an unusually high occurrence of the disease. This could be attributed to the Spanish Influenza weakening people's immune system in general, or strict rations towards the end of World War I, as further investigated in Subsection 7.2. See Table A1-A3 in the Appendix for descriptive statistics for infant mortality, tuberculosis and diarrhea in different regions.

Figure A2 in the Appendix show the accumulated absolute number of municipalities with hydroelectricity plants in each year of our period of interest. The rollout of hydroelectricity plants began in the late 1800s, and picked up speed at the turn of the century. The growth escalated around 1907-1908, and declined at the outbreak of World War I, explained by the fact that a number of construction parts were imported from Germany, and the war resulted in delays and inconsistencies in machine part deliveries (Kristiania Elektricitetsverk, 1916). Construction seems to pick up towards the end of the period, after the end of World War I. Note that the graph only shows the hydroelectricity plants included in our data set – those with a known construction year. Graphs showing the development in each region can be found in Figure A3 in the appendix.

6 Empirical Strategy

The goal of the empirical analysis is to identify the causal effect of the electrification of households on various health outcomes; infant mortality, diarrhea and tuberculosis rates. The strategy exploits differences in timing of when municipalities got hydroelectricity plants and it was made possible for households to connect to an electricity grid. We assume that while municipalities might differ for many reasons, the *timing* of the opening of a hydroelectricity plant mainly depends on technological progress and geographic suitability and will produce an exogenous shock that affects our outcome variable, health. However, as discussed in Section 4, the sign of the effect of hydroelectricity on health is uncertain. To explore the effect of interest, we perform a staggered differences-in-differences (DiD) analysis, making use of the rolling out approach to utilize the full variation in the aggregate data set. In order to understand the rollout strategy, which in reality is an advanced form of a DiD analysis, deducing the DiD strategy is useful.

6.1 Differences-in-Differences

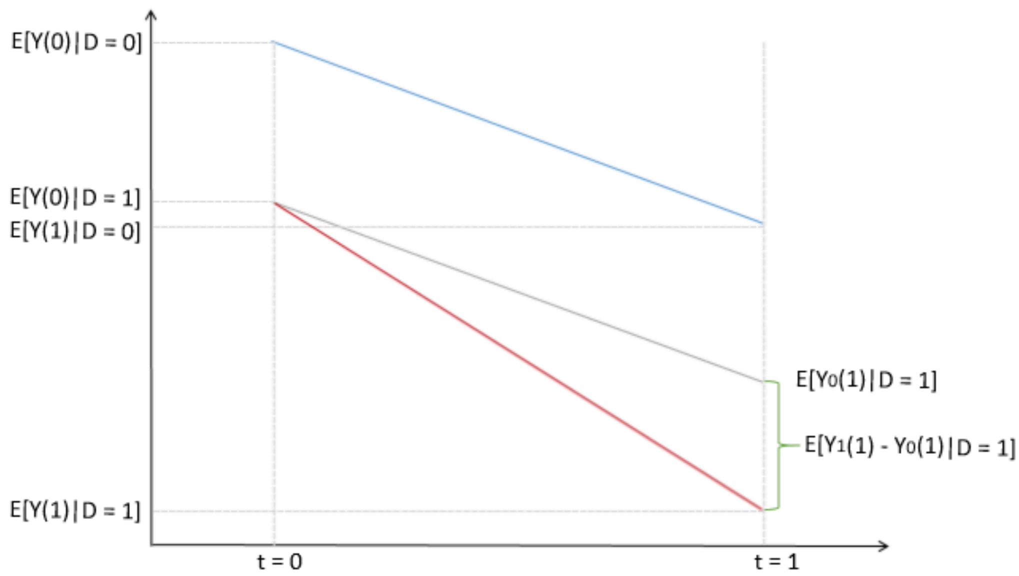
In an ideal set-up, the health status of a municipality that received a hydroelectricity plant (the treated unit) would be contrasted with the health status of the same municipality without a hydroelectricity plant (its true counterfactual). However, as the true counterfactual is unobservable, we need to introduce control groups, which are other municipalities that we assume have the same evolution in trends as the treated group in the absence of treatment. One such method would be the DiD strategy (Angrist and Pischke, 2015).

DiD in its simplest set-up with one treatment and one control group, and with two time periods, the time of treatment (t = 0) and after (t = 1), is shown in Figure 3. In the figure, D = 1 denotes the treatment group, while D = 0 is the control group. Y(0) is the outcome variable at t = 0, and Y(1) is the outcome variable at t = 1. Y_1 is the treated Y, and Y_0 is Y in the absence of treatment. Here, the estimator of interest is $E[Y_1(1) - Y_0(1)|D = 1]$, namely the difference for the treated municipality at time t = 1 when it has received treatment, $E[Y_1(1)|D = 1]$, and in the absence of treatment, $E[Y_0(1)|D = 1]$. However, as mentioned, the latter is unobservable. Therefore the estimator of interest, $E[Y_1(1) - Y_0(1)|D = 1]$, is derived by:

$$(1) \quad E[Y_1(1) - Y_0(1)|D = 1\,] = \{E[Y(1)|D = 1\,] - E[Y(1)|D = 0]\} - \{E[Y(0)|D = 1\,] - E[Y(0)|D = 0]\}$$

$\{E[Y(1)|D = 1] - E[Y(1)|D = 0]\}$, denotes the difference between the treated and the control group at t = 1. $\{E[Y(0)|D = 1] - E[Y(0)|D = 0]\}$ denotes the difference between the same groups at t = 0. The difference between the two gives the DiD estimator, and is an approximation of the difference between the treated group and its true counterfactual. The change in the control group from t = 0 to t = 1 is assumed to be what the change in the treatment group would have been in the absence of treatment. Any additional change in the development in the treatment group is therefore assigned to the treatment itself.

Figure 3: Differences-in-Differences illustration



Notes: The figure illustrates DiD in its simplest set-up with one treatment and one control group, and with two time periods, the time of treatment (t = 0) and after (t = 1). The grey line illustrate the counterfactual outcome when the parallel trend assumption holds.

For this to be true, the parallel trend assumption for DiD has to hold. This assumption associates the average change in the control municipality with the counterfactual change in the treatment municipality in the absence of treatment. It is a strong assumption, as there are several factors that could affect municipalities differently at different points in time.

However, it can be a reasonable assumption for taking account of pre-treatment differences in levels. The divergence of a post-treatment path from the trend established by a comparison group might signal a treatment effect.

When we evaluate the aggregate effect of hydroelectricity on health, and want to utilize the entire variation in our data set, normal DiD is insufficient. The usual DiD setting with a binary treatment variable allows for only one common treatment time for all treated municipalities. However, the implementation of new technologies across a country often occurs with spatial and temporal variation. This is also the case with hydroelectricity, as explained in Subsection 2.1. Hence, the basic DiD-model needs to be extended. This is done by exploiting this staggered implementation in our regression analysis, as described in the following subsection.

6.2 Rolling Out

The rollout method takes into consideration that electrification occured at different times in different municipalities, and allows for finding an average effect (for examples where rollout is utilized, see Akerman, Gaarder, and Mogstad, 2015; Bütikofer, Løken, and Salvanes, 2016; and Kose, Kuka, and Shenhav, 2014). All municipalities enter the analysis simultaneously, regardless of treatment time, and we will thereby utilize the entire variation in the data set. The control group is comprised of municipalities that at a given time during the period had not received their first hydroelectricity plant, while the treatment group is comprised of the remaining municipalities, i.e. those that received or had received the first hydroelectricity plant at that same time. In particular, we estimate the following model:

$$(2) \quad y_{m,r,t} = \alpha + \gamma D_{m,t} + \beta X_{m,t} + \delta_r + \sigma_t + \varepsilon_{m,r,t}$$

Where $y_{m,r,t}$ is the outcome of interest for municipality m in region r at time t. α is a constant. $D_{m,t}$ is an indicator variable equal to zero for a municipality before it receives a hydroelectric power plant, and equal to one after one is built. $X_{m,t}$ is a set of municipality specific characteristics that are time variant, including doctors per capita, the number of people in poverty, and taxable income per taxpayer. δ_r is a set of indicator variables at regional level, allowing for differences in levels between the five regions. If unobserved time invariant effects are also correlated with the outcome variable, it will lead to an omitted variable bias. An example of such fixed effects is the geographical features of the counties or regions. Northern Norway will, for example, have different geographical and cultural features than

Eastern Norway for the entire duration of the data set. δ_r controls for problems caused by this by removing time invariant factors, including the unobserved ones, thereby eliminating the bias at regional level. σ_t is a set of year dummies to control for common time shocks. An example of such shocks could be epidemics striking the entire country one year, or economic shocks affecting standards of living. γ is the variable of interest and shows the effect on health of a hydroelectric plant opening.

Including control variables, $X_{m,t}$, in the model specification controls for effects that could cause treatment and control municipalities to differ after treatment, and help avoid omitted variable bias (OVB) by contributing to the conditional mean independence assumption that the error term is independent of the variables of interest. This is given by $E(\varepsilon|D,X) = E(\varepsilon|X)$. Under this assumption, OLS will give unbiased estimates for the variable of interest, but not for the coefficients of the covariates. Thus, controlling for relevant covariates makes it more likely that we will be able to find conclusive results from a regression.

The control variables should not themselves be possible outputs. Such variables are known as bad controls, and can create problems. Generally, variables that are measured after the treatment may be determined in part by the treatment, in which case they are themselves outcomes (Angrist and Pischke, 2015). We will test for this by removing the control variables to see whether the baseline estimates change significantly. A significant change might imply that there could be a problem with bad controls.

The panel structure of the data set can potentially pose issues that need to be addressed. When estimating models using panel data, the error terms can possibly be serially correlated, meaning that the values of variables for nearby periods are likely to be similar. When this is the case, usual standard errors cannot be used for inference. In the health data, an example where this might happen is if an epidemic that occur one year continues into the next period, causing sickness rates to be higher than average in both years. When the dependent variable is serially correlated, the standard errors should be adjusted accordingly (Angrist and Pischke, 2015). This issue will be addressed by specifying the standard errors as clustered at municipality level, and robust to heteroscedasticity, which allows for correlated data within each municipality[8].

In order to draw inference from a rollout analysis, the key identification assumption that the timing and location of the opening of a new hydroelectric plant is independent of other factors affecting health, has to hold. We are not able to observe precisely what influence the decisions

[8] See Appendix A2 for a discussion of clustering at a higher level.

that lead to the timing and location of the establishment of each hydroelectricity plant. However, in Norway, the plants' location was dependent on access to waterfalls (Hveding, 1992) and technological feasibility (Hveding, 1992; Øberg and Bråten, 2014; Nissen, 1951) and not transitory local characteristics. Receiving a plant could therefore be viewed as an exogenous shock to the municipalities that received it. We test this assumption under Subsection 7.1.1.

6.3 Event Study

In order to test the assumption of independence in timing and location of an opening, it is of interest whether the decision was influenced by specific preopening trends, for example whether hydroelectricity plants were constructed in municipalities where health status is increasing or decreasing (Bütikofer et al., 2015). A method of examining this assumption visually is provided by an event study specification. This is done by looking at lead values of the variables of interest.

Furthermore, when estimating our model of interest we want the parallel trend assumption to hold, as described under Subsection 6.1. This entails that municipalities not receiving a hydroelectricity plant must represent a plausible counterfactual of the outcomes in municipalities receiving a plant. Therefore, differential pre-opening trends among municipalities correlated with the construction of hydroelectricity plants, which could in turn influence health outcomes, could be threats to the identification (Kose et al., 2015). Statistically significant lead values in the event study specification below could be an indication that the parallel trend assumption does not hold.

The following event-study specification is utilized, which allows us to estimate the differential effects of receiving a hydroelectricity plant (see Bailey and Goodman-Bacon, 2015):

$$(3)\quad y_{m,r,t} = \alpha + \sum_{\tau=k}^{-2} \omega_\tau E_m 1(t - T_m^* = \tau) + \sum_{\tau=0}^{q} \psi_\tau E_m 1(t - T_m^* = \tau) + \beta X_{m,t} + \delta_r + \sigma_t + \varepsilon_{m,r,t}$$

In the specification, E_m is a binary indicator variable that is equal to one if the municipality ever received a hydroelectric plant and zero otherwise. The event-year dummy variables, $1(t - T^*_m = \tau)$, are equal to one when the observation is $\tau = k,\dots,0,\dots,q$ years from the year T^*_m when a municipality received a hydroelectricity plant. This allows for k anticipatory effects prior to the opening of a plant, and q post-opening effects. The year prior to the opening year will be omitted as control year. ω_τ will show the development in the outcome of interest in municipalities that are about to get a hydroelectric plant in the years prior to construction, while ψ_τ shows the development in the outcome τ years after an opening of a hydroelectricity plant after adjusting for control variables.

Thus, when carrying out the event study, we ideally want ω_τ to be insignificant. This indicates that there are no differential trends in the outcome variable in the treatment municipalities, as compared to the control municipalities. Such an observation would entail that the parallel trend assumption likely holds. On the contrary, if the anticipatory effects are significantly different from zero, or show clear signs of trends, it implies that there could be differential trends between treatment and control groups.

7 Empirical Analysis

In the following section, we estimate the average effect from a hydroelectricity plant on infant mortality, tuberculosis and diarrhea. The analysis consists of two main parts. First, we present the results of our baseline estimations, followed by a test of the key identifying assumption. We then conduct an event study to test for anticipatory effects, and to check whether there is a lagged effect from hydroelectricity plants to health outcomes. Further, we conduct a variety of robustness tests to check the sensitivity of our results, before we conclude the section by summarizing the results.

7.1 Main Results

The main estimated results for infant mortality, tuberculosis and diarrhea using the rolling out approach in Equation 2, are shown in column (2) in Table 1. All estimates have standard errors clustered at municipality level.[9]

Overall, we find that exposure to a hydroelectricity plant leads to a decrease in infant mortality by -2.931 per 1000 born infants, significant at a 5 percent level. Compared to the pre-construction average for municipalities receiving a hydroelectricity plant, constituting a 4.74 percent decrease in infant mortality. This indicates that the rollout of hydroelectricity might explain part of the decrease in infant mortality in Norway in the early 1900s, and is part of the explanation for Norway's low infant mortality as compared to the US. The effect on tuberculosis cases is estimated to decrease by -0.234 per 1000 inhabitants. This translates to a 8.84 percent decrease in tuberculosis rates, compared to the pre-construction average. Both results indicate a considerable gain in a municipality's health status when receiving a hydroelectricity plant. For diarrhea cases, the estimated effect is –0.038 per 1000 inhabitants, constituting a 0.60 percent decrease as compared to the pre-construction average. However, this effect is small and not significant, indicating that the effect from receiving a hydroelectricity plant could be negligible. Based on the suggestive mechanisms from Section 4, which sign the effect would take is uncertain. Therefore, obtaining a decrease in the outcome variables is interesting.

[9] For a discussion about standard errors clustered at a higher level, see the Appendix A2.

Table 1: Rolling out estimates

	(1) Mean Pre-Construction	(2) Baseline Results	(3) No Control Variables
Infant Mortality	61.807	-2.931**	-2.170*
		(1.124)	(1.141)
Region and Year Dummies		Yes	Yes
No. of Clusters		714	714
Observations	2,425	15,302	15,302
Tuberculosis	2.646	-0.234***	-0.0730*
		(0.054)	(0.0553)
Region and Year Dummies		Yes	Yes
No. of Clusters		714	714
Observations	1,791	12,423	12,423
Diarrhea	6.300	-0.038	-0.797*
		(0.330)	(0.388)
Region and Year Dummies		Yes	Yes
No. of Clusters		714	714
Observations	1,795	12,800	12,800

Standard errors in parentheses

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

Notes: Each parameter is from a separate regression of the outcome variables on the access to a hydroelectricity plant, based on the model in Equation 2. Outcome variables are infant mortality (the number dead infants between age 0 and 1 per 1000 born), tuberculosis rates (the number of tuberculosis cases per 1000 inhabitants), and diarrhea cases (the number of diarrhea cases per 1000 inhabitants. Control variables include doctors per 1000 inhabitants, tax income per tax payer and the number of poor people per 1000 inhabitants, as well as indicator variables for regions and years. The standard errors are clustered at municipality level and robust to heteroscedasticity. Our period of interest is 1900 – 1921 for infant mortality, and 1903 – 1921 for tuberculosis and diarrhea cases. Column (1) show the mean outcome values for all municipalities receiving a hydroelectricity plant during our period of interest prior to construction; Column (2) show the baseline results from Equation 2; Column (3) show the estimates when control variables are excluded.

Estimating the model specification without control variables will challenge the validity of our identification strategy, and help determine whether we estimate the effect of interest, as explained in Subsection 6.2. The results are reported in column (3) in Table 1. For infant mortality, the estimated effect is smaller when there are no control variables. However, the estimate is not significantly different from the baseline effect. For tuberculosis, the estimate is significantly different from the baseline estimate, and the effect is much lower. A possible reason could be that more doctors could increase the quality of care and the overall hygienic conditions, which again could reduce contagion. For diarrhea, the estimated effect is larger, and goes from statistically insignificant at 10 percent with control variables, to significant without control variables. One explanation could be that much of the previous observed effect of opening hydroelectricity plants on diarrhea might be attributed to the prevalence of poverty. Across the different dependent variables, removing control variables does not alter the sign of our estimates.

The estimated results in Table 1, column (2) are intention-to-treat (ITT) estimates. These estimates are based on whether a household has the *opportunity* to connect to the electricity grid when a hydroelectricity plant opened, and gives the effect of the offer of treatment (Angrist and Pischke, 2009). However, the real uptake rate, namely the share of households that installed electricity when given the opportunity, is likely to be lower. Thus, our estimated health effect in column (2) is likely too small relative to the average treatment effect for households that indeed installed electricity (Angrist and Pischke, 2015). Therefore, the ITT effects gives the lower bounds on our estimates.

In order to obtain the effect of treatment on the treated (TOT), the ITT estimate is divided by the uptake rate. In 1920, 64 percent of Norway´s population lived in houses with electricity, as stated in Subsection 2.1. This gives some indication of the uptake rate, and we use this to calculate an *upper bound* to the treatment effect. To see that this is an upper bound, note that 64 percent of the *entire* population had installed electricity. Hence, in municipalities with a hydroelectricity plant, the uptake rate was likely to be higher (see Subsection 8.3.1 for a discussion). The resulting TOT effect shows that infant mortality was reduced by 4.58 deaths per 1000 born, and tuberculosis and diarrhea cases were reduced by 0.37 and 0.06 per 1000 inhabitants, respectively, in the households that actually installed electricity.

7.1.1 Test of the Key Identifying Assumption

Since the rollout strategy hinges on the assumption that the timing of the opening of a hydroelectricity plant is uncorrelated with other determinants of health outcomes, we ideally want to test whether this identifying assumption is true. This is not possible, and instead we test whether access to electricity is correlated with trends in the observable characteristics of medical districts and municipalities. Although the absence of such a correlation is not proof of the identifying assumption, it does provide some suggestive evidence that it could be true.

The results are reported in Table A5 in the Appendix. Column (1) and (2) show the results from analyzing whether 1900 characteristics can predict the opening and opening years of hydroelectricity plants. The results indicate that the observable characteristics hold some predictive power over the opening of hydroelectricity plants. Specifically, the results indicate that municipalities with a higher share of doctors per capita and less economic activity in 1900 were more likely to receive a hydroelectricity plant. However, the test results do not establish a causal relationship between economic activity and opening of hydroelectricity plants, but might indicate that plants were built where there were waterfalls, which oftentimes were in rural areas. In columns (3) and (4) we analyze whether early changes in characteristics can predict opening and opening year. We find that doctors per 1000 inhabitants do no hold any predictive power over opening or opening year of hydroelectricity plants. The economic variables are still significant.

Together, the test results could indicate that there is in fact a correlation between the timing of the opening of a hydroelectricity plant and other determinants of changes in health outcomes. Thus, our assumption of random assignment of hydroelectricity stations across municipalities could be violated, which is worrisome because it would threaten the validity of our results.

We include the economic variables as controls in the regression equation as we go forward. We also include doctors, since it is significant for the opening of hydroelectricity plants when using 1900 characteristics. We continue with the argument that conditional on our control variables, year and region fixed effects, spatial and temporal variation in the roll out of hydroelectricity plants across municipalities is plausibly exogenous. See Subsection 8.1 for a discussion.

7.1.2 Event Study Results

To further test the assumption that the timing of the opening of hydroelectricity plants is uncorrelated with trends in health, we utilize the event study specification in Equation 3. The specification will give an indication of whether the parallel trend assumption holds, as described in Subsection 6.3. Furthermore, the event study will allow us to see whether there is a lagged effect from receiving a hydroelectricity plant, and whether receiving one has a positive or negative effect on health over time.

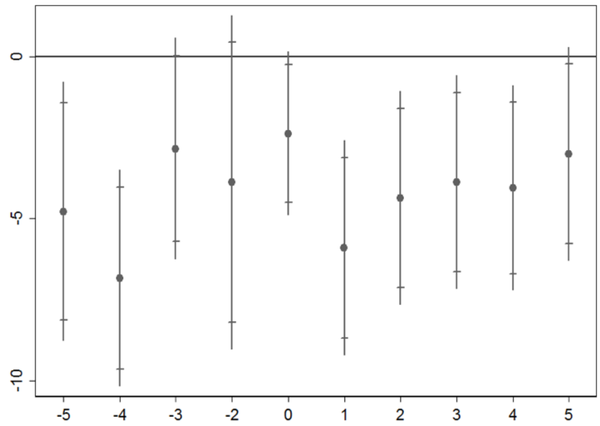
Figure 4 plots the event study estimates for infant mortality rates, tuberculosis cases and diarrhea cases, as well as the 90 percent and 95 percent confidence intervals when including five anticipatory and five post-opening effects. All control variables and the full sample is included in the estimates.

For infant mortality the effect is negative and significant in the opening year. The effect on infant mortality persist until year 5. The strengthening of the effect in the years post-opening could imply that it takes time for households to connect to the electricity grid, making the effect strengthen over time. It is also noteworthy that the effect is larger a year after the opening, than in the opening year. This result match our expectations, as it seems reasonable to assume that a portion of the hydroelectricity plants were constructed late in the year, making the effect in the year of construction small. It is also in line with the findings in Severnini (2014), which found that infant mortality increased the longer an unborn baby had been exposed to higher pollution levels, and that the effect on infant mortality was highest when the mother had been exposed during all nine months of the pregnancy. Based on these findings it is possible that the reduction in infant mortality increased with the share of the pregnancy the mother spent in a less polluted indoor environment, so that the reduction is largest when indoor pollution is lower during all nine months of pregnancy. Further, there seems to be significant anticipatory effects in year 4 and 5 leading up to the opening of a hydroelectric plant. The significant negative results in year 4 and 5 pre-opening could be an indication that the parallel trend assumption is violated and that there are certain differences in pre-treatment trends. This suggests that our results might not be an accurate representation of reality.

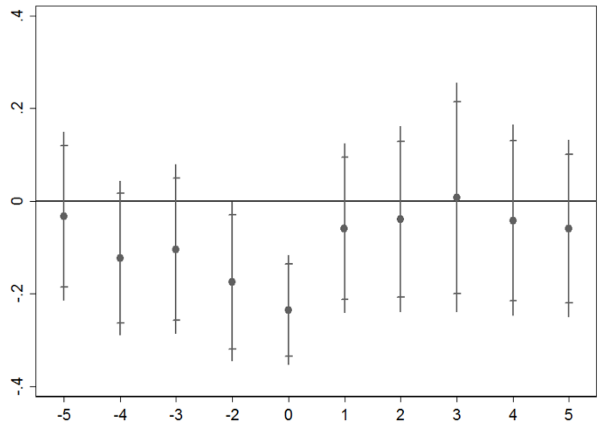
For tuberculosis there is a significant decrease in cases in the opening year, but no significant post-opening effects. The anticipatory effects are somewhat ambiguous, as the results are negative but insignificant in year 3, 4 and 5 pre-opening, while the effect is significant 2 years before opening. This could indicate that the parallel trend assumption is violated. No effect in the post-opening years could result from tuberculosis infections being know as "class-less", as described in Subsection 2.2, even if the socio-economic position of the infected definitely had a high impact on the development to a clinic disease and death (Liestøl et al., 2007). Thus, it seems reasonable that a change in standard of living from a hydroelectricity plant has a small effect on tuberculosis cases. It could be that mortality data for tuberculosis could paint a somewhat different picture.

For diarrhea, the results show no anticipatory effects at 95 percent, indicating that the parallel trend assumption is not violated. There is no significant effects in the opening year, which could have the same explanation as for infant mortality - it takes time to connect to the electricity grid, especially if the hydroelectricity plant was constructed late in the year. However, in the years after construction, the impact on diarrhea cases is negative, significant and persistent. This could be explained by the increased ability to do household chores at night and to spot dirt, as well as a generally cleaner indoor environment from fewer polluting light sources.

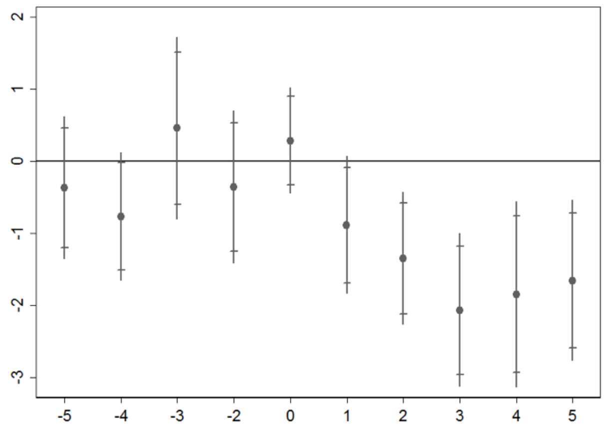
Figure 4: Event study estimates of the impacts of hydroelectric plants on infant mortality, tuberculosis cases and diarrhea cases



(a) Infant Mortality



(b) Tuberculosis cases



(c) Diarrhea cases

Notes: The figures show the post-treatment effects, ψ_τ, and anticipatory effects, ω_τ, from the event-study specification in Equation 3 as well as the 90 percent and 95 percent confidence intervals. Tax income per tax payer, poverty levels and the number of doctors per 1000 inhabitants are included as control variables. The standard errors are cluster robust at municipality level. The period of interest is 1900 – 1921 for infant mortality (the number of dead infants between age 0 and 1 per 1000 born) and 1903 – 1921 for diarrhea and tuberculosis cases per 1000 inhabitant. Figure (a) shows the effect on infant mortality, figure (b) shows the effect on tuberculosis, while figure (c) shows the effect on diarrhea.

7.2 Sensitivity Analysis

A sensitivity analysis might expose weaknesses in our model and our estimated results, and the robustness of our empirical results is a key issue. In the following, we first test whether our results are robust to the exclusion of various outliers from our sample. Then we preform two tests where we remove municipalities without hydroelectricity stations during our period of interest, and with unknown construction year, respectively. Finally, we allow for differential linear time trends between regions in our model specification to see whether such trends need to be controlled for.

7.2.1 Excluding Outliers

Excluding outliers will help identify whether there are certain observations that drive our results. However, since our sample is relatively small, we must be careful not to remove too many observations, as this could erode the significance of our study. In the following section, we exclude potentially worrisome observations in our data set, before we investigate the effect when we remove counties and regions one at a time. Equation 2 is the regression specification for all tests.

Excluding extreme observations

Table A6 in the Appendix examines the robustness of our estimates when sequentially excluding various extreme values. Column (2) excludes the highest observations of each dependent variable. Taken together, the estimation results indicate that the results from the baseline regressions are not driven by these observations.

Column (3) reports the estimates when the four largest cities are excluded. We suspect that they had different health trends compared to more rural municipalities due to higher pollution levels. For infant mortality and tuberculosis rates, the estimated effect is slightly more negative and statistically significant, but not distinguishable from our baseline results. The effect indicates that the reduction is smaller for the largest cities. For diarrhea rates, the effect is slightly more negative than the baseline results, but not significant nor distinguishable from the baseline results. Overall, the test show that our baseline regressions are robust to the exclusion of the largest cities.

Column (4) shows the regressions when Finnmark is excluded. We suspect Finnmark has differences that could affect disease and infant mortality rates. Firstly, Finnmark has a colder climate than the other counties (Norwegian Meteorological Institute, 2016). Secondly, in the period of interest around 45 percent of the population in Finnmark was of Sami or Kven origin, with a different culture and at lifestyle than other Norwegians (Ryymin, 2008). Lastly, the economic conditions were particularly difficult in Finnmark (Ryymin, 2008).

The estimation results for both diseases are in line with the results from our baseline specification. Interestingly, the estimated effect on infant mortality is almost halved and significant at a 10 percent level. The estimate is not within the standard errors from the baseline regression, which might indicate that Finnmark drives our results. During our period of interest only 5 hydroelectricity plants, 0.7 percent of the total number of hydroelectric plants in Norway, are constructed in Finnmark. Of these, three opened before 1903, and two after 1916. Thus, during a large part of our period of interest, the treatment group in Finnmark is very small. This, combined with the previously discussed differences in county characteristics, could explain why our baseline estimate is not robust to the exclusion of this county. Despite the results we will not exclude Finnmark in the further sample, as we are interested in the overall effect for Norway and excluding it would imply excluding a large number of municipalities. However, it is somewhat worrisome that our baseline results are not very robust to the exclusion of Finnmark.

Excluding counties one at a time

Based on the difference in estimation results for infant mortality when Finnmark is excluded from the sample, we also want to investigate whether excluding other counties alter the estimated results. As discussed earlier, the effect of hydroelectricity is ambiguous, and there could be geographical, cultural or economic features that affect health outcomes. It seems likely that these could differ between counties. Furthermore, we control for fixed effects at region level, so it is interesting to see whether there are effects at the lower county level that could drive our results. The results are reported in Figure A4 in the Appendix.

For infant mortality, the results are relatively consistent across counties. However, Finnmark, as explained above, and Nord-Trøndelag are clear outliers from the sample. When excluding Nord-Trøndelag, the estimation results are no longer significantly different from zero at 90 percent. However, we do not consider Nord-Trøndelag a problematic outlier, as Nord-Trøndelag had 8.2 percent of the total plants in our data set by 1921. Thus, identification results from Nord-Trøndelag is an important part of our identification and dropping this county will remove too much variation. For tuberculosis and diarrhea cases, the estimation results are relatively consistent across all counties.

Excluding historic events

The health effects from receiving a hydroelectricity plant might differ during years with especially strong external shocks, compared to the effect in municipalities that got a plant in years before or after the shock. Thus, we exclude historic events that we suspect influence the population's general health. Three such events are recognized as possible external shocks in our period: the Spanish Influenza, World War I and the end of the union with Sweden. The results are shown in Table A7 in the Appendix.

The estimated results when the last two years of World War I is excluded, as shown in column (2). Norway remained neutral during the war, and incurred a low number of fatalities. However, the warring nations were subject to a mutual embargo, by which Norway was affected trough increased prices on several goods and services (Brazier and Kirkhusmo, 2016). The embargo and rationing was most severe in the last two years of the war, and therefore only these are excluded. The estimates for infant mortality and tuberculosis show no significant changes, and they are not distinguishable from the baseline results. The sign for diarrhea changes from negative to positive, but the estimate resides within the standard error of the baseline estimate.

Column (3) shows the results when excluding the Spanish Influenza. Between 13,000 and 15,000 are estimated to have died from the Spanish Influenza in Norway, and more than one million were infected (Borza, 2016). The disease struck Norway in three epidemics: the summer 1918, the fall 1918 and the winter 1918-1919. Thus, 1918 and 1919 are excluded. The estimated effects are all in line with the baseline results.

Finally, the union with Sweden was dissolved in 1905, and the estimated results when the period up to 1905 is excluded is presented in column (4). The union is regarded as having few practical implications for Norway other than a common king and Sweden dealing with matters of foreign policy (Sejersted, 2015). However, towards the end of the union, there was an escalating conflict of interest, and even thought the dissolution was relatively peaceful, there could be severe differences before and after. For tuberculosis and diarrhea cases, the results are in line with the baseline results, which can be expected, as data is only recorded from 1903 for these variables. However, for infant mortality, the estimated results change to -4.118, significant at 1 percent level and outside the standard error of the baseline estimate.

It is likely that the effect on infant mortality from receiving a hydroelectric plant accelerate into the 1900s, as increasingly more households gained access to electricity, as explained in section 2.1. We are interested in the average effect over the entire period, and the increased effect is more likely to stem from electric light becoming more common rather than from the Norwegian-Swedish union. We therefore choose not to exclude 1900 to 1905.

7.2.2 Removing Municipalities without a Hydroelectricity Plant

Norway has big geographical differences and is elongated, resulting in differences in culture, nutrition, and ways of living. When this is the case, the DiD parallel trend assumption is particularly strong, and we suspect that treated and untreated municipalities might differ. We already have indications that this could be the case from Subsection 7.1.2. As a robustness test we limit the sample to include only municipalities that received a hydroelectricity plant in our period of interest. Thus, all municipalities that received a hydroelectrcity plant after 1921, or before 1900, are excluded. A reasonable assumption could be that municipalities receiving a hydroelectricity plant around the same period share more characteristics than those that did not receive one. Thus, when excluding municipalities without a hydroelectricity plant, the treatment and control municipalities might become more equal, and the results can be driven by the randomness in construction timing, rather than more fundamental geographical or cultural differences.

A further argument for removing these municipalities is that we expect some measurement error in our treatment from the high occurrence of missing construction years in the official statistics of hydroelectricity plants. It is therefore likely that there is hydroelectricity supply in some of the municipalities in our control group, which could attenuate our estimated results up until this point.

The results are given in Table A8 in the Appendix and show a sizable change to all estimates. Infant mortality has changed sign, and indicate that infant mortality increase by 0.803 per 1000 live birth. The effect on tuberculosis infections is reduced to –0.178, and the effect on diarrhea has changed sign, indicating that the number of diarrhea cases increase by 0.192 per 1000 inhabitants when a hydroelectric plant is opened. However, all results are statistically insignificant, so we cannot draw inference from the estimates or say they are different from the baseline.

The new estimates reduce our confidence in the baseline estimates, given what they say about the control group as a counterfactual. However, it is ambiguous what these numbers indicate for the overall validity of our results, as we drop more than two thirds of the observations, eroding much of the variation. This could explain why the regressions yield insignificant results. Furthermore, in the end of the period of interest, there will be very few control municipalities left, as all municipalities in this sample selection will have a hydroelectricity plant by 1921. We know that many of the municipalities which had not received a hydroelectricity plant by 1921, did so few years after our period of interest (NVE, 1930). This might make them good control municipalities. Hence, we choose to continue with our baseline sample.

7.2.3 Removing Municipalities with Unknown Construction Years

As discussed in Subsection 5.2, we have treated hydroelectricity plants with unknown construction year as not existing, under the assumption that they are very small or otherwise insignificant. If municipalities have other hydroelectricity stations built prior to one with a missing construction year, our assumption will not affect in which group (treatment or control) the municipality is placed. However, if a hydroelectricity plant with unknown construction year was the first in a municipality, then it will be wrongly placed in the control group. This could lead to an under estimation of the real effect from receiving a hydroelectricity plant. Furthermore, it might create a sample selection bias, given that municipalities that did not report construction years also have other similar characteristics (Woolridge, 2014). For example, it could be that these municipalities have less resources than those that did report construction year. Thus, restricting the sample might alter the estimated effect.

We test this theory by running our baseline regression on a sample where all municipalities with a hydroelectricity plant with unknown construction year are excluded from the sample. The results are reported in Table A9 in the Appendix.

The estimated results for infant mortality and tuberculosis show a slightly increased effect compared to the baseline estimates, the results are still statistically significant, and not distinguishable from the baseline results. Thus, our findings might indicate that our simplifying assumption does not create a noteworthy bias to our results. For diarrhea, we obtain a smaller effect than in the baseline regression, and the effect is different from the baseline results. However, the estimation result is not statistically significant, and we cannot draw any inference.

7.2.4 Differential Time Trends

Including a region-specific time trend might allow to distinguish the effect of the opening of a hydroelectricity plant from differential persistent region specific time trends. Such time trends could occur if there are time driven underlying trends in health in each region that differ between the regions. For example, if the rate of hygiene improvement is better in one region as compared to another, or if the rate of constructing health facilities is higher in some areas. The data set includes a sufficient number of municipalities and years, such that introducing a degree of nonparallel evolution in outcomes between regions is viable. Thus, including region-specific time trends is done to test robustness for such trends, and the specification will be as follows:

$$(4) \quad y_{m,r,t} = \alpha + \gamma D_{m,t} + \beta X_{m,t} + \delta_r + \rho_r T + \sigma_t + \varepsilon_{m,r,t}$$

ρ_r is the coefficient of a region-specific time trend multiplied with a linear time trend variable, T. All other variables are as specified in Equation 2. If the specification yield results that are significantly different from the baseline, there might be differential secular time trends between regions that should be controlled for.

In Table A10 in the Appendix, we see that there are no significant or sizable changes to the baseline estimates from including a differential time trend on infant mortality, diarrhea or tuberculosis. Hence, we conclude that there are no such trends affecting the development of these variables.

7.3 Summary of the Results

Our findings suggest that receiving a hydroelectricity plant has a positive effect on health. Specifically, we find that the construction of hydroelectricity plants lead to a decrease in infant mortality by -2.931 per 1000 born infants, significant at 5 percent. For tuberculosis cases the reduction is estimated to –0.234 per 1000 inhabitant, significant at 1 percent. The number of diarrhea cases are reduced by –0.038 per 1000 inhabitant, but the results are not significant. These findings imply a reduction in infant mortality by 4.74 percent, tuberculosis cases decrease by 8.84 percent and diarrhea cases decrease by 0.60 percent, as compared to the pre-construction average. The estimates are ITT effects, and the estimated reductions increase when we adjust for the uptake rate to find the TOT effect. The effects are likely to lie between the ITT and TOT effects.

The estimate for infant mortality is not significantly different from the baseline regression when the control variables are excluded from the model. Across the different dependent variables, removing control variables does not alter the sign of the estimates.

Furthermore, in the event study, we see indications that the effect for infant mortality and diarrhea persist and increase in the years after intervention. For tuberculosis, the effect bounce back to zero one year post intervention. This could be an indication that the real effect of electrification on tuberculosis is negligible. Tests for the parallel trend assumption and the assumption that the timing of the opening of a hydroelectricity plant is uncorrelated with other determinants of health outcomes yield results that are somewhat worrisome. This will be further discussed in Subsection 8.1.

We have performed a variety of robustness test. Overall, they show that the baseline estimates are not sensitive to changes in the sample or specification. However, there are certain problems with all outcome variables, occurring in different tests for each variable. For infant mortality, we have problems with the exclusion of the pre-union period, Finnmark and Nord-Trøndelag, and municipalities without a hydroelectricity plant. Removing municipalities without a hydroelectricity plant changes the estimated results for both diarrhea and tuberculosis, and removing municipalities with unknown construction years changes the estimate for diarrhea. However, the estimates are not significant, and we cannot draw inference that they are different from the baseline results. We argue that none of the test results change the estimates considerably enough to make the baseline results invalid. However, the problems will be further discussed in Section 8.

8 Discussion

Our findings suggest that receiving a hydroelectricity plant has positive health externalities. In the following, we will discuss possible shortcomings of the estimated results and the data set. Then, we will look at the limitations imposed on our study as a result of the chosen estimation strategy, as well as some possible improvements of the model. Finally, we discuss how our study can contribute to the existing literature.

8.1 Discussion of the Results

As tested under Subsection 7.1.1, there are some issues with the assumption of independence between the opening of a hydroelectricity plant and other determinants of health, and we therefore include these variables in our specification. The possible correlation between early municipality economic characteristics and the timing of the opening of hydroelectricity plants, are the most worrisome. It does not seem unlikely that there could be a correlation between the economic conditions in a municipality and the establishment of a plant. However, as historical sources claim that hydroelectricity plants were built close to water resources, and that these were spread around the country, we have chosen to go forward with the specification where these municipality characteristics are controlled for in our model.

Furthermore, under 7.1.2, we see indications that the parallel trend assumption might be violated for infant mortality rates, as the anticipatory effects four and five years prior to opening is significant. This could pose problems for the validity of our results. However, we do not expect anticipatory effects from hydroelectricity on infant mortality before opening, as infant mortality should not change from the prospects of receiving a plant in the future. Thus, the opening year should be a good indication of the first effect of hydroelectricity plants.

As mentioned in Subsection 7.2.3, the decision to treat hydroelectricity plants with unknown construction years as non-excising can lead to a sample selection bias. Our analysis in the same section suggest that the size of this bias is unnoteworthy. However, there might be misreported data also in the medical reports, which could lead to a selection bias in the health data. It seems likely that, on average, medical statistics would be underreported, either due to inadequate monitoring of disease rates or infant mortality rates. However, more accurate data on health conditions in the medical districts does not exist. Thus, it is not possible for us to further test whether a bias of this sort exists.

As stated in Section 2, the rollout of hydroelectricity plants coincided with a period of significant expansions in the health care system. This could be a problem to our research design if substantial health care expansions in a municipality were conducted in the same year as a hydroelectricity plant opening. However, knowing that the factors driving the decision to expand health care were very different from the decision to build a hydroelectricity plant, we assume that the healthcare rollout is uncorrelated with the electricity rollout on average.

8.2 Limitations to the Data Set

As mentioned in Subsection 4.1, a person's individual characteristics and behaviors is a key determinant of health. We have not controlled for this as we have not sampled data on individual level. For example, education could serve as a control for ability. However, SSB reports school data in another geographical unit than any of our other data. The lack of controls for individual characteristics might cause OVB in our estimated model as it seems likely that both health and the decision to construct a hydroelectricity plant could be, to some extent, correlated with abilities. High levels of individual abilities is expected to better an individual's health, while low ability levels could have an adverse effect on health.

The geographical units might cause problems, as we have hydroelectricity data reported at municipality level, and health data at medical district level. Thus, we observe the effects from a hydroelectricity plant on a higher level than municipality level. This might cause a measurement error in our estimates, as we assume the hydroelectricity plant supplies only the municipality in which it resides. The error will be larger for big medical districts than for small. However, we have not mitigated this problem, as the composition of medical districts change a lot from year to year, making it difficult to establish medical districts as a separate level in our analysis.

Misreporting by medical personnel might be a potential issue in the health statistics. For example, it could be that in medical districts where one or few doctors cover a larger area, the further away from the doctor inhabitants live, the less likely it is that their infections are reported. In these cases, underreporting might lead to a false impression of good medical conditions. If, however, underreporting is constant over time, it should not bias our results. Underreporting could also be an issue if the physician moves from a medical district, and this results in a breach in reporting. It could look like the health condition improves, while really the statistics is underreported. If this occures in the same year as the opening of a hydroelectricity plant, it might be a confounding factor. However, this problem is assumed to be only minor, and is not possible to account for in the analysis.

We assume in the analysis that hydroelectricity plants in one municipality supplied electricity to households in the same municipality only, due to the poor transmission mechanisms, providing incentives to supply electricity to households residing near the power plants (Oslo Lysverker, 1952). However, a spillover effect can occur if households in a municipality which has not received a hydroelectricity plant can connect to a plant in a neighboring municipality, and get the positive effects this entails while being wrongfully placed in the control group. This is possible if the municipality was small in size, or the plant was located close to a municipality border. Also, it is likely that spillover effects increased towards the end of the period as technology improved. We expect the effect to bias our estimates towards zero, as the health effects in control and treatment municipalities become systematically more equal when spillovers occur. Controlling for this would require checking the area supplied by each hydroelectricity plant, an extensive job that require looking into different historical sources for each hydroelectricity plant. We have done this for the four largest cities, but not for any other municipality.

8.3 Limitations to the Estimation Strategy

At least three problems with the estimation strategy can be identified. First, the decision to get a house connected to the electricity grid could be correlated with unobservable determinants of health, which might cause endogeneity problems. For example, job creation from the hydroelectricity plant might cause a family with increased income to install electricity.

Second, we have assumed that a hydroelectricity plant opening must be a strong predictor of homes getting access to electricity, and that the rate of household connection to the electricity grid is equal across municipalities. This could pose problems, as the uptake rate would vary depending on municipality and household characteristics, distance to the hydroelectricity plant, as well as the size of the plant. Without further information on these issues, we cannot give more precise estimates of the uptake rate. As explained in subsection 7.1, we have calculated the TOT effect by dividing the baseline effects by the total uptake rate of electricity for the entire population in 1920. This effect is likely to be too high, but the effect on health will be between the ITT and the TOT estimate.

Third, treatment and control municipalities should not be subject to structural health shocks that is not common to both groups in the treatment period. Examples could be local health reforms at county or municipality level which is not picked up by the region fixed effects. Structural processes driving the development of health infrastructure at the time was decentralized, and most hospitals had private ownership (Grønlie, 2015). For example, the hospital system grew based on local needs before 1930, characterized by voluntary and cost-sharing initiatives, both locally and nationally. The result was a multitude of hospitals of varying sizes and quality (Grønlie, 2015). Hence, tracking the structural processes driving health infrastructure must happen on a case-by-case basis for each municipality, and has not been done for this thesis. This might cause the health variables to occasionally change due to structural changes alone. A concurrence between a structural change and an opening of a hydroelectricity plant in a treatment municipality would likely cause an overestimation of the effect on health.

8.4 Possible Alternative Models

A possible improvement to the model could be to include a more specific uptake rate of electricity. For example, distinct uptake rates for each municipality, or increasing uptake rates in line with the establishment and size of new hydroelectricity plants. However, without further information on the size of the turbines of each hydroelectricity plant, we have no indication of the likely number of households a given plant could supply. Obtaining an accurate measure would also require knowledge about the span of transmission lines, ideally observed at different points in time, as they gradually developed during our period of interest. For the purpose of this thesis, we did not have resources to collect these data. However, information on the production capacity of each hydroelectricity station is recorded by NVE, and could be possible to collect and digitize.

Another alternative measure could be to register a rate of electrification at medical district level in the same way as described for the uptake rate above, and in that way get a more consistent measure within the same geographical unit. This will mitigate the measurement error resulting from the geographical unit problem discussed in Subsection 8.2, and could be done by including variables indicating how many hydroelectricity plants reside within a medical district at any given time. However, as mentioned, the medical districts change often and arbitrarily, so it would be hard to obtain accurate measures that do not change as often as the medical districts.

8.5 Implications of the Study

Our estimated effects are relatively small compared to studies that analyze health effects when modern household appliances were introduced or when the establishment of hydroelectricity plants resulted in a decrease in outdoor pollution. However, since access to electricity in our period neither implied access to electric utilities other than light bulbs, nor a shift in outdoor air pollution levels, the estimated effects are in line with our expectations. Also, our estimated effects are ITT effects, whereas not all households in a municipality that got a hydroelectricity plant installed electricity. Taken together, we still deem them comparable to those of other studies.

Our findings also support previous research on the negative health effects associated with polluting indoors energy sources. It is likely that much of the positive health effects observed stem from an improved indoor environment. This is in line with previous literature suggesting that changing from a polluting to a clean energy source has positive health effects.

The findings further adds to other positive effects from gaining access to electric light found in the literature. For example, UNICEF (2015) finds that poor lighting is associated with reduced time available for children's education, and that electrification help bridge the gender gap, as girls are able to study after sunset. Thus, knowing that access to electric light also has an effect on infant mortality rates strengthens the argument for investing in electrification programs in order to achieve electricity access for the 15 percent of the world population that still do not have it (World Bank, 2012).

In Subsection 3.3 we emphasized that the estimated results from US studies might not be similar to effects found in Norway. In the same way, it is possible that the estimated effects from Norway are not directly transferrable to other countries. Therefore, it would be interesting to conduct similar studies of the first access to electricity elsewhere. It would also be interesting to analyze whether similar estimates could be obtained in other countries, given that electricity was mainly used for lighting. However, this might prove difficult, as Norway was one of few countries where access to electricity became widespread before electrical household appliances were introduced.

9 Conclusion

This thesis aims to analyse the research question "*Did access to electricity from hydroelectricity plants affect health outcomes in Norway in 1900-1921?*" In order to answer the question at hand, we have used a staggered differences-in-differences approach, utilizing the near randomness in timing and location of the rollout of hydroelectricity plants in Norway.

We utilize newly digitized data from historical sources on health and hydroelectricity in Norway in 1900-1921. Specifically, we have collected variables on construction year for hydroelectricity plants at municipality level, as well as infant mortality rates, diarrhea cases and tuberculosis cases at medical district level.

The results imply that access to a clean energy source has positive societal impacts. Specifically, we find indications that infant mortality decrease by 4.74 percent, tuberculosis cases decrease by 8.84 percent and diarrhea cases decrease by 0.60 percent in the year of opening. These estimates are intention-to-treat effects, and the estimated reductions increase when we adjust for the uptake rate. Overall, the estimated results pass several robustness tests.

This thesis contributes to prior research on health effects of electrification by isolating the effect from gaining access to simple, low-level electric appliances, mainly electric lights. Due to the low development levels of Norway at the time of our sample, we argue that the sign of the effect could be comparable in developing countries today.

In the early 1900s, Norway was relatively poor as compared to the US, but observed significantly lower infant mortality rates. Also, Norway was among the earliest adaptors of electricity. Therefore, the positive effect on infant mortality from hydro electrification could contribute to explaining the low infant mortality rates observed in Norway as compared to other western nations.

Our results also have implications for current policies, as about 15 percent of the world population lack access to electricity, according to the World Bank (2012) - almost all of whom live in developing countries. Taken together, our results imply that investing in hydroelectricity or other clean energy sources in developing countries will have positive externalities on health.

10 References

Arceo, E., Hanna, R., and Oliva, P. (2016). Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. *The Economic Journal*, 126(591), pp. 257-80.

Akerman, A., Gaarder, I., and Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *The Quarterly Journal of Economics*, 130(4), pp. 1781-824.

Angrist, J. D., and Pischke, J. (2009). Mostly Harmless Econometrics – An Empiricist's Companion. *1st ed*. Princeton and Oxford: Princeton University Press.

Angrist, J. D., and Pischke, J. (2015). Mastering 'Metrics – The path from cause to effect. *1st ed*. Princeton: Princeton University Press.

Apple, J., Vicente, R., Yarberry, A., Lohse, N., Mills, E., Jacobson, A., and Poppendieck, D. (2010). Characterization of particulate matter size distributions and indoor concentrations from kerosene and diesel lamps. *Indoor Air*, 20(5), pp. 399-411.

Backer, J. (1961). Dødeligheten og dens årsaker i Norge 1856 til 1955. Retrieved from Statistics Norway on 8 April, 2016, from https://www.ssb.no/a/histstat/sos/sos_010.pdf

Bailey, M., and Goodman-Bacon, A. (2015). The War on Poverty's Experiment in Public Medicine: The Impact of Community Health Centers on the Mortality of Older Americans, *American Economic Review*, 105(3), pp. 1067-104.

Barnett, E., and Casper, M. (2001). A Definition of "Social Environment". *American Journal of Public Health,* 91(3), pp. 465.

Borza, T. (2001). Spanskesyken i Norge 1918 – 19. *Tidsskrift for Den norske legeforeningen*, 121(30), pp. 3551-4.

Brazier, E., and Kirkhusmo, A. (2016). Første Verdenskrig. Retrieved from Store Norske Leksikon on 19 April, 2016, from https://snl.no/F%C3%B8rste_verdenskrig

Bhargava, A., Jamison, D., Lau, L., and Murray, C. (2001). Modeling the effects of health on ecnomic growth. *Journal of Health Economics*, 20(3), pp. 423-40.

Bruce, N., Perez- Padilla, R., and Albalak, R. (2000). Indoor air pollution in developing countries: a major environmental and public health challenge. *Bulletin of the World Health Organization*, 78(9), pp. 1078-92.

Bütikofer, A., Løken, K., and Salvanes, K. (2015). Long-Term Consequences of Access to Well-Child Visits. Discussion Paper 9546, IZA.

Cameron, A.C., and Miller, D. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2), pp. 317-72.

Cattaneo, M., Galiani, S., Gertler, P., Martinez, S., and Titiunik, R. (2009). Housing, Health, and Happiness. *American Economic Journal: Economic Policy*, 1(1), pp. 75-105.

Clay, K., Lewis, J., and Severnini, E. (2015). Canary in a Coal Mine: Impact of Mid-20th Century Air Pollution Induced by Coal-Fired Power Generation on Infant Mortality and Property Values. Retrieved 26 May, 2016, from https://www.webdepot.umontreal.ca/Usagers/lewisj/MonDepotPublic/clay_lewis_severnini_pollution.pdf

Daware, K. (2016). Hydroelectric Power Plant: Layout, Working and Types. Retrieved 23 May, 2016, from http://www.electricaleasy.com/2015/09/hydroelectric-power-plant-layout.html

Eika, T. (2008). Det svinger i Norsk økonomi. *Samfunnsspeilet/Statistics Norway*, 22 (5-6/2008), pp. 98-108. Retrieved 2 May, 2016, from https://www.ssb.no/a/samfunnsspeilet/utg/200805/ssp.pdf

European Commission. (1995). Externalities of Energy - Volume 6: Wind & Hydro. Brussels: European Commission, pp.135 - 290. Retrieved 16 March, 2016, from http://www.externe.info/externe_d7/sites/default/files/vol6.pdf

Faugli, P. (2012). Vann- og energiforvaltning – glimt fra NVEs historie. Retrieved from NVE on 23 February, 2016, from http://webby.nve.no/publikasjoner/rapport/2012/rapport2012_26.pdf

Gapminder. (2016 I). Data: Income per person (GDP/capita, PPP$ inflation-adjusted). Retrieved 22 May, 2016, from http://www.gapminder.org/data/

Grønlie, T. (2015). Helsereformer. Retrieved from Store norske leksikon on 16 March, 2016, from https://snl.no/helsereformer

Hveding, V. (1992). Vannkraft i Norge. Trondheim: Sats, repro og trykk AS.

Joseph, A. (2006). The Impact of Light on Outcomes in Healthcare Settings. Retrieved from Concord: The Centre of Health Design 11 May, 2016, from https://www.healthdesign.org/sites/default/files/CHD_Issue_Paper2.pdf

Judkvam, D. (1999). Historisk oversikt over endringer i kommune- og fylkesinndelingen. Retrieved from SSB on 20 February, 2016, from https://www.ssb.no/a/publikasjoner/pdf/rapp_9913/rapp_9913.pdf

Kjeldstadli, K. (2015). Da spytting kunne være dødelig. Retrieved from University of Oslo on 2 May, 2016, from https://www.norgeshistorie.no/forste-verdenskrig-og-mellomkrigstiden/hus-og-hjem/1605-da-spytting-kunne-vaere-dodelig.html

Kose, E., Kuka, E., and Shenhav, N. (2015). Women's Enfranchisement and Children's Education: the Long-Run Impact of the U.S. Suffrage Movement. UC Davis mimeo.

Kristiania Elektricitetsverk (1916), Beretning og regnskap for 1915, Oslo: Kristiania Elektricitetsverk.

Lago-Peñas, S., Cantarero Prieto, D., and Blázquez Fernández, C. (2013). On the relationship between GDP and health care expenditure: A new look. *Economic Modelling*, 32, pp.124-9.

Lam, N., Smith, K., Gauthier, A., and Bates, M. (2012 I). Kerosene: A Review of Household Uses and their Hazards in Low- and Middle- Income Countries. *Journal of Toxicology and Environmental Health, Part B*, 15(6), pp. 396-432.

Lam, N., Chen, Y., Weyant, C., Venkataraman, C., Sadavarte, P., Johnson, M., Smith, K., Brem, B., Arineitwe, J., Ellis, J., and Bond, T. (2012 II). Household Light Makes Global Heat: High Black Carbon Emissions From Kerosene Wick Lamps. *Environmental Science & Technology*, 46(24), pp. 13531-8.

Larsen, Ø. (2003). Har helsevesenet sin egen historie?. *Tidsskrift for Den norske legeforening*, 123(12), pp.1718-20.

Liestøl, K., Tretli, S., Tverdal, A., and Mæhlen, J. (2007). Hvem fikk tuberkulose - og var de generelt skrøpelige? På liv og død: Helsestatistikk i 150 år. *Statistisk Sentralbyrå: Historisk Helsestatistikk*, 266, pp. 120-34.

Linder, F., and Grove, R. (1947). Vital Statistics Rates in the United States 1900-1940. In *Federal Securuty Agency United States Health Services - National Office of Vital Statistics*, pp. 572. Washington: Government Printing Office.

Lu, X., and White, H. (2014). Robustness checks and robustness tests in applied economics. *Journal of Econometrics*, 178(1), pp. 194-206.

Lewis, J. (2014). Fertility, Child Health, and the Diffusion of Electricity into the Home. Retrieved 26 May, 2016 from https://www.webdepot.umontreal.ca/Usagers/lewisj/MonDepotPublic/fertilityhealth.pdf

Lyse. (2016). Stor produsent av vannkraft. Retrieved 21 April, 2016, from http://www.lysekonsern.no/vannkraft/category836.html

Miller, G. and Urdinola, B. P., (2010). Cyclicality, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Colombia. *Journal of Political Economy*, 118(1), pp. 113-55

Nordhagen, R., Major, E., Tverdal, A., Irgens, L., and Graff Iversen, S. (2014). Folkehelse i Norge 1814 - 2014. Retrieved from the Norwegian Insitute of Publich Health on 8 April, 2016, from http://www.fhi.no/publikasjoner-og-haandboker/folkehelserapporten/folkehelse-i-historien

Norwegian Meteorological Institute. (2016). Nord-Norge. Retrieved 23 May, 2016, from http://met.no/Nord-Norge.b7C_wljO3i.ips

Norwegian Water Resources and Energy Directorate (NVE). (1922). Utbygget Vannkraft i Norge. Kristiania: H. Aschehoug & co.

Norwegian Water Resources and Energy Directorate (NVE). (1930). Utbygget Vannkraft i Norge. Oslo: H. Aschehoug & co.

Norwegian Water Resources and Energy Directorate (NVE). (2013). Kulturminner i norsk kraftproduksjon. *NVE-rapport,* 2013(52). Oslo: NVE.

Norwegian Water Resources and Energy Directorate (NVE). (2015). *Kraftverk: Oltedal*. Retrieved 11 May, 2016, from https://www.nve.no/vann-vassdrag-og-miljo/kulturminner/kraftverk/oltedal/

NSD. (1900). NSDs Kommunedatabase: Fattigstatistikk 1900: Hovedpersoner i alt. Retrieved March 20, 2016, from https://trygg.nsd.uib.no/kdbbin/search.exe

Norwegian Centre for Research Data (NSD). (1905). NSDs Kommunedatabase: Fattigstatistikk 1905: Hovedpersoner i alt. Retrieved March 20, 2016, from https://trygg.nsd.uib.no/kdbbin/search.exe

Norwegian Centre for Research Data (NSD). (1910). NSDs Kommunedatabase: Fattigstatistikk 1910: Hovedpersoner i alt. Retrieved March 20, 2016, from https://trygg.nsd.uib.no/kdbbin/search.exe

Norwegian Centre for Research Data (NSD). (1915). NSDs Kommunedatabase: Fattigstatistikk 1915: Hovedpersoner i alt. Retrieved March 20, 2016, from https://trygg.nsd.uib.no/kdbbin/search.exe

Norwegian Centre for Research Data (NSD). (1920). NSDs Kommunedatabase: Fattigstatistikk 1920: Hovedpersoner i alt. Retrieved March 20, 2016, from https://trygg.nsd.uib.no/kdbbin/search.exe

Nøkleby, H. and Feiring, B. (2006). Det norske vaksinasjonsprogrammet. *Tidsskrift for Den norske legeforeningen*, 126(19), pp. 2538-40.

Oslo Lysverker. (1952). Oslo Elektrisitetsverk 1892 – 1952, Oslo: Oslo Lysverker.

Pedersen, A. (2003). 100 år med redusert spedbarnsdødelighet. *Samfunnsspeilet/Statistics Norway,* 17(3/2003), pp. 34-9. Retrieved 10 February, 2016, from https://www.ssb.no/a/samfunnsspeilet/utg/200303/ssp.pdf

Riksantikvaren (2013). Hammeren Kraftverk - Nordmarkveien Kraftverk. Kulturminnesøk. Retreived 21 April, 2016, from http://www.kulturminnesok.no/Lokaliteter/Oslo/Oslo/Kraftverk-Kristiania-Elektricitetsverk-Hammeren-kraftverk-Nordmarkveien-2-og-6

Rosvold, K (2010). Vannfallsenergi. Retrieved from Store Norske Leksikon on 27 May, 2016, from https://snl.no/vannfallsenergi

Rosvold, K. (2012). Frøland kraftverk. Retrieved from Store Norske Leksikon on 11 May, 2016, from https://snl.no/Frøland_kraftverk

Ryymin, T. (2008). Forebygging av tuberkulose i Finnmark 1900 – 60. *Tidsskrift for Den norske legeforening*, 128(24), pp. 2864-7.

Sandberg J. (1951). Trekk fra elektrisitetsforsyningens utvikling, Del 1, *1st ed*. Oslo : Norske elektrisitetsverkers forening.

Savitha, M., Nandeeshwara, S., Pradeep Kumar, M., Farhan-ul-haque, and Raju, C. (2007). Modifiable risk factors for acute lower respiratory tract infections. *Indian Journal of Pediatrics*, 74(5), pp. 477-82.

Sejersted, F. (2015). Unionsoppløsningen i 1905. Retrieved from Store Norske Leksikon on 25 April, 2016, from https://snl.no/Unionsoppl%C3%B8sningen_i_1905

Severnini, E. R. (2014) Air Pollution, Power Grid, and Infant Health: Evidence from the Shutdown of TVA Nu.clear Power Plants in the 1980s. Retrieved 2 May, 2016, from http://repository.cmu.edu/cgi/viewcontent.cgi?article=1380&context=heinzworks

Sovacool, B., and Vera, I. (2014). Electricity and education: The benefits, barriers, and recommendations for achieving the electrification of primary and secondary schools. Retrieved from UNDESA on 22 April, 2016, from http://www.un-energy.org/sites/default/files/share/une/undesa-energy-education-121714-final.pdf

Statistics Norway (SSB). (1901). Meddelelser fra Det Statistiske Centralbureau. Attende Bind. 1900. Kristiania: SSB. Retrieved 19 April, 2016, from https://www.ssb.no/a/histstat/sm/sm_1900.pdf

Statistics Norway (SSB). (1902). Beretning om Sundhetstilstanden og Medicinalforholdene i Norge, 1900. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_iv_055.pdf

Statistics Norway (SSB). (1903). Beretning om Sundhetstilstanden og Medicinalforholdene i Norge, 1901. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_iv_077.pdf

Statistics Norway (SSB). (1904). Beretning om Sundhetstilstanden og Medicinalforholdene i Norge, 1902. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_iv_103.pdf

Statistics Norway (SSB). (1905). Beretning om Sundhetstilstanden og Medicinalforholdene i Norge, 1903. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_iv_128.pdf

Statistics Norway (SSB). (1906 I). Sundhetstilstanden og Medicinalforholdene 1904. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_v_023.pdf

Statistics Norway (SSB). (1906 II). Meddelelser fra Det Statistiske Centralbureau. Treogtyvende Bind. 1905. Kristiania: SSB. Retreived 19 April, 2016, from: https://www.ssb.no/a/histstat/sm/sm_1905.pdf

Statistics Norway (SSB). (1908). Sundhetstilstanden og Medicinalforholdene i Norge 1905. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_v_055.pdf

Statistics Norway (SSB). (1908). Sundhetstilstanden og Medicinalforholdene 1906. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_v_072.pdf

Statistics Norway (SSB). (1909). Sundhetstilstanden og Medicinalforholdene 1907. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_v_098.pdf

Statistics Norway (SSB). (1910). Sundhetstilstanden og Medicinalforholdene 1908. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_v_122.pdf

Statistics Norway (SSB). (1911 I). Sundhetstilstanden og Medicinalforholdene 1909. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_v_152.pdf

Statistics Norway (SSB). (1911 II). Meddelelser fra Det Statistiske Centralbyraa. Otteogtyvende Bind. 1910. Kristiania: SSB. Retreived 19 April, 2016, from https://www.ssb.no/a/histstat/sm/sm_1910.pdf

Statistics Norway (SSB). (1912). Sundhetstilstanden og Medicinalforholdene 1910. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_v_181.pdf

Statistics Norway (SSB). (1913). Sundhetstilstanden og Medicinalforholdene 1911. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_v_216.pdf

Statistics Norway (SSB). (1914). Sundhetstilstanden og Medicinalforholdene 1912. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vi_019.pdf

Statistics Norway (SSB). (1915). Sundhetstilstanden og Medicinalforholdene 1913. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vi_056.pdf

Statistics Norway (SSB). (1916). Meddelelser fra Det Statistiske Centralbyraa. 33te Bind - 1915. Kristiania: SSB. Retreived 19 April, 2016 from https://www.ssb.no/a/histstat/sm/sm_1915.pdf

Statistics Norway (SSB). (1917). Sundhetstilstanden og Medicinalforholdene 1914. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vi_094.pdf

Statistics Norway (SSB). (1918). Sundhetstilstanden og Medicinalforholdene 1915. Kristiania: Direktøren for det civile Medicinalvæsen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vi_133.pdf

Statistics Norway (SSB). (1920). Sundhetstilstanden og Medisinalforholdene 1916. Kristiania: Direktøren for det civile Medisinalvesen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vi_186.pdf

Statistics Norway (SSB). (1921 I). Sundhetstilstanden og Medisinalforholdene 1917. Kristiania: Direktøren for det civile Medisinalvesen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vii_003.pdf

Statistics Norway (SSB). (1921 II). Meddelelser fra Det Statistiske Centralbyrå. 38te Bind - 1920. Kristiania: SSB. Retreived 19 April, 2016, from https://www.ssb.no/a/histstat/sm/sm_1920.pdf

Statistics Norway (SSB). (1922). Sundhetstilstanden og Medisinalforholdene 1918. Kristiania: Direktøren for det civile Medisinalvesen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vii_058.pdf

Statistics Norway (SSB). (1923). Sundhetstilstanden og Medisinalforholdene 1919. Kristiania: Direktøren for det civile Medisinalvesen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vii_108.pdf

Statistics Norway (SSB). (1924). Sundhetstilstanden og Medisinalforholdene 1920. Kristiania: Direktøren for det civile Medisinalvesen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vii_138.pdf

Statistics Norway (SSB). (1925). Sundhetstilstanden og Medisinalforholdene 1921. Kristiania: Direktøren for det civile Medisinalvesen. Retrieved 7 January, 2016, from http://www.ssb.no/a/histstat/nos/nos_vii_152.pdf

Statistics Norway (SSB). (2001). Elektrisitet 1937-1961 - Det lyste dog som en stjerne. Samfunnsspeilet/*Statistics Norway*. (1/ 2001), pp. 94-107.

Statistics Norway (SSB). (2016). Statistikkbanken: Konsumprisindeksen. Retreived 19 April, 2016, from https://www.ssb.no/statistikkbanken/selectvarval/Define.asp?subjectcode=&ProductId=&MainTable=KpiAarHist&nvl=&PLanguage=0&nyTmpVar=true&CMSSubjectArea=priser-og-prisindekser&KortNavnWeb=kpi&StatVariant=&checked=true

Statnett. (2013). Første vannkraftverk i Norge og første i Vesterålen. Retreived 22 February, 2016, from http://www.statnett.no/Media/Nyheter/Sentralnettet-pa-kryss-og-tvers/Blogginnlegg/Forste-vannkraftverk-i-Norge-og-forste-i-Vesteralen/

Stene-Larsen, G. (2006). 1880 - 2005 - fra fattigdomssykdommer til overflodslidelser. *Tidsskrift for Den norske legeforening*, 126(1), pp. 38-43.

Swift, R. (2011). The relationship between health and GDP in OECD countries in the very long run. *Health Economics,* 20(3), pp. 306-22.

Trondheim Energiverk. (2001). Glimt fra Trondheim Energiverks Historie. Retrieved 21 April, 2016, from http://www.strindahistorielag.no/Wikibilder/TEV%2520historien%2520kort.pdf

Trueman, C. N. (2016). Antibiotics. Retrieved 30 May, 2016 from http://www.historylearningsite.co.uk/a-history-of-medicine/antibiotics/

UNICEF. (2015). Why sustainable energy matters to children: The critical importance of sustainable energy for children and future generations. Retrieved 16 May, 2016, from http://www.unicef.org/environment/files/UNICEF_Sustainable_Energy_for_Children_2015.pdf

Vasskrafta. (2016). Historie - Vasskrafta: Hammeren Kraftverk. Retrieved 11 May, 2016, from http://www.vasskrafta.no/kraftverk/hammeren/historie/

Vogt, J. (1971). Elektrisitetslandet Norge: Fra norsk vassdrags- og elektrisitetsvesens historie. Oslo: Universitetsforlaget,

Woolridge, J.M. (2014). Introduction to Econometrics. *Europe, Middle East & Africa Edition*. Hampshire: Cengage Learning EMEA.

World Bank. (2012). Data: Access to electricity (% of population). Retreived 2 May, 2016, from http://data.worldbank.org

World Health Organization (WHO). (2016 I). Infant mortality. Retrieved 22 May, 2016, from http://www.who.int/gho/child_health/mortality/neonatal_infant_text/en/

World Health Organization (WHO). (2016 II). Tuberculosis. Retrieved 23 February, 2016, from http://www.who.int/mediacentre/factsheets/fs104/en/

World Health Organization (WHO). (2016 III). Diarrhoeal disease. Retrieved 22 May, 2016, from http://www.who.int/mediacentre/factsheets/fs330/en/

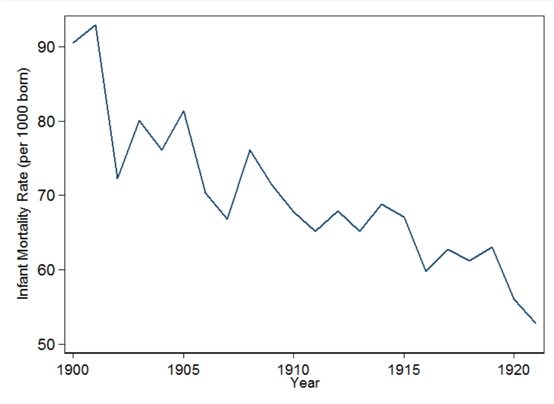
World Health Organisation (WHO). (2016 IV). The determinants of health. Retrieved 15 March, 2016, from http://www.who.int/hia/evidence/doh/en/

Øberg, M., and Bråten, J. (2014). Samfunnsoppdrag - det var en gang. Retrieved from Statnett on 15 May, 2016, from http://www.statnett.no/Samfunnsoppdrag/Neste-generasjon-kraftsystem/Det-var-en-gang/

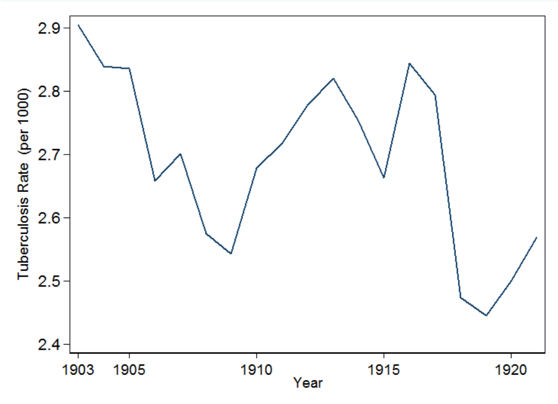
11 Appendix

A1: Descriptive Statistics

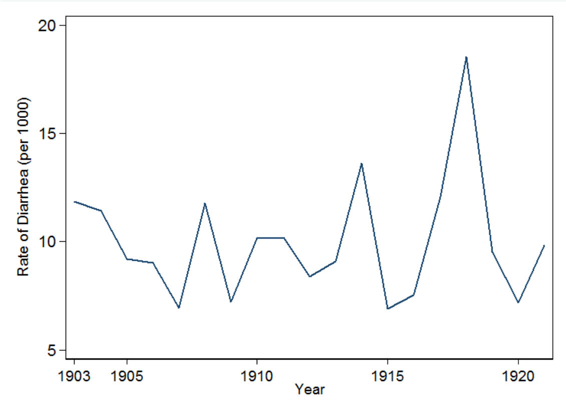
Figure A1: Developments in infant mortality, tuberculosis cases and diarrhea cases in Norway



(a) Infant mortality



(b) Tuberculosis cases



(c) Diarrhea cases

Notes: The figures show the development in each of the dependent variables during the period of interest. The figures show (a) infant deaths age 0-1 per 1000 born 1900-1921, (b) the number of tuberculosis cases per 1000 inhabitant 1903-1921 and (c) the number of diarrhea cases per 1000 inhabitant 1903-1921.

Table A1: Summary statistics: Infant mortality per 1000 born for different regions

	(1) Observations	(2) Mean	(3) Std.Dev.	(4) Min	(5) Max
Eastern Norway	4,699	61.84	24.79	0	474.3
Southern Norway	1,670	60.06	23.25	0	180
Western Norway	4,468	54.62	22.36	0	166.7
Trøndelag	1,110	61.59	22.94	0	148.2
Northern Norway	3,355	75.48	41.18	0	500

Notes: The table shows (1) the number of observations, (2) the mean value, (3) the standard deviation from the mean value, (4) the minimum value, and (5) the maximum value for the number of infant deaths age 0 to 1 per 1000 born in Norway 1900-1921.

Table A2: Summary statistics: Tuberculosis in different regions

	(1) Observations	(2) Mean	(3) Std.Dev.	(4) Min	(5) Max
Eastern Norway	3,839	2.400	2.706	0	38.56
Southern Norway	1,361	2.828	1.134	0	7.636
Western Norway	3,640	2.465	1.160	0	15.36
Trøndelag	897	4.009	1.710	0	10.50
Northern Norway	2,686	2.924	1.719	0	17.30

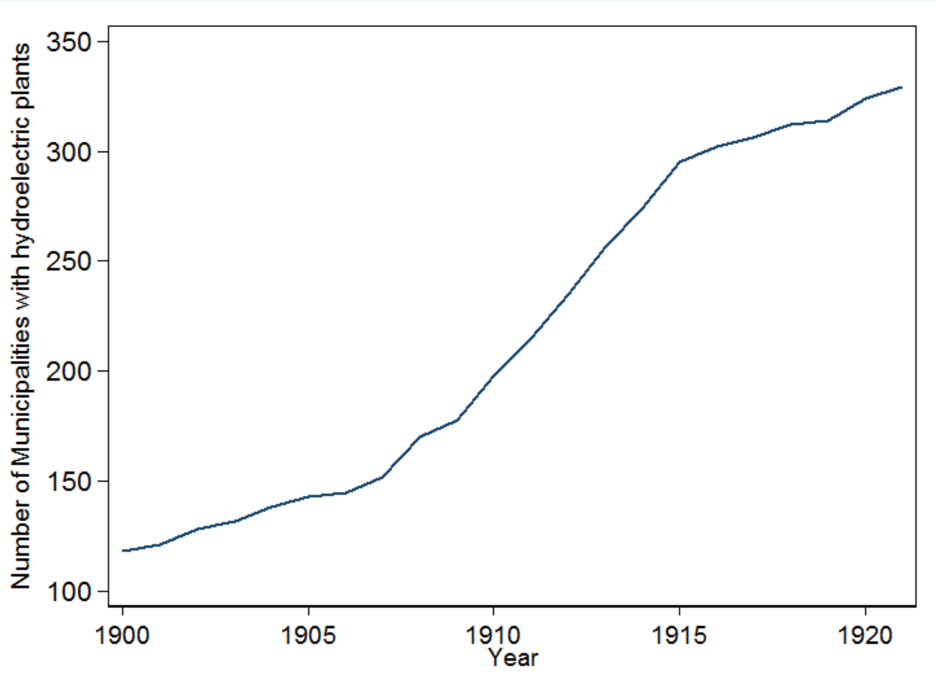
Notes: The table shows (1) the number of observations, (2) the mean value, (3) the standard deviation from the mean value, (4) the minimum value, and (5) the maximum value for the number of tuberculosis cases per 1000 inhabitants in Norway 1903-1921.

Table A3: Summary statistics: Diarrhea in different regions

	(1) Observations	(2) Mean	(3) Std.Dev.	(4) Min	(5) Max
Eastern Norway	4,023	11.72	16.72	0	269.7
Southern Norway	1,395	6.090	4.810	0	48.68
Western Norway	3,743	5.147	5.929	0	63.78
Trøndelag	899	5.647	6.226	0	64.44
Northern Norway	2,740	6.568	7.198	0	79.27

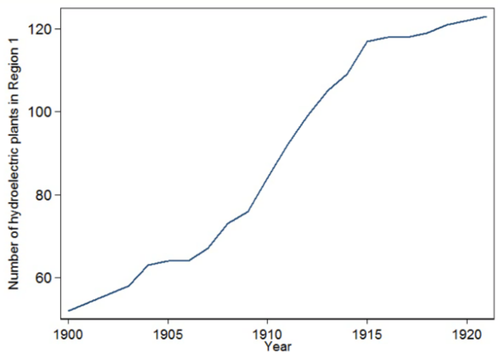
Notes: The table shows (1) the number of observations, (2) the mean value, (3) the standard deviation from the mean value, (4) the minimum value, and (5) the maximum value for the number of diarrhea cases per 1000 inhabitants in Norway 1903-1921.

Figure A2: The number of municipalities with a hydroelectricity plant

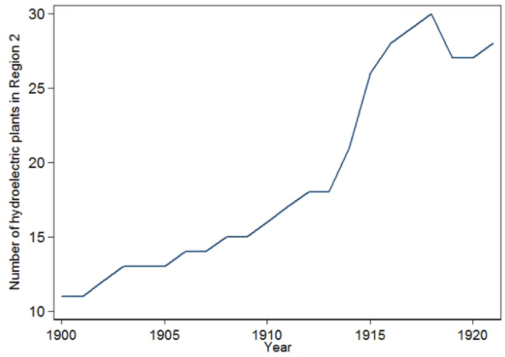


Notes: The figure shows the accumulated number of municipalities with hydroelectricity plants with known construction years between 1900 and 1921.

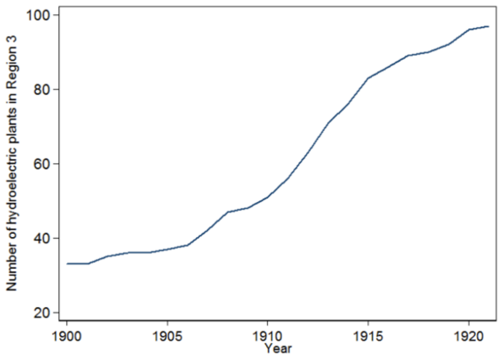
Figure A3: Number of municipalities with hydroelectricity plants per region



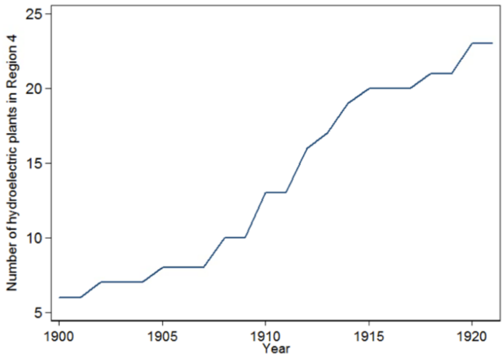
(a) Eastern Norway



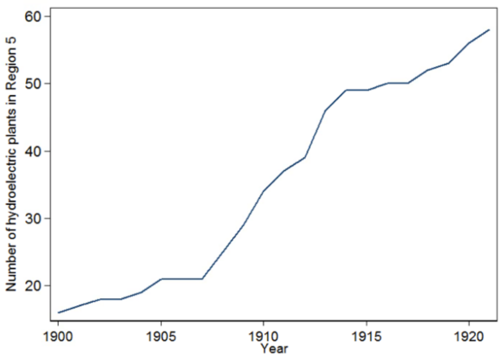
(b) Southern Norway



(c) Western Norway



(d) Trøndelag



(e) Northern Norway

Notes: The figures show the development in the number of municipalities with hydroelectricity plants opened within our period of interest per region. Only plants with a known construction year is included, so the graphs show the number of plants in our data set and not the real number of plants in Norway.

A2: Standard Errors Clustered at a Higher Level

Standard errors could be correlated across municipalities over time if there are area-year shocks. For example, municipalities in the same county might have similar health outcomes because they have the same environmental and cultural influences. Furthermore epidemics that last for a long period and spread across municipalities within a county could cause such correlation. Clustering at a higher level allows for the standard errors of municipalities within the same group to be correlated with each other and over time.

In order to test for clustering of standard errors at a higher level, we cluster at county level. The results can be seen in Table A4 below, and show that the standard errors of the results for tuberculosis and diarrhea are not sizably altered. The level of significance for infant mortality is reduced from 5 percent to 10 percent. However, clustered standard errors are unlikely to be reliable with few clusters, as it might lead to underestimation of either the serial correlation over time in a random shock, or the correlation between municipalities (Angrist and Pischke, 2009). This could again lead to a "over-rejection" of the estimated values, which could explain the lower significance level for infant mortality rates (Cameron and Miller, 2015). There are ways of mitigating the problem with few clusters, see Angrist and Pischke (2015, pp.320-322) for a description.

Another option would be to cluster at medical district level, which would allow for a higher number of clusters. This could make sense e.g. because all municipalities within the same medical district share doctors. Thus, clustering at medical district level could be optimal, as these municipalities might share common health characteristics. We have not established medical districts as its own level of analysis in our data set, since there are significant boarder adjustments and name changes in the medical districts during our period of interest.

Table A4: Standard errors clustered at county level

	(1) Infant Mortality	(2) Tuberculosis	(3) Diarrhea
Main	-2.931*	-0.234***	-0.038
	(1.525)	(0.068)	(0.280)
No. of Clusters	20	20	20
Observations	15,302	12,423	12,800

Standard errors in parentheses

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

Notes: Each parameter is from a separate regression of the outcome variables on the access to a hydroelectricity plant, based on the model in Equation 2. Control variables include doctors per 1000 inhabitants, tax income per tax payer and the number of poor people per 1000 inhabitants, as well as indicator variables for regions and years. The standard errors are clustered at county level and robust to heteroscedasticity. The period of interest is 1900 – 1921 for infant mortality, and 1903 – 1921 for tuberculosis and diarrhea cases. Column (1) show the results for infant mortality (the number dead infants between age 0 and 1 per 1000 born); Column (2) show the results for the number of tuberculosis cases per 1000 inhabitants; Column (3) show the number of diarrhea cases per 1000 inhabitants.

A3: Test of the Key Identifying Assumption

Table A5: Test of identifying assumption: The effect of municipality characteristics on the timing of hydroelectricity plant opening

	1900 Municipality Characteristics		Changes in Municipality Characteristics (1900 to 1910)	
	(1)	(2)	(3)	(4)
	Opening 1900-1921	Opening Year	Opening 1900-1921	Opening Year
Doctors	0.0574***	-8.419**	-0.0129	-4.037
	(0.00741)	(3.487)	(0.0283)	(3.281)
Tax Income	-0.000224***	0.00768***	0.0000835	-0.00702***
	(0.0000786)	(0.00271)	(0.0000748)	(0.00196)
Poverty	0.00811***	0.0179	-0.0120**	-0.0709
	(0.00301)	(0.0523)	(0.00505)	(0.138)
No. of Clusters	714	714	714	714
Observations	591	591	591	591

Standard errors in parentheses

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

Notes: Columns (1) and (3) represent linear probability models of the likelihood of a hydroelectricity station being built between 1900 and 1921 in relation to observed municipality and medical district characteristics, or early changes in these characteristics. Columns (2) and (4) represents regressions of the year a hydroelectricity station were built on our observed municipality and medical district characteristics, or early changes in these characteristics. Doctors are calculated as the number of doctors per 1000 inhabitants. Tax Income is calculated as tax income per taxpayer. Poverty is calculated as the number of poor people per 1000 inhabitants.

A4: Results from the Sensitivity Analysis

Table A6: Extreme observations excluded

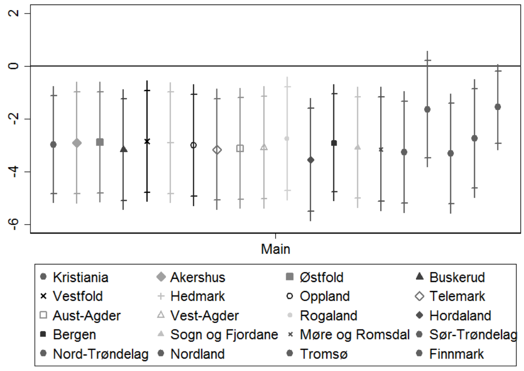
	(1) Baseline Results	(2) Excluding Largest Observations	(3) Excluding the 4 Largest Cities	(4) Excluding Finnmark
Infant Mortality	-2.931**	-2.908***	-3.021***	-1.557*
	(1.124)	(1.120)	(1.130)	(0.831)
Observations	15302	15295	15214	14796
Tuberculosis	-0.234 ***	-0.193***	-0.222***	-0.191***
	(0.054)	(0.0458)	(0.0538)	(0.0523)
Observations	12423	12401	12351	12056
Diarrhea	-0.038	0.138	-0.0526	-0.0594
	(0.330)	(0.321)	(0.328)	(0.331)
Observations	12800	12786	12724	12410
No. of Clusters	714	714	714	714

Standard errors in parentheses

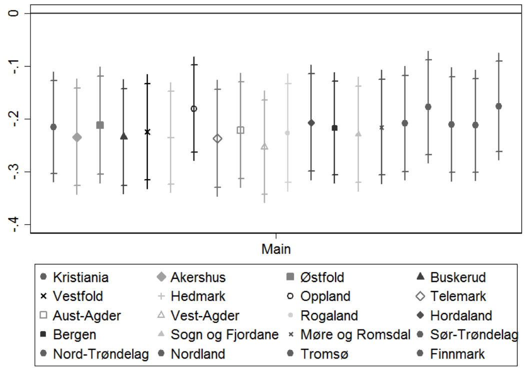
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Each parameter is from a separate regression of the outcome variables on the access to a hydroelectricity plant, based on the model in Equation 2. Outcome variables are infant mortality (the number dead infants between age 0 and 1 per 1000 born), tuberculosis rates (the number of tuberculosis cases per 1000 inhabitants), and diarrhea cases (the number of diarrhea cases per 1000 inhabitants. Control variables include doctors per 1000 inhabitants, tax income per tax payer and the number of poor people per 1000 inhabitants, as well as indicator variables for regions and years. The standard errors are clustered at county level and robust to heteroscedasticity. The period of interest is 1900 – 1921 for infant mortality, and 1903 – 1921 for tuberculosis and diarrhea cases. Column (1) show the baseline for all outcome variables where all observations are included; Column (2) show the estimates for each outcome variable when the largest observations of the outcome variable is excluded. We exclude the observations of infant mortality above 400 deaths per 1000 live born, the observations of tuberculosis rates above 30 incidents per 1000 inhabitants, and the observations of diarrhea above 100 incidents per 1000 inhabitants; Column (3) show the estimates for each outcome variable when Oslo, Bergen, Stavanger and Trondheim is excluded; Column (4) show the estimates for each outcome variable when Finnmark is excluded.

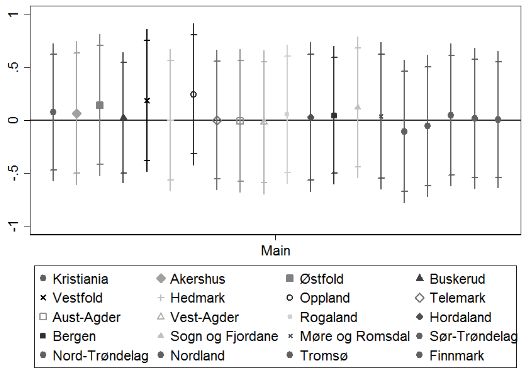
Figure A4: Leaving all counties out one at a time



(a) Infant Mortality



(b) Tuberculosis case



(c) Diarrhea cases

Notes: Each plot in the figures stems from a separate regression, where each county is excluded one at a time. The county excluded is indicated. The plots show the estimated results, as well as the 90 percent and 95 percent confidence intervals. Tax income, poverty levels and the number of doctors per 1000 inhabitants is included as control variables. The standard errors are robust at municipality level. The period of interest is 1900 – 1921 for infant mortality rates and 1903 – 1921 for Diarrhea and Tuberculosis cases per 1000 inhabitant. Figure (a) show the effect on infant mortality per 1000 born, figure (b) show the effect on tuberculosis cases per 1000 inhabitants, while figure (c) show the effect on diarrhea cases per 1000 inhabitants.

Table A7: Excluding historic events

	(1) Baseline	(2) 1917-1918, World War I	(3) 1918-1919, The Spanish Influenza	(4) Pre 1905, The Swedish-Norwegian union
Infant Mortality	-2.931***	-2.340**	-2.235*	-4.118***
	(1.124)	(1.143)	(1.151)	(1.119)
Observations	15,302	13,920	13,925	11,108
Tuberculosis	-0.234***	-0.221***	-0.229***	-0.256***
	(0.0540)	(0.0497)	(0.0534)	(0.0568)
Observations	12,423	11746	11,746	10,344
Diarrhea	-0.0383	0.196	0.0718	-0.0979
	(0.331)	(0.292)	(0.310)	(0.340)
Observations	12,800	11,542	11,506	10,721
No. of Clusters	714	714	714	714

Standard errors in parentheses

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

Notes: Each parameter is from a separate regression of the outcome variables on the access to a hydroelectric plant. All specifications include standard errors robust at municipality level, and tax income, poverty levels and the number of doctors per 1000 inhabitant is included as control variables. Our period of interest is 1900 – 1921 for infant mortality (the number dead infants between age 0 and 1 per 1000 born), and 1903 – 1921 for diarrhea cases (described as the number of cases per 1000 inhabitants) and tuberculosis cases (described as the number of cases per 1000 inhabitants). Column (1) show the baseline for all outcome variables where all observations are included; Column (2) show the estimates for each outcome variable when the last two years of World War I is excluded; Column (3) show the estimates for each outcome variable when the Spanish Influenza is excluded; Column (4) show the estimates for each outcome variable when the Norwegian-Swedish union is excluded.

Table A8: Estimates when removing municipalities without a hydroelectricity plant

	(1) Infant Mortality	(2) Tuberculosis	(3) Diarrhea
Effect from Hydroelectricity	0.803 (2.102)	-0.178 (0.111)	0.192 (0.626)
No. of Clusters	714	714	714
Observations	4832	3940	4076

Standard errors in parentheses

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

Notes: Each parameter is from a separate regression of the outcome variables on the access to a hydroelectricity plant, based on the model in Equation 2. Control variables include doctors per 1000 inhabitants, tax income per tax payer and the number of poor people per 1000 inhabitants, as well as indicator variables for regions and years. The standard errors are clustered at county level and robust to heteroscedasticity. The period of interest is 1900 – 1921 for infant mortality, and 1903 – 1921 for tuberculosis and diarrhea cases. Column (1) shows the results for infant mortality (the number dead infants between age 0 and 1 per 1000 born); Column (2) shows the results for tuberculosis rates (the number of tuberculosis cases per 1000 inhabitants); Column (3) shows the results for diarrhea rates (the number of diarrhea cases per 1000 inhabitants).

Table A9: Excluding municipalities with hydroelectricity plants with unknown construction year

	(1)	(2)
	Baseline	No Unknown Construction Years
Infant Mortality	-2.931***	-3.181**
	(1.124)	(1.472)
Observations	15302	12869
Tuberculosis	-0.234***	-0.289***
	(-0.054)	(0.0669)
Observations	12423	10436
Diarrhea	-0.038	-0.000285
	(0.330)	(0.000385)
Observations	12800	10729
No. of Clusters	714	714

Standard errors in parentheses

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

Notes: Each parameter is from a separate regression of the outcome variables on the access to a hydroelectricity plant, based on the model in Equation 2. All regressions include standard errors robust at municipality level, and tax income per tax payer, number of poor people per 1000 inhabitants and the number of doctors per 1000 inhabitant is included as control variables, as well as indicator variables for regions and years. Our period of interest is 1900 – 1921 for infant mortality (the number of dead infants between age 0 and 1 per 1000 born), and 1903 – 1921 for diarrhea cases (described as the number of cases per 1000 inhabitants) and tuberculosis cases (described as the number of cases per 1000 inhabitants). Column (1) show the baseline for all outcome variables where all observations are included; Column (2) show the estimates for each outcome variable the sample is restricted to municipalities with known construction years only.

Table A10: Estimates with region-specific linear time trends

	(1) Infant Mortality	(2) Tuberculosis	(3) Diarrhea
Effect from Hydroelectricity	-3.021***	-0.235***	-0.0682
	(1.124)	(0.0541)	(0.329)
No. of Clusters	714	714	714
Observations	15,302	12,423	12,800

Standard errors in parentheses

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

Notes: Each parameter is from a separate regression of the outcome variables on the access to a hydroelectricity plant, based on the model in Equation 4. Control variables include doctors per 1000 inhabitants, tax income per tax payer and the number of poor people per 1000 inhabitants, as well as indicator variables for regions and years. The standard errors are clustered at county level and robust to heteroscedasticity. Our period of interest is 1900 – 1921 for infant mortality (the number of dead infants between age 0 and 1 per 1000 born), and 1903 – 1921 for diarrhea cases (described as the number of cases per 1000 inhabitants) and tuberculosis cases (described as the number of cases per 1000 inhabitants). Column (1) show the results for infant moratlity, Column (2) show the results for tuberculosis, Column (3) show the results for diarrhea.