Water Improvement and Health: Historical Evidence on the Effect of Filtering Water on Urban Mortality

Daniel Knutsson*

April 19, 2017

Abstract

In this paper, I analyse how technologies for cleaning and distributing water can affect health using new historical data from Sweden. The city of Stockholm introduced a slow filter water cleaning system and piped distribution network in 1861 enabling parts of the population in-house access. The historical context allows me to analyse these technologies without sewerage access as no major sewerage system was constructed at the same time. By using detailed information on water access through contemporary contract lists I can measure access to clean water with great precision. My findings suggest large beneficial effects of having access to clean in-house water. This effect is apparent for the general population but not as precise for infants and in line or even larger than previous estimates. I document heterogeneity in infant mortality with respect to gender where girls seem to have benefited more.

Keywords: Water; Piped water; Filtered water; Infant Mortality; Mortality; Public Health

JEL classifications: I15; I18; N33; Q53; J11

*I am grateful to Per Pettersson-Lidbom, Anna Tompsett, Grant Miller, Nathan Nunn and participants at seminars and conferences at NHH Bergen, SDU Odense, ULB-ECARES and Stockholm University for helpful comments. Contact: Daniel Knutsson, Stockholm University, Department of Economics, 10691 Stockholm, Sweden. Email: daniel.knutsson@ne.su.se.

1 Introduction

In this paper, I analyse how technologies for cleaning and distributing water can affect mortality using new historical data from Sweden. The city of Stockholm introduced a slow filter water cleaning system and piped distribution network in 1861 enabling parts of the population in house access. The effect of access to these technologies is identified using the sudden introduction within - and differences in exposure intensity over time between - the geographical parishes comprising Stockholm at this time. The historical context allows me to analyse these technologies without sewerage access as no major sewerage system was constructed at the same time. By using detailed information on water access through contemporary contract lists I can measure access to clean water with great precision.

Understanding how urban water technologies can affect health is not only of historical relevance[1]. Still today, many people globally lack access to improved water sources and diarrhoea is still the second leading cause of mortality for children below five (WHO, 2013). More than 50% of the population world wide now live in cities and this number is expected to rise even further. 30% of these city dweller live in slums with lower access to improved water and sanitation (UN, 2015). Over the last two decades, low income countries have seen the highest urbanization growth rates. These are also the countries where clean water access is most scarce (UN, 2016). Understanding how urban water technologies can affect health is important to support low income countries in expanding, maintaining and providing access to water and sanitation technologies as urbanization continue. At the same time, research on the health effects of water and sanitation improvements in a present day context has produced surprisingly inconclusive results (Schmidt, 2014). Historical evidence, in contrast to experimental or intervention studies, can potentially be more informative in how large scale urban water technologies can improve health.

There are several ways that increased water quality and quantity can affect health. Improvements in water quality can reduce exposure to water born pathogens like Cholera and Typhoid fever (Ferrie and Troesken, 2008). Increasing access to water (availability and quantity) can affect sanitary behaviour which could decrease exposure to contaminants. The reduced time and effort of gathering water can also be used to increase positive health investments for especially women and their children (DFID, 2013). Infants and small children are more affected by water quality than adults (Kremer et al., 2011; Bhalotra and Venkataramani, 2013). Breastfeeding patterns could be important for how infant respond to higher quality water. While breastfeeding, infants are protected from direct exposure to water borne pathogens (Butz et al., 1984). Since many diseases are spread through contact between individuals, there could be meaningful negative externalities of not providing full access to water and sewerage technologies to a population (Duflo et al., 2015). Access to better water quality, quantity and sanitation could also be compliments to each others. Passive sewerages based on elevation differences require a steady

[1]This paper relates to an historical literature where the fundamental causes of the epidemiological transition in developed countries has yet to be explained. The jointly determined increases in purchasing power within a given population and the introduction of different health promoting technologies are hard to disentangle (James, 2003; Bruce and Jo, 2002).

stream of water to dispose sewer. At the same time, using water for cleaning streets, lavatories or homes require a water irrigation system to transport away dirty water from residential areas. Sewer transportation away form habited areas could be important to decrease recontamination. Sewerages could therefore have a positive effect on health on their own (Kesztenbaum and Rosenthal, 2014). They could also act as complements to piped water and hence be more effective if both technologies are jointly supplied (Alsan and Goldin, 2015).

To estimate the effect of clean water and the distribution network on health I use parish level data retrieved from church death records and contract lists from the public water utility in Stockholm between 1850 and 1872. The sudden introduction within parishes and differences in expansion over time between parishes allow me to identify the effect of in-house access to clean water on mortality in a difference-in-differences model.

A paper close to the present study investigates if clean water innovations in the US between 1900 and 1936 had an impact on mortality (Cutler and Miller, 2005). They use a selection of major cities and the temporal variation in implementation of sewerages, filtered water and chlorinated water. The authors identify the effect of these technologies on mortality. They find that filtration could explain nearly half of both the total mortality reduction and the infant mortality reduction in these cities. In a similar setting, Alsan and Goldin (2015) evaluates the effect of introducing clean water and sewerage technologies in Massachusetts between 1880 and 1915 on infant mortality. The authors find that the two technologies seem to behave as complements and have no important effect by them selves. They also find that the introduction of these technologies where responsible for about 44% of the decrease in infant mortality during that time. The present paper differ from these papers in that; (i) it investigates only how filtered, piped water affected mortality in a higher mortality setting and (ii) use only detailed within city variation on mortality and exposure intensity. The main benefit is that a specific technology can be evaluated where the water quality is the same for all with access.

I find that the water system introduced, substantially reduced mortality for all age groups. The full population response in my preferred specifications suggest a 30 log points decrease in population mortality and a 40 log points decrease in infant mortality. Low precision for the estimated effect on infant mortality can be explained by gender heterogeneity. Boys seem to have been unaffected while girls benefited greatly. Using monthly data I provide some evidence of seasonal heterogeneity supporting a water related mechanism. Mortality during summer months, when deaths from water borne pathogens are much more common (Cheney, 1984), seem to have been more affected than winter months.

The main results are robust to a number of different specifications and a placebo intervention in 1849. I further explore the gender difference in the clean water effect on infant mortality. Here, I find suggestive evidence that girls where much more affected during the first month of life but that the effect is more similar during the rest of the first year. I argue that there are two main mechanism consistent with the data. First, as I find that clean water decreased still births there could have been gender specific selection into life. The second explanation suggest that boys where preferentially treated with respect to breastfeeding. If boys where breastfed more

during their first month in life they would already be, to a large extent, protected from water contaminants during that time.

I structure the rest of the paper in the following way. In section 2 of the paper I provide a background to historical setting at hand. Section 3 details the data and empirical strategy. The following section (4) presents the results and provides a discussion. Section 5 concludes the paper.

2 Background

City life in the mid 19th century US and Europe was often crowded and short. In New York City, one of the most disease prone cities in the US at the time, the mortality rate was just above 3% in 1863. Well above other American cities such as Boston or Philadelphia and even higher than many European cities such as London[2]. Figure 7 shows that Stockholm did not fare any better than New York City during the 1850s with mortality rates well above 3% on average. The only European city to compare with Stockholm in terms of mortality during the 1850s and 1860s was Russian St: Peterborough (Lindberg, 1980). Stockholm was the largest city in Sweden, attracting many rural labourers as agricultural work was reorganised and migration laws where eased. Poverty was high and housing construction did not keep pace in this early stage of industrialization. At this time Cholera was common during the summer months and especially violent outbreaks hit the city in 1853 and 1857. Stockholm had the highest infant mortality rate in the country and had twice the mortality rate to the national average (Lindberg, 1980). During this time, about four out of ten children born in Stockholm did not survive to see their first birth day (see fig 7). Life expectancy at birth followed the high infant mortality rate and was around 20 for men and 26 for women during the 1850s (Lindberg, 1980). Inadequate water quality is a natural candidate to explain parts of this difficult health situation.

Before the introduction of piped water in Stockholm its inhabitants relied on wells and nearby lakes for their daily drinking and washing water supply. There where more than 300 private and around 25 public wells at the time but the water quality and reliability during summer droughts where reported to be bad in general (Cronström, 1986). The wealthy inhabitants could always buy shipped water from out of city springs or wells where water quality was much better. The public wells where investigated by the local health commission in 1867 and only six out of 27 was deemed moderately suitable to drink from. Results where based primarily on the amount of organic compounds and water hardness, the metrics used at this time to assess water quality. The commission urged the public not to use well water for drinking but to rely on the piped water now in place in large parts of the city[3]. The water quality might have been even worse before piped water was introduced as high pressure water cleaning of the streets was not an option. The increased usage of piped water could also have been beneficial for removing trash and excrement from city streets. Bacteriological analysis of the drinking water in Stockholm was

[2]See Evelynn Hammonds, "Infectious Diseases in the 19th-Century City," Link

[3]Information retrieved from the city archive in Stockholm. Available at request.

not systematically undertaken until 1884 when a chemist was employed by the city (Cronström, 1986).

Sewage irrigation in 1850s Stockholm consisted mainly of street ditches of which some where covered by wood. These allowed for ground infiltration and released its waste in nearby lakes and streams and had a tendency to become clogged up during the dry summer months. In many major cities of continental Europe, underground self contained sewage systems had been constructed already in the 1840s. But in Stockholm it would take until 1875 until a large scale system was started to be built. Piped water was seen as a prerequisite as the sewage needed continuous water flow not to clog up. A few public sewerages where build during the 1860s and in 1872 there where four main lines. Locally, these could service a few percent of the population at most (Cronström, 1986, p. 100). In 1875 the responsibility of city planning and construction was put on a newly created institution, The construction office (Byggnadskontoret), that had more access to the resources needed than its predecessor (Tredje Drätselnämnden). City wide plans to build sewage irrigation systems could then be approved. In 1876 it also became mandatory for house owners to have a local sewage pipe to the facility when possible (if there was a main sewage drum near by). As of then, large underground sewage irrigation construction with in house connections rapidly took place.

Before underground sewerages where build in Stockholm, human excrement where put in buckets that was transported and emptied away from the city. Boats where to take closed buckets and depart on a daily basis. In practice the buckets where often emptied on the ships which departed when full (more on a weekly basis)(Lindman, 1911, p. 272). This system of excrement disposal was improved in 1859 when the city council took over responsibility from criticised private contractors (Linroth, 1897).

In the mid 19th century, water cleaning and distribution technologies where widely discussed in Stockholm but conservative forces opposed the implementation primarily on fiscal grounds. The proponents argued that the sanitary environment would be improved, especially around poor neighbourhoods but there was also a strong group arguing for fire safety reasons. The older parts of Stockholm where built densely and wooden houses where not uncommon inside the city. Captain F. W. Leijonnacke, from the Swedish Royal Road and Water Construction Corps., presented a water cleaning and distribution plan with full city coverage in 1853 that caught the interest of local government (Hansen, 1897). During a study visit to Germany and the UK he had assembled enough information to put forward a complete plan for water cleaning and distribution in Stockholm. Before the plan was enacted a second opinion was demanded from foreign expertise. The British water technology expert and engineer Thomas Hawksley was hired as a consultant and reviewed the project plan with approval (Cronström, 1986). Although the relationship between Cholera and water quality was unclear at the time, contemporary writers related disease to garbage and faecal waste on the streets which could be cleaned using pressurised water. Local authorities where further convinced about making the investment in a water distribution network as Cholera struck Stockholm in 1853 and again in 1857.

The fire insurance department, physician association and a group of well borne

residents of Stockholm was supporting the water construction plan. The opinions of the physical parishes of Stockholm where pivotal in deciding weather to go ahead or not. City bonds where required for funding and representatives of the parishes of Stockholm had the final say on those matters. After some concessions to the more sceptical parishes (building owners where for example held accountable to the bonds that the water construction company required) the plan was approved. One important feature was that further construction, beyond a limited initial plan, had to be financed with revenue, not through debt. To make the enterprise of full access to water possible and fast, the initial plan was set up to give as many as possibly access, increasing the probability of income flows to the company enabling further construction. Scarcely inhabited parishes such as Kungsholmen had to wait for access and the flow of revenue one year would be important for the expansion of the net the following years. Although not explicitly stated, there might also have been a fairness aspect between parishes in access. As all parishes where accountable and held ownership to some extent, construction would have to have been spread out somewhat disregarding strict revenue considerations.

In 1858 the plan got its final approval and the Stockholm Water Company was set up. In June, 1861 the first 30 km of pipes had been laid and in October, for the first time clean, piped water was available to the inhabitants of Stockholm. Due to financial reasons, the initial suggestion from 1853 had been sliced heavily. The geographical parishes had reduced the capital needed for the full access plan by around 25%. In 1872 the distribution network had increased to 80 km in length and there where approximately 100′000 users with in-house access to water[4](Kommunalförvaltning, 1871). To increase availability, a few public fountains where set up where clean water could be collected free of charge.

The first water plant took its water from a near by bay (Årstaviken), connected to lake Mälaren. At the time it was considered to provide adequate water quality as well as technical benefits (i.e. close to the cleaning water plant). Before deciding on the location of water intake, Mälaren was examined on several places for organic compounds, where Årstaviken got the best results and was the strongest candidate.

A concern regarding the water quality, mentioned in 1876, was that a city ditch, from the southern parts of Stockholm exited in Årstaviken. The content included sewage from a local hospital. Water was taken from an other part of Årstaviken but still had contemporary writers and physicians worried about the possibility of ditch water entering the city water supply. Later on, as new sources of water was considered, more thorough measures of Årstaviken and the supplied water was undertaken. With the city ditch in mind, local expertise was consulted, whom where somewhat sceptical towards the quality of the piped water. In 1879 a British renowned water chemist was consulted (professor E Frankland). He approved the water, which in terms of faecal residue was superior to the water in London, leading to the decision to stay with Årstaviken as the water source (Cronström, 1986).

The water cleaning technology was inspired by the newest plants in the UK. Large pools where built to facilitate slow filtration in adequate amounts for the population of Stockholm. The plant was built at Skanstull in the southern parts of Stockholm and initially had three slow filter pools with a total area of 1600 square

[4]Stockholm at this time had around 140′000 inhabitants.

meters. The sand layers in the pools where in total seven feet deep with eight different layers of sand and stone, from fine sand to a layer of stones the size of coconuts (Cronström, 1986). The top layer fine sand was partially removed and washed around every sixth day during summer and more rarely during winter time. When the top layer was at a minimum, new clean sand was added. The complete sand filter was replaced every other year (Hansen, 1897).

For the main distribution net, pipes where made of cast iron coated in tar to prevent oxidation. The service pipes in to buildings where made of lead which where easy to shape in ways that made access around existing infrastructure easy. Already at this time there where concerns about health effects of led on people but it was dismissed as experimentation with other materials had shown that there existed at the time no functional alternative material (Cronström, 1986, p. 71). This conclusion was to some extent also based on experiences from other cities in Europe which introduced piped water with lead service pipes and the benefits seemed to out weight the potential health concerns.

Historical evidence from the US during the early 20th century suggest that lead exposure through tap water had a meaningful effect of infant mortality (Clay et al., 2014). How lead in pipes diffuse into the water depends on the water acidity (pH), the amount of lead in the pipes and to a large extent on how old the pipes are. As the pipes age, oxidation inside the pipes will create a sort of barrier to the lead. New pipes with high acidity water would release the most lead (Troesken, 2008). Lead is also believed to be released from the skeleton during pregnancy and lactation making already sensitive foetuses and infant more exposed (Gulson et al., 2003).

In 1864, Sten Stenberg, a professor at Karolinska Institutet (at the time Carolinska Medico-Kirurgiska Institutet), investigated the lead content of the piped water using 14 different water samples[5]. In most samples he noted minor lead content and showed that the amount was on average $8 - 9$ times lower than for example in Claremont where the piped water had been extensively examined. He concluded that overnight still water from led pipes had higher lead content and showed that boiling lake water for an hour using kitchen pots available at the time (a copper pot with lead mixed tin coating) would in it self provide about the same amount of lead as the piped water. Still, he recommended that over night water be flushed out before use. [6].

Any city using lake water at this time had problems with water temperatures during the summer months, also Stockholm. Reports from the time note that the piped water was not popular to drink during the warmest time of the year as it became warm (Lindman, 1911, p. 274). The temperature would naturally vary within Stockholm depending on the time spent underground. Still, we can note that usage increased dramatically during the studied time period [7]. In 1872, daily per capita consumption was approximately 40 litres, which can be compared to

[5]Source: Print copy of report from Stockholm City Archives available at request. Dated 1864.

[6]The overnight water sample was measured to 16ppm while today it is recommended that tap water do not exceed 0.015ppm, see `https://www.epa.gov/`.

[7]Since water was not initially measured at the house level there is no way to assess household per capital consumption during the first ten years of operation

around 180 litres today[8].

The water plant was situated in the southern part of Stockholm (see figure 5) and main pipes where laid north, crossing Södermalm and continuing north through Old town (the small island in the middle) and over to Norrmalm. The map shows the initial network, constructed before 1861 (in blue), which is heavily concentrated to Old town and the southern parts of Norrmalm. When access to a main pipe was available, house owners had to decide on financing a in-house supply line and to put the water bill on those renting the rooms in the building. House owners paid the bill that was charged per room in the house. After some initial scepticism, water access became popular and more houses got connected. In 1872, it was estimated that 70% of the population in Stockholm had in-house access.

How infant health will respond to changes in the water quality depend heavily on the dominant breastfeeding practices (Butz et al., 1984). There is very limited information about how common breastfeeding was in Sweden during this time but some information was gathered in the early 1870s. Due to worries raised by Berg (1869) regarding differences between regions in Sweden in infant deaths within the first year of life, a request was sent out from the Medical board at the time (Medicinalstyrelsen) to all local physicians under its authority. In this request, physicians where told to report back their impression of the state of breastfeeding in their district. Using this information, (Brändström, 1984) constructed a breastfeeding map over Sweden for all districts where information was returned (see figure 3).

The map is divided into three categories. If breastfeeding was common, if alternative infant feeding practices existed along with breastfeeding and if cow milk and other practices where the dominant form of infant feeding. It is clear from the map that most areas practised breastfeeding to some extent in 1870 but for Stockholm, there is no information available. Since more than half of the population in Stockholm at this time where born outside of Stockholm we can expect that practices varied. Most migration to Stockholm came from surrounding districts where breastfeeding either where where practiced to some extent, or was common. It is likely that many women in Stockholm practised breastfeeding but that it was not practised exclusively.

3 Data and Empirical Strategy

Although there are limitations to using very old data there are also some appealing benefits. In the historical context interventions can be evaluated at scale, such as providing water technologies to an entire city, which might not be feasible today. Sweden has a long history of collecting and saving data which makes it an appealing context for historical studies.

Data

Since health is not observed in the population at hand, I use mortality as a proxy for health. This has been used extensively in the literature (see e.g. Almond et al.,

[8]See `http://www.drickkranvatten.se/virtuellt-vatten`

Table 1: Data Sources

Measure	Source	Frequency	By Sex	Availability
Infant mortality	Tabellverket	Yearly	Yes	1845 − 1859
Infant mortality	Statistics Sweden	Yearly	Yes	1860−
All age mortality	Tabellverket	Yearly	Yes	1850 − 1859
All age mortality	Statistics Sweden	Yearly	Yes	1860−
Still births	Tabellverket	Yearly	No	1850 − 1859
Still births	Statistics Sweden[1]	Yearly	No	1860 − 1872
All age mortality	Tabellverket	Monthly	Yes	1850 − 1859
Born/Population	Tabellverket	Yearly	Yes	1850 − 1859
Born/Population	Statistics Sweden[1]	Yearly	Yes	1860 − 1872
Wedlock birth	Tabellverket	Yearly	Yes	1850 − 1859
Wedlock birth	Statistics Sweden[1]	Yearly	Yes	1860 − 1872

Note: (1). Summaries documented in the archives of Statistics Sweden. Can be found at the Swedish archives (Riksarkivet) `www.ra.se`. All variables defined at the parish level.

2012; Bhalotra and Venkataramani, 2011), but in the historical context there are some additional advantages. In a high mortality setting with few medical technologies available, such as antibiotics or even knowledge about how infectious diseases operate, mortality is likely to respond in a measurable and reliable way to even moderate changes in the health environment. Today, this might not be the case due to lower mortality rates, availability of information on prevention and access to medical services.

The mortality data used between 1850 and 1859 comes from parish summaries produced by Tabellverket, the predecessor to Statistics Sweden. These summaries are based on mortality tables put together by parish officials (often religious representatives) which where sent to Tabellverket for processing on a yearly basis from 1749 and onwards. Yearly data comes with great detail but monthly data is grouped in more fixed ways. In 1860 Statistics Sweden was created and from then, I have access to individual death records for each parish [9]. This individual data does not have date of death information but is sorted only into years. Still the original death books are sorted by month so I have manually inferred month breaks into the data. These breaks are more or less noisy depending on the consistency of monthly order in the death records but should reflect the monthly death count fairly accurate. Table 1 describe the sources and availability of mortality data from Tabellverket and Statistics Sweden.

To get monthly Water data, contract records was retrieved from the Stockholm City archives of the Old Water Cleaning Company (the predecessor of today's water cleaning and distribution utility, Stockholm Vatten). The files in the archive included a contract book from 1861 to 1872 which have been the source of water exposure data. The contract data includes day, month and year of contract, address, parish, block and number of rooms debited. The contract was set up to the house (building) owner and (s)he was then charged for how many rooms the building con-

[9]All mortality data after 1859 is available from Swedish archives scanned. See Riskarkivet

tained. Water access was shared by residents through in-house tap in a common area of the building. Some discrepancy between the number of rooms with access to clean water in the data and contemporary sources of household access emerged as the data was compiled. There may have been some problems in reading the contracts from primary sources to the contract list used (the primary sources where not interpretable), some double booking of houses changing owner and contracted houses belonging to firms. I treat the data as a random sample and scale it to align with the estimates of total household access derived from contemporary sources (Kommunalförvaltning, 1871). I can further get a measure of individual exposure by using data from 1860 on houses in each parish, population in each parish and average number of rooms per house in each parish [10]. At the parish level I calculate:

$$Population_{1860} / \left[(Rooms/House)_{1861-1872} * Number\ of\ Houses_{1860} \right] \approx Inhabitants\ per\ room$$

Scaling exposed rooms by this estimate, I get a proxy for individual exposure. Pre treatment variables was gathered for 1860 from auxillary sources[11]. When applicable, these control variables are entered into the model as linear expansions from 1860 and onwards as shown below. Only the slope differ between parishes for these variables suggesting that if mortality trends, post treatment was determined by parish pre treatment characteristics (and linear), we would expect them to explain variation between the outcome and treatment variable.

$$Demographic_{pt} = Demographic_{p}^{1860} \cdot 1(Year > 1860) \cdot t$$

It is interesting to consider the amount of total variation in access to clean water which to some extent can be approximated from figure 5. To be more specific I have plotted the treatment intensity over time for the eight parishes under study (figure 8). There is substantial variation both within and between parishes over time.

Identification

To identify the effect of access to in-house clean water, I rely on the sudden introduction and differences in access within parishes over time. Since all parishes where treated at almost the same time, but with different intensity, they all contribute to control for each other. As I do not know the individual take up, I can only estimate the intent-to-treat effect which often is the policy relevant question as forcing people to use for example new health technologies might be infeasible or unwanted.

There are two main potential sources of selection in this setting; (i) that the construction of the main water grid (indirect access) was executed in a way related to the response to treatment and (ii) that the choice of the building owner to connect to the main grid (direct access) was endogenous. If the main pipes where

[10] Houses and population in each parish see Statistisk Tidskrift 1860 part 1 pp.187. Rooms per contract/house come from own data and calculations.

[11] The 1860 census (see statistiscs sweden archive at `www.ra.se`), Statistisk Tidskrift (1860, `http://www.scb.se/`) and Kommunförvaltningen berättelser 1872 (See Kommunalförvaltning (1871))

constructed in a random way (without care for population characteristics) and take up at the house level was random (or complete), the intent-to-treat (ITT) effect would be the average population response to a complete increase in access. Below I will discuss the plausibility of these assumptions.

The initial plan, from 1853 was made to facilitate access to the full population of Stockholm. Fire safety, hygiene and sanitation where the main reasons for introducing clean piped water. This plan stretched all inhabited parts of Stockholm, included hundreds of fire posts and was calculated to cost 1405076 Riksdaler banco (approx 10 million USD in 2015 values). While presenting the proposal, the plan architect argued that clean water could reduce epidemics, improve general hygiene, benefit the poor, benefit industry, improve fire safety and increase the comfort of the wealthy Cronström (1986, p. 18). In the final approved plan, the budget had been cut and the initial full coverage grid had to be slimmed down to only service the most populous parts of the city in a first stage. It was decided that further expansions of the grid, beyond its first part would have to be financed internally by revenue. Hence there was a strong incentive to supply dens parts of the city early on to get as many contracts as possibly from the start. Parish borders where not a concern but since all parishes where financing this venture, there could have been some incentives to provide access to all of them as soon as possibly.

The population of Stockholm was to some extent mixed. Southern parts of Stockholm was in general poorer while different income groups often shared the same building with the wealthier living on the lower floors (Cronström, 1986). The direction of the possible selection due to grid access is not obvious. Denser parts of the city could mean both wealthier, many people live where demand for housing is high (i.e. expensive), and poorer as poor people lived more crowded all else equal. House owners had little incentives not to connect to the water supply if available. Many of them lived in their buildings and the low contract cost could be put on the other residents. The initial cost per room debited was 2.6 Riksdaler which deflated to current prices would yield a yearly cost per room of 18USD (156SEK). This system of payment further benefited poor families since they did neither pay per user nor usage, but for each room. A room which in some cases could support both a family and several lodgers while wealthy residents, with fewer people per room, paid relatively more.

House owners still had to provide the service line from the main pipe in to the house which could have been a substantial investment. The selection story is not clear here either. Given that there where a substantial cost of connecting to the main grid, wealthier tenants might demand access to improved water technologies. At the same time demand for clean water might be dependent on the quality of the alternative water source (e.g. the local well) which could be correlated with both rent and income in the building. For example the parish Nicolai had very low pre treatment mortality rates, got full main pipe access in 1861 but take up was fairly low during the whole time period. Adequate quality alternative water sources at a close distance (and high income) might have played an important role in deciding not to take up treatment. In 1872 more than half of the houses in Stockholm had main water pipe access and more than two thirds of the population had access to in house water suggesting that take up at the house level was significant

(Kommunalförvaltning, 1875).

To investigate the relationship between underlying health and clean water access in this setting I would optimally like to estimate the following type of equation

$$Health_i = \alpha + \beta \cdot Clean\ Water\ Consumption_i + u_i \tag{1}$$

Which, if the treatment was randomised, identifies the dose-response effect of clean water on health. Still, externalities due to non-complete take up could still be a problem. The data I have allows me to estimate a similar intent-to-treat equation at the parish level where mortality act as a proxy measure for health

$$log(Mortality)_{pt} = \alpha + \beta \cdot Clean\ Water\ Access_{pt} + v_{pt} \tag{2}$$

Here, mortality at the parish level is defined as number of cases divided by the parish population in 1860 (pre treatment). Clean water access is the population fraction with access as described above and v is a parish level random error term. If *Clean Water Access* was randomly assigned I would in principle be able to identify the average effect at the parish level. That kind of randomness is not a credible assumption in this context so I rely on a Difference-in-Differences (DiD) specification, where the main identifying assumptions include common pre treatment trends and no omitted variables bias during the treatment period.

$$log(Mortality)_{pt} = \beta \cdot Clean\ Water\ Access_{pt} + \delta_t + \gamma_p + v_{pt} \tag{3}$$

Equation 3 augments the previous specification with time and parish fixed effects and is important as a starting point. But with many treatment intensities, the identifying assumption can be hard to assess (i.e. parallel trends). To somewhat relax the strict parallel pre treatment trends assumption I include parish specific linear time trends in the following DiD model.

$$log(Mortality)_{pt} = \beta \cdot Clean\ Water\ Access_{pt} + \delta_t + \gamma_p + \rho_p \cdot t + v_{pt} \tag{4}$$

In this equation δ, γ and ρ are nuisance parameters of fixed parish and time effects and parish linear time trends. The inclusion of group specific trend variables is no "quick-fix" to get around the main identifying assumption of the design but impose different assumptions regarding the pre treatment trends (in this case parallel growth) (Mora and Reggio, 2012). Although parallel trends might be a strong assumption, constant growth might be more credible. If the parish time trends are roughly linear but the slopes are not parallel, including linear trends will not only mitigate bias (if trends are correlated with treatment) but also increase power (if trends have random slopes). If trends are indeed parallel, inclusion of linear time trends does not bias the estimated effect (see Mora and Reggio (2012)) but power could be reduced. Under the assumption of parallel growth in trends and that clean water access over time is as good as random, equation 4 identifies the population average β.

The expansion might have been correlated with pre treatment characteristics which also could determine the outcome. To assess the randomness of the expansion I add a set of control variables measured pre treatment and operationalized as

described above. These are mean mortality pre treatment, fraction of children born out of wedlock, fraction of women married, fraction of inhabitants born in Stockholm and two measures of density, houses per area and average number of inhabitants per room. I stack these in $\mathbf{X}$ and include them in the following model

$$log(Mortality)_{pt} = \beta \cdot Clean\ Water\ Access_{pt} + \mathbf{X}'\sigma + \delta_t + \gamma_p + \rho_p \cdot t + \upsilon_{pt} \qquad (5)$$

If the roll out of the intervention was related to any of these variables, the effect of clean water could be modified by their inclusion.

Using only eight groups can cause problems with inference using the cluster-robust variance estimator as described by e.g. Wooldridge (2010). As the estimator provides asymptotic consistency in the number of groups, the approximation might be poor and simulation evidence suggest that clustering tend to over reject at "few" clusters (or "few" treated/non-treated groups) (Cameron and Miller, 2015). The risk of over rejection based on few treated (untreated) groups is probably not very prominent in the present setting as all groups provide controls for each other (netting out group specific shocks), but the consistency of the estimated "clustered" covariance matrix with only eight groups can surely be a concern.

Interpretation, Exposure and Externalities

Waterborne pathogens have multiple ways of being transmitted to a human host. Direct exposure through food and drinks as well as indirect contact exposure via an other human host (contact spread). Depending on the relative importance of these transmission mechanisms the population response to clean water can look very different. If (indirect) contact drives disease propagation and clean water is not important in reducing this type of spread, full population access of clean water might be required to observe a meaningful protective effects against these water borne diseases (Duflo et al., 2015). If on the other hand recontamination of drinking sources by already infected humans is the main driver of disease spread, having clean water could imply sufficient protection for each individual. Sewerages could in this setting be pivotal in in reducing disease spread. In both settings increased supply of clean water could have positive externalities by acting as a vaccine. By not using re-contaminated water sources and practising better hygiene, those with access could have a low probability of being infected by a water borne pathogen. And where higher coverage could prevent a disease from firmly entering a human population.

What would this mean for estimating the dose-response intent-to-treat effect of clean water on mortality? First, if we view Stockholm in aggregate, we would expect general reductions in mortality. More people with access reduce the probability of an initial infection. This is also true for the parish sub populations. The amplitude, or intensity with which a pathogen propagates in the population is dependent on how transmission occurs and will respond to marginal increases in access to clean water given that it "vaccinates" people from obtaining the disease (recontamination) or reduce contact spread (better hygiene). In this framework we can infer that how water born pathogens spread in this environment is important for if and

how protective clean water access is. And also that externalities should attenuate mortality differences between parishes.

4 Results

In the following section I present regression output estimating the intent-to-treat effect of access to clean water on mortality for different age groups and by sex. As population size and number of births could be endogenous post treatment I scale the crude mortality counts by these measures in 1860[12]. Also the treatment variable is scaled by the parish specific population (as they where in 1860) such that we can interpret the estimates as log point changes in the mortality rate providing access to all inhabitants in an average parish. Standard errors are clustered at the parish level. All parameters are estimated using population (or births) in 1860 as weights if not explicitly otherwise stated.

I use both yearly and monthly data to estimate the effect of clean water on mortality. The benefit of using data on a yearly level is that I have access to more and better data as well as larger cell sizes. Using monthly data provides more variation (as the treatment variable varies at an even lower level) and allows me to investigate seasonal variation in the clean water effect[13]. Still, the lower quality, reduced availability of measures and smaller cell sizes of the data inclines me to put most emphasis on yearly data in my analysis. The regression tables referred to herein can be found in Appendix 1*a* (section 6).

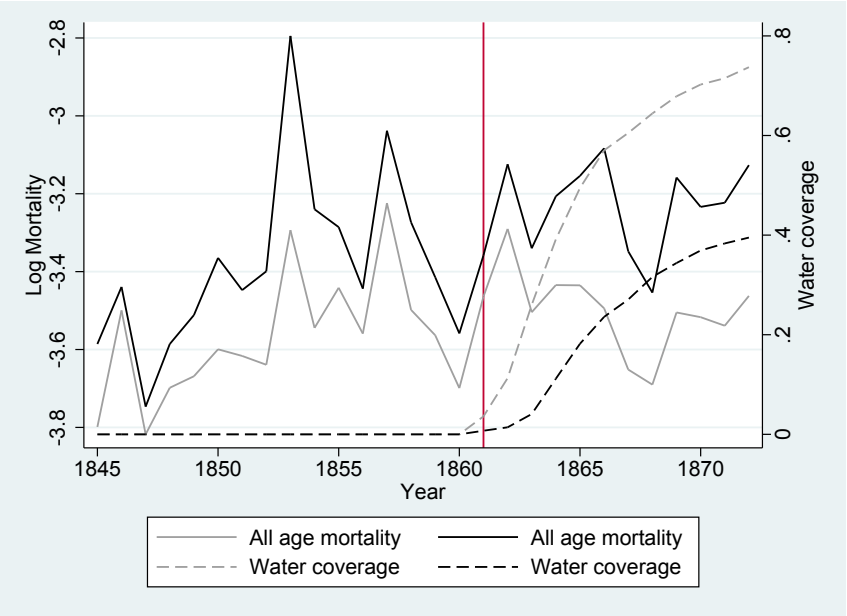
Figure 1 provides some graphical support for an effect of clean water access on log mortality. The graph shows how log mortality developed between the four parishes with the lowest 1872 water coverage compared with the four highest coverage parishes, together with their respective average treatment intensity. A divergence in log mortality is somewhat apparent after differences in clean water access emerge. To confirm this suggestive evidence I use OLS regression models as described above to analyse the data.

The main results are presented in table 3. There, I use a specification with year and parish fixed effects as well as linear time trends and control variables. The models in the first three columns use log mortality as the outcome and separate by gender. I find that full access to clean water reduced mortality by around 30 log points and maybe a bit more for women than for men (although there is no statistical difference between point estimates). The next three column repeat the same exercise using log infant mortality as the outcome. Here I measure infant mortality as deaths occurring during the first year of life. There is a lot less precision analysing this measure (with much smaller cell sizes). The pattern is though similar to total mortality but with more pronounced differences by gender. The estimated effect for all infants is a reduction in mortality of around 40 log points. None of the estimates are significant but the difference in response to clean water access between infant boys and girls as given by the point estimates are surprising. The general

[12]Results are robust to using yearly measures.

[13]Although it changes the implicit weighting of OLS somewhat. As mortality shocks at this time where often limited to a few months, much of the yearly fluctuations comes from these shocks.

Figure 1: Mortality in high vs. low clean water access parishes



Note: All age mortality in the four parishes with the highest clean water coverage rate in 1872 and those with the lowest rate. Right hand scale shows water coverage, dashed lines in graph, and left hand scale shows mortality (solid lines). Both measures scaled by population in 1860.

pattern suggest that clean water had a meaningful negative effect on mortality, for adults and infants.

To assess if water is a likely mechanism for the effect of the intervention on mortality, I use monthly data and analyse it by season (see table 4). I can not use infant mortality as that data was not available[14]. In this setting, I expect that clean water should be more effective during the warm months when conditions for water borne pathogens are more favourable. The pattern is not extremely clear but the imprecise point estimates point to larger effects during summer. I also use a more flexible specification with treatment-month interactions. The results is presented in figure 2, where I show that although there are no significant differences in a statistical sense between estimates, a peak in the effect can be seen at late summer when the temperature is high and rain is scarce.

To further understand the age-gender differences in response to clean water access, I estimate a model with mortality cells including deaths above the age of three (table 5, columns $1-3$). There seem to be little difference in mortality for older children and adults between men and women. I also use additional outcomes that could indicate if in-utero health was affected. In table 5, columns $4-5$, I use the log stillbirth rate and the fraction girls divided by total births each year as outcomes. The large estimated negative effect of clean water on the stillbirth rate, almost a 90 log points decrease, should be considered with cell size in mind. Still born infants are only about 5 percent of total number of births in the data and represent very small cell sizes with a lot of variability. Still, the estimate indicate that not only the health of adults, children and possibly infants where improved but

[14]I would also have had many more zero value cells when disaggregating to monthly level.

also pre-natal foetal health. An other indication of in-utero health is the fraction of girls being born relative to that of boys. Male foetuses seem to be more sensitive to their health environment, at least late during pregnancy (Orzack et al., 2015). This fraction is hence expected to decrease given a more advantageous environment where relatively more boys are born. This phenomena could potentially explain parts of the gender difference in infant mortality. Implying a negative selection of boys into life and an attenuated effect on infant mortality. The estimated effect is negative, indicating that this might be the case, but precision is very low.

By using data from 1860 and onwards I can further assess if selection into life or some other type of gender biased behaviour was important. To do this I decompose the infant mortality data into deaths during the first month and deaths occurring later during the first year in life [15]. In table 6, I show how the models compare using the full data and using the restricted. The model employed here is without time trends but with control variables as this specification had the best fit[16]. Although the estimates are not identical, the general pattern is the same. Finally, in table 7 I decompose the restricted data on infant mortality into two groups. Deaths within the first and later months during the first year of life. The first month in life shows clear gender differences with a large negative effects for girls and a positive insignificant effect for boys. Later months shows a more similar pattern with large negative effects for both genders. The magnitudes estimated here using even less data might not be very reliable but the pattern is interesting. Part of this difference could be due to improved in-utero conditions negatively selecting male foetuses into life. But it is unlikely that this could explain the full differences given the rare event of a stillbirth.

In table 8 I further probe the preferred model to different specifications. Clustering on parish is not optimal in this setting with only 8 parishes but as can be seen from the 1st row, it is more conservative than not clustering at all and not very different from clustering on both parish and year (2nd row). The 3rd row shows unweighed estimates. Here, I find in general larger estimates suggesting that the effect was larger for smaller parishes. Still, I believe that individuals are the relevant unit of analysis here as treatment was not delivered at the parish level. The 4th row shows how the estimates respond to a lagged dependent variable as an additional regressor. This specification is not consistent but still provides some assurance that the change on change relationship between water access and mortality has some bearing. In the 5th row, I include a measure of year to year net migration as a control variable[17]. This measure could naturally be an outcome if migration patterns to Stockholm was determined by yearly connections to the water system at the parish level (and here I also have mortality on both sides of the equation). Still, it is reassuring to see that it is not the flow of people, in this time of rapid urbanisation, that drives the relationship. Taking the control variables away increase

[15]Since I only have aggregated summaries prior to 1860, this exercise is not feasible using the whole data. From 1860 and onwards I have the individual entries from the death records.

[16]With "best fit" I mean that the reduced form data produce the most similar estimates as the full data.

[17]Net migration is calculated as the difference in population between two ascendant years net of mortality and natality

the point estimate somewhat. I also try a specification without time trends that give very attenuated results suggesting that the pre trends bias the estimated effect towards zero. The increased standard errors in this specification further points to violations from the parallel trends assumption and the validity of including time trends variables. Including quadratic time trends with control variables take away all the effect but without controls the estimates are very close to the base line specification. Excluding the severe mortality shock years (cholera years) of 1853 and 1857 increase the point estimates. Since the DiD model is sensitive to the functional form of the outcome I also use the raw mortality rates as outcome. These estimates are comparable in size to the estimates using natural logs. In the last row of table 8 I use a bit longer data and introduce the clean water in 1849 instead of the true introduction in 1861. In this way I can let the expansion roll out over an untreated time period as a placebo. It is comforting that the placebo effect on log mortality is close to zero but somewhat larger on infant mortality. In tables 9 and 10 I investigate if there where compositional changes due to the introduction of clean water. As there was a cost associated with having in-house access, the inhabitants of Stockholm might have resorted into different parishes. The results in table 9 suggest that clean water access did not affect poverty, population or the number of births in any meaningful way. Here I use the log fraction of children born out of wedlock as a proxy for poverty. Taking this measure (by gender) as a control variable, I do not find that these measures are biasing the results on infant mortality (see table 10). The clean water intervention did not seem to have affected population composition in this dimension at the parish level (children born out of wedlock has an average of 37% in the sample).

5 Conclusion

In this paper I have shown that clean water seem to have had a large impact on mortality in Stockholm during the 1860s. Base line estimates suggest that full population access to the water technologies reduced mortality by around 30 log points for men and women (around 35% reduction using $(\exp(\hat{\beta}) - 1) * 100$). For infant mortality, my main specification suggest an effect in the order of 40 log points reduction. For infant mortality I do not have enough precision to be confident in the estimates but the pattern with respect to the sign and magnitude is consistent throughout most specifications. There where some indication that the effect was more pronounced during summer months, supportive of a water related mechanism. Using auxiliary outcomes measures there was evidence of improved in-utero health. I also found suggestive evidence of heterogeneity in the effect for infants with respect to gender. Gender differences found elsewhere in the literature on clean water and early life interventions[18] are supported. The results indicated that infant boys did not benefit from clean water during their first month of life, but females did. Possible (albeit speculative) explanations could relate to gender specific negative selection into life as described above. Another explanation could point to gender

[18]That girls and women are more effected by clean water has been suggested by Kremer et al. (2011) and for early life/in-utero exposure see e.g. Nilsson (2016)

biased breast feeding practices. Given the volatile nature of the data and the small sample size, it is not obvious that this difference is very meaningful.

As the service water pipes where made of lead and new, it is likely that lead intake increased for the population with access (Troesken, 2008). Measuring health using mortality records, it is also more likely that infant health was affected more in a negative way due to the lead than adults (Clay et al., 2014). As lead accumulate in the skeleton and is released during pregnancy and lactation, this negative effect on primarily infants could have increased over time initially and later decreased as the pipes oxidised (Gulson et al., 2003). The estimated effect, especially for infants, could hence be seen as lower bounds on the true effect. Taken together, the results from the present paper suggest a clean water effect at parity or even larger than previously has been documented is similar contexts both for total mortality and for infant mortality. Filtration using the technologies available at the time was an effective method to improve health. Even without access to sewerages.

References

Almond, D., Currie, J., and Herrmann, M. (2012). From infant to mother: Early disease environment and future maternal health. *Labour Economics*, 19(4):475 – 483. European Association of Labour Economists 23rd annual conference, Paphos, Cyprus, 22-24th September 2011.

Alsan, M. and Goldin, C. (2015). Watersheds in infant mortality: The role of effective water and sewerage infrastructure, 1880 − 1915. Technical report, Princeton University. Accessed 10th of April 2015.

Berg, F. (1869). Om dödligheten i första lefnadsåret. *Statistisk tidskrift*, 23:435–493.

Bhalotra, S. and Venkataramani, A. (2011). The captain of the men of death and his shadow: Long-run impacts of early life pneumonia exposure. Discussion paper 6041, IZA.

Bhalotra, S. R. and Venkataramani, A. (2013). Cognitive development and infectious disease: Gender differences in investments and outcomes. IZA Discussion Paper 7833, IZA, Bonn.

Brändström, A. (1984). *De kärlekslösa mödrarna: spädbarnsdödligheten i Sverige under 1800-talet med särskild hänsyn till Nedertorneå*. Dissertation, Umeå University.

Bruce, G., L. and Jo, C., P. (2002). Mckeown and the idea that social conditions are fundamental causes of disease. *American Journal of Public Health*, 92(5):730–732.

Butz, W. P., Habicht, J.-P., and Davanzo, J. (1984). Environmental factors in the relationship between breastfeeding and infant mortality: The role of sanitation and water in malaysia. *American Journal of Epidemiology*, 119(4):516–525.

Cameron, C. A. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Cheney, R. A. (1984). Seasonal aspects of infant and childhood mortality: Philadelphia, 1865-1920. *The Journal of Interdisciplinary History*, 14(3):pp. 561–585.

Clay, K., Troesken, W., and Haines, M. (2014). Lead and mortality. *The Review of Economics and Statistics*, 96(3):pp. 458–470.

Cronström, A. (1986). *Stockholms tekniska historia: Vattenförsörjning och avlopp*. Liber, Almqvist & Wiksell, Uppsala.

Cutler, D. and Miller, G. (2005). The role of public health improvements in health advances: The twentieth-century united states. *Demography*, 42(1):pp. 1–22.

DFID (2013). Water, sanitation and hygiene evidence paper. Report, Department for international development, UK Government.

Duflo, E., Greenstone, M., Guiteras, R., and Clasen, T. (2015). Toilets can work: Short and medium run health impacts of addressing complementarities and externalities in water and sanitation. Working Paper 21521, National Bureau of Economic Research.

Ferrie, J. P. and Troesken, W. (2008). Water and chicago's mortality transition, 1850–1925. *Explorations in Economic History*, 45(1):1 – 16.

Gulson, B. L., Mizon, K. J., Korsch, M. J., Palmer, J. M., and Donnelly, J. B. (2003). Mobilization of lead from human bone tissue during pregnancy and lactation—a summary of long-term research. *Science of The Total Environment*, 303(1–2):79 – 104. Lead Remediation Issue.

Hansen, F. V. (1897). Stockholms vattenledning. In Dahlgren, E. W., editor, *Stockholm : Sveriges hufvudstad skildrad med anledning af allmänna konst- och industriutställningen 1897*, pages 322–352. Beckmanska Boktryckeriet, Stockholm.

James, L. (2003). Rethinking mckeown. *American Journal of Public Health*, 93(7):1032–1032.

Kesztenbaum, L. and Rosenthal, J.-L. (2014). Income versus sanitation; mortality decline in paris, 1880-1914. Working papers, HAL.

Kommunalförvaltning, S. (1871). *Berättelse angående Stockholms kommunalförvaltning : jämte statistiska uppgifter för samma och föregående tid*. Beckmanska Boktryckeriet, Stockholm.

Kommunalförvaltning, S. (1875). *Berättelse angående Stockholms kommunalförvaltning : jämte statistiska uppgifter för samma och föregående tid*. Beckmanska Boktryckeriet, Stockholm.

Kremer, M., Leino, J., Miguel, E., and Zwane, A. P. (2011). Spring cleaning: Rural water impacts, valuation, and property rights institutions*. *The Quarterly Journal of Economics*, 126(1):145–205.

Lindberg, F. (1980). *Växande stad. Stockholms stadsfullmäktige 1862-1900.* Liber, Almqvist & Wiksell, Uppsala.

Lindman, C. (1911). *Sundhets och befolkningsförhållanden i sveriges städer 1851-1909*. Schmidts Boktryckeri, Hälsingborg.

Linroth, K. (1897). Stockholms hälso och sjukvård. In Dahlgren, E. W., editor, *Stockholm : Sveriges hufvudstad skildrad med anledning af allmänna konst- och industriutställningen 1897*, pages 477–568. Beckmanska Boktryckeriet, Stockholm.

Mora, R. and Reggio, I. (2012). Treatment effect identification using alternative parallel assumptions. Working Paperr 12-33, Universidad Carlos III de Madrid.

Nilsson, J., P. (2016). Alcohol availability, prenatal conditions, and long-term economic outcomes. *Journal of Political Economy*. Forthcomming.

Orzack, S. H., Stubblefield, J. W., Akmaev, V. R., Colls, P., Munné, S., Scholl, T., Steinsaltz, D., and Zuckerman, J. E. (2015). The human sex ratio from conception to birth. *Proceedings of the National Academy of Sciences*, 112(16):E2102–E2111.

Schmidt, W.-P. (2014). The elusive effect of water and sanitation on the global burden of disease. *Tropical Medicine & International Health*, 19(5):522–527.

Troesken, W. (2008). Lead water pipes and infant mortality at the turn of the twentieth century. *The Journal of Human Resources*, 43(3):553–575.

UN (2015). Water for a asustainable world. Report, The United Nations World Water Assessment Programme.

UN (2016). World cities report. Technical report, United Nations Human Settlements Programme.

WHO (2013). World health organization, fact sheet 330. Accessed: 2017-04-18.

Wooldridge, J., M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT press.

6 Appendix 1a, Tables and Graphs

Table 2: Descriptive Statistics

Variables	Mean	St. dev	Min	Max	Obs
Panel A. Yearly Treatment					
Clean Water Access pc	0.22	0.30	0.00	1.01	184
Number of Rooms With Access	2329	3453	0	13004	184
Panel B. Yearly Outcomes					
Infant Mortality Rate Girls	0.27	0.11	0.09	0.76	184
Infant Mortality Rate Boys	0.32	0.13	0.14	0.82	184
Mortality Rate Female	0.03	0.01	0.01	0.08	184
Mortality Rate Male	0.04	0.02	0.02	0.10	184
Dead Born Rate	0.05	0.02	0.01	0.09	184
Fraction Girls Born	0.49	0.03	0.37	0.55	184
Panel C. Yearly Outcomes: >1859					
First Month Mortality Girls	0.06	0.02	0.01	0.11	104
First Month Mortality Boys	0.07	0.02	0.02	0.12	104
$2-12$ Month Mortality Girls	0.19	0.07	0.04	0.47	104
$2-12$ Month Mortality Boys	0.21	0.08	0.08	0.54	104
Panel D: Demographics					
Population	14202	4664	3704	23767	184
Children born	408	142	103	707	184
Net Migration	0.03	0.03	-0.07	0.29	176
Panel E. Demographics in 1860					
Population in 1860	12420	3966	4599	17102	8
Children born in 1860	418	135	168	592	8
Mean Mortality Pre Reform	0.03	0.01	0.02	0.06	8
Houses per Area in 1860	0.06	0.08	0.01	0.24	8
Inhabitants Per House 1860	29.48	5.57	23.05	41.27	8
Inhabitants Per Area in 1860	1.71	2.00	0.17	6.37	8
Fraction Married Women in 1860	0.22	0.03	0.17	0.25	8
Inhabitants Born in Sthlm PC 1860	0.47	0.04	0.40	0.54	8
Fraction born out of wedlock 1860	0.35	0.08	0.24	0.48	8
Panel F: Monthly Treatment					
Clean Water Access Rooms pc	0.19	0.28	0.00	1.01	2208
Number of Rooms With Access	2075	3242	0	13004	2208
Panel G: Monthly Outcomes					
Mortality Rate Female	0.03	0.02	0.00	0.44	2207
Mortality Rate Male	0.04	0.03	0.01	0.51	2207

Table 3: Main Regression Results

Outcome in logs:	Mortality			Infant Mortality		
	All	Women	Men	All	Girls	Boys
	(1)	(2)	(3)	(4)	(5)	(6)
Clean Water Access	-0.300*	-0.402**	-0.213	-0.403	-0.892	-0.012
	(0.139)	(0.156)	(0.156)	(0.305)	(0.486)	(0.261)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes
Linear Time Trends	Yes	Yes	Yes	Yes	Yes	Yes
Outcome mean	0.036	0.032	0.041	0.295	0.272	0.319
R square	0.956	0.947	0.938	0.878	0.848	0.818
Observations	184	184	184	184	184	184

Note: Each column present different regressions with respect to the outcome variable. All outcomes are analysed in logs but I also provide unweighted outcome population means.

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the parish level in parentheses.

Table 4: All Age Log Mortality Rate by Season

	All	Summer 6 m	Winter 6 m	Summer 3 m	Winter 3 m
	(1)	(2)	(3)	(4)	(5)
Clean Water Access	-0.432**	-0.483**	-0.377*	-0.807**	-0.304
	(0.141)	(0.180)	(0.169)	(0.318)	(0.197)
Controls	Yes	Yes	Yes	Yes	Yes
Year*Month FE	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes
Linear Trends	Yes	Yes	Yes	Yes	Yes
Outcome mean	0.036	0.037	0.034	0.035	0.034
R square	0.783	0.802	0.764	0.795	0.751
Observations	2207	1104	1103	552	552

Note: Each column present different regressions using different samples of the monthly data. Summer six months (column 2) use only the months April until September while summer three months use only June through August. The three winter months are December to February. To make estimates somewhat comparable between yearly and monthly data, I have scaled the mortality measure with the yearly population divided by 12.

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the parish level in parentheses.

Table 5: Selection and Older Age Mortality

Outcome in logs:	Mortality > 3 Years			Stillbirths	Fraction
	All	Women	Men	All	All
	(1)	(2)	(3)	(4)	(5)
Clean Water Access	-0.326	-0.370**	-0.318	-0.875**	-0.068
	(0.185)	(0.140)	(0.302)	(0.306)	(0.053)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes
Linear Trends	Yes	Yes	Yes	Yes	Yes
Outcome mean	0.021	0.021	0.020	0.047	0.487
R square	0.893	0.864	0.876	0.614	0.301
Observations	184	184	184	184	184

Note: In column 5, the outcome is defined as $ln\left(\frac{Girls}{Girls+Boys}\right)$. All outcomes are analysed in logs but I also provide unweighted outcome population means.

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the parish level in parentheses.

Table 6: Comparison of Long and Short Data

	Full Sample			> 1859 Data Only		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Boys	Girls	All	Boys	Girls
Clean Water Access	-0.324	-0.032	-0.684	-0.510	-0.145	-0.986
	(0.342)	(0.280)	(0.463)	(0.386)	(0.266)	(0.580)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes
Outcome mean	0.295	0.319	0.272	0.298	0.323	0.274
R square	0.862	0.792	0.837	0.878	0.807	0.848
Observations	184	184	184	104	104	104

Note: The first three columns show full sample regression without linear trends. The estimates are in line with those using linear trends although smaller and less precise. The last three columns use data for > 1859 only for the same specification and show the same pattern although somewhat more negative. This specification, with controls showed to have the best properties when using the restricted data only.

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the parish level in parentheses.

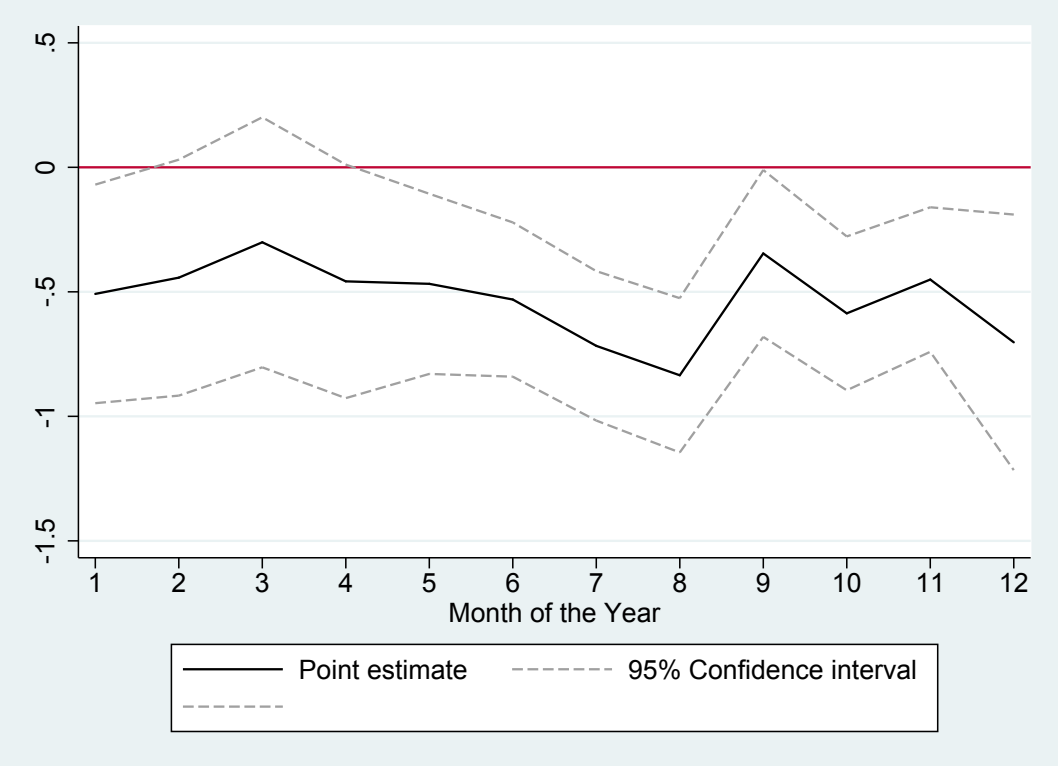
Table 7: Decomposing Infant Mortality by Age at Death

	First Month in Life			$2-12$ Months Mortality		
	(1) All	(2) Boys	(3) Girls	(4) All	(5) Boys	(6) Girls
Clean Water Access	-0.344	0.452	-1.452***	-0.698**	-0.537**	-1.016*
	(0.195)	(0.314)	(0.367)	(0.283)	(0.210)	(0.503)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes
Outcome mean	0.064	0.070	0.058	0.201	0.214	0.188
R square	0.612	0.408	0.513	0.856	0.789	0.774
Observations	104	104	104	104	104	104

Note: In this table I separate between mortality in the first month (columns $1-3$) and other infant mortality (columns $4-6$). The sample is restricted to $1860-1872$ data as there was no data on this level for earlier time periods.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the parish level in parentheses.

Figure 2: Monthly Treatment Variables



Note: Graphs show point estimates with corresponding confidence intervals at 95% for one regression including month times treatment interaction variables. The regression include month*year and parish fixed effects as well as linear parish specific time trends. Standard errors clustered at the parish level.

Table 8: Specification Checks

	Log Mortality	Log Infant Mortality
Robust Standard Errors	-0.300**	-0.403
	(0.138)	(0.259)
Clustering Parish and Year	-0.300*	-0.403
	(0.135)	(0.334)
Unweighted data	-0.361**	-0.581
	(0.134)	(0.320)
Lagged dependent variable	-0.293	-0.427
	(0.168)	(0.319)
Control for Net Migration	-0.501*	-0.621
	(0.215)	(0.348)
No Controls	-0.439**	-0.427
	(0.169)	(0.322)
No Time Trends	0.036	-0.324
	(0.236)	(0.342)
Quadratic time trends	-0.046	-0.185
	(0.237)	(0.379)
Quadratic time trends - no controls	-0.316*	-0.361
	(0.144)	(0.269)
Excluding year 1853	-0.424**	-0.422
	(0.156)	(0.320)
Excluding year 1857	-0.440**	-0.516
	(0.169)	(0.328)
Outcome not in logs	-0.011*	-0.093
	(0.006)	(0.084)
Placebo intervention in 1849	-0.086	-0.284
	(0.321)	(0.172)

Note: Each row in this table shows point estimates of clean water access on log mortality and log infant mortality for different specifications. The underlying model includes control variables, parish and year fixed effects as well as linear time trends where standard errors are clustered at the parish level and each cell is weighted by population (births) in 1860. Deviations from this base line is described to the left of the point estimates. The placebo intervention in the last row use data from 1845 and counterfactually assume that the clean water intervention started in 1849 and ended in 1860. Total number of observations for this model is 128.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the parish level in parentheses.

Table 9: Compositional Changes Correlated With Clean Water

Log	Poverty	Births	Population
	(1)	(2)	(3)
Clean Water Access	-0.081	0.125	0.088
	(0.278)	(0.135)	(0.076)
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes
Linear Trends	Yes	Yes	Yes
Outcome Mean	0.37	407.91	14201.86
R square	0.856	0.972	0.996
Observations	184	184	184

Note: Each column present different regressions with respect to the outcome variable. The outcome in the first column is the log fraction of children born out of wedlock, a proxy for poverty at the parish level. All regression are WLS using births measured in 1860.

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the parish level in parentheses.

Table 10: Bias From Poverty - Infant Mortality

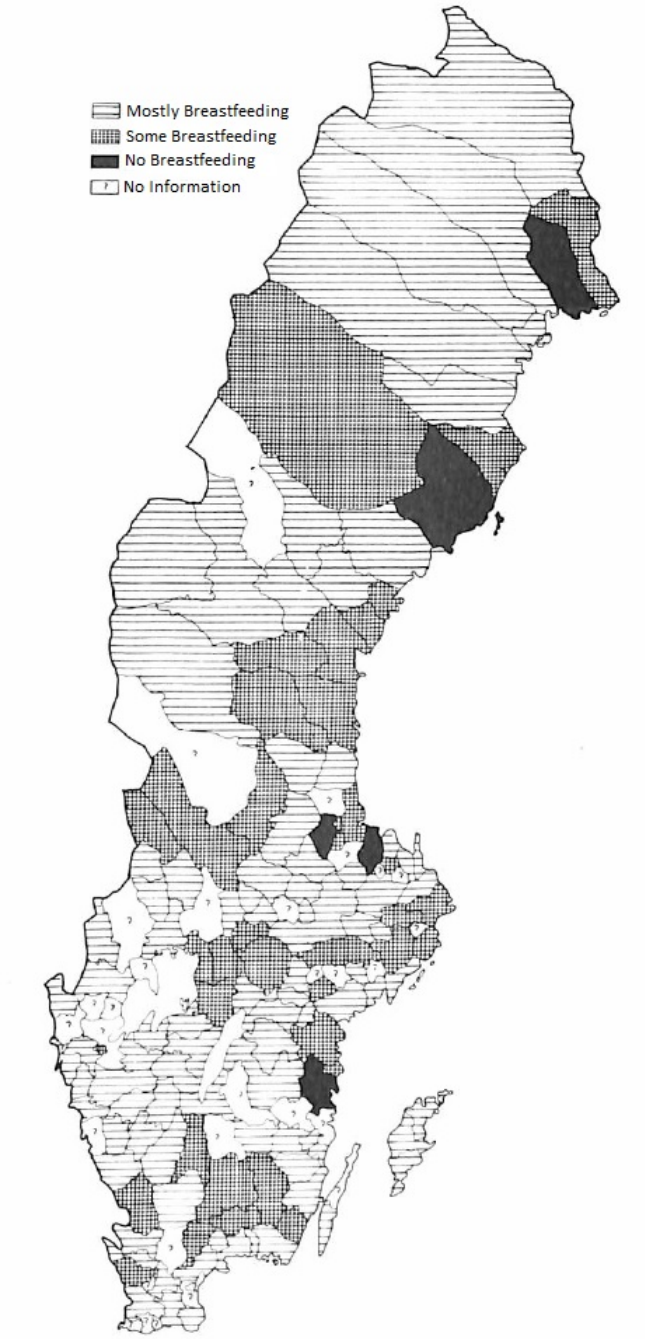
Control for Poverty	All		Boys		Girls	
	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)
Clean Water Access	-0.417	-0.384	-0.071	-0.068	-0.854	-0.808
	(0.309)	(0.322)	(0.255)	(0.249)	(0.481)	(0.496)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes
Linear Trends	Yes	Yes	Yes	Yes	Yes	Yes
Outcome mean	0.295	0.295	0.319	0.319	0.272	0.272
R square	0.880	0.885	0.826	0.827	0.850	0.854
Observations	184	184	184	184	184	184

Note: Each column present different regressions with respect to the outcome variable. The outcomes are log infant mortality for all children (columns $1-2$), boys only (columns $3-4$) and girls only (columns $5-6$). For each outcome, I show treatment effects of the clean water intervention without control for poverty (first column in each pair) and with (second column in each pair). I measure poverty here as the log of the fraction of children born out of wedlock. A measure closely related to poverty at this time. The estimates without the poverty measure will be slightly different from previous estimates as I am excluding the 1860 control "born out of wedlock" as a separate control variable.

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the parish level in parentheses.

7 Appendix 2, Descriptive Maps and Figures

Figure 3: Breastfeeding Practices in Sweden circa 1870

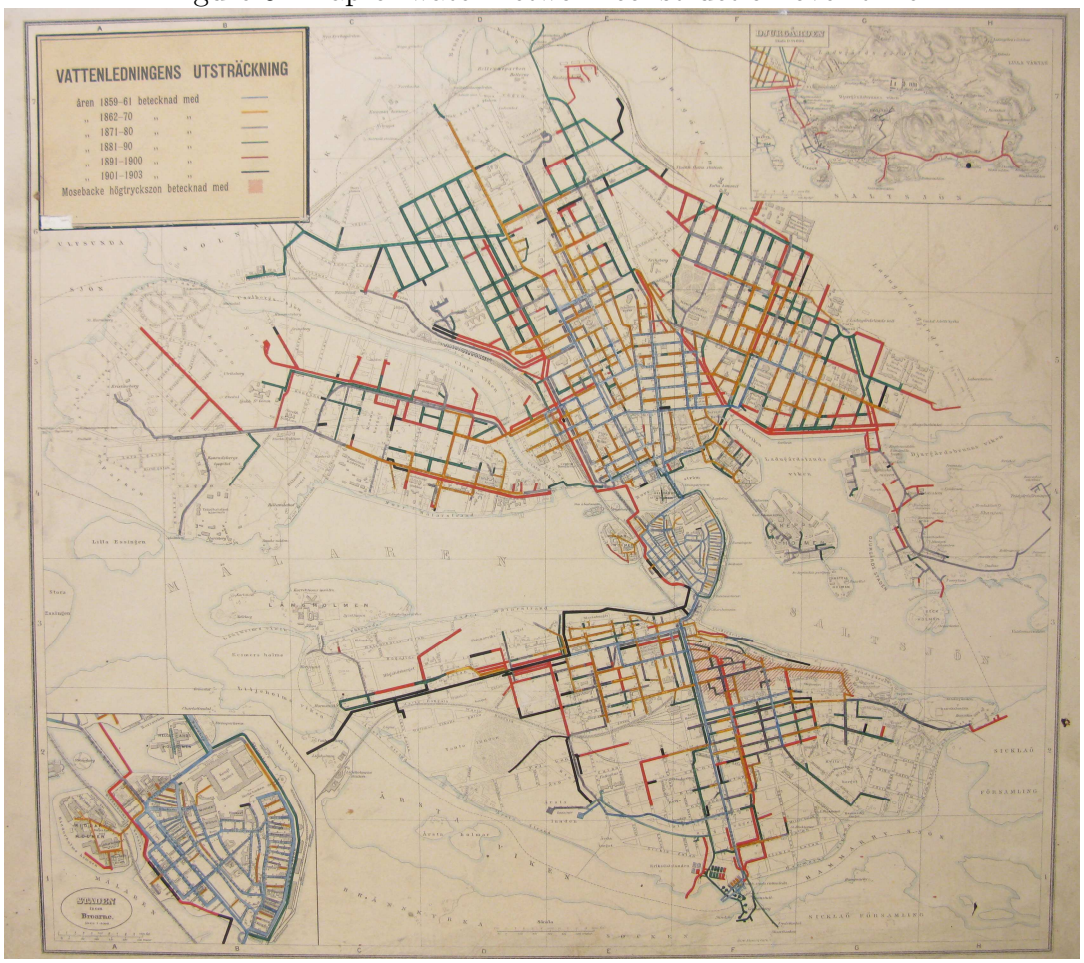


Note: Source: A. Brändström, "De kärlekslösa mödrarna" Dissertation at Umeå University, 1984.

Figure 4: Parishes in Stockholm as of 1905

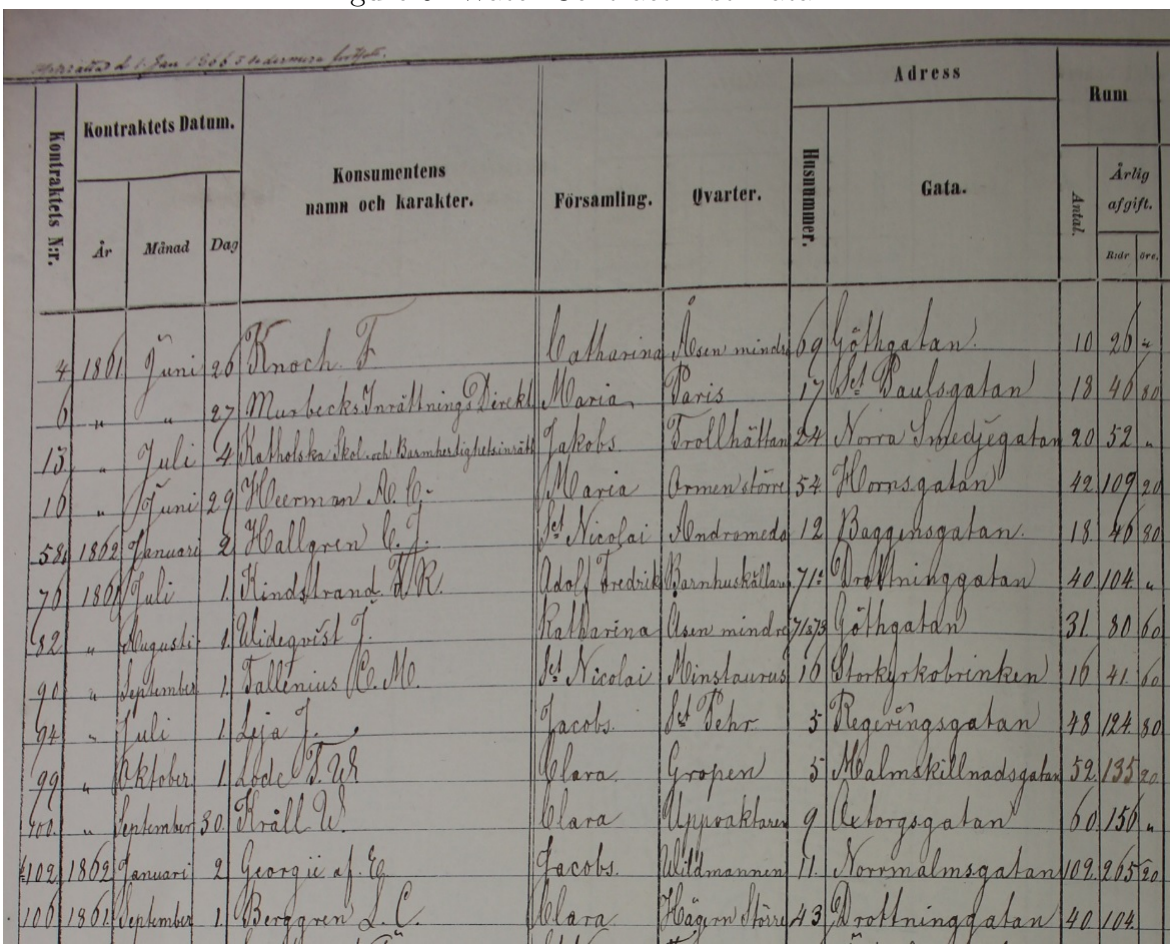
Note: The map has a later origin than the period studied here but gives a good approximation to the parish borders at the earlier time. These borders where not changed substantially. Skeppsholmen (the small island to the right of Old Town, small mid island) was a parish of its own originally but had very few inhabitants. Bold numbering $1-8$ shows the parishes that are separated by thick lines. Parish names are placed in the map. The eight geographical parishes of Stockholm complemented the 24 other non-geographical small parishes existing at that time. These 24 consisted of non protestant confessional groups, military installations, parts of the royal court and other institutions such as hospitals. Together they only accounted for a small fraction of the total population and did not have any geographical boundaries. The geographical parishes included these others as inhabitants but did not have responsibility to register them . Södermalm consisted of two parishes divided almost in the middle from north to south (called Maria and Katarina). The small island in the middle, Old town, was a parish of its own. The large island to the west is Kungsholmen which was separated from the rest through water but connected by bridges. Normalm, the northern part of Stockholm, was divided in four parishes with somewhat arbitrary borders (Klara, Jakob and Johannes, Adof Fredrik and Hedvig Eleonora). Source: Statistical year book for the city of Stockholm `http://www.stockholmskallan.se/Soksida/Post/?nid=18212`.

Figure 5: Map of water network construction over time



Note: Blue lines represent the distribution network at the time of the introduction in 1861. Yellow lines show the extra pipes laid during the 1860s and the other colors represent later construction. Source: Stockholm Water Company archive at Stockholm City Archives. Accessed with permission from Stockholm Vatten.

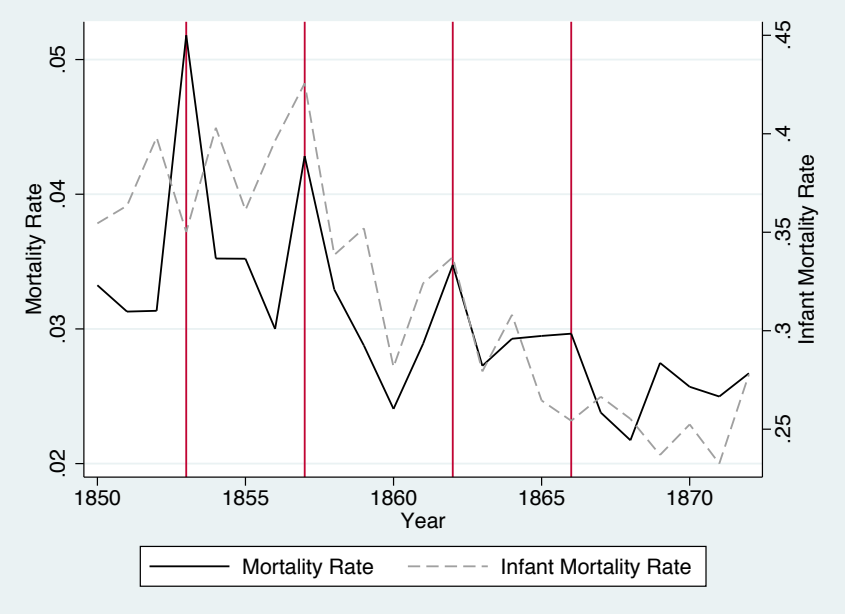
Figure 6: Water Contract List Data



Kontraktets Nr.	Kontraktets Datum. År	Månad	Dag	Konsumentens namn och karakter.	Församling.	Qvarter.	Adress. Husnummer.	Gata.	Rum. Antal.	Årlig afgift. Riksdr	Öre.
4	1861	Juni	26	Knoch F.	Catharina	Åsen mindre	69	Götgatan	10	20	4
6	"	"	27	Murbecks Inrättnings Direkt.	Maria	Paris	17	S:t Paulsgatan	18	40	80
15	"	Juli	4	Katholska Skol- och Barmhertighetsinrätt.	Jakobs	Trollhättan	24	Norra Smedjegatan	20	52	"
16	"	Juni	29	Heerman & Co.	Maria	Ormen stora	54	Klarasgatan	42	109	20
58	1862	Januari	2	Hallgren C. J.	S:t Nicolai	Andromeda	12	Baggensgatan	18	46	80
70	1861	Juli	1	Lindstrand A. R.	Adolf Fredrik	Barnhuskällan	71	Drottninggatan	40	104	"
82	"	Augusti	1	Widegvist J.	Katharina	Åsen mindre	7 1/3	Götgatan	31	80	60
90	"	September	1	Tallenius C. M.	S:t Nicolai	Minotaurus	16	Storkyrkobrinken	16	41	60
92	"	Juli	1	Leja J.	Jacobs	S:t Peter	5	Regeringsgatan	48	124	80
99	"	Oktober	1	Lode P. A.	Clara	Gropen	5	Malmskillnadsgatan	52	135	20
101	"	September	30	Kroll W.	Clara	Apparaten	9	Regeringsgatan	60	156	"
102	1862	Januari	2	Georgii af E.	Jacobs	Wildmannen	11	Norrmalmsgatan	102	265	20
106	1861	September	1	Berggren L. C.	Clara	Hägern Stora	43	Drottninggatan	40	104	

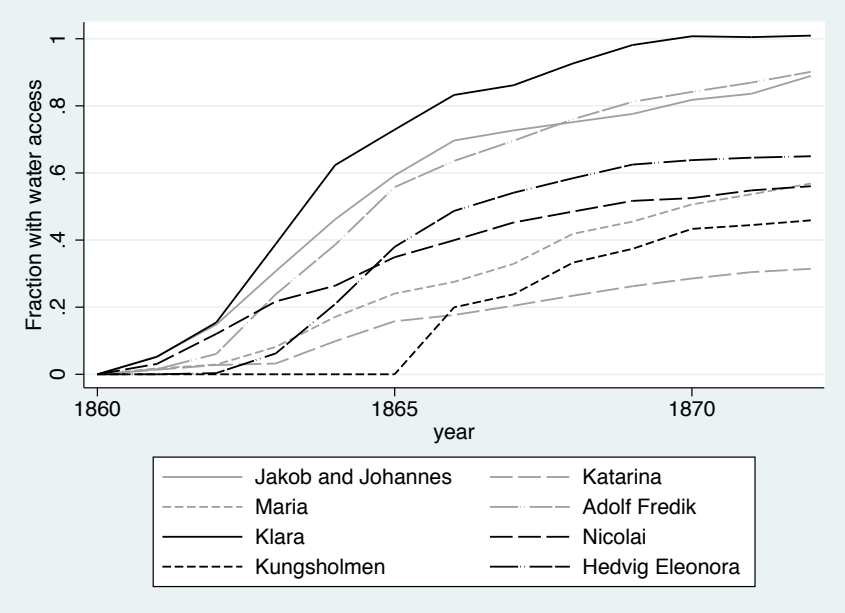
Note: Contract lists between 1861 and 1872. From the left the data includes contract number, contracting date, contract holders name, resident parish, block name, street address and number of rooms debited (total yearly cost). Source: Stockholm Water Company archive at Stockholm City Archives.

Figure 7: Mortality and Infant Mortality Rate in the Geographical Parishes of Stockholm 1850 – 1872



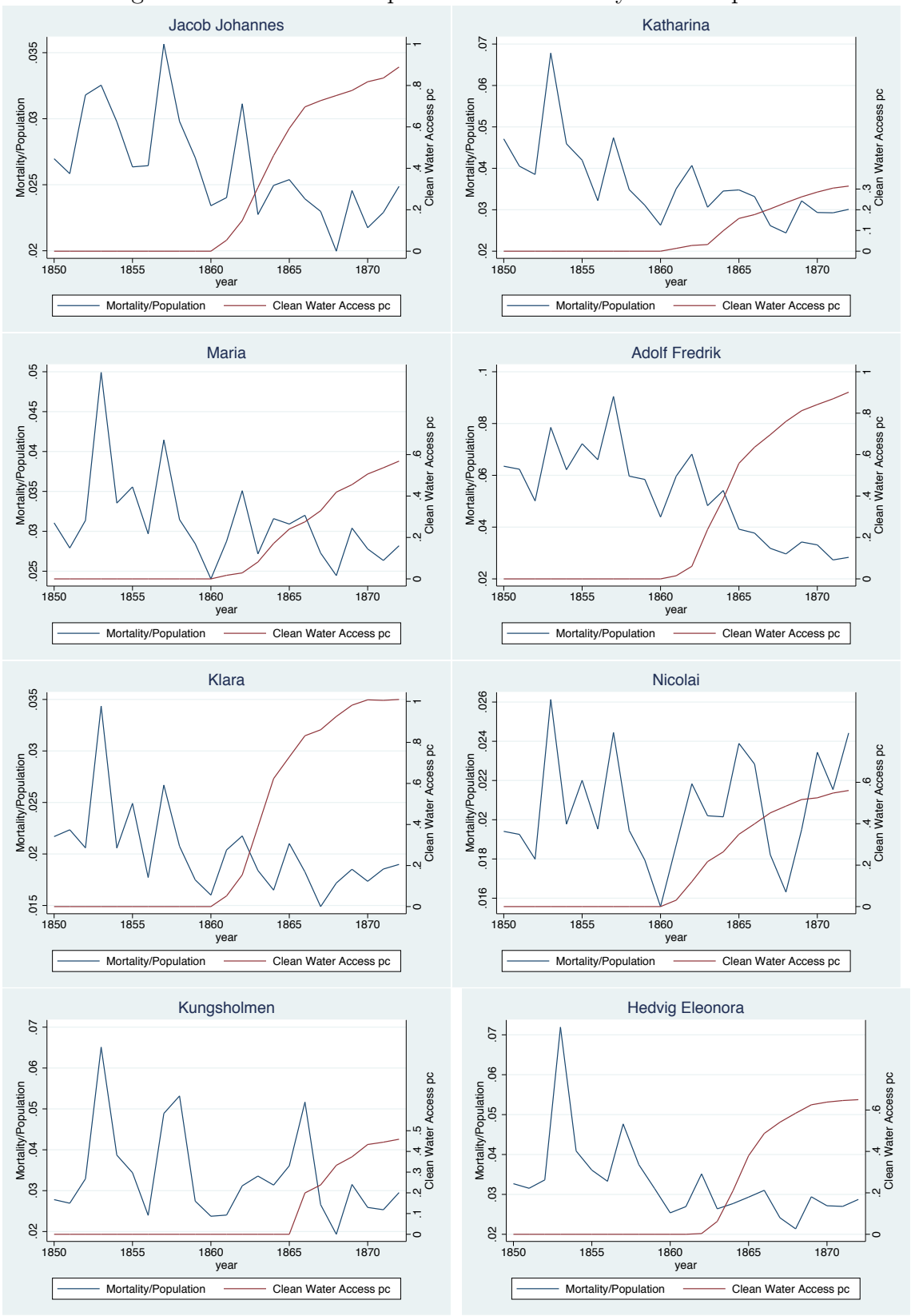
Note: Mortality rate calculated as total number of fatalities a given year divided by the tax registered population. Infant mortality is calculated as the number of deceased children between ages zero and one in a given year divided by the number of alive born children. Vertical lines represent years with disease outbreaks. Cholera in 1853, 1857, 1866 and measles in 1861

Figure 8: Variation in treatment between parishes in Stockholm 1861 – 1872



Note: Number of rooms with water access retrieved from contract list from Stockholm water company. The crude numbers have been scaled on a yearly basis, keeping the between parish variation, to align with official reports of private household rooms with access. The discrepancy could be generated by firm contracts or double reporting as buildings where sold. Number of rooms are divided by population in each parish in 1860.

Figure 9: Clean Water expansion and mortality for each parish



Note: Mortality rate and clean water access for the 8 parishes under investigation in this paper. Mortality rate on the left hand axis and clean water access on the right hand. Left hand axis differ between parishes but the right hand is fixed to be equal for all.