The Value of Piped Water and Sewers: Evidence from 19th Century Chicago

Michael Coury,[*] Toru Kitagawa,[†] Allison Shertzer,[‡] Matthew A. Turner [§]

17 December 2021

Abstract: *We estimate the impact of piped water and sewers on property values in mid-19th century Chicago. The cost of sewer construction depends sensitively on imperceptible variation in grade, and such variations in grade delay water and sewer service to part of the city. This delay provides quasi-random variation for causal estimates. We extrapolate ATE estimates from our natural experiment to the area treated with water and sewer service during 1874-1880 using a new estimator. Water and sewer access increases property values by more than a factor of two. This exceeds costs by about a factor of 60.*

JEL: O18, R3, L97, N11

Keywords: Piped water and sewer access, Infrastructure, Extrapolation

[*]University of Pittsburgh, Department of Economics. email: mrc93@pitt.edu.

[†]Brown University, Department of Economics, Box B, Brown University, Providence, RI 02912. email: toru_kitagawa@brown.edu.

[‡]University of Pittsburgh ,Department of Economics. email: shertzer@pitt.edu. Also affiliated with the NBER. Coury and Shertzer gratefully acknowledge the support of the National Science Foundation (SES-1918554) which funded much of this research.

[§]Brown University, Department of Economics, Box B, Brown University, Providence, RI 02912. email: matthew_turner@brown.edu. Also affiliated with PERC, IGC, NBER, PSTC, S4. Turner gratefully acknowledges the support of a Kenen felelowship at Princeton University during some of the time this research was conducted.

1 Introduction

We estimate the impact of piped water and sewers on land values in mid-19th century Chicago. To conduct this estimation, we rely on novel, purpose-collected data describing Chicago land transactions in the late 19th century, and detailed annual maps of piped water and sewer networks. To identify causal effects, we exploit the fact that the construction cost for sewers varies sensitively with variations in grade that are otherwise imperceptible and, therefore, affect land values only through their effect on the timing of piped water and sewer access. We propose a new estimator to extrapolate treatment effects from the small region where we can defend our natural experiment to a region that is more relevant for cost-benefit analysis. In our most conservative estimate, we find that access to piped water and sewers more than doubles the value of residential land in Chicago. Aggregating this increase over affected parcels and comparing to construction costs, we find that the benefits of piped water and sewer infrastructure exceed costs by about a factor of 60.

These results are of interest for several reasons. First, according to the World Bank, about 15% of the world's urban population did not have access to safely managed drinking water in 2020. A larger share, about 40%, did not have access to safely managed sanitation facilities.[1] Given the likely impact of safely managed water and sanitation on health and mortality, the provision of such services would seem to be a priority. Yet, many cities also lack other basic services such as decent roads, sufficient public transit, adequate schooling and reliable electricity. Thus, trade-offs inevitably arise. By providing estimates of the benefits of piped water and sewer access, we hope to inform policy makers facing such trade-offs.

Second, our estimates inform us about an important aspect of the development of the American economy during the late 19th and early 20th centuries. Economic historians have long emphasized the importance of public health infrastructure for the development of American cities (Ferrie and Troesken, 2008). The existing literature on sanitation investments relies almost entirely on time series or panel data relating city-level changes in health and mortality to changes in the availability of particular public health interventions

[1] `https://data.worldbank.org/indicator/SH.H2O.SMDW.UR.ZS` and `https://data.worldbank.org/indicator/SH.STA.SMSS.UR.ZS`, Accessed December 15, 2021.

(e.g., Cutler and Miller (2005), Alsan and Goldin (2019)). However, this time period also saw changes in food purity laws, improvements in water and sewer access and quality, widespread acceptance of the germ theory of disease, and dramatic increases in income that could confound estimates based on time-series variation, and results in Anderson et al. (2018) suggest that this concern is not purely hypothetical. We contribute to this debate by estimating the value of piped water and sewer infrastructure using a novel cross-sectional identification strategy to provide new evidence for the importance of capital-intensive public health interventions in the development of American cities.

Third, we pioneer a new identification strategy for estimating the causal effects of sewers. The effects of sewer access on the development of cities and the well being of their inhabitants have been much less studied than have the effects of other types of infrastructure such as electrification or transportation. This partly reflects the intrinsic difficulty of observing underground pipes. But it also reflects the lack of a compelling identification strategy. We hope that our research design will prove portable, and will facilitate research on the effects of sewer and water infrastructure in cities of the modern world.

Finally, building on the marginal treatment effect model proposed by Carneiro et al. (2011), we develop a method for extrapolating treatment effects from a quasi-experimental region to a more economically relevant region. The reliance on small, carefully constructed samples to identify the effects of location specific policies is common, and our hope is that our technique will permit researchers using such designs to extrapolate their results to more relevant samples in a principled way.

2 Literature

The implementation of public health initiatives in the late 19th and early 20th century has been carefully studied. In a landmark study, Cutler and Miller (2005) look at changes in mortality rates following the start of water filtration and chlorination in 13 large American cities between 1900 and 1936. They estimate that water filtration reduced child mortality by 0.46 log points and infant mortality by 0.43 log points. These are, respectively, declines of 37% and 35% against 1900 baseline rates of 28/1000 and 190/1000. Chlorination alone has no measurable effect but chlorination has a small beneficial impact when interacted with filtration.[2] Comparing the value of foregone infant mortality to the cost of water infrastructure, Cutler and Miller (2005) conclude that the ratio of benefits to costs for water filtration programs was about 23:1.

The subsequent literature typically relies on differences in sanitation conditions within a city or metropolitan area. Ferrie and Troesken (2008) estimate the effects of improvements to the quality of the municipal water supply on the crude death rate in Chicago between 1853 and 1925, a period when this rate declined from 27/1000 to 11/1000. Comparing mortality rates before and after three major improvements, they conclude that these improvements were responsible for between 32 and 52% of the total decline in crude death rates. Alsan and Goldin (2019) examine the effect on infant mortality rates of a series of interventions to protect drinking water quality in the Boston Harbor watershed between 1880 and 1915. They estimate that these interventions caused a decline in infant mortality rates of 0.21 log points, 19%, from an 1880 level of 163/1000. Kesztenbaum and Rosenthal (2017) examine the effect of the increasing availability of sewers in Paris between 1880 and 1915 and finds that a 10% increase in neighborhood sewer connections increases neighborhood mean life expectancy, conditional on reaching age one, by 0.13 years. Beach (2021) argues that the various innovations in municipal sanitation and water supply were responsible for the elimination of typhoid in American cities between 1900 and 1930.

All of the papers in this literature are similar in that identification of treatment effects relies on the comparison of outcomes before and after a public health innovation. However, as Haines (2001) documents, the late 19th and early 20th century saw the widespread adoption of vaccination, the development of the germ theory of disease, the increasing availability of refrigeration, and the widespread adoption of food purity standards, particularly for milk. It is thus is natural to suspect that some of the effects that the papers described above attribute to improved water and sanitation are actually due to one of these other factors. Indeed, a literature in economic history has argued that rising incomes in the late 19th and early 20th century reduced mortality by improving nutrition

[2]This is what one would expect on the basis of the modern understanding of water treatment. Chlorination is more effective at eliminating pathogens once other suspended solids have been removed (Drinan et al., 2000, p.111).

and resistance to disease (McKeown (1976), Fogel et al. (2004), Eli (2015)).

In an attempt to address this possibility, Anderson et al. (2018) performs a reanalysis of the data used in Cutler and Miller (2005). Relative to Cutler and Miller (2005), they make two changes. First, they correct a small number of errors in the the data on which Cutler and Miller (2005) is based. Second, in addition to water filtration and chlorination, they consider measures of clean water projects, sewage treatment projects, and two distinct milk safety standards. They find that water filtration decreases infant mortality by 0.13 log points, or 9%. This is about one quarter of the effect estimated by Cutler and Miller (2005). In addition, Anderson et al. (2018) find that the joint effect of all public health initiatives on the crude death rate cannot be distinguished from zero.[3]

Summing up, the literature examining the effects of public health innovations in late 19th and early 20th century America overwhelmingly relies the comparison of outcomes before and after implementation. At present, the papers in this area have produced estimates of the value of water and sewer infrastructure ranging from modest, as in Anderson et al. (2018), to enormous, as in Cutler and Miller (2005) or Alsan and Goldin (2019). In light of this uncertainty, our introduction of a new, cross-sectional identification strategy should be particularly helpful.

Unlike the extant literature, our outcome is land prices rather than a measure of morbidity or mortality. Calculating benefits from estimates of such health impacts requires the extra step of valuing illness or lost lives, a particularly difficult exercise for historical settings (Costa and Kahn (2004)). The focus on mortality also precludes an evaluation of non-health benefits from sewers and piped water, and contemporary descriptions of urban life without sewers suggest that these benefits were important. Thus, our focus on land prices allows an immediate calculation of benefits and reflects all place-specific benefits. Using our research design to study health and mortality is a natural topic for further research, should suitable data become available.

Water and sewer infrastructure in developing countries is also the subject of a small literature. Ashraf et al. (2017) find that interruptions to piped water

[3]Anderson et al. (2018) and Cutler and Miller (2005) also rely on different population data for part of the study period. This difference makes an important contribution to the difference between the two papers. See Cutler and Miller (2020) and Anderson et al. (2019) for more detail.

supplies in urban Lusaka have a significant impact on the incidence of diarrhea and typhoid, and are associated with an increased time at chores and decreased time at study for young women. Devoto et al. (2012) find that randomly assigned help obtaining credit for piped water connections significantly increases time allocated to leisure activities in an RCT conducted in Tangiers in 2007. Kremer et al. (2011) examine the effect of randomized improvements of springs in rural Kenya around 2005. They find that these improvements lead to reductions in childhood diarrhea but that the willingness to pay for such improvements is low. Galiani et al. (2005) examine the effects of privatizing the provision of municipal water supplies in Argentina in the 1990s and conclude that the resulting improvements in service quality reduced child mortality by 8%.

Finally, Gamper-Rabindran et al. (2010) investigate the relationship between increased access to piped water and sewers in Brazil between 1970 and 2000. During this period, the share of households with piped water increased from 15 to 62% and the infant mortality rate fell from 125/1000 to 34/1000. On the basis of a panel data estimation, they conclude that each percentage increase in piped water access decreases infant mortality by 0.48/1000. Thus, the realized expansion in piped water access decreased infant mortality by $(62-15) \times 0.48 \approx 22/1000$, about 25% of the total decrease of 91/1000. Gamper-Rabindran et al. (2010) also examine the effects of increased sewer access and find no effect.

Summing up, the available literature on the effects of improvements to water supply or sewer access broadly supports the hypothesis that such innovations are important contributors to health, particularly of children, and to well being more broadly defined. However, only Gamper-Rabindran et al. (2010) provides an analysis of policies to construct urban water and sewer networks, and ours is the only examination of the effect of piped water and sewer infrastructure on land prices. Given the ability of land prices to capitalize place specific benefits, this means that our estimates provide a unique foundation for the evaluation of the benefits of piped water and sewer construction projects.

In addition to our primary object of estimating the effects of piped water and sewer infrastructure on land prices, we develop a new method for extrapolating estimates based on a quasi-experiment to a more economically relevant sample for which quasi-random assignment of the treatment is not available. Our

Figure 1: Land transactions in the Chicago Tribune

SATURDAY'S TRANSFERS.

The following instruments were filed for record Saturday, April 10:

CITY PROPERTY.

Walnut st, 120 ft e of Western av, s f, 30x 126 ft, dated April 10 (A. E. and C. M. Hemler to John T. Shannon)............$ 2,025

West Superior st, 49 4-10 ft e of Lincoln, n f, 25x128 ft, dated April 10 (B. F. Crosby to O. B. Olson).................. 600

Cleaver st, 225 ft s of Bradley, w f, 37½ x125 ft, dated April 8 (Mat Schillo et al. to M. Kufel et al.)..................... 750

West Madison st, 428 ft w of Staunton, s f, undivided ⅝ of 24x126 ft, dated April 6 (Mary J. Seymour to C. L. Wehe).................................. 2,400

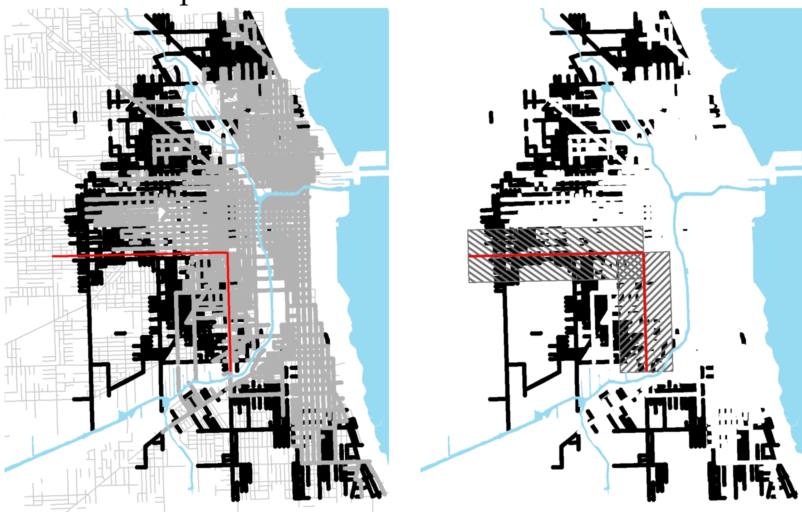
Note: *An example of listings of land transactions in the Chicago Tribune. Our land transaction data results from digitizing all transactions reported on Saturday between 1873 and 1889. Note that each record reports the nearest intersection, price, and area. Most records also report if the parcel is "improved" or "corner."*

approach to this problem builds on the marginal treatment effects estimator developed by Heckman and Vytlacil (2005) and Carneiro et al. (2010) but extrapolates to units not in the original estimation sample. Other methods for extrapolating causal effects to populations other than the sampled population include Hotz et al. (2005), Angrist and Fernández-Val (2013), Andrews and Oster (2019), and Dehejia et al. (2021). There is also a small literature (Angrist and Rokkanen (2015), Rokkanen (2015), and Cattaneo et al. (2020)) considering the related question of extrapolating treatment effects estimated using an RDD design to points away from the discontinuity. The possibility of extrapolation from quasi-experimental samples to more economically relevant samples based on marginal treatment effect estimates has not been previously considered.

3 Data

Our main empirical exercise requires two main types of data, a measure of land values and a measure of piped water and sewer access. For econometric purposes, we also require a description of the attributes of transacted parcels. To complete our cost benefit analysis, we must also measure construction costs. We here describe the data we use for each purpose.

Figure 2: Extent of piped water and sewer network, Southwest Triangle, and Quasi-experimental samples



(a) (b)

Note: *(a) Sewers before 1874, during 1874-1880, after 1880, and boundaries of the Southwest triangle. (b) "Relevant' sample area (1874-1880 expansion) and "Quasi-experimental' sample areas.*

Between 1873 and 1889, the Chicago Tribune reports every land parcel transaction filed with the municipal title office on the previous day. We collect all transactions listed in the Sunday edition, which is usually the day of the week with the largest number of listings. This results in about 700 observations per year in the 1870s and 1000 per year in the 1880s.[4]

The Tribune consistently reports; price, parcel dimensions, either a street address or the nearest intersection, and whether the parcel is "improved." Figure 1 illustrates a sample of transaction listings. Because the Tribune separately indicates transactions with a "premises", that is, parcels with a structure, we are confident that our data describe land transactions only. The newspaper does not define "improved" and it is clear from the data it does not refer to water and sewer access or to the presence of a structure. We believe that "improved" indicates that the parcel fronts a paved road.

We geocode our sample parcels in two steps. First, we attempt to match the "nearest intersection" reported by the Tribune to an intersection in the

[4]The Tribune still published parcel transactions after 1889, but the coverage is limited to parcels with a value of at least $1000 (nominal value).

contemporary street grid described by the Google Maps API. When we cannot match a reported intersection to the contemporary street grid, we attempt to match it to an intersection in the circa 1880 street map created by Logan et al. (2011). This process allows us to geocode about 77% of transactions by assigning them the coordinate of their nearest intersection.[5]

We rely on historical GIS maps describing the block-by-block expansion of the sewer network from 1830-1930 Fogel et al. (2014). These maps derive from the annual reports of the Chicago Department of Public Works and report both the location and opening date for each segment of the sewer network. Water and sewer service was almost always installed simultaneously, and so we rely exclusively on sewer maps.

We say a transaction "has water and sewer access" if the nearest intersection to the transaction is within 75 feet of an operating sewer line in the transaction year. Visual inspection of the matching process indicated that this rule resulted in an accurate matching of *intersections* to sewers. One can imagine situations in which a *parcel* without access to sewer and water matches to an *intersection* where access is available, though such situations should be rare.[6] False negatives are harder to imagine.

The left panel of figure 2a illustrates the expansion of piped water and sewer access during the post-Civil War period. In this figure, the heavy, light gray lines indicate water and sewer lines predating our 1874-1880 study period. Unsurprisingly, these lines tend to be close to the center of the city. Heavy black lines indicate water and sewer lines constructed during our 1874-1880 study period. Also unsurprisingly, these lines are mostly located on the periphery of the previous network. Finally, the fine gray lines indicate sewer and water lines built after the end of our study period; these lines are also peripheral to the 1880 network and often extend beyond the boundary of the figure.

We calculate a number of control variables from GIS data layers. For each parcel, we calculate distance to the CBD as the distance to City Hall in 1873 (now known as the Rookery Building). We calculate distance to the lake as distance to

[5] Addresses are not universally reported for our transactions and Chicago undertook a complete renumbering of addresses in 1909. This rules out the geocoding of addresses.

[6] A parcel on a street without water and sewer service could match to an intersection where the cross-street has water and sewer access.

the modern lakeshore,[7] and calculate distance the Chicago River similarly. Finally, we calculate distance to a horse car line and a major street using contemporaneous maps of the two networks.[8]

To estimate the cost of piped water and sewer expansion, we rely on reports of annual expenditures on water and sewer construction in the Annual Reports of the Chicago Department of Public Works (accessed through Hathi Trust). Expenditures vary year to year but are increasing in the early 1870s and decline during the recession of the late 1870s. Waterworks, including pumping stations, were typically the largest category of expenditure, with sewer construction second. Sewer maintenance costs, including manual flushing (discussed below), were stable and relatively small throughout the period. Expansions to the sewer and water system were primarily financed by bonds, and nineteenth-century Chicago had a large tax base of valuable land on which to levy the property taxes were the primary source of revenue to service these bonds.[9]

4 Background

The Census reports Chicago's population as 300,000 in 1870 and above one million in 1890. The Great Fire of 1871 destroyed the central business district and much of the city, but barely checked this growth. The city continued to expand throughout the 1870s and 1880s, particularly in the band of mostly unsettled land a few miles from the downtown where our study area lies. This rapid growth was driven by immigrants from Europe and by internal migration. Chicago provided relatively high-wage employment opportunities for unskilled workers. The average income per laborer in the city of Chicago was as high as

[7]The hydro file was obtained from Cook County Government Open Data, see https://datacatalog.cookcountyil.gov/GIS-Maps/Historical-ccgisdata-Lakes-and-Rivers-2015/kpef-5dtn.

[8]The 1880 horse-drawn streetcar routes were digitized using a map from the Illinois State Grain Inspection Department. The street network in 1880 was digitized by John Logan, see https://s4.ad.brown.edu/Projects/UTP2/39cities.htm

[9]Special assessments and connection fees also helped to finance sewer and piped water infrastructure. However, the Sewerage Board was reluctant to rely too heavily on fees and user charges because the resulting negotiations with building owners slowed down the expansion process (Melosi, 2000, p. 98).

$650 in 1880 dollars or $17,000 in 2021 dollars.[10]

Hoyt (2000) describes Chicago's land market between 1830 and 1930. He reports rapid growth in the value of land in the early 1870s. Farms that sold for $25 to $100 an acre were platted into town lots that sold for $400 to $1000 immediately thereafter (Hoyt, 2000, p. 108). Prices declined from their peak after the panic in 1873 and the value of the land within city limits declined 50 percent by 1877. Speculative landlords had "their cup of misery filled to the brim" in 1877 when the largest savings banks in the city of Chicago also failed (Hoyt, 2000, p. 123).[11] Economic conditions improved in the early 1880s and Chicago's land values recovered to their peak value of 1873 by 1882 (Hoyt, 2000, p. 140). Population growth and land prices were both relatively stable during the following decade. In short, our 1874-1880 study period spans a major recession (1873-1877) and recovery (1878-1882). Several years of moderate growth followed. Population growth was robust throughout the whole period from 1870-1890.

Chicago's infant mortality rate in the 1870s was a staggering 74 per 1000. This is similar to contemporaneous rates reported in other US cities, e.g., Alsan and Goldin (2019) or Haines (2001), and also current rates in poor developing countries like Sierra Leone or Somalia.[12] Most deaths were caused by infectious disease and occurred predominantly among the young (Ferrie and Troesken, 2008).

In the 1850s, the quality of Chicago's drinking water was notably poor. Most residents drank from backyard wells. These wells were often near privy vaults and these vaults were seldom tight. Households with access to the city water system found it contaminated by industrial pollutants and minnows from Lake Michigan. Water quality improved as the city moved the water intakes further

[10] From estimates of wages per non-agricultural worker for the state of Illinois taken from (Easterlin, 1960, 73-140) ($627 per year) and Hoyt's (2000, pp.118-119) estimates of wages for workers in the city of Chicago during the 1870s ($3 a day for unskilled laborers). These values were inflated to 2021 price levels using CPI estimates from Sahr (2009) for 1880-1912 and the BLS CPI series for 1913-.

[11] Hoyt used 1879 prices to proxy for the bottom of the market in 1877 because it was difficult for him obtain data for this year. Our data reports transactions in 1877 and 1878.

[12] Estimate for Chicago taken from Ferrie and Troesken (2008) and for Africa from the UN Inter-agency Group for Child Mortality Estimation (UNICEF, WHO, World Bank, UN DESA Population Division) at childmortality.org.

out into Lake Michigan and reduced the volume of waste dumped in the lake. Specifically, water quality improved with the completion of the Two Mile crib (1867), the Four Mile crib (1892), and the complete reversal of the Chicago River in 1900 (Ferrie and Troesken, 2008). Importantly, our study period (1874-1880) is located entirely within the Two Mile crib period.

The condition of the City's poorly drained streets was grim. The well-known Chicago history, (Asbury, 1940, p.23) reports that the "gutters [run] with filth at which the very swine turn up their noses in supreme disgust. . . ". When storms washed these wastes into Lake Michigan or private wells, cholera and dysentery epidemics followed. Such events killed hundreds of people in both 1852 and 1854, prompting the city to begin planning the improvements to its water and sewer infrastructure that we discuss below.

Typical gravity fed sanitary sewers require a grade of about 1:200 to prevent suspended solids from settling and blocking the pipe. The precise required grade is sensitive to the details of the system; the rate of flow, pipe size and cross-sectional shape, and the smoothness of interior walls. For details see, e.g., Mara (1996). Importantly, variation in grade that is critical for sewer construction is practically beyond human perception. Aldous (1999) reports that people begin to perceive a playing field as sloped at a grade of about 1:70. Variation in grade is less relevant to piped water networks.

Our research design will be organized around transactions that occurred in the area around Tyler Street, currently the Eisenhower Expressway, and extending West about three miles from Halsted Street. The present day corner of Halsted and Tyler streets is about two miles from and twelve feet above the level of Lake Michigan, a grade of about 1:880. This is much too flat for conventional gravity-fed sanitary sewers. Indeed, such grades are so flat that water generally does not drain away. Rainfall either evaporates or is absorbed into the ground. Chicago's unusually flat terrain contributes to the benefits of sewers as well as to the difficulty of constructing them.

Chicago hired noted engineer Ellis Chesbrough to design a sewer system capable of operating in Chicago's flat topography, and substantially followed the proposal he submitted in 1855. Chesbrough proposed what is now known as a "combined" sewer system to manage household sewerage and street runoff. Chesbrough's plan called for continuous mechanical flushing, although the city ultimately adopted a system under which sewer mains were manually flushed using water delivered by horse-drawn carts.[13] This systematic manual flushing allowed sewer mains to operate at a grade of 1:2500, far shallower than conventional sewers.

To function, even Chesbrough's sewers require large enough flows of water that they are only practical if piped water is available. For this reason, sewers cannot be installed before piped water. In fact, drainage in Chicago was so poor, that the increased volume of wastewater that accompanied piped water caused cesspools to overflow (Melosi, 2000, p. 91), so that installing piped water without sewer access was also impractical. For these reasons, the provision piped water and sewer access almost always coincided.

Because water and sewer service are almost always provided together, we estimate their joint value. With this said, the discussion above points out that water and sewer service were highly complementary, so that providing one without the other would probably have had much less value.

Construction of Chesbrough's sewers required a massive program of regrading in order to raise streets to the required grades. The process for constructing sewers involved first laying sewer and water pipes at the required grade, whether above or below ground, and then filling in the space around them with earth as required. The newly raised streets were then sometimes paved over to conclude the process. Because street paving could independently contribute to property values, this raises the possibility that our estimates reflect the joint value of water, sewer and street paving. We address this possibility by controlling for improved status in our estimations.

Buildings, particularly those built out of stone and brick, were raised in the downtown to match the new street level as the sewer system expanded. These well-known feats of engineering predate our 1874-1880 study period. Our analysis focuses on vacant lots in outlying areas.

Chicago issued its original plan for sewerage in 1855. This document describes the street grades in each region of the city needed to accommodate the proposed sewer system (Plan of Sewerage, Chicago Board of Sewerage Commissioners, 1855). Subsequent ordinances were issued at regular intervals as

[13] As late as 1940, horse-drawn tanks were still used to manually flush certain sewer lines in Chicago (Cain, 1978, p. 32).

the sewer system expanded beyond the streets covered in this initial report. The sewer ordinances describe the details of the regrading operation and list, block by block, the planned elevation of each street intersection relative to the level of the lake. The 1855 plan states, "It will be necessary to raise the grades of streets an average of eighteen inches per 2500 feet going West." To get a sense for the scale of this undertaking, it requires about 8300 cubic yards of fill to raise a 2,500 foot segment of a 20 foot wide street by 18 inches. At about 1.5 tons per cubic yard, this is almost 12,500 tons of fill per 2500 foot segment of road.

The historical record suggests that municipal authorities knew which streets had the worst drainage and were anxious to sewer them as soon as the network reached them. From the Chicago Tribune (June 25th, 1873, page 4):

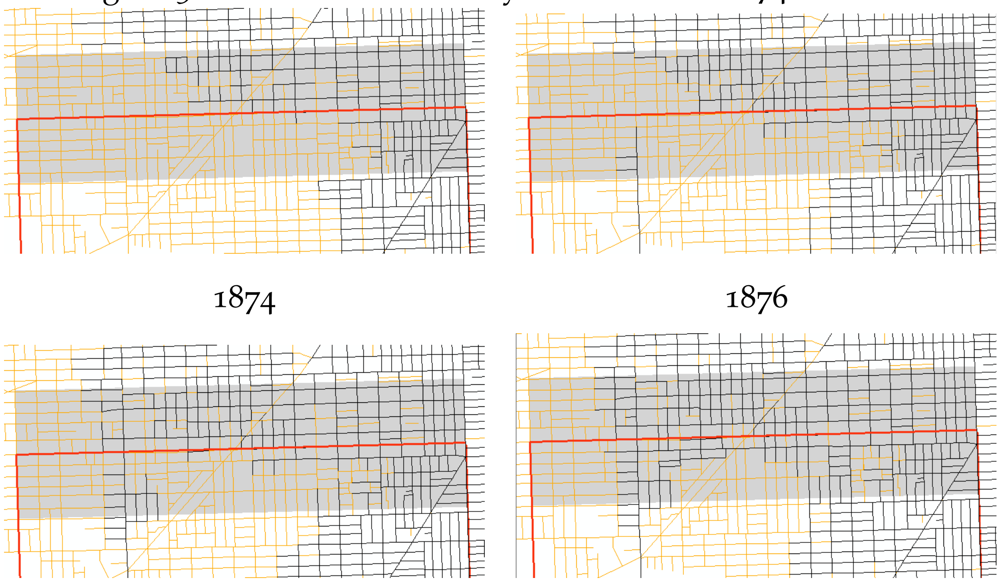
> "The Mayor points out the various localities where this sewerage is the most needed. It so happens that the unsewered portion of the city is that which, of all others, most needs it. ... These neighborhoods are densely populated by people who have not the means to adopt any sanitary measures."

Thus, there is no reason to believe that the assignment of sewers to neighborhoods and streets was independent of land value.

The 1855 ordinance describes a "triangle" southwest of the downtown that was at a slightly lower elevation than the rest of the city. Chesbrough wrote of this region, South of Tyler Street (now the Eisenhower Expressway) and West of Halsted Street: "The extreme south-west part of the city [is] too low [to sewer], "as the depth of filling required to raise streets over it would average two feet" (p. 16). Recalling that the plan calls for streets to be raised "an average of eighteen inches per 2500 feet going West", this means that the marginal 6" of fill required in this region was decisive. Chesbrough concludes by writing, "[a]s this part of the city may not be improved for several years, it is deemed sufficient for present purposes to state the general depth of filling that would be required" (p. 15).

Figure 2 illustrates the expansion of the Chicago sewer system that occurred between 1870 and 1890. In both panels, thick light grey lines indicate the extent of the sewer network prior to 1874, thick black lines indicate the expansion that occurred between 1874 and 1880, and, thin light gray lines indicate post-1880

Figure 3: Sewer extent in study area between 1874 and 1880



1874 1876

1878 1880

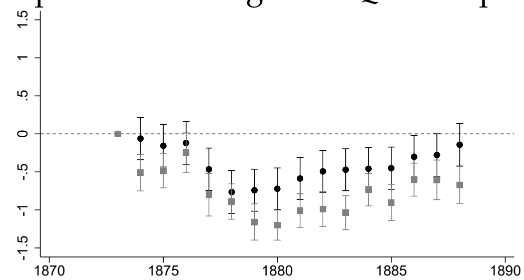
Note: *Tan indicates the 1930s street network and red indicates boundaries of the Southwest Triangle. Light gray indicates the area within 2000 feet of Tyler street running 14,000 feet West from Halsted Street. Black lines indicate the sewer network. There is more sewer coverage in the Northern half of our study area than the southern half during the 1874-80 study period.*

expansion. Red lines indicate the northern and eastern border of the Southwest Triangle, Tyler and Halsted streets.

While the 1855 plan refers to "a triangle", it specifies only northern and eastern borders. We draw a western boundary near the limit of the 1880 sewer network, 14,000 feet west of Halsted street, and a southern boundary at the Chicago River. We exclude parcels exactly on Tyler street, i.e., those matching to intersections within 75′ of Tyler Street, for two reasons. First, the 1855 plan is ambiguous about whether or not Tyler street lies inside or outside the Southwest Triangle. Second, our data does not allow us to determine whether parcels matching to Tyler Street lie north or south of the street. Thus, we cannot determine whether parcels matching to Tyler street are inside or outside the Southwest Triangle.

The black region in Figure 2b illustrates the entire region that received sewer and water access between 1874 and 1880. This is the region for which we observe

Figure 4: Land prices in Chicago and Quasi-experimental sample



Note: *Mean* $\ln(Price)$ *by year in Quasi-experimental sample (Gray) and all of Chicago (Black). Controls:* $\ln(miles\ to\ CBD)$, *improved, corner,* $\ln(Area)$.

construction costs and it is the economically relevant area for the purpose of policy evaluation. We often refer to a sample drawn from this area as a "Relevant sample." Our estimation of causal effects is primarily based on the region within 2000 feet of the northern boundary of the Southwest Triangle, Tyler Street. We refer often refer to a sample drawn from this area as a "Quasi-experimental sample". We sometimes consider the effect of sewers in the area within 2000′ of the northern *or* eastern boundary of the Southwest Triangle, Tyler and Halsted streets. We often refer to a sample drawn from this area as an "Extended-quasi-experimental sample." Figure 2b illustrates all three regions.

Figure 3 highlights the evolution of the sewer network in the Quasi-experimental sample. This figure makes it clear that, even 20 years after the adoption of the 1855 sewer ordinance, the construction of sewers south of Tyler street lags the northern side of the street by several years. It is this north-south difference in sewer assignment on which we base our estimates of the causal effects of piped water and sewer access.

5 Description

Our Quasi-experimental sample is a set of 351 transactions occurring between 1874-1880 within 2000′ of Tyler Street, west of Halsted. This is the sample where the case for quasi-random assignment of sewer and water access as a function of membership or exclusion from the Southwest Triangle is strongest.

Gray squares in figure 4 report mean log transaction price by year (after controlling for improved and corner status, log of parcel area, and log miles to the CBD), for all transactions falling in the Quasi-experimental region at any time between 1873 and 1880. Black points show the corresponding prices calculated for the entire city of Chicago. Whiskers indicate 95% confidence intervals. Unsurprisingly, annual means are more precise for the whole city than for the smaller sample drawn from the Quasi-experimental region.

This figure shows the same basic patterns described in Hoyt (2000). Prices fall between 1873 and 1880, before beginning a slow recovery. Figure 4 also shows that prices in the Quasi-experimental region follow those in the city as a whole. That is, the Quasi-experimental region is a small part of a large, liquid land market. This suggests that the assignment of sewers and piped water (or not) to parcels in the Southwest Triangle should not affect prices outside of the Southwest Triangle. On the basis of this observation, we ignore the general equilibrium price effects in our analysis of the Quasi-experimental sample.

Table 1 presents sample means for the Quasi-experimental sample. The first column describes transactions inside the Southwest Triangle, i.e., south of Tyler Street, the second, transactions outside the Triangle, i.e. north of Tyler Street. As the 1855 Ordinance prescribes, and as figure 3 shows, piped water and sewer incidence is lower inside the Southwest Triangle than outside. About half of transactions in the Southwest Triangle have water and sewer access during 1874-80 and access is almost universal outside. Consistent with a large effect of water and sewer access on value, unconditional prices are 0.72 log points or 105% higher outside of the Southwest Triangle than inside. The frequency of corner parcels is the same on both sides of the boundary. Improved parcels are more frequent outside the Southwest Triangle indicating the importance of this control. Parcels outside the Southwest Triangle are at most slightly larger than those inside. Parcels outside the Southwest triangle are on average one city block closer to the nearest horse car line, though both sides of Tyler street are well integrated with the horse car network. Major streets in Chicago occur at one mile intervals, or every eight blocks. Parcels on either side of Tyler street are on average one to two blocks from the nearest major street. The region inside the Southwest Triangle is marginally further from the CBD than the region outside,

Table 1: Summary Statistics 1874-1880

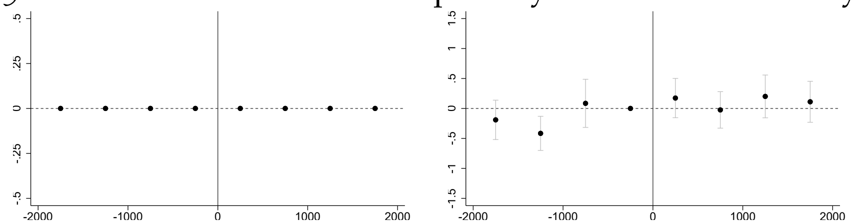
	(1) SW△ $=1$	(2) SW△ $=0$	(3) t-test	(4) Relevant
Share Sewered	0.47	0.92	11.04	0.70
	(0.50)	(0.27)		(0.46)
Log Price	7.70	8.42	8.44	7.41
	(0.86)	(0.76)		(0.91)
Log Distance to CBD	9.13	9.10	-0.89	9.49
	(0.38)	(0.38)		(0.25)
Log Area	8.12	8.26	1.88	8.17
	(0.62)	(0.69)		(0.54)
Share Improved	0.11	0.23	2.99	0.15
	(0.31)	(0.42)		(0.36)
Share Corner	0.11	0.13	0.42	0.14
	(0.32)	(0.33)		(0.34)
Distance to Horsecar	884	427	-9.53	1757
	(573)	(335)		(1351)
Distance to Major Street	564	475	-2.13	441
	(427)	(363)		(372)
Year	1877.18	1877.45	1.14	1877.60
	(2.19)	(2.17)		(2.26)
N	150	211		1358

Note: *Means and standard deviations of parcel characteristics. Column 1 reports on parcels in the Quasi-experimental sample (within 2000′ of Tyler St. west of Halsted) that are in the Southwest Triangle (south of Tyler Street). Column 2 reports on parcels that are not in the Southwest Triangle (north of Tyler Street). Column 3 reports the t-statistic for the difference between the first two columns. Column 4 presents parcel means and standard deviations for all parcels in the Relevant sample. In all columns, we restrict attention to parcels transacted during 1874-1880.*

and so transactions outside are nearer the CBD than those inside by construction.

The fourth column of table 1 highlights one of our main econometric challenges. It reports sample means from the Relevant sample. On average, these parcels are less expensive and further from the CBD than parcels in the Quasi-experimental sample. If we are to apply estimates of the effects of water and sewer access based on the quasi-random study region to this larger policy relevant area, we should consider the possibility that treatment effects may vary

Figure 5: Sewer and water share and price by distance to boundary, 1886-9



(a) (b)

Note: *(a) x-axis is distance to Tyler Street boundary, with $x < 0$ displacement South, "inside" and conversely. y-axis is share of transactions sewered between 1886-89, controlling for year indicators, $\ln(Area)$, and $\ln(mi.\ to\ CBD)$) by 500′ long bins. (b) Same as left panel but y-axis is $\ln(Price)$, controlling for the same set of covariates. Piped water and sewer access and prices are both the same at the border after sewer and water provision is completed in the Southwest Triangle.*

systematically between the two samples.

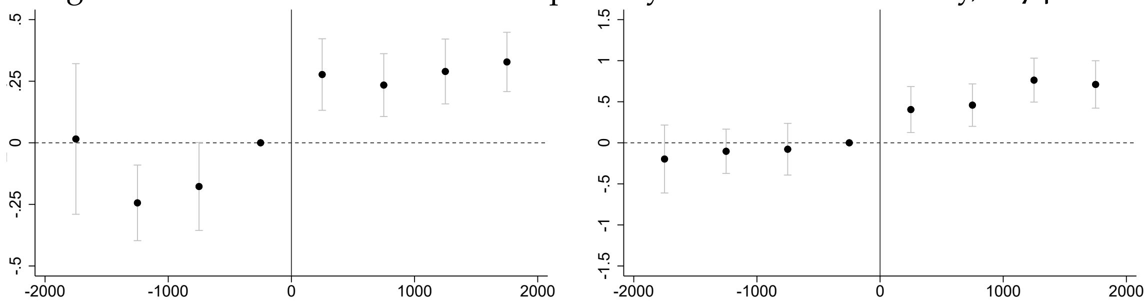
Ideally, to check that unobservable determinants of value are the same on both sides of Tyler Street, we would check land prices before piped water and sewer service was available on either side of the border. However, such data are not available.[14] Instead, we compare land prices on either side of Tyler street a short time after our study period when piped water and sewer access was universal.

Table A2 describes transactions occurring in the Quasi-experimental region during 1886-8, six to nine years after the end of main study window. This table replicates the first three columns of table 1 for the later time period. This table indicates that the same basic patterns present in the data during 1874-80 largely persist into 1886-9, with two notable exceptions. Piped water and sewer access is universal during the later period, and the difference between prices inside and outside the Southwest Triangle that shows so clearly in Table 1 is no longer present in the later period.

Figure 5a illustrates piped water and sewer access in our experimental study area during 1886-9 as a function distance to Tyler Street. The x-axis of this figure is distance from Tyler Street. Negative distances indicate displacement into the Southwest Triangle, and conversely for positive values. The y-axis indicates

[14]The Tribune began reporting transactions only in 1873, and 1860 census did not ask about home values or about the value of vacant land.

Figure 6: Sewer incidence and land price by distance to boundary, 1874-80



(a) (b)

Note: *Same as figure 5, but for transactions occurring between 1874 and 1880.*

piped water and sewer share relative to the share in the bin just inside the Southwest Triangle. Sewerage is universal across the boundary by 1886.

Figure 5b is similar, but reports on transaction prices. The y-axis indicates log price relative to the bin just inside the Southwest Triangle. Mean log price in each bin is calculated controlling for year indicators, ln(area), and ln(mi. to CBD). Whiskers indicate 95% confidence intervals. Table 1 indicates a 105% difference in prices across this boundary during 1874-80. Figure 5 indicates that this difference is completely erased in less than 9 years, once sewer incidence across the border equalizes. This confirms what we see in the unconditional means presented in table A1.

Table 2 shows that parcels in the Southwest Triangle were less valuable during our study period. There is evidence that such initial disadvantages often "lock-in" and lead to long run differences between places (e.g., Bleakley and Lin (2012) or Ambrus et al. (2020)). Poor places stay poor and rich places stay rich. Given this, our finding that price differences largely disappear with the elimination of the difference in sewer access is surprising. The available evidence suggests that path dependence works against the price equalization that we see in figure 5. We suspect this reflects the dynamic nature of the Chicago real estate, the pervasiveness of cheap, short-lived structures, and our focus on vacant lots.

The descriptive evidence provided so far is consistent with the following narrative. Parcels in the Southwest Triangle are less likely to have access to piped water and sewers because of a nearly imperceptible change in elevation that affected costs of constructing gravity fed sewers. There is no a priori reason to suspect that parcels on opposite sides of Tyler street are systematically different, except that parcels inside the Southwest Triangle are slightly more remote from the CBD. This suggests that conditional on controls, a comparison of changes in prices and sewer access across Tyler street should yield an unconfounded estimate of the effect of water and sewer access on prices.

Figure 6 performs this comparison. Panel (a) shows changes in sewer incidence across the Tyler street border of the Southwest Triangle and panel (b) shows the corresponding changes in log price. The construction of this figure is the same as figure 5, except that it is based on data from our main study period, 1874-1880. Consistent with the unconditional means presented in table 1, we see that piped water and sewer incidence and land prices are lower in the Southwest Triangle. These figures illustrate the variation on which our estimates are based. The left panel is a first-stage regression, the right panel is a reduced form. The ratio of the two cross-boundary gaps, averaged over the four interior and exterior bins, yields (approximately) a local average treatment effect for the whole Quasi-experimental sample.

We note that Figure 6 suggests the possibility of implementing a fuzzy-RD design. Given our already small sample, this research design would rely heavily on a tiny set of observations. To avoid this, we abstract from the spatial structure of the data and base our estimates on an instrumental variable design using whole Quasi-experimental sample. Note that our Quasi-experimental study region is narrow enough to walk across in less than 20 minutes and lies in an a priori homogeneous landscape. We can reasonably hope to have restricted attention to parcels with on average identical unobserved determinants of land price. Nevertheless, to the extent our sample allows, we investigate the possibility of confounding spatial trends in unobservables in our regression analysis.

6 Estimation

Let Y_i be the log of parcel i's transaction price observed in the data. Let X_i denote a vector of observable parcel attributes drawn from, *transaction year indicators*, $\ln($*miles to CBD*$)$, $\ln($*Parcel Area*$)$, *Corner* and *Improved indicators*, *distance to horsecar line* and *distance to major street*. Let D_i be a treatment

Table 2: OLS, First Stage, Reduced form, and TSLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: OLS								
Sewer=1	.413***	.39***	.4***	.328***	-.018	.194***	.276***	.239***
	(.086)	(.082)	(.084)	(.139)	(.101)	(.08)	(.081)	(.078)
R^2	0.386	0.502	0.504	0.567	0.598	0.505	0.376	0.439
B: Red. Form								
$\text{SW}\triangle = 0$	.657***	.568***	.714***	.439***	.292*	.3***	.336***	.332***
	(.072)	(.069)	(.073)	(.093)	(.151)	(.068)	(.063)	(.059)
R^2	0.486	0.568	0.591	0.606	0.602	0.527	0.397	0.462
C. 1st Stage								
$\text{SW}\triangle = 0$	.432***	.443***	.451***	.323***	.194**	.443***	.259***	.259***
	(.039)	(.04)	(.043)	(.057)	(.097)	(.04)	(.031)	(.031)
R^2	0.451	0.455	0.455	0.456	0.474	0.455	0.333	0.335
F-stat	119.729	125.018	110.664	32.311	3.992	125.018	71.711	71.283
D. IV								
Sewer=1	1.522***	1.283***	1.582***	1.36***	1.501	.678***	1.296***	1.283***
	(.22)	(.191)	(.209)	(.352)	(1.067)	(.164)	(.277)	(.266)
Year FE & ln(Area)	Y	Y	Y	Y	Y	Y	Y	Y
ln(mi. CBD)	Y	Y	Y	Y	Y	Y	Y	Y
Imp. & Corner		Y	Y	Y	Y	Y		Y
H.car & Maj. St.			Y					
Sample	Q.E.	Q.E.	Q.E.	Q.E. 1k′	Q.E.	Q.E.	E.Q.E.	E.Q.E.
Observations	351	351	351	172	351	351	533	533

Note: *All results based on transactions during 1874-80. Columns 1-3, 5 rely on the Quasi-experimental sample, 7 and 8 on the Extended-quasi-experimental sample, and column 4 restricts attention to the subset of the Quasi-experimental sample within 1000′ of Tyler Street. (A) Reports OLS regressions of log transaction price on the treatment indicator. (B) Reports reduced form regressions log transaction price on the instrument. (C) Reports first stage regressions of treatment on instrument. (D) Reports TSLS estimate of the effect of water and sewer access on log parcel price. The bottom panel of the table indicates controls for all regressions in the column above. Robust standard errors in parentheses.* $*, **, ***$ *indicates* 10%, 5%, 1% *significance.*

indicator, with $D_i = 1$ if and only if parcel i has piped water and sewer access. Let Z_i be a binary variable indicating $Z_i = 1$ if and only if the parcel is *not* in the Southwest Triangle. We view Z_i as an instrumental variable and assume that it shifts the cost of access to piped water and sewage without directly affecting the land price, fixing the controlling covariates. By defining Z so that $Z_i = 1$ outside of the Southwest Triangle, we assure a conventional positive relationship between instrument and treatment.

We adopt the convention of indicating potential outcomes with a subscript, so that Y_{1i} is the price of parcel i in a state of the world where it is treated, and Y_{0i} is the untreated price. Let U_1, U_0, U_D denote three error terms to be defined later. Finally let P denote our Quasi-experimental sample and, abusing notation slightly, the joint distribution of $(Y_1, Y_0, X, Z, D, U_1, U_0, U_D)$ drawn from this sample.

We are also interested in the corresponding quantities drawn from the Relevant sample, all transactions in the area receiving water and sewer access during 1874-80. We indicate these quantities with an asterisk. For example, Y_i^* is a transaction price drawn from this sample, and P^* denotes the distribution of $(Y_1^*, Y_0^*, X^*, Z^*, D^*, U_1^*, U_0^*, U_D^*)$.

We would like to estimate the average treatment effect on the economically relevant sample, that is, $ATE^* \equiv E(Y_1^* - Y_0^*)$. This treatment effect permits an immediate evaluation of a realized policy and matches neatly to available data on costs. Estimating ATE^* requires that we address the conventional problem of estimating ATEs rather than LATEs. In addition, we must find a way to extrapolate our estimated treatment effect from the Quasi-experimental to the Relevant sample.

We first estimate local average treatment effects of piped water and sewer access with TSLS.[15] We next implement the local IV framework proposed by Carneiro et al. (2010). This framework allows the explicit calculation of an average treatment effect and tests for heterogeneity of treatment effects with respect to observable and unobservable characteristics. The LIV/MTE framework also provides a foundation for a novel, principled approach to the extrapolation of treatment effects. We develop and implement this method in the final stage of our analysis.

---

[15] Under instrument exclusion, exogeneity, and monotonicity (no-defier condition) conditional on X, we can interpret the linear TSLS estimand as a weighted average of the local average treatment effects aggregating complier's conditional average causal effects given X. See Angrist and Imbens (1995), Kolesár (2013), and Słoczyński (2021).

Local Average Treatment Effects Table 2 presents four sets of estimates. For reference, Panel A presents OLS regressions of the form

$$Y_i = A_0 + A_1 D_i + A_2 X_i + \varepsilon_i$$

These regressions show a significant positive association between piped water and sewer access, and transaction prices. In the first column, we control for year indicators and log miles to the CBD. In the second column, we add indicators for corner lot and improved status. In the third column, we add controls for distance to horsecar and distance a major street. In each case, transaction prices are about 0.4 log points higher for parcels with water and sewer access. We postpone a discussion of the remaining columns.

Panel B presents the corresponding reduced form regressions of transaction price on the instrument,

$$Y_i = A_0 + A_1 Z_i + A_2 X_i + \varepsilon_i.$$

We see in column 1 that being in the Southwest triangle decreases transaction prices by about 0.6 log points. This effect is estimated precisely and varies only slightly as we add control variables in columns 2 and 3. Column 3 uses the same controls as we used in figure 6b, and so the estimated effect approximately corresponds to the average price difference between inside and outside parcels that we see in this figure.

Panel C presents first stage regressions,

$$D_i = A_0 + B_1 Z_i + B_2 X_i + \mu_i.$$

Conditional on control variables, being in the Southwest triangle reduces the probability of piped water and sewer access by about 40%. Again, this effect corresponds approximately to the mean difference between inside and outside parcel sewer access in figure 6a. First stage F statistics are above critical values for conventional weak instrument tests (e.g., Stock and Yogo (2002)).

Panel D presents TSLS estimates of the effect of piped water and sewer access on transaction prices. IV estimates range between about 1.3 and 1.5 log points, estimated precisely. This treatment effect is enormous. A 1.3 log point increase in parcel price is a factor of 3.7.

Comparing IV to OLS results suggests that the equilibrium process assigns piped water and sewer service to parcels that are less valuable after conditioning on observable controls. This is consistent with anecdotal evidence presented earlier.

Figure 6 illustrates an increase in piped water and sewer access and transaction prices that occurs when we cross Tyler street to leave the Southwest triangle. These changes appear to occur sharply in the figure. Nevertheless, we are concerned that this increase may reflect a confounding trend correlated with treatment and transaction prices. To address this concern, in column 4 of table 2 we restrict the sample to a narrower window which includes parcels within 1000 ft. of Tyler street. The magnitudes of the reduced form and first stage are reduced, but the IV estimate is unchanged. In column 5, we include controls for distance to Tyler street in our regression of column 2, where we allow the slope of this trend to change at Tyler street. Once again these controls reduce the magnitude of first stage and reduced form effects by about half, but leave the IV point estimate unchanged, although the standard error increases to just above the 10% significance threshold.

To refine this test, we consider the impact of a hypothetical confounding trend in land prices across Tyler Street, the trend that we observe across the Tyler Street boundary during 1886-9, after piped water and sewer access is universal on both sides of the border. Implicitly, we suppose that the entire (small) trend we observe in 1886-9 is due to confounding unobservables rather than path dependence on an otherwise homogeneous landscape. Appendix Table A2 is similar to panel C of table 2, and reports this trend in column 3. We then subtract this trend from transaction prices, the dependent variable, in our 1874-80 sample in column 6 of table 2. Unsurprisingly, this leads to a smaller estimated treatment effect, but one that is estimated precisely and is still nearly 0.7 log points.

Summing up, the validity of our research design rests on four pieces of evidence. First, the sensitivity of sewer construction costs to otherwise imperceptible changes in grade supports the a priori argument that the instrument affects outcomes only through its effect on the likelihood of treatment. Second, the near disappearance of price differences across Tyler street after water and sewer access equalizes across this boundary suggests that, except for piped water and sewer access, the distribution of parcel prices is the same on both sides of the boundary. Third, the difference between OLS and IV estimates is consistent with what one would predict from anecdotal evidence about the assignment process, the equilibrium assignment process favors cheaper parcels. Finally, the robustness of results to various permutations of control variables, and to correction for a confounding spatial trend, suggests that omitted variables correlated with instrument and outcome are not driving our estimates.

The estimates in table 2 are LATEs for our Quasi-experimental sample. We now turn our attention to whether this estimate differs from the ATE in this sample and whether we can extrapolate to the Relevant sample.

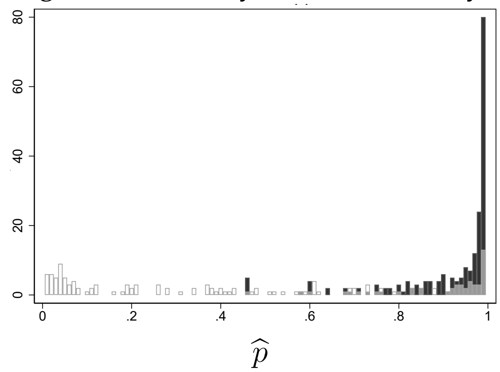
To begin, columns 7 and 8 of table 2 re-estimate the specifications of columns 1 and 2 on the Extended-quasi-experimental sample. That is, the sample of transactions drawn from within 2000′ of the Northern or Eastern boundary of the Southwest Triangle.

A Local Average Treatment Effect coincides with the Average Treatment Effect if treatment effects are the same for all units. By expanding our sample, we change the set of compliers, and hence the sample of units over which the LATE is estimated. We observe that coefficients in columns 7 and 8 are statistically indistinguishable from their counterparts estimated on the smaller Quasi-experimental sample. This suggests either that treatment effects are not very heterogeneous, or that the distributions of treatment effects in the two samples of compliers are similar.

We would ultimately like to extrapolate our estimate to the Relevant sample. The Extended-quasi-random sample has a larger support for X and presumably, a larger support for unobservable determinants of treatment and treatment effects. In this sense, less extrapolation is required from the Extended-quasi-experimental sample to the Relevant sample, than from the smaller Quasi-experimental sample.

We note that the validity for our research design is easier to defend on the smaller Quasi-experimental sample than the Extended-quasi-experimental. Figures A2 in the appendix reproduces the border plots of figure 6 for the larger sample. Neither prices nor sewer access change as sharply at the boundary of

Figure 7: Density of treatment by $\widehat{p}$



Note: *Density of treated and untreated parcels by propensity score. The propensity score distribution is skewed toward one, but conditional on a mass of propensity scores, treated and untreated parcels both occur. Based on column 2 of table 3.*

the Southwest Triangle in the larger sample.[16] This increases our concern about the possibility of a confounding trend across the border and motivates our preference for estimates bed on the smaller Quasi-experimental sample.

Marginal and Average Treatment Effects The LIV/MTE framework developed in Heckman and Vytlacil (2005) and Carneiro et al. (2010) offers a method to estimate treatment effect heterogeneity and a framework to evaluate the difference between LATEs and ATEs. Moreover, as we will show, this framework provides a foundation for extrapolating our estimates from the Quasi-experimental to the Relevant sample under a weaker assumption than "no heterogeneous treatment effects".

The LIV/MTE framework recasts the potential outcome framework as a Roy model. Each unit selects into treated or untreated status on the basis of a third selection equation. Formally,

$$\begin{aligned} Y_1 &= X'\delta_1 + U_1 \\ Y_0 &= X'\delta_0 + U_0 \\ D &= \mathbb{1}[v(X,Z) - U_D \geq 0], \end{aligned} \tag{1}$$

[16]This is because, 20 years after the 1855 ordinance, both sides of the Eastern boundary of the Southwest Triangle have sewer service, see figure 2.

Table 3: LIV Regression Test Statistics

	(2)	(4)	(6)	(8)	(10)
χ^2	220	221	237	243	245
H0: $\delta_1 - \delta_0, \gamma_1, \gamma_2, \gamma_3 = 0$	0	0	0	.005	.002
H0: $\delta_1 - \delta_0 = 0$	.108	.07	.074	.298	.205
H0: $\gamma_2, \gamma_3 = 0$	.002	0	.001	.656	.498
H0: $\delta_1 - \delta_0,\ \gamma_2, \gamma_3 = 0$	.001	.001	.001	.15	.076
ATE	1.04***	.72**	.8***	1.31*	1.31**
	(.4)	(.35)	(.32)	(.69)	(.65)
ATE^*	1.04***	.75***	.89***	1.05**	.87**
	(.31)	(.27)	(.36)	(.46)	(.41)
Carr & Kitagawa	0.156	0.154	0.434	0.792	0.916
Year FE & ln(Area)	Y	Y	Y	Y	Y
ln(mi. CBD)	Y	Y	Y	Y	Y
Improved and Corner		Y	Y		Y
Horsecar and Major Street			Y		
Sample	Q.E.	Q.E.	Q.E.	E.Q.E.	E.Q.E.
Observations	351	351	351	533	533

Note: *Various test statistics based on estimates of the LIV model of equation* (4) *and estimates of ATE and ATE^* based on equations* (6) *and* (9). *Complete report of coefficient estimates is in table A3. All estimations based on transactions during 1874-80. Columns 2,4, and 6 rely on the Quasi-experimental sample, 8 and 10 on the Extended-quasi-experimental sample. Omitted odd numbered columns report first stage Logit coefficients in appendix table A3. Bottom panel indicates controls for the regression above. Bootstrapped standard errors in parentheses.* $*, **, ***$ *indicates* 10%, 5%, 1% *significance.*

where Y_1 denotes treated potential outcome and Y_0 not treated. We assume that the controls enter the potential outcome equations linearly with coefficients δ_1 and δ_0, and make the "practical independence" assumption as in Carneiro et al. (2010),

$$(X,Z) \perp (U_1,U_0,U_D) \tag{2}$$

U_D measures unobserved "resistance to treatment," in our context, unobservable determinants of the cost of piped water and sewer access for each parcel. We assume that U_D is continuously distributed.

Let $p = F(X,Z) \equiv P(D = 1|X, Z)$ be the propensity score in the Quasi-experimental sample. A normalization of the unobservable U_D common in the literature transforms it through its cdf $F_{U_D}(\cdot) \equiv P_{U_D}(U_D \leq \cdot)$,

$$D = 1\{\underbrace{F_{U_D}(U_D)}_{\widetilde{U}_D \sim Unif(0,1)} \leq \underbrace{F_{U_D}(v(X,Z))}_{F(X,Z)}\} = 1\{\widetilde{U}_D \leq F(X,Z)\}. \tag{3}$$

The cdf-transformed unobserved heterogeneity $\widetilde{U}_D = F_{U_D}(U_D) \sim Unif(0,1)$ ranks the individuals in the population P according to the unobservable cost of access to piped water and sewage, i.e., $\widetilde{U}_D$ is smaller as unobserved costs of piped water and sewer access are smaller.

Define marginal treatment effects, MTE, for each conditioning covariate value X and $\widetilde{U}_D \in [0,1]$ as

$$MTE(X,\widetilde{U}_D) \equiv E(Y_1 - Y_0 | X, \widetilde{U}_D)$$

That is, MTE describes how the causal effects vary with observable characteristics, X, and with the unobservable $\widetilde{U}_D$.

To estimate MTEs, we run the local IV regression

$$\begin{aligned} p &\equiv \Pr(D = 1|X,Z) = F(X,Z), \\ Y &= X'\delta_0 + \widehat{p}X'(\delta_1 - \delta_0) + K(\widehat{p}) + \varepsilon. \end{aligned} \tag{4}$$

The first equation is a first stage binary regression of treatment status on the instrument and controls. In our case, we specify a Logit regression with linear index in (X,Z) for the first stage. The second equation is a structural equation with a control function in $\widehat{p}$, where the additive functional form follows from our specification (1) and the practical exogeneity restriction (2). In light of our small sample size, we restrict attention to the case with parametric cubic specification for $K(\cdot)$,

$$K(\widehat{p}) = \gamma_1\widehat{p} + \gamma_2\widehat{p}^2 + \gamma_3\widehat{p}^3.$$

Heckman and Vytlacil (2005) show that the derivative of the local IV regression with respect to the propensity score identifies the marginal treatment effect, and that taking the expectation of MTE over $(X,\widetilde{U}_D)$ identifies the average treatment effect. That is,

$$MTE(X,\widetilde{U}_D) = X'(\delta_1 - \delta_0) + \gamma_1 + 2\gamma_2\widetilde{U}_D + 3\gamma_3\widetilde{U}_D^2 \tag{5}$$

$$ATE = E(X)'(\delta_1 - \delta_0) + \gamma_1 + \gamma_2 + \gamma_3. \tag{6}$$

Equation (5) allows explicit tests for heterogeneity of treatment effects. If $\delta_1 - \delta_0 \neq 0$ then the marginal treatment effects vary with unit observables. If γ_3 or $\gamma_2 \neq 0$ then the marginal treatment effects vary with unobserved resistance to treatment. Rejecting both sorts of treatment heterogeneity means that LATE, any weighted average of MTEs, and ATE are all equal. In this case, we can interpret the conventional linear TSLS estimator for the coefficient of endogenous D as a consistent estimator for ATE.

We estimate equation (4) for specifications corresponding to those in columns 1,2,3, 7, and 8 of table 2. Because equation (4) is quite long, we relegate a complete report of parameter estimates and bootstrapped standard errors to appendix table A3. Table 3 reports estimates of ATE derived from these regressions, along with several hypothesis tests.

The first row of table 3 reports a χ^2 test of significance of our instrument in the first stage Logit regression. As in our TSLS estimations, we easily reject the hypothesis that our instrument does not affect treatment.

The second row of table 3 reports p-values of the tests of the hypothesis that all terms involving the propensity for treatment are zero. That is, that treatment effects are different from zero. This is rejected at in all specifications. Piped water and sewer almost surely affect land prices in our Quasi-experimental and Extended-quasi-experimental samples.

The third row tests the hypothesis that there is heterogeneity in treatment effects by observables. The fourth row tests whether there is heterogeneity in treatment effects by unobservables. The fifth row tests the joint hypothesis of either sort of treatment effect heterogeneity.

The results of these tests vary with sample. In our Quasi-experimental sample, columns 1,2 and 3, we see clear evidence of treatment heterogeneity on unobservables, somewhat weaker evidence for treatment effects on observables, and clearly reject the hypothesis of no treatment heterogeneity at all. Columns 4 and 5, we consider the larger Extended-quasi-experimental sample. Here, we reject the hypothesis of any treatment effect heterogeneity at the 7 or 15% level, depending on specification, but we cannot reject the hypothesis of treatment heterogeneity by observables or by unobservables alone. Inspection of appendix table A3 suggests that treatment effects likely vary by year in all specifications, though there is no clear pattern in the coefficients across years.

The sixth row of table 3 calculates the average treatment effect given in equation (6) along with bootstrapped standard errors. Comparing to the LATEs estimated in table 2 we see that ATEs are marginally smaller than LATEs in the Quasi-experimental sample, [0.72,1.04] versus [1.28,1.52] and both are estimated precisely. In the larger Extended-quasi-experimental sample, ATE and LATE are statistically indistinguishable. Even the smallest of these ATE estimates is still very large; $e^{0.72} \approx 2$, so these estimates indicate that piped water and sewer access at least doubles land values.

The differences between between LATE and ATE estimates are consistent with other results in rows 3 to 5 of table 3. Heterogeneous treatment effects are necessary if ATE and LATE are to diverge.

Figure 7 presents a standard diagnostic for the LIV regression presented in column 2 of tables A3 and 3. Figure 7 is a histogram showing the frequency of treated and untreated transactions as a function of $\widehat{p}$. As we expect from table 1, the distribution of parcels is heavily skewed toward "treated"; 0.47 of the Quasi-experimental sample South of Tyler street has piped water or sewer access, and this share is even higher to the North. With this said, conditional on this skewed distribution, the histograms for treated and untreated parcels are similar, although there is more mass left of 0.6 for untreated parcels. The corresponding histograms for other specifications reported in A3 (not reported) are qualitatively similar.

Figure A1 is a second standard diagnostic figure. Figure A1 plots marginal treatment effects as a function of resistance to treatment, $\widetilde{U}_D$, and lets us to visualize the importance of treatment heterogeneity on unobservables. In light of the hypothesis test presented in column 2, row 4 of table 3, that this figure suggests marginal treatment effects change with unobservables is unsurprising. Because most of the probability mass of treated and untreated parcels has $\widehat{p}$ of at least 0.6, the region of panel (a) to the left of 0.6 should be understood as extrapolation from the larger values.[17]

The final row of Table 3 presents the p-value for the instrument validity test

[17] Identification of $MTE(X,\widetilde{U}_D)$ without a parametric control function $K(\cdot)$ is possible for values of $\widetilde{U}_D$ supported by the distribution of propensity scores. Figure 7 indicates that observations with propensity scores near 1 largely contribute to the estimation of cubic $K(\cdot)$. MTE estimates for the range of $\widetilde{U}_D$'s without much probability mass extrapolate using the functional form of $K(\cdot)$.

proposed in Carr and Kitagawa (2021). This test evaluates the joint null hypothesis of practical exogeneity (2), instrument monotonicity, and the functional form specification for the potential outcome equations (1). p-values consistently above 15% indicate that the data do not reject the assumptions on which our MTE and ATE estimates rely.[18]

Extrapolation to Relevant sample While our LIV estimation does not offer conclusive evidence for the importance of heterogeneous treatment effects, neither does it offer much reassurance that they are not important. Given this, we consider the problem of extrapolating our ATE estimates under both assumptions, that treatment effects are heterogeneous, and that they are not.

In the absence of treatment heterogeneity, extending our treatment effect estimates from the Quasi-experimental to the Relevant sample is straight forward. Estimates in table 2 can be interpreted as Average Treatment Effects, and provided treatment effects remain constant on the larger support of the Relevant sample, these estimates apply immediately to units in the larger sample.

However, table 3 suggests that concern about treatment heterogeneity is warranted. Given this, we develop a method for extrapolating treatment effects in the presence of treatment heterogeneity.[19]

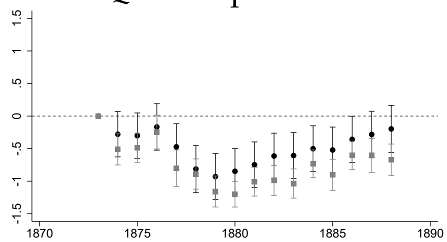
This extrapolation requires that equations (1) and (2) continue to hold on the Quasi-experimental sample. In addition, we assume

$$\begin{aligned} Y_1^* &= X^{*\prime}\delta_1 + U_1^* \\ Y_0^* &= X^{*\prime}\delta_0 + U_0^* \\ D^* &= \mathbb{1}[v(X^*,Z^*) - U_D^* \geq 0]. \end{aligned} \tag{7}$$

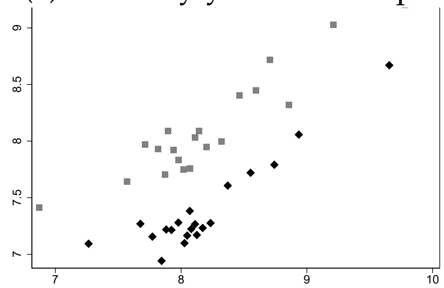
[18]We also apply the IV validity test of Mourifié and Wan (2017). This test evaluates the strict exogeneity of instrument (i.e., Z is also independent of X) rather than conditional exogeneity. We do not reject the null of instrument validity at 5% significance level for the Quasi-experimental sample. However, we do reject the null at the same level for the Extended-quasi-experimental sample. Taken together with the results of the Carr & Kitagawa test reported in table 3, this means that we reject the strict exogeneity of of our instrument, but fail to reject conditional exogeneity. It follows that controlling for conditioning covariates is necessary for the estimation of causal effects in our model, particularly in the Extended-quasi-experimental sample.

[19]We note that simply conducting our TSLS regressions on the Relevant sample offers a particularly simple solution to this problem. However, and unsurprisingly, our instrument is not relevant on this larger more heterogeneous sample.

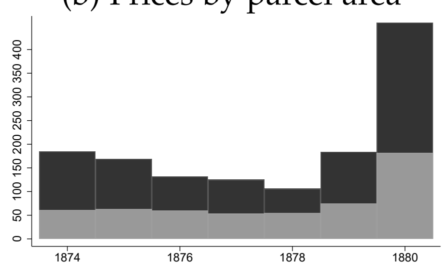
Figure 8: Comparison of Quasi-experimental and Relevant samples.



(a) Prices by year and sample



(b) Prices by parcel area



(c) Transaction frequency by year and sample

Note: *(a) Mean log transaction price by year in main Quasi-experimental (gray) sample and the Relevant (black) sample. Conditional on:* $\ln(Area)$, $\ln(miles\ to\ CBD)$, *improved, corner. Means and variances of* Y *in the two samples are similar conditional on year. (b) Mean log transaction price by parcel area. (c) Transactions by year and sample. Relevant sample is larger, but distribution of transactions across years is similar in Quasi-experimental and Relevant samples. The spike in 1880 reflects a change in sampling effort, not in transaction volume.*

and that

$$P^*_{U_1,U_0,U_D} = P_{U_1,U_0,U_D}. \tag{8}$$

In words, we assume that the same econometric model governs the effects of treatment in the Relevant sample as in the Quasi-experimental sample and that the joint marginal distribution of residuals is the same across the two samples.

These conditions would be satisfied, for example, if the mechanism and magnitude of the causal effect are the same in both samples, and unobserved resistance to receive the treatments are identically distributed between them.

In our data, the cost shock Z is observed on the Quasi-random sample and latent on the Relevant sample. In addition, we can credibly assume that Z is randomized in the Quasi-experimental sample, but it is probably not randomized in the Relevant sample. Our approach to extrapolation does not require that the joint distributions of observable characteristics and the instrument are identical for the Quasi-experimental and Relevant samples.

Assuming equations (1), (2), (7) and (8), we can extrapolate MTE estimates from the Quasi-experimental to the Relevant sample and use them to calculate an average treatment effect on the Relevant sample as follows,

$$ATE^{*} = E(X^{*})'(\delta_1 - \delta_0) + \gamma_1 + \gamma_2 + \gamma_3. \tag{9}$$

Appendix B provides a proof.

In words, the average treatment effect for the Relevant sample is the same as for the Quasi-experimental sample, except that we must adjust for differences in the distributions of observable controls between the two samples. If the structural equations that govern treatment effects and assignment are the same across samples, and if the distribution of unobservables is the same, then we can extrapolate MTE estimates. This result holds even if the instrument is latent or dependent on the unobservables in the Relevant sample, or if the support of observable controls differs across samples. This result seems intuitive and, to our knowledge, no similar result exists in the literature.

The seventh row of table 3 presents our estimates of ATE^{*} for each of our specifications, along with bootstrapped standard errors. All are estimated precisely enough that they may easily be distinguished from zero. These estimates of ATE^{*} range from 0.75 to 1.04, across all samples and specifications. There is even less variation in ATE^{*} across samples and specifications than we saw for ATE, but in no case is the ATE^{*} statistically distinguishable from the corresponding ATE.

Conditional on the validity of our estimates of ATE, the validity of our estimates of ATE^{*} hinges on equations (7) and (8). Ideally, we would be able to test whether these equations hold in our data. We have not been able to define such a test, and our investigations suggests that a test may not exist except in the uninteresting case where there is no treatment heterogeneity. In the absence of a formal test, we provide informal evidence that the Quasi-random and Relevant samples are both governed by the same basic economic logic.

Figure 8 compares the Quasi-random and Relevant samples. Panel (a) of figure 8 reports mean log prices by year in the Relevant and Quasi-experimental samples, conditional on: ln(Area), ln(miles to CBD), improved and corner. Panel (b) reports mean log prices by parcel area in both samples, conditional on year indicators, ln(miles to CBD), improved and corner. Finally, panel (c) gives counts of transactions by year and sample. None of these figures obviously contradicts the hypothesis that the same basic economic forces are at work determining prices in the Quasi-experimental and Relevant samples.

7 The value of piped water and sewer access

We can now calculate the effect of piped water and sewer access on land values in the relevant area. We proceed in four steps. First, we calculate the area affected by the piped water and sewer expansion of 1874-80. Second, we calculate average price per square foot of an untreated parcel in this region. Third, we calculate the increase in price per square foot that results from piped water and sewer access. Fourth, multiplying this increase by the area affected gives the total increase in land value resulting from piped water and sewer expansion during 1874-80.

An average residential lot in any of our samples is about 125 feet deep. If we assume that every sewer serves lots on both sides of one street, then each linear foot of sewer serves 250 ft^2 of land area. Our shapefiles of the sewer network then allow us to calculate that about 138m ft^2 of land received piped water and sewer access during 1874-80.

During 1874-80, 384 untreated parcels transacted in the Relevant sample area. The total area of these parcels was about 1.8m ft^2, and their aggregate value was about 0.81m 1880 dollars. Dividing, the average price per ft^2 of land in the Relevant area was about 0.45 dollars.

We must now decide whether to use estimates of ATE based on an assumption that there are no heterogeneous treatment effects, or based on our estimates of treatment heterogeneity. Our LIV estimates do not strongly support either hypothesis, and so we proceed using the smallest estimates, 0.75, from column 4 of table 3.

Applying this treatment effect to the price per square foot of untreated land in the Relevant sample area, we calculate that piped water and sewer access increases the value of land in this area by $0.45 \times (e^{ATT^*} - 1) = 0.50\$/\text{ft}^2$. That is, using our most conservative estimate, piped water and sewer access increases the value of land by about 110%. Multiplying this increase by the area affected, the total value of the piped water and sewer expansion was slightly above 69m 1880 dollars.

This estimate requires several comments. First, this calculation reflects our smallest estimate of the average treatment effect. If, as we might do on the basis of column 8 of table 3, we reject the hypothesis of heterogeneous treatment effects, then the LATEs we estimate in Table 2 can be defended as ATEs and extended to the relevant sample. In this case, using the column 7 in table 2, the analog of column 8 of table 3, we have $ATE = 1.3$. Using this estimate to evaluate the effects of piped water and sewer access gives about 164m 1880 dollars.

Second, an average parcel in the Quasi-experimental sample receives piped water and sewer service about four years after it is sold. Thus, our estimates reflect the flow value of four years of piped water and sewer access, not the full asset value. Hoyt (2000) reports that interest rates were about 8% during our study period. If we denote our estimated aggregate value by V^* and assume that this flow value arrives every four years for perpetuity, then the full asset value of piped water and sewer access is $\sum_{t=0}^{\infty} \left[\left(\frac{1}{1.08} \right)^4 \right]^t V^* \approx 3.8 V^*$. Thus, we should multiply by about 3.8 to scale up our about four year flow value of piped water and sewer access to an asset value. Applying this adjustment to our 69m dollar estimate of the four year flow value, we have an asset value of about 262m 1880 dollars.

Third, as we noted earlier, piped water and sewer expansions were largely paid for with bonds that were serviced by property taxes (Chicago Board of Public Works, 1873). If there is any sort of capitalization of piped water and sewer construction costs into transaction prices, then this would bias our estimates of treatment effects downward.

Finally, while it seems reasonable to ignore general equilibrium effects in our estimates of treatment effects based on the relatively small Quasi-experimental sample, this assumption seems difficult to defend when we extend our estimates to the Relevant area, the entire area that receive piped water and sewer access between 1874-80. Given this, our estimates of the value of piped water and sewer expansion should be understood as a basis for evaluating a marginal counterfactual change in the extent of the treated area, or as being net of general equilibrium effects.

With our estimates of the value of piped water and sewer access in place, we turn to estimates of its cost. We digitized expenditures on water and sewer for the 1874-80 period (Chicago Board of Public Works, 1873). Construction costs during this time were: Sewer Construction, $1.5m; Maintenance, $0.4m; Waterworks construction, $2.4m. Summing, we have a total of $4.3m.

Our estimate of the four year flow value of piped water and sewer access was about $69m, about 16 times as large as construction costs. Our estimate of the total asset value piped water and sewer access is $262m, about 60 times as large as costs. Both of these calculations are based on our smallest estimate of average treatment effects. If we use one of our larger (but still defensible) estimates of ATE, these ratios approximately triple.

8 Conclusion

While tremendous progress has been made in providing safe water and modern sanitation for the relatively poor recent immigrants to developing world cities, access is far from universal. A large body of evidence suggests that in the absence of modern public health and sanitation infrastructure, people living in dense cities make each other sick. Thus, increasing access to high quality drinking water and modern sanitation would seem to call for a crisis response. However, relatively poor developing world cities face a portfolio of crises. Not only do their residents need more and better water and sewer infrastructure, they also need more and better roads, public transit, electricity supply and distribution, education, and housing. Trade-offs will inevitably need to be evaluated and made.

With this in mind, piped water and sewer access are conspicuously understudied. There is now a large active literature evaluating various improvements to transportation infrastructure, both in the developed and developing world. Electricity generation and distribution has also received attention. The literature on piped water and sewer access is much less developed. Indeed, as a result of conflicting estimates presented in Cutler and Miller (2005) and Anderson et al. (2018), recent research has served to increase our uncertainty about the importance public health policy. In this light, our results are doubly important. We are the first to evaluate the effect of piped water and sewer access on land prices, a comprehensive revealed preference measure of value, and our results suggest a value of piped water and sewer access that is large, even relative the large estimates of Cutler and Miller (2005).

This generally supports a high priority for water and sewer infrastructure. It also highlights the importance of further research on the the issue. The disease environment in modern Latin America and Africa is clearly different than it was in 19th century Chicago, so the desirability of studies conducted in these places is high. An important obstacle to such research has been the absence of a credible research design for estimating causal effects. We are hopeful that some variant of the research design we develop can help to address this issue.

Our results also inform the ongoing inquiry into the development of the American economy. Up until now, almost all evidence for or against the importance of piped water and sewer infrastructure reflects changes in mortality rates, and is estimated by comparing outcomes before and after a particular intervention. By offering a novel research design, and a different outcome, we provide independent evidence for the importance piped water and sewer infrastructure. Our most conservative estimate indicates that piped water and sewer access more than doubled land prices.

Finally, we propose a technique for the principled extrapolation of treatment effects from a quasi-experimental study area to an area that is more relevant for economics analysis. The practice of restricting attention to small areas, carefully chosen so that a quasi-experimental research design may be defended, is a pervasive practice in applied micro-economic analyses. Thus, so to is the problem of extrapolating to more economically interesting regions. We hope that our technique for extrapolating treatment effects will, therefore, find wide use among other applied researchers.

References

Aldous, D. (1999). *International turf management handbook*. CRC Press.

Alsan, M. and Goldin, C. (2019). Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, 127(2):586–638.

Ambrus, A., Field, E., and Gonzalez, R. (2020). Loss in the time of cholera: Long-run impact of a disease epidemic on the urban landscape. *American Economic Review*, 110(2):475–525.

Anderson, D. M., Charles, K. K., and Rees, D. I. (2018). Public health efforts and the decline in urban mortality. Technical report, National Bureau of Economic Research.

Anderson, D. M., Charles, K. K., and Rees, D. I. (2019). Public health efforts and the decline in urban mortality: Reply to Cutler and Miller. *Available at SSRN 3314366*.

Andrews, I. and Oster, E. (2019). A simple approximation for evaluating external validity bias. *Economics Letters*, 178:58–62.

Angrist, J. and Imbens, G. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Associations*, 90:431–442.

Angrist, J. D. and Fernández-Val, I. (2013). *ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework*, volume 3 of *Econometric Society Monographs*, page 401–434. Cambridge University Press.

Angrist, J. D. and Rokkanen, M. (2015). Wanna get away? regression discontinuity estimation of exam school effects away from the cutoff. *Journal of the American Statistical Association*, 110(512):1331–1344.

Asbury, H. (1940). *Gem of the prairie: An informal history of the Chicago underworld*. AA Knopf.

Ashraf, N., Glaeser, E., Holland, A., and Steinberg, B. M. (2017). Water, health and wealth. Technical report, National Bureau of Economic Research.

Beach, B. (2021). Water infrastructure and health in US cities. *Regional Science and Urban Economics*, page 103674.

Bleakley, H. and Lin, J. (2012). Portage and path dependence. *The Quarterly Journal of Economics*, 127(2):587–644.

Cain, L. P. (1978). *Sanitation strategy for a lakefront metropolis*. Northern Illinois University Press.

Carneiro, P., Heckman, J. J., and Vytlacil, E. (2010). Evaluating marginal policy changes and the average effect of treatment for individuals at the margin. *Econometrica*, 78(1):377–394.

Carneiro, P., Heckman, J. J., and Vytlacil, E. (2011). Estimating marginal returns to education. *American Economic Review*, 101(6):2754–2781.

Carr, T. and Kitagawa, T. (2021). Testing instrument validity with covariates.

Cattaneo, M. D., Keele, L., Titiunik, R., and Vazquez-Bare, G. (2020). Extrapolating treatment effects in multi-cutoff regression discontinuity designs. *Journal of the American Statistical Association*, 0(0):1–12.

Chicago Board of Public Works (1873). *Annual Report of the Board of Public Works to the Common Council of the City of Chicago*. The Board of Public Works.

Costa, D. L. and Kahn, M. E. (2004). Changes in the value of life, 1940–1980. *Journal of Risk and Uncertainty*, 29(2):159–180.

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

Cutler, D. M. and Miller, G. (2020). Comment on "re-examining the contribution of public health efforts to the decline in urban mortality". *Available at SSRN 3312834*.

Dehejia, R., Pop-Eleches, C., and Samii, C. (2021). From local to global: External validity in a fertility natural experiment. *Journal of Business Economics and Statistics*, 39(1):217–243.

Devoto, F., Duflo, E., Dupas, P., Parienté, W., and Pons, V. (2012). Happiness on tap: Piped water adoption in urban Morocco. *American Economic Journal: Economic Policy*, 4(4):68–99.

Drinan, J., Drinan, J. E., and Spellman, F. (2000). *Water and wastewater treatment: A guide for the nonengineering professional*. Crc Press.

Easterlin, R. (1960). Interregional differences in per capita income, population, and total income, 1840-1950. Technical report, National Bureau of Economic Research.

Eli, S. (2015). Income effects on health: evidence from union army pensions. *The Journal of Economic History*, 75(2):448–478.

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

Fogel, R., Costa, D., Villarreal, C., Bettenhausen, B., Hanss, E., Roudiez, C., Yetter, N., and Zemp, A. (2014). Historical urban ecological data set. Technical report, Center for Population Economics, University of Chicago Booth School of Business, and The National Bureau of Economic Research.

Fogel, R. W. et al. (2004). *The escape from hunger and premature death, 1700-2100: Europe, America, and the Third World*, volume 38. Cambridge University Press.

Galiani, S., Gertler, P., and Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *Journal of Political Economy*, 113(1):83–120.

Gamper-Rabindran, S., Khan, S., and Timmins, C. (2010). The impact of piped water provision on infant mortality in brazil: A quantile panel data approach. *Journal of Development Economics*, 92(2):188–200.

Haines, M. R. (2001). The urban mortality transition in the United States, 1800-1940. In *Annales de Démographie Historique*, number 1, pages 33–64. Belin.

Heckman, J. J. and Vytlacil, E. (2005). Structural equations, treatment effects, and econometric policy evaluation 1. *Econometrica*, 73(3):669–738.

Hotz, V. J., Imbens, G. W., and Mortimer, J. H. (2005). Predicting the efficacy of future training programs using past experiences at other locations. *Journal of Econometrics*, 125:241–270.

Hoyt, H. (2000). *One hundred years of land values in Chicago: The relationship of the growth of Chicago to the rise of its land values, 1830-1933*. Beard Books.

Kesztenbaum, L. and Rosenthal, J.-L. (2017). Sewers' diffusion and the decline of mortality: The case of Paris, 1880–1914. *Journal of Urban Economics*, 98:174–186.

Kolesár, M. (2013). Estimation in an instrumental variables model with treatment effect heterogeneity. *unpublished manuscript*.

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.

Mara, D. (1996). *Low-cost sewerage*. John Wiley London.

McKeown, T. (1976). *The modern rise of population*. London (UK) Arnold Pub.

Melosi, M. V. (2000). *The sanitary city: Urban infrastructure in America from colonial times to the present*. Johns Hopkins University Press Baltimore.

Mourifié, I. and Wan, Y. (2017). Testing local average treatment effect assumptions. *Review of Economics and Statistics*, 99(2):305–313.

Rokkanen, M. A. (2015). Exam schools, ability, and the effects of affirmative action: Latent factor extrapolation in the regression discontinuity design.

Sahr, R. (2009). *Inflation conversion factors for dollars 1774 to estimated 2019*. University of Oregon Working Paper Series.

Słoczyński, T. (2021). When should we (not) interpret linear IV estimands as LATE? *unpublished manuscript*.

Stock, J. H. and Yogo, M. (2002). Testing for weak instruments in linear IV regression. Technical report, National Bureau of Economic Research Cambridge, Mass., USA.

Appendix A Supplementary Results

Table A1: Summary Statistics 1886-1889, after piped water and sewer construction

	(1) $SW\triangle = 1$	(2) $SW\triangle = 0$	(3) t-test
Share Sewered	1.00	1.00	.
	(0.00)	(0.00)	
Log Price	8.35	8.56	1.56
	(0.94)	(0.78)	
Log Distance to CBD	9.08	8.98	-1.46
	(0.35)	(0.48)	
Log Area	8.29	8.19	-0.99
	(0.67)	(0.51)	
Share Improved	0.22	0.15	-1.11
	(0.42)	(0.36)	
Share Corner	0.09	0.10	0.34
	(0.29)	(0.31)	
Distance to Horsecar	751	374	-5.50
	(527)	(314)	
Distance to Major Street	512	438	-1.11
	(431)	(390)	
Year	1887.19	1887.35	0.95
	(0.95)	(1.07)	
Observations	68	86	

Note: *Means and standard deviations of parcel characteristics. Column 1 reports on parcels in the Quasi-experimental sample (within 2000′ of Tyler St. west of Halsted) that are in the Southwest Triangle (south of Tyler Street). Column 2 presents corresponding values for parcels that are not in the Southwest Triangle (i.e., north of Tyler Street). Column 3 reports a the t-statistic for the difference between the first two columns. In all columns, we restrict attention to parcels transacted during 1886-1889.*

Table A2: Reduced form regressions after completion of piped water and sewer network.

	(1)	(2)	(3)	(4)	(5)	(6)
Reduced Form: ln(Price)						
SW $\triangle = 1$	-.174	-.233***	.165	-.183*	-.146	-.164*
	(.119)	(.096)	(.225)	(.105)	(.1)	(.09)
Miles to Boundary			1.03			
			(.539)			
R^2	0.364	0.580	0.590	0.598	0.330	0.454
Year FE & ln(Area)	Y	Y	Y	Y	Y	Y
ln(mi. CBD)	Y	Y	Y	Y	Y	Y
Improved and Corner		Y	Y	Y		Y
Horsecar and Major Street				Y		
Sample	Q.E.	Q.E.	Q.E.	Q.E.	E.Q.E.	E.Q.E.
Observations	143	143	143	143	213	213

Note: *All results based on transactions during 1886-9. Columns 1-4 rely on the Quasi-experimental area, 5 and 6 on the Extended-quasi-experimental area. Regressions are reduced form regressions of log transaction price on the instrument and, in column (3), distance to the Tyler Street. Bottom panel of the table indicates control variables. Unlike the 1874-80 period, the entire Southwest Triangle has piped water and sewer access by 1886-9 and the price difference across the Tyler Street boundary is small economically and statistically.*
Robust standard errors in parentheses. $*, **, ***$ *indicates* 10%, 5%, 1% *significance.*

Table A3: (a) LIV Regression Results

	(1) 1st Stage	(2) 2nd Stage	(3) 1st Stage	(4) 2nd Stage	(5) 1st Stage	(6) 2nd Stage	(7) 1st Stage	(8) 2nd Stage	(9) 1st Stage	(10) 2nd Stage
Z	3.95***		4.08***		5.55***		2.76***		2.74***	
ln(Area)	-.08	.72***	.01	.63***	-.02	.63***	-.34	.72***	-.33	.67***
	(.29)	(.22)	(.33)	(.21)	(.35)	(.21)	(.23)	(.2)	(.25)	(.2)
$\mathbb{1}$(Year = 1875)	.56	.45**	.6	.42**	.57	.35*	.21	.38*	.24	.42*
	(.64)	(.2)	(.65)	(.19)	(.72)	(.19)	(.54)	(.23)	(.53)	(.22)
$\mathbb{1}$(Year = 1876)	.95	.39	.99	.37	.89	.29	.42	.35	.44	.38
	(.66)	(.26)	(.68)	(.27)	(.75)	(.28)	(.54)	(.32)	(.54)	(.31)
$\mathbb{1}$(Year = 1877)	1.41*	.52	1.59**	.58	1.73**	.47	1*	.42	.89	.38
	(.72)	(.36)	(.74)	(.39)	(.8)	(.38)	(.57)	(.37)	(.58)	(.33)
$\mathbb{1}$(Year = 1878)	3.06***	.32	3.31***	.38	3.6***	.23	1.58***	.29	1.38**	.21
	(.83)	(.43)	(.89)	(.44)	(.93)	(.38)	(.66)	(.5)	(.69)	(.43)
$\mathbb{1}$(Year = 1879)	2.45***	-.08	2.66***	.03	2.86***	-.03	1.15**	-.38	1.05*	-.27
	(.73)	(.49)	(.76)	(.44)	(.81)	(.49)	(.56)	(.58)	(.57)	(.53)
$\mathbb{1}$(Year = 1880)	3.65***	-.63	3.86***	-.26	4.09***	-.59	2.72***	-1.54	2.6***	-1.21
	(.71)	(.63)	(.75)	(.51)	(.79)	(.57)	(.53)	(.94)	(.54)	(.74)
ln(mi. CBD)	-5.83***	.31	-5.93***	.03	-8.3***	.09	-5.41***	.85	-5.38***	1.2
	(.91)	(.64)	(.93)	(.57)	(1.32)	(.58)	(.71)	(.79)	(.71)	(.76)
$\mathbb{1}$(Improved)			-.6	.43	-.7	.51			.66	.52
			(.63)	(.52)	(.64)	(.46)			(.5)	(.66)
$\mathbb{1}$(Corner)			-.52	.53*	-.6	.43			.12	.35
			(.64)	(.29)	(.7)	(.29)			(.49)	(.34)
Year FE & ln(Area)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
ln(mi. CBD)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Improved and Corner			Y	Y	Y	Y			Y	Y
Horsecar and Major Street					Y	Y				
Sample	Q.E.	Q.E.	Q.E.	Q.E.	Q.E.	Q.E.	E.Q.E.	E.Q.E.	E.Q.E.	E.Q.E.
Observations	351	351	351	351	351	351	533	533	533	533

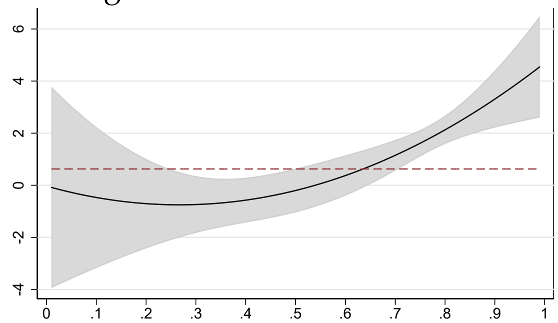
Note: *Table continued next page*

Table A3: (b) LIV Regression Results

	(1) 1^{st} Stage	(2) 2^{nd} Stage	(3) 1^{st} Stage	(4) 2^{nd} Stage	(5) 1^{st} Stage	(6) 2^{nd} Stage	(7) 1^{st} Stage	(8) 2^{nd} Stage	(9) 1^{st} Stage	(10) 2^{nd} Stage
$\widehat{p}$		.74		1.21		1.3		2.39		3.59
		(2.84)		(2.73)		(2.8)		(2.91)		(2.92)
$\widehat{p}^2$		-3.56		-3.04		-2.74		-.94		-1.71
		(4.83)		(4.41)		(4.23)		(4.51)		(4.1)
$\widehat{p}^3$		3.81		3.65		3.26		1.05		1.59
		(3.03)		(2.77)		(2.62)		(2.72)		(2.5)
$\widehat{p}\ln(\text{Area})$		-.1		.02		.02		.09		.16
		(.23)		(.23)		(.22)		(.23)		(.23)
$\widehat{p}\mathbb{1}(\text{Year} = 1875)$		-.97***		-.93***		-.77***		-.66*		-.69*
		(.33)		(.32)		(.29)		(.37)		(.36)
$\widehat{p}\mathbb{1}(\text{Year} = 1876)$		-.64*		-.6		-.39		-.35		-.38
		(.39)		(.4)		(.38)		(.46)		(.46)
$\widehat{p}\mathbb{1}(\text{Year} = 1877)$		-1.4***		-1.66***		-1.4***		-.93*		-1.02**
		(.54)		(.56)		(.49)		(.5)		(.46)
$\widehat{p}\mathbb{1}(\text{Year} = 1878)$		-1.24**		-1.58***		-1.18***		-1.04*		-1.19**
		(.54)		(.55)		(.44)		(.6)		(.53)
$\widehat{p}\mathbb{1}(\text{Year} = 1879)$		-1.09*		-1.43***		-1.17**		-.36		-.64
		(.59)		(.54)		(.55)		(.67)		(.61)
$\widehat{p}\mathbb{1}(\text{Year} = 1880)$		-.51		-1.2*		-.62		.78		.21
		(.72)		(.62)		(.62)		(1.01)		(.83)
$\widehat{p}\ln(\text{mi. CBD})$		-.11		.14		.07		-.57		-.92
		(.68)		(.61)		(.62)		(.85)		(.81)
$\widehat{p}\mathbb{1}(\text{Improved})$				.38		.28				0
				(.56)		(.51)				(.69)
$\widehat{p}\mathbb{1}(\text{Corner})$				-.14		-.01				-.05
				(.36)		(.34)				(.39)
Year FE & ln(Area)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
ln(mi. CBD)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Improved and Corner			Y	Y	Y	Y			Y	Y
Horsecar and Major Street					Y	Y				
Sample		Q.E.		Q.E.		Q.E.		E.Q.E.		E.Q.E.
Observations		351		351		351		533		533

Note: *Estimates of the LIV model of equation* (4). *Odd columns are Logit first stage coefficients and even columns are corresponding second stage. Specifications and samples match those reported in the same columns of table 3. Bottom panel indicates controls for the regression above. Bootstrapped standard errors in parentheses.* $*, **, ***$ *indicates* 10%, 5%, 1% *significance.*

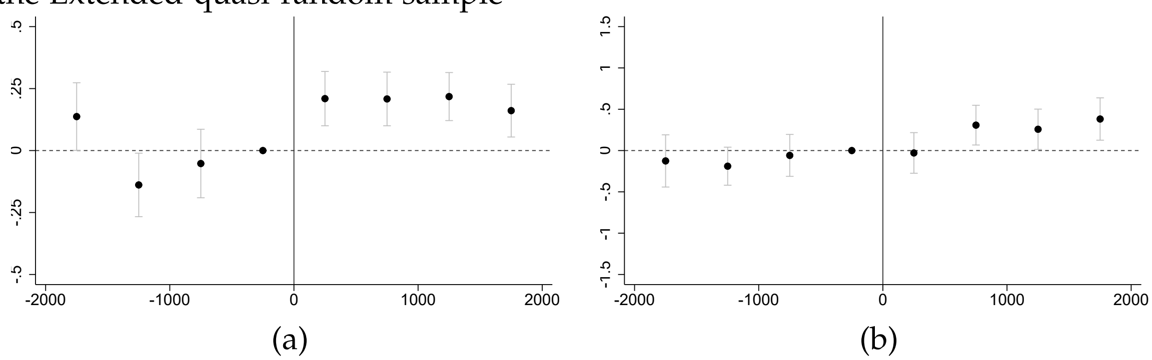
Figure A1: Marginal Treatment Effect as a function of $\widetilde{U}_D$



Marginal Treatment Effects

Note: *Expected MTE as a function of* $\widetilde{U}_D$*. Dashed line shows ATE for this sample/specification and sample average* X*'s. Based on column 2 of Table 3.*

Figure A2: Sewer incidence and land price by distance to boundary, 1874-80, for the Extended-quasi-random sample



Note: *(a) Share of parcels sewered 1874-80 by 500′ bins of distance to SW△ boundary,* $x < 0$ *is "inside".* $x \in [-500{,}0]$ *is* y *intercept. Conditional on year,* $\ln(\textit{area})$, $\ln(\textit{mi. to CBD})$. *(b) Same as left panel but* y*-axis is* $\ln(\textit{Price})$.

Appendix B Derivation of equation (9)

We maintain the MTE model with the semiparametric potential outcome equations introduced in the main text; see (1) in the main text. We also maintain the key restriction of practical exogeneity; see (2) in the main text. With propensity score $p = F(x,z) = P(D = 1|X = x, Z = z)$ introduced in the main text and the normalized unobserved heterogeneity in the selection process,

$\widetilde{U}_D \sim Unif[0,1]$, the selection equation can be represented as

$$D = 1\{\widetilde{U}_D \leq F(X,Z)\}. \qquad \text{(Appendix B.1)}$$

Under the cubic polynomial specification of the control function $K(p)$ in (4), MTE at each conditioning covariate value X and $\widetilde{U}_D \in [0,1]$ is given as in (5), and averaging $(X,\widetilde{U}_D)$ for the population of the Quasi-experimental sample leads to ATE in the Quasi-experimental sample (6).

Our interest is to obtain an estimate for ATE for the population of Relevant sample P^* as denoted by ATE^* in the main text. To analyze formally how to extrapolate to the Relevant sample, we assume that a unit in the Relevant sample admits the same structural equations (7) with the same parameter values as a unit in the Quasi-experimental sample. Importantly, even though we assume that a binary cost shifter Z^* is present and measures the cost of access to sewage in the same scale for each unit in the Relevant as in the Quasi-experimental sample, Z^* is not observed for any unit of the Relevant sample. In addition, unlike in the Quasi-experimental sample, Z^* is not need not be randomly assigned and an analogue of the instrument exgeneity assumption $Z^* \perp (U_1^*,U_0^*,U_D^*)$ may fail in P^*.

The following assumption describes what is necessary, and what is not, for feasible extrapolation from P to P^*.

Assumption EX: (The relationship between P and P^*)

(i) The equations of potential outcomes and selection given in (1) are identical between the Quasi-experimental and Relevant samples (other than that Z^* is not observed in P^*). Furthermore, the distribution of (U_1,U_0,U_D) and (U_1^*,U_0^*,U_D^*) are common.

(ii) The joint distribution of observable covariates X and cost shifter (instrument) Z in the Quasi-experimental sample and the joint distribution of X^* and Z^* in the Relevant sample can be different.

Under (i), we can normalize U_D^* of (7) to define the uniform random variable $\widetilde{U}_D^* = F_{U_D^*}(U_D^*)$ such that for $\widetilde{U}_D$ defined in (Appendix B.1), $\widetilde{U}_D^* = \widetilde{U}_D$ is equivalent to $U_D^* = U_D$. In other words, a unit in the Relevant sample and a unit in the Quasi-experimental sample that share the values of $\widetilde{U}_D^*$ and $\widetilde{U}_D$ have identical unobservables in the selection equation. In addition, Assumption EX (i) implies that the control function term $K(\cdot)$ in the LIV regression (4) are common between the two samples because the control function term is determined only by the distribution of $(U_1,U_0)|U_D$ and this does not vary between the two samples by the assumption. As a result, for MTE in the Relevant sample $MTE^*(X^*,\widetilde{U}_D^*)$, $MTE(X,\widetilde{U}_D) = MTE^*(X,\widetilde{U}_D^*)$ holds whenever $X = X^*$ and $\widetilde{U}_D = \widetilde{U}_D^*$ hold. We hence obtain

$$MTE^*(X^*,\widetilde{U}_D^*) = (X^*)'(\delta_1 - \delta_0) + \gamma_1 + 2\gamma_2\widetilde{U}_D^* + 3\gamma_3\widetilde{U}_D^{*2}. \qquad \text{(Appendix B.2)}$$

Taking the expectation with respect to X^* and $\widetilde{U}_D^* \sim Unif[0,1]$, we obtain the equation of (9) in the main text, where $E(X^*)$ is directly identified by the data of Relevant sample. Note that this argument does not require any condition for Z^*.