Channeling Density? Sewers and the Growth of 19th-Century Paris

Ignacio Banares-Sanchez[†] Yoshiki Wiskamp[‡]

(Job Market Paper)

Click **here** for latest version.

November 3, 2026

Abstract

Urban density promises agglomeration benefits but also exacerbates pollution and disease transmission. Sanitation infrastructure may therefore be essential for cities to sustain dense concentrations of economic activity. To study how sanitation investments shape urban development, we compile a novel address-level dataset covering the construction of the Parisian sewer network in the nineteenth century. Exploiting within-neighbourhood variation in the timing of sewer connections in an event-study design, we find that connecting a parcel to the sewer network increases both population and firm counts. Sewer connections in nearby locations generate additional increases, with larger spillovers in more densely populated areas. We incorporate these findings into a quantitative urban model and estimate that sewer access increases residential amenity values by 2.5% and firm productivity by 5% in connected locations. The analysis further reveals that, in unsewered areas, the marginal effect of additional people on others' utility turns negative as neighbourhoods densify and pollution costs intensify. Sewers eliminate this downturn. Counterfactuals indicate that sewerage investments account for 30% of Paris' population growth in the second half of the nineteenth century. One-quarter of this effect operates through reductions in negative sanitation externalities, while the remainder reflects the private benefits of sewer access. These findings highlight how addressing waste pollution can lower the costs of proximity and sustain urban growth.

We thank Oriana Bandiera, Gharad Bryan, Robin Burgess, and Daniel Sturm for their fantastic advice. For helpful feedback, we also thank Clare Balboni. We also benefitted from comments by Sree Ayyar, Johannes Boehm, Jonathan Colmer, Swati Dhingra, Maitreesh Ghatak, Edward Glaeser, Matthew Gordon, Allan Hsiao, Felix Iglhaut, Marta Morando, and numerous seminar participants at LSE. We are very grateful to Lionel Kesztenbaum and Matthijs Korevaar for sharing information about data sources. Kartika Herscheid provided excellent research assistance.

[†]UC Berkeley, Department of Agricultural and Resource Economics: i.banares-sanchez@berkeley.edu

[‡]London School of Economics, Department of Economics: y.wiskamp@lse.ac.uk

1 Introduction

Sanitation infrastructure is among the major public investments undertaken in cities throughout the urbanisation process. In many places, these investments are yet to take place, with about one-third of the world's urban population still lacking access to safely managed sanitation (WHO and UNICEF 2025). In its absence, population density gives rise to pollution and disease transmission, which counteract the economic benefits of urban proximity (Glaeser 2011). While there is ample evidence on the public health benefits of sanitation (e.g. Cutler and Miller 2005; Alsan and Goldin 2019), we know surprisingly little about its importance for sustaining economic activity in cities.

Sanitation investments may be transformative for cities. By reducing the negative externalities that people impose on one another, they may relax constraints on the willingness of households and firms to locate close to each other. This proximity is what allows for economic activity to flourish and the productivity gains from agglomeration to emerge. A long literature has highlighted the benefits from fostering economic interactions by reducing the costs of distance, either via transport (e.g. Donaldson 2018; Faber 2014) or communication networks (e.g. Steinwender 2018). Much less is known about how the spatial organisation of cities and urbanisation depend on lowering the costs of proximity.

In this paper, we seek to quantify the role of public sanitation investments in allowing economic activity in cities, and to understand the channels through which they operate. This requires data linking infrastructure expansion to changes in local economic outcomes, which are often incomplete for the periods in which these investments take place. We therefore turn to the expansion of the Parisian sewer network in the second half of the nineteenth century, which is documented in unusually rich historical records. Moreover, sewerage investments in Paris were centrally planned and publicly funded, providing a setting to study the effects of a textbook public intervention to address externalities.

We make three contributions. First, we construct a novel address-level dataset covering the sewer network expansion in nineteenth-century Paris. We combine newly digitised sewer maps, population registers, property transactions and health records with existing business directories. Second, we use these data to provide empirical evidence on how sewers transformed neighbourhoods. Using a difference-in-differences design that compares parcels within the same neighbourhood that were connected at different points in time, we find that receiving a sewer increased the number of residents and firms. Sewer connections in nearby locations generated positive spillovers that increased with local density, consistent with sewers improving neighbourhood conditions more strongly where more people live within close proximity. Third, we quantify the aggregate effects of sanitation in a quantitative urban

model in which sewers raise local amenities and productivity and reduce the negative externalities that residents impose on one another in the absence of sanitation. We estimate the model using simulated method of moments to match our empirical estimates. Through counterfactuals, we examine the contribution of sewers to Paris's population growth and the channels through which they operate. We further study complementarities between sanitation and other public investments, asking whether reducing sanitation externalities allows more people to move to locations that received these investments.

Travelling back in time, sanitary conditions in early nineteenth-century Paris were reportedly poor. The historic centre only had a few sewers dating back to medieval times. In their absence, wastewater accumulated in the streets, creating damp, malodorous environments that provided breeding grounds for disease. A major centralised modernisation effort began in the 1850s, making Paris one of the first cities in the world to undertake a large-scale expansion of its sewerage network. Large collectors were built to carry wastewater beyond Paris, following topographically favourable routes. Smaller pipes were spread out from these collectors to expand coverage across the city as quickly as possible. By the end of the century, the network covered the vast majority of Paris.

We construct a novel address-level dataset to study the effects of the Parisian sewer expansion. We digitise historical sewer maps to reconstruct the network over time and determine where and when infrastructure investments took place. To trace population responses, we digitise electoral registers preserved in the Paris archives, photographing over 50,000 pages. These registers record individual-level information on the male population of voting age, including residential addresses, occupations and places of birth. We complement these data with existing digitised business directories, which record firm locations and industries, property transaction records to measure house prices, and historical maps of typhoid deaths, which we digitise to locate deaths from this waterborne disease.

At the core of our analysis is a framework that distinguishes between two channels through which sewers may increase the attractiveness of locations to residents and firms. First, sewer access provides a private benefit to those located in connected areas by making it easier to dispose of wastewater and drain rainwater around their properties. Second, sewers may reduce the negative externalities that people impose on one another by removing wastewater that would otherwise accumulate in densely populated areas. Quantifying the strength of these two mechanisms is important because they have both common and distinct implications. Both mechanisms affect the relative attractiveness of locations within the city, shaping the spatial distribution of residents and firms, and may also increase the overall attractiveness of the city, encouraging in-migration. However, only if sewers reduce the cost of proximity that each additional urban resident imposes on others do they change how strongly the

city responds to forces that attract additional residents and firms, such as improvements in productivity, amenities or public infrastructure.

Our empirical analysis proceeds in two steps. We first study how sewers transform the locations in which they are introduced. Exploiting within-neighbourhood variation in the timing of sewer connections, we compare small parcels connected during the first expansion wave (1857–1872) with parcels connected later (1873–1896). Event-study estimates show no differential pre-trends, with effects emerging around the timing of sewer connections. We find that receiving a sewer increases the number of voters by 10% and the number of firms by 18%. Our estimates remain similar when controlling for contemporaneous changes in access to public fountains, railway stations and parks. To better understand the drivers of this transformation, we examine heterogeneity across firm types and changes in household composition. The increase in firm numbers is driven by customer-facing businesses in the retail and service industries, suggesting that cleaner surroundings were particularly valuable for businesses that rely on attracting customers. On the household side, we construct measures of social class using individuals' qualifications recorded in the electoral registers and identify migrants using information on their birthplace. We find no evidence of changes in either the social-class composition of residents or the share of migrants. These findings suggest that sewer expansion did not lead to detectable changes in neighbourhood composition along these dimensions, potentially reflecting the rapid expansion of the network across the city. Finally, we also study a later phase of network expansion (1880–1904), for which we observe typhoid deaths, and find that typhoid deaths declined in locations that received sewers, consistent with improved sanitary conditions.

Our second empirical exercise investigates more directly whether sewers reduce negative externalities. In particular, we examine whether nearby sewerage investments generate positive spillovers for a given location, and whether these spillovers vary with surrounding population density. For each parcel, we measure nearby sewerage investments as the share of addresses within 250 metres that gained access to sewers during the first expansion wave (1857–1872). We measure surrounding density as the number of addresses within the same radius, which we show to be strongly correlated with the number of residents and firms. Using a triple-difference design, we find that nearby sewerage investments have larger effects on the number of residents in more densely populated areas, with positive spillovers emerging mainly at higher levels of density. These findings suggest that, at lower levels of density, sewerage primarily provides a private benefit to those who gain direct access to the network. To sustain high levels of population density, however, sewerage also becomes important by reducing the negative externalities that residents impose on one another through wastewater pollution. For firms, we do not find a comparable density gradient. Instead, firms' location

choices appear to depend on whether the parcel itself has access to sewers, allowing them to operate in cleaner surroundings where they attract customers.

Building on the empirical analysis, we leverage an urban model to quantify the effects of sewerage investments. The model builds on the frameworks of Ahlfeldt et al. (2015) and Heblich et al. (2020), in which workers jointly choose residential and workplace locations to maximise utility, while firms choose their labour input to maximise profits. We incorporate sewers into this framework by allowing them to affect both local residential amenities and firm productivity through two distinct channels. First, sewers generate private benefits by increasing the amenity value of locations for residents and productivity for firms. Second, they mitigate negative sanitation externalities by reducing the extent to which nearby population density erodes residential amenities and firm productivity. We measure density as the number of residents and workers within 250 metres, with the associated sanitation externalities depending on the share of this population connected to the sewer network. By reducing these externalities, sewers allow for denser concentrations of residents and firms to be sustained. Finally, we allow Paris's population to adjust endogenously through in-migration from the rest of France, enabling us to quantify the aggregate contribution of sewerage to the city's population growth.

We estimate the effects of sewerage on amenities and productivity using Simulated Method of Moments (SMM), which is an indirect inference approach that selects model parameters to match estimates from the empirical analysis. The estimation proceeds in two steps. First, we invert the model to recover location-specific amenities and productivity as structural residuals that rationalise the observed distribution of residents and workers as the equilibrium outcome of their location choices. We do this for 1872, the post-treatment period, as this is the year for which we observe the distribution of both residents and workers across the entire city. Second, for each candidate set of sewer parameters, we remove the effects of sewer connections built between 1856 and 1872 from the recovered amenities and productivity, then solve the model under the 1856 sewer network to obtain the implied distribution of residents and workers. We run the same regressions on the simulated pre-period and observed post-period as in the empirical analysis. The SMM estimates are the parameter values that best match two sets of empirical moments. The first comprises the difference-in-differences estimates of the effects of a location's own sewer access on the number of residents and workers. The second captures the effects of nearby sewer access on the number of residents across bins of local density, allowing us to identify how the benefits of nearby sewers vary with the number of people living nearby.

The parameter estimates themselves already provide valuable insights into the effects of sewers. We find that sewer access raises the amenity value of a location by 2.5% and

its productivity by 5%, capturing the private benefits for residents and firms in sewered locations. In addition, sewers mitigate a sanitation externality whose severity increases sharply with density. In areas where nobody within a 250-metre radius has access to sewers, this externality is negligible at the median density in Paris in 1872 but reduces amenities by 30% at the 95th percentile. Combined with the positive agglomeration benefits estimated by Ahlfeldt et al. (2015), this implies an inverse U-shaped relationship between population density and the attractiveness of unsewered neighbourhoods. At low densities, additional residents make a location more attractive on net as positive spillovers dominate. Beyond approximately the 80th percentile of the 1872 density distribution, however, each additional resident reduces its attractiveness to others as sanitation costs increasingly outweigh the benefits of agglomeration. Sewers remove this downturn, allowing cities to sustain the high densities at which agglomeration benefits can be realised, rather than reaching a point at which further growth becomes self-defeating. We believe that these findings also offer broader insights into urban development. Cities can begin to emerge before sanitation infrastructure is provided at scale, but sustaining dense concentrations of population may eventually require investments that limit the cost of proximity.

We use the estimated model to examine the aggregate effects of sewerage investments. In the first counterfactual, we quantify the contribution of the sewer network to Paris's growth by facilitating in-migration. For each census year between 1872 and 1896, we calibrate the outside option of living elsewhere in France so that the model with the sewer network in place in that year reproduces Paris's observed population. We then reduce the sewer network to its 1856 extent and solve for the new equilibrium, in which the utility of living in Paris is equalised with that of living elsewhere in France. Without the sewer expansion, Paris would have had 12% fewer residents in 1896, accounting for approximately 30% of the city's population growth between 1856 and 1896. About a quarter of this effect arises from the reduction in sanitation externalities, with this share increasing over time as urban growth makes the costs of proximity in unsewered neighbourhoods increasingly severe.

In the second counterfactual, we quantify the extent to which sanitation amplifies the effectiveness of other public investments. By reducing the costs that residents impose on one another, sewers may allow more people to live closer to locations where other investments take place. We study this complementarity by simulating reductions in commuting times equivalent to those that the opening of the Metro (at its 1910 extent) would have induced. We conduct this exercise for 1896, when much of Paris was already covered by the sewer network but before the actual opening of the Metro in 1900. We first reduce travel times and measure the resulting increase in Paris's population as improved accessibility makes the city more attractive. We then repeat the exercise in a hypothetical version of the same city

without any sewers, in which the observed distribution of residents and workers is sustained by the same fundamentals, but sanitation externalities operate in all locations. Comparing the population responses, we find that without sewerage, the population gain from the Metro would have been about one-eighth smaller. These complementarities suggest that sanitation infrastructure can increase the returns to subsequent public investments, making its early provision an important part of a city's development.

Finally, we compare the benefits of the sewer expansion with its costs. Measured by the increase in the value of the building stock, the expansion of the network between 1856 and 1872 generated benefits of about 760 million francs, against municipal expenditure of 45.5 million francs, a benefit–cost ratio of about 17. By comparison, the metro generated a benefit–cost ratio of about 6, so that per franc spent, the sewer expansion yielded about 2.7 times the benefits of the metro, a comparison that likely understates the returns to sewers, since part of the gains from the metro relied on the sewer network already being in place.

Contribution to the literature. Our paper contributes to three literatures. First, we contribute to the literature on the effects of water and sanitation infrastructure on urban development. Several studies examine the effects of water and sanitation improvements, primarily on mortality, using variation across cities (e.g. Cutler and Miller (2005); Anderson et al. (2022)) and within cities (e.g. Costa and Kahn (2015); Kesztenbaum and Rosenthal (2017); Alsan and Goldin (2019)), as well as experimental evidence from developing countries (e.g. Galiani et al. (2005); Kremer et al. (2011)).[1] More recent work shows that sanitation also affects house prices (Coury et al. 2024), the long-run spatial distribution of poverty following a cholera epidemic (Ambrus et al. 2020), and population counts (McCulloch et al. 2025). This literature has primarily examined the effects of sanitation on outcomes in locations that receive infrastructure. We instead quantify its aggregate implications using a general equilibrium model that accounts for both the reallocation of economic activity within the city and migration into the city. We distinguish the direct benefits of sewer access from reductions in sanitation externalities, quantifying their effects on residents, firms, and overall city size.

Second, we contribute to the quantitative spatial literature on the forces that determine the concentration of economic activity (Ahlfeldt et al. 2015; Allen and Donaldson 2020; Heblich et al. 2020; Redding and Sturm 2024). Existing quantitative spatial models typically capture positive and negative spatial externalities through the elasticities of amenities and productivity with respect to density (Allen and Arkolakis 2014; Ahlfeldt et al. 2015). Our

[1] For Paris specifically, Kesztenbaum and Rosenthal (2017) document a positive association between sewer access and neighbourhood-level life expectancy between 1880 and 1913.

analysis highlights that public infrastructure can affect not only the level of location-specific amenities and productivity, but also the strength of these externalities. The determinants of these elasticities, and whether they can be altered by public policy, remain less well understood. We show that sanitation is an important source of density-related externalities and that sewerage investments mitigate them, allowing cities to sustain denser concentrations of residents and firms. This mechanism may be particularly important in cities at early stages of development, where sanitation infrastructure is yet to expand (Bryan et al. 2020; Henderson and Turner 2020).

Third, we contribute to the literature on the effects of pollution on economic activity in cities. A large body of work shows that air and water pollution reduce health, productivity, and property values (e.g. Chay and Greenstone (2005); Graff Zivin and Neidell (2012); Currie et al. (2015); Greenstone and Hanna (2014); Keiser and Shapiro (2019)). Historical pollution has also shaped the long-run sorting of residents within cities (Heblich et al. 2021) and urban growth (Hanlon 2020). We show that pollution from human waste generated a density-dependent dispersion force that constrained urban concentration. By quantifying how sewerage relaxed this constraint, we demonstrate how public investment can mitigate the cost of pollution.

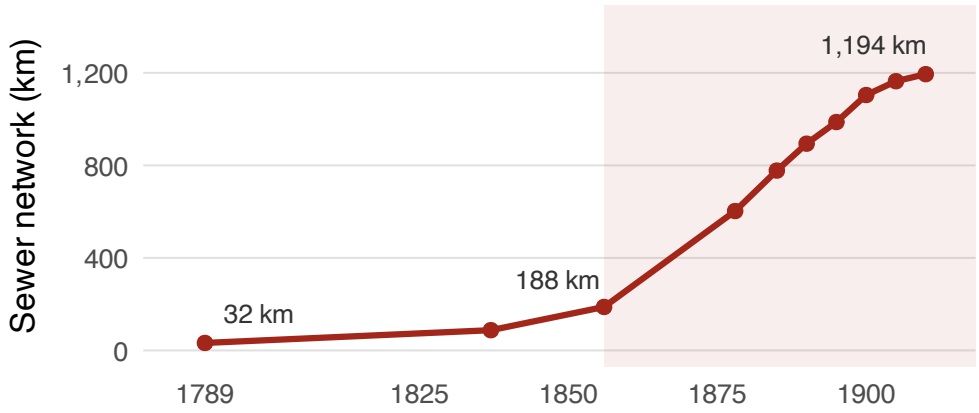
The remainder of the paper is organised as follows. Section 2 provides historical and institutional background on the expansion of the Parisian sewer network. Section 3 describes the construction of our address-level dataset. Section 4 presents our empirical evidence on the effects of sewers. Section 5 develops the quantitative spatial model. Section 6 describes the estimation strategy and results. Section 7 presents our counterfactual analyses. Section 8 concludes.

2 Historical Background

Sewers are among the most fundamental pieces of urban infrastructure, affecting the living conditions of millions of people in cities around the world. Yet large-scale expansions of sewer networks for which detailed data exist on both the infrastructure itself and the response of population and economic activity at a fine spatial scale are rare. We therefore turn to the past. The construction of the Parisian sewer network in the second half of the nineteenth century provides a suitable setting, as it combines one of the largest sanitation programmes of its time with unusually rich historical records.

Sanitary conditions in the early nineteenth century. In the early nineteenth century, Paris had only a few medieval sewers, largely confined to the historical core of the city, and sanitary conditions were poor. In the absence of sewers, wastewater and rainwater

Figure 1: Length of the Parisian sewer network



Notes: Length of the sewer network of Paris in kilometres. Sources: historical sewer maps (1789, 1837, 1856, 1878 and 1885) and the *Annuaire Statistique de la Ville de Paris* (1890–1910). The shaded area marks the expansion from 1856 onwards.

remained in the streets, which were wet and malodorous places. Stagnant water provided a breeding ground for bacteria and facilitated the spread of contagious diseases, and open defecation was common out of necessity. By 1852, the city had only around 142 kilometres of sewers, narrow passages that could carry only liquid waste and that emptied much of their contents into the Seine. Solid waste was collected at night by workers known as *vidangeurs* and brought to dumps on the outskirts of the city. These conditions had already proved lethal, most notably during the cholera epidemics of 1832 and 1849.

The expansion of the sewer network. As shown in Figure 1, the large-scale expansion of the sewer network began in the 1850s and was part of the renovation of Paris under Napoleon III, led by the prefect of the Seine, Baron Haussmann, and his chief engineer, Eugène Belgrand. The aim of the programme was to connect the entire city to the network. Its layout was shaped by engineering constraints. Since wastewater was evacuated by gravity, the placement of the large collectors in the plan of Belgrand was largely determined by the topography of the city, and the network of smaller branch sewers spread out from these collectors (see Appendix Figure A25). Construction began in 1857 with the *collecteur général d'Asnières*, which gathered wastewater from the street sewers and discharged it into the Seine downstream of the city. The sewers were built by excavating the street, constructing the tunnel and covering it again. The programme was financed by the city. Sewerage investment continued throughout the second half of the nineteenth century, and by its end most parts of Paris had access to sewers.

What sewers changed. The main effect of the new sewers was to remove wastewater from the streets and thereby to reduce pollution at the level of the neighbourhood. Direct connections between buildings and the sewer network remained the exception during our

period of study. Instead, an intermediate toilet technology emerged, the *tinette filtrante*, a filtering tub that allowed liquid waste to drain into the sewer while retaining solid waste for collection. According to the *Annuaire Statistique de la Ville de Paris*, the number of *tinettes filtrantes* rose from about 17,000 in 1882 to almost 35,000 in 1890. At the same time, about 81,000 cesspits (*fosses*) were still in use, while fewer than 2,000 toilets were directly connected to the sewer network. Direct connections only spread in the following decade, reaching about 23,000 in 1900 and 42,000 in 1906, when they overtook the cesspits. Connecting buildings directly to the sewer network only became compulsory for landlords in 1894 (Kesztenbaum and Rosenthal 2017), and night soil collection remained widespread until the 1890s. Our estimates therefore capture the effects of sewers on the cleanliness of neighbourhoods, rather than the full gains from modern flush toilets connected to the sewer network that cities experience today.

Other changes in the city. Infrastructure investments typically do not take place in isolation. In 1860, Paris annexed its surrounding communes, which tripled the area of the city and reorganised it into 20 arrondissements and 80 quartiers. We use variation both within the historical core and in the annexed areas, where we expect sewers to operate through the same channels. In the historical core, sewers mainly upgraded existing neighbourhoods, while in the annexed areas they prepared neighbourhoods for future development. The renovation under Haussmann also included the construction of new boulevards and parks, which we account for in our empirical analysis (Section 4). At the same time, the transport network of the city evolved. Throughout the second half of the nineteenth century, public transport relied on horse-drawn omnibuses, until the first line of the metro opened in 1900. We account for these changes in travel times to capture how access to jobs and residents changed across locations.

These investments coincided with a period of rapid growth. Between 1856 and 1896, the population of Paris increased by about 64%, from 1.55 million to 2.54 million, and the number of establishments listed in the commercial directories roughly doubled, from about 81,000 to 163,000. In this paper, we study how the expansion of the sewer network contributed to these changes.

3 Geolocated historical database

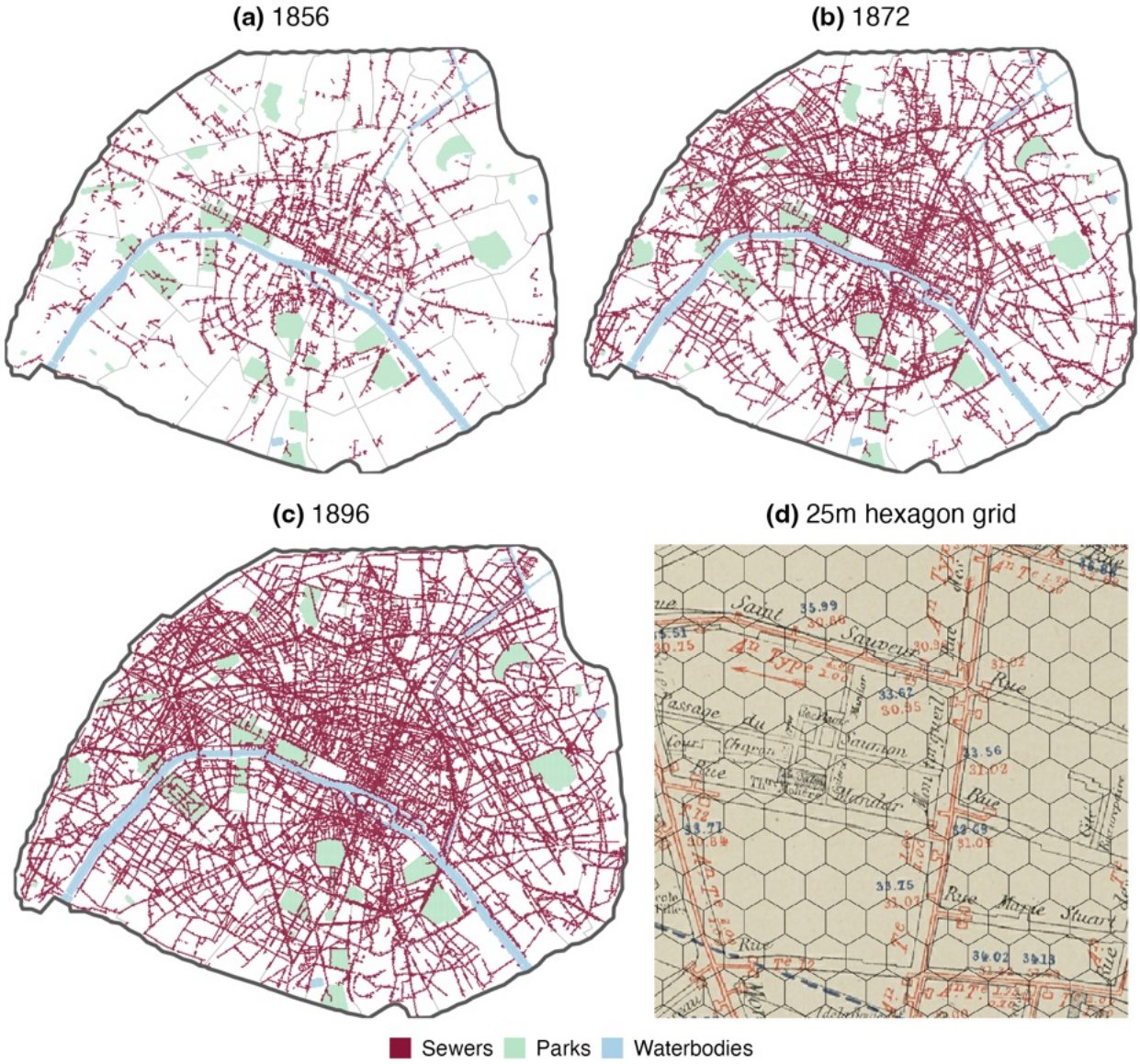
To understand the potentially localised effects of sewers, we construct a novel granular historical dataset covering Paris in the second half of the nineteenth century. We digitise historical maps of the sewer network, which allow us to date when individual parcels gained access to sewers. We combine these with population registers, business directories, property

transactions and records of typhoid deaths, most of which we digitise ourselves, to trace how residents, firms, land values and health responded to the expansion of the network. Together, these sources provide a comprehensive picture of a major infrastructure programme and how it shaped the city. The remainder of this section describes the main data sources and how we digitise them, while Appendix A.1 provides a more detailed discussion.

Geography. There are two levels of spatial aggregation that are relevant for our analysis. The first is the administrative division of Paris after the annexation of 1860 into 20 arrondissements and 80 quartiers (see Figure 2, Panel A). We use quartiers to define neighbourhoods, and this is the level at which we control for neighbourhood-wide changes in our empirical analysis. The second level is much finer and allows us to capture the fact that, in the early stages of the network's construction, access to sewers varied across streets even within neighbourhoods. Since our data are recorded at the address level, we exploit this variation by overlaying Paris with a grid of hexagons with a width of 25 metres between opposite sides, which yields more than 146,000 hexagons covering the city. We refer to these hexagons as parcels and aggregate our address-level data to this level for most of the analysis. We define a parcel as having access to sewers in a given year if it is crossed by a sewer on the map of that year, a definition that the fine spatial unit makes precise (see Figure 2, Panel D).

Sewer network. To trace the expansion of the sewer network throughout the nineteenth century, we digitise historical maps of the network. These maps are not available on an annual basis but provide sufficient coverage of the expansion periods we focus on, with maps for 1856, 1872, 1880, 1889 and 1896. Appendix Table A1 provides an overview of their sources. We digitise the maps using a combination of manual referencing and an automated colour-detection routine, leveraging the fact that sewer lines were typically painted in colours such as red, while other features (e.g. roads) were printed in black. After manually georeferencing the maps, we collect a sample of the RGB values of pixels corresponding to sewer lines in order to construct a colour range for the relevant sewer colour. We then automatically detect all pixels in the image that fall within this range. We subsequently apply post-processing procedures to filter out noise and to remove sewer labels that are commonly printed adjacent to the lines in the same colour. We then convert the detected pixels into lines and harmonise the networks of all years to the 1896 map, so that small differences in georeferencing across map editions do not appear as changes in the network. We classify a hexagon as having access to sewers in a given year if it is crossed by a line of the harmonised network, and we treat access as permanent once a hexagon has been connected. Appendix A.1.1 describes the digitisation in detail and compares it with a manual tracing of the network.

Figure 2: The sewer network and the hexagon grid



Notes: Panels (a) to (c) show the 25m wide hexagons that are crossed by a sewer in the historical maps of 1856, 1872 and 1896 (burgundy), within the boundary of Paris after the annexation of 1860. Grey lines mark the boundaries of the 80 quartiers. Parks are shown in green and waterbodies in blue. Panel (d) overlays the hexagon grid on an extract of a historical sewer map, on which sewers are drawn in red.

Figure 2 shows how access to sewers changed over time. In 1856, the network was still largely confined to the historical centre of the city and consisted to a large extent of sewers that predated the expansion programme. Between 1856 and 1872, a first wave of construction extended the network considerably. Our empirical analysis focuses on this first wave of the expansion, as it is the first expansion wave for which we observe both the sewer network and the distribution of residents. In 1872, a substantial share of locations had not yet been connected, which provides the variation in the timing of sewer access that we exploit. In the counterfactual analysis, we also consider the later expansion of the network up to 1896, by when most parts of Paris were covered.

Residents. Our main digitisation effort focuses on population records in order to construct changes in population density at a granular level. As census data are only reported at the more aggregate quartier level, and the original census lists are not preserved until the 1920s, we turn to the best available data source for this period: the electoral lists. These lists were compiled at the arrondissement level and record all male citizens above the age of 21, comprising approximately one-third of the total population of Paris at the time. They include individuals' names, dates and places of birth, qualifications, and, importantly, their addresses. This information allows us to geolocate individuals and construct measures of population counts at the hexagon level, as well as indicators of socioeconomic composition based on recorded qualifications. To digitise these historical documents, we visited the Archives de Paris to photograph the original electoral registers. We took pictures of over 50,000 pages covering the years 1853, 1859/60, 1871, 1880, and 1892. We transcribe the handwritten entries with a vision–language model (Claude Opus 5, Anthropic), which reads each page and returns one record per individual entry.[2] Of the 1.24 million records, we are able to geolocate up to 77.2% per year to a house number in the address register described below. In our main empirical analysis we examine changes between 1859/60 and 1871 in the six arrondissements for which both lists are available (1st, 2nd, 3rd, 7th, 9th and 12th).

Addresses. To geolocate the addresses recorded in the electoral lists, we construct a reference register of the streets and house numbers of Paris in 1880, the midpoint of our analysis period. We first geolocate streets. The *Atlas administratif des égouts de la ville de Paris* of 1880 covers the city in 16 sheets, on which street names are clearly legible. Starting from the contemporary street network, we verify the names of more than 6,000 streets as of 1880. To identify streets that had been renamed before 1880, we draw on the official list of former street names published by the City of Paris and our own data, in which

[2] In an earlier pipeline, we combined an object-detection model (YOLOv8) to isolate table cells with a handwriting recognition model (Kraken). That pipeline lost more records, because errors compounded across its steps, so we use the model-based transcription throughout (Appendix A.1.3).

most electors recorded on a street in 1859/60 reappear at the same house numbers on a street with a different name in 1871. This pattern suggests that the street was renamed rather than its residents having moved. We then locate house numbers within streets using contemporary addresses, adjusting the numbering of buildings along streets that have since been merged or split (Appendix A.1.2). Together, these steps allow us to geolocate the addresses of residents listed in the electoral lists. To validate the register, we compare it with addresses recorded on cadastral maps from the second half of the nineteenth century, which we manually georeference at the address level. Across 689 addresses, the median location error is 2.3 metres (Appendix Figure A6), providing support for the accuracy of our geolocation approach.

Firms. To measure the location of economic activity, we draw on the annual commercial directories of Paris (*Didot* and *Didot-Bottin*), the equivalent of historical "Yellow Pages". These directories list businesses operating in the city together with a short description of their activity and their address. The directories have been digitised and geolocated by the Social Dynamics in Urban Context (SoDUCo) project, which also assigned each business description to the NAICS industry categories. The number of listed businesses rose from about 81,000 in 1856 to 163,000 in 1896. We use these data to measure the number of firms in each parcel over time and to study which industries responded most strongly to sewer access, distinguishing in particular between customer-facing businesses, such as shops, restaurants and health practitioners, and other sectors.

Typhoid deaths. To examine how sewers affected public health, we use records of deaths from typhoid fever, a waterborne disease that spreads through contamination with faecal matter and is therefore directly linked to the disposal of wastewater. For each year between 1880 and 1904, the *Annuaire statistique de la ville de Paris* published a map of the city on which every death from typhoid is marked by a dot at the address of the deceased (Appendix Figure A15). We digitise all 25 maps by extracting the dots with an automated thresholding routine and correcting the output manually, and count deaths in each parcel. The resulting panel contains almost 23,000 deaths. Annual deaths fell from 1,777 in 1880 to 335 in 1904. Since these maps start only in 1880, we use them to study the later expansion of the sewer network rather than the 1856–1872 period of our main analysis (Appendix A.1.5).

Property transactions. To measure how sewer access was capitalised into property values, we use the *sommier foncier*, the land registry of Paris, which records every transfer of a property together with its date and value to verify tax declarations. Eichholtz et al. (2021) digitised a random sample of this registry, concentrated in the historical centre of the city. We extend their data by digitising 34 further register volumes covering the outer arrondissements between 1859 and 1880, the period of the most substantial expansion of the

sewer network, and geolocate the properties in both datasets with our address register. Since only a fraction of properties change hands in any given period, these data are considerably sparser than our data on residents and firms. We therefore measure property values as average transaction prices at the level of street segments and use them as supporting evidence rather than as a main outcome (Appendix A.1.4).

Transport network. In the class of urban models we employ, households jointly choose where to live and where to work, and the transport network determines the cost of commuting between the two. To measure commuting costs, we digitise the horse-drawn omnibus and tram network of Paris from historical maps for four cross-sections between 1860 and 1895, and the metro network of 1910 for our counterfactual analysis. We trace all networks along the same road network and compute travel times between every pair of parcels as the minimum of walking the whole way and combining walking with a ride on one of these modes of transport, assuming that omnibuses and trams travelled at around twice walking speed, in line with the evidence for 19th-century London in Heblich et al. (2020) (Appendix A.1.6). Since commuting costs enter the model directly and vary over time, our quantitative analysis accounts for the improvements in the transport network over the period, so that the estimated effects of sewers are not confounded with changes in access to jobs.

4 Empirical Evidence

Two mechanisms underlie our theoretical framework in Section 5. First, sewers make the locations they serve more attractive to residents and firms. We hypothesise that, at a very fine spatial scale (see our spatial units in Figure 2, Panel (d)), much of this benefit is local to the treated location, or largely private. As such, this effect does not depend on the number of people living nearby. A sewer running beneath a street drains wastewater and rainwater, reduces local waste pollution and unpleasant smells, and enables households to connect directly to the network. Second, sewerage may generate additional benefits by reducing the negative externalities associated with population density. In the absence of adequate sanitation, more people living nearby generate more waste pollution, imposing costs on surrounding locations. By mitigating these costs, sewers may make densely populated areas more attractive places to live and do business. We refer to these density-dependent spillovers as externality-reducing effect of sewers.

We start our empirical strategy by documenting how the arrival of sewers shaped the locations that received them. We use a difference-in-differences design that compares 25-metre hexagons connected to the sewer network between 1856 and 1872 with hexagons that were connected only later, but before the end of the 19th century. Our main outcomes are

voter and firm counts. We further study the composition behind these effects, comparing firms across industries and voters across social classes and birth location in France. We further examine how the arrival of sewers affected typhoid death incidences in later periods for which these data are available (1880–1904).

Through a second exercise, we estimate spillover effects, distinguishing between the effect of sewer access in a location itself and the effect of sewer access in nearby locations, allowing the latter to vary with surrounding population density. We are particularly interested in estimating to what extent sewers reduce the disutility that people impose on one another in the absence of sanitation. Since this disutility grows with the number of people nearby, the benefit of nearby sewers should be larger in denser neighbourhoods if sewers reduce congestion forces. We test this with a triple-difference design that compares locations where different shares of the surrounding population newly gained sewer access, across areas with different baseline population densities.

To shed light on complementarities between sanitation improvements and other infrastructure investments, we examine how the returns to changes in market access via the omnibus network depend on sewerage access. Locations better connected to the rest of the city are more attractive places to live and to set up a business. Congestion forces, such as higher land rents, floorspace constraints, or pollution, push in the other direction and can partially offset these gains. If sewers alleviate dispersion forces, they should yield differential gains in firm growth from market access increases. We test this hypothesis leveraging our firm-level data, and the joint expansion of the sewerage and omnibus networks.

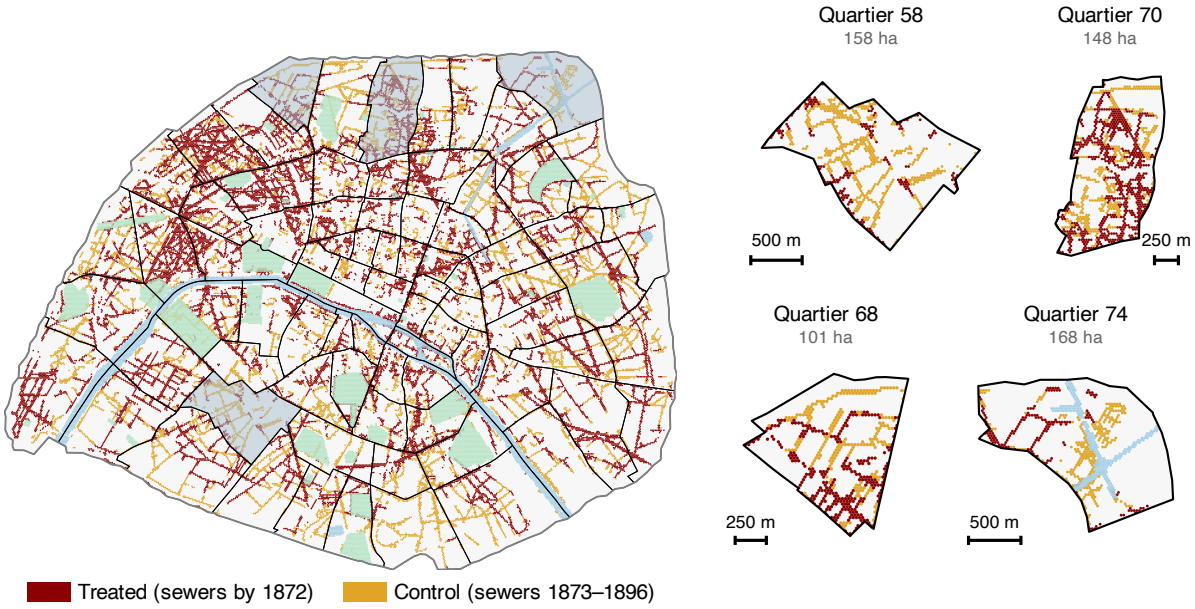
4.1 Effect of sewer access in own location

We start by examining how voter and firm counts responded to the arrival of sewers in the locations that received them. We aggregate our address-level data to the 25-metre hexagon grid illustrated in Figure 2, Panel (d), and restrict the sample to hexagons that had no sewer access in 1856 but were connected by 1896. By this time, the vast majority of streets in Paris had been connected to the network. Figure 3 illustrates the distribution of treatment and control and the scale of the sewerage expansion during this period.

We use an event-study design, which traces the difference between treated and control hexagons year by year. This dynamic specification allows us to investigate whether differential changes between treated and control locations coincide with the timing of sewer network expansion. We estimate:

$$y_{ht} = \delta_h + \delta_{qt} + \sum_{\tau \neq t_0} \beta_\tau \, \text{Treated}_h \times \mathbf{1}[t = \tau] + \varepsilon_{ht} \tag{1}$$

Figure 3: Treatment and control parcels



Notes: The map on the left shows the distribution of treated hexagons (in maroon) and control hexagons (in orange) for the whole of Paris. The maps on the right show treated and control hexagons within four selected quartiers.

where δ_h and δ_{qt} are hexagon and quartier-year fixed effects, Treated_h equals one if the hexagon was connected to the sewer network between 1856 and 1872, and t_0 is the omitted baseline year before the expansion. The coefficients β_τ trace the difference between treated and control hexagons in each year relative to the baseline. We estimate (1) for our key variables of interest, residents and firm counts, as well as for house prices. Our firm data from the business directories are available annually. For residents, the time dimension is more limited. The electoral lists record voters' addresses only from 1852 onwards, and the standardised table formatting used consistently over time was introduced in 1853. We pool the 1859 and 1860 lists of six arrondissements (1st, 2nd, 3rd, 7th, 9th and 12th) as the baseline and add 1871 and 1892 (the same six arrondissements) and 1880 (four arrondissements) as post-periods. As a pre-period, we use the 1853 lists, which we digitised for the 7th and 8th arrondissements. For the 8th arrondissement we do not observe the 1859/60 lists, but it contributes to the estimation through its 1853, 1871 and 1892 lists. Finally, we compute the log of the average property transaction price within ten-year periods at the street-segment level. For voter and firm counts, we estimate (1) by Poisson pseudo-maximum likelihood (PPML), which accommodates the many hexagons with zero counts, so that the coefficients are approximately proportional effects. For house prices, which enter in logs, we use OLS.

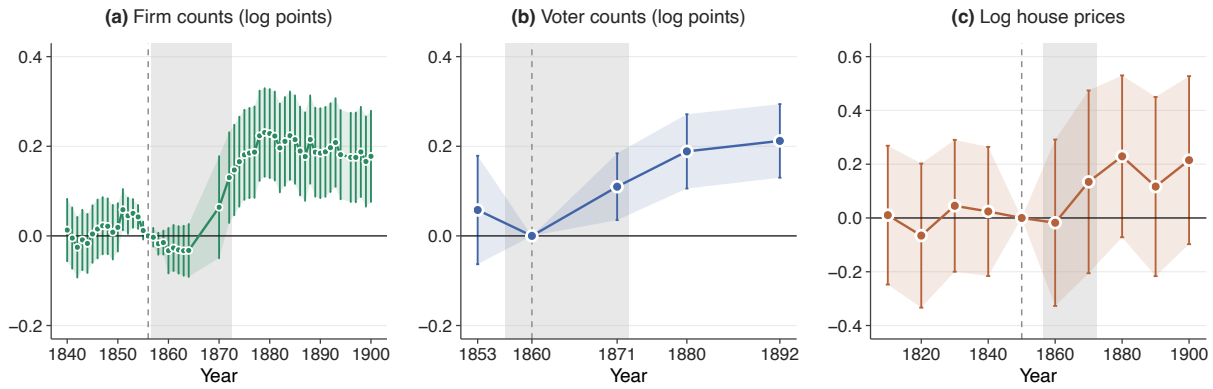
The identifying assumption is that, in the absence of sewers, voter and firm counts in treated and control hexagons within the same quartier, conditional on the included controls, would have evolved in the same way. The main threat to this assumption is that the expansion of the sewer network coincided with the broader Haussmann renovation of Paris at this same level of spatial granularity.

Several features of our setting and design mitigate concerns regarding violations of our identifying assumption. First, the fine spatial granularity of our data allows us to absorb all time-invariant differences between small locations with hexagon fixed effects. Second, quartier-year fixed effects absorb any shock common to a neighbourhood, including the broad transformation of the city under Haussmann. The remaining variation is the timing with which individual hexagons within the same quartier were connected to the network. This timing was governed by the incremental progress of construction, which extended the network street by street, and is unlikely to coincide with other shocks operating at the same fine spatial scale. Third, we directly control for other contemporaneous changes, namely changes in access to public water fountains, parks and railway stations (Appendix Table A10). Access to fountains is balanced between treated and control hexagons, while access to parks and railway stations improved slightly more for control hexagons, which, if anything, works against finding an effect of sewers. Our estimates remain positive and significant when we add these controls.

4.2 Results

Figure 4 reports the estimates for our three outcomes. The first panel shows the estimates for voter counts. The baseline 1853 coefficient is small and statistically insignificant, consistent with the absence of differential pre-trends. Voter counts in treated hexagons then rise steadily. By 1871, they are about 11 log points (12%) higher than in control hexagons, by 1880 about 19 log points (21%), and by 1892 about 21 log points (24%). The effect thus persists even after most control hexagons had been connected to the network by the 1890s. The second panel shows the estimates for firm counts. Before the expansion, the coefficients are small and mostly statistically indistinguishable from zero, with no sign of a systematic differential pre-trend. Firm counts in treated hexagons start to rise during the expansion period (1857–1872). By 1872, they are about 13 log points (14%) higher than in control hexagons, and by 1880 about 23 log points (26%). The effect then remains stable at around 18–20 log points until 1900. Finally, in the third panel, we report estimates for the house price index. Before the expansion, house prices in treated and control segments follow similar trends. During and after the expansion, prices in treated locations rise more strongly. By 1880, the point estimate implies a relative increase of about 25%. The house

Figure 4: Event study results



Notes: The figure reports event study estimates of the effect of sewer access on (a) the number of firms and (b) the number of voters, both measured on a 25 m hexagonal grid; (c) log house prices, measured at the street segment level. The estimating sample is restricted throughout to units that had no sewer access in 1856 but had gained it by 1896. A unit is treated if it had sewer access by 1872. Each panel plots the coefficients on the interaction of the treatment indicator with year dummies, relative to the omitted baseline year marked by the dashed vertical line (1856 for firms, 1860 for voters, 1850 for house prices). Voter counts come from the digitised electoral lists: 1853 (7th and 8th arrondissements), 1859/60 (1st, 2nd, 3rd, 7th, 9th and 12th), 1871 and 1892 (the same six plus the 8th) and 1880 (1st, 3rd, 9th and 12th). All specifications include spatial unit fixed effects and quartier-year fixed effects. Firm and voter counts are estimated by PPML, so coefficients are in log points. House prices are estimated by OLS on log prices, weighted by the number of underlying transactions. Standard errors are clustered at the unit level. The shaded band and the capped vertical bars both show 90% confidence intervals. The grey vertical rectangle marks the 1857–1872 sewer expansion window.

price sample is, however, an order of magnitude smaller than the firm sample, and the confidence intervals are wide. We cannot reject the absence of an effect at conventional levels. Taken at face value, the magnitude is in line with Coury et al. (2024), who find that house prices in 19th-century Chicago roughly doubled once properties were connected to the water and sewer network, and suggests that the increased attractiveness of sewered locations was capitalised into land values.

4.2.1 Static specification and Haussmann controls

We next report results from the following static difference-in-differences specification, where we introduce a set of hexagon-time varying controls that reflect other urban improvements during the Haussmann renovation of the city.

$$y_{ht} = \delta_h + \delta_{qt} + \beta \, \text{Treated}_h \times \text{Post}_t + \mathbf{X}'_{ht}\gamma + \varepsilon_{ht} \tag{2}$$

For voters and firms y_{ht} reflects two outcomes. At the extensive margin, we use a binary variable indicating whether there is at least one voter (firm) in a hexagon h in a given year t. At the intensive margin, we use the number of voters (firms) in a hexagon in a given year. Treated_h equals one if the hexagon was connected to the sewer network between 1856 and

Table 1: Difference-in-differences estimates for the overall effect of sewers

	Voters				Firms				House prices	
	Presence		Count		Presence		Count		Log price	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Post	0.040***	0.038***	0.099**	0.080*	0.040***	0.036***	0.163***	0.159***	0.128	0.104
	(0.010)	(0.010)	(0.046)	(0.047)	(0.005)	(0.005)	(0.052)	(0.052)	(0.236)	(0.204)
Observations	13,434	13,434	6,492	6,492	71,816	71,816	22,884	22,884	192	192
Control mean (pre)	0.323	0.323	30.461	30.461	0.145	0.145	0.886	0.886	10.846	10.846
Estimator	OLS	OLS	PPML	PPML	OLS	OLS	PPML	PPML	OLS	OLS
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Estimates of specification (2). The sample is restricted to units that had no sewer access in 1856 but were connected by 1896; treated units were connected by 1872. Columns (1)–(4) compare voters in 1871 with the 1859/60 baseline (1st, 2nd, 3rd, 7th, 9th and 12th arrondissements), using voter counts from the digitised electoral lists; columns (5)–(8) compare firms in 1872 with 1861. *Presence* is an indicator for a hexagon having at least one voter (firm), estimated by OLS. *Count* is the number of voters (firms) per hexagon, estimated by PPML. Columns (9)–(10) report OLS estimates for log house prices at the street-segment level. Even-numbered columns add the Haussmann controls, each interacted with Post: market access (MA) to public water fountains, and the change in market access to railway stations and to parks. All specifications include hexagon (street segment for house prices) and quartier-year fixed effects. Standard errors, clustered at the hexagon (street-segment) level, are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

1872 and zero if it was connected between 1873 and 1896. Post_t indicates the post-expansion period. We compare 1872 to 1861 for firms and 1871 to 1859/60 for voters. δ_h are hexagon fixed effects and δ_{qt} quartier-year fixed effects. $\mathbf{X}_{ht}$ is a vector of hexagon-level controls for other elements of the Haussmann renovation, namely access to public water fountains, parks and railway stations. We introduce time-invariant controls interacted with Post_t. Table 1 reports the results.

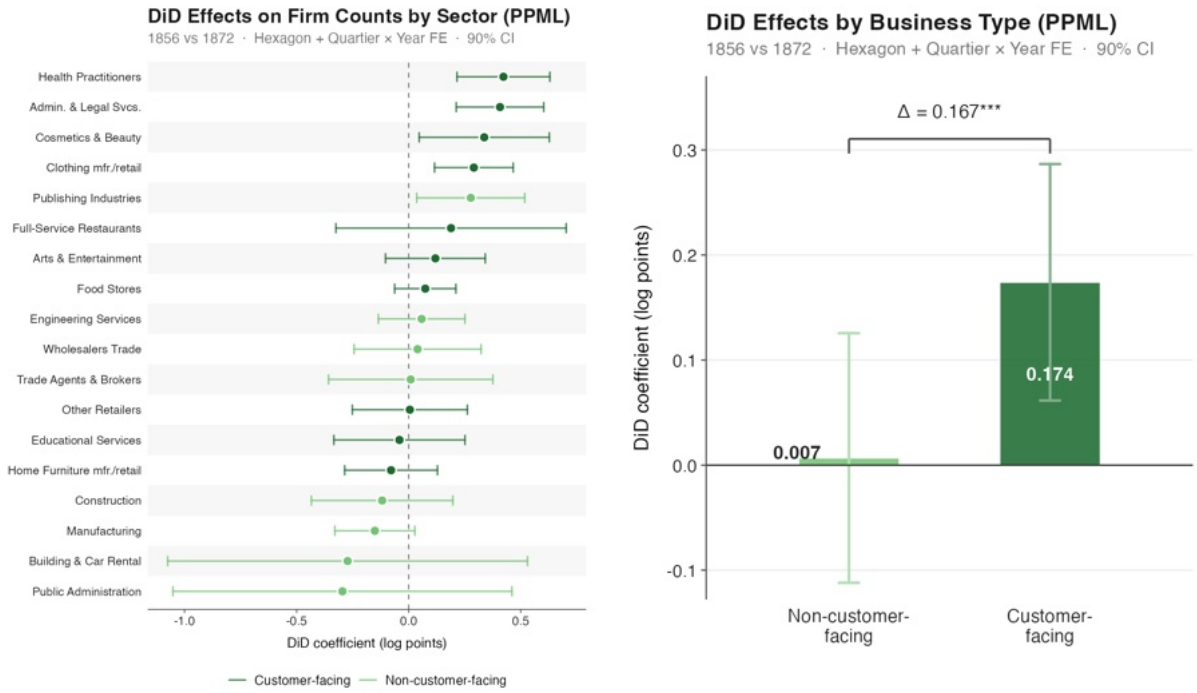
Without the Haussmann controls, sewer access increased economic activity along both margins. Treated hexagons became 4.0 percentage points more likely to host any voters and 4.0 percentage points more likely to host any firm (columns 1 and 5), increases of about 12% and 28% relative to the control-group means. The number of voters and firms rose by about 10% and 18%, respectively (columns 3 and 7). House prices in treated street segments increased by about 14%, but the house price sample is small and the estimate is imprecise (column 9). Adding the Haussmann controls shifts our estimates only slightly. The presence estimates fall to 3.8 and 3.6 percentage points and the count estimates to 8% for voters and 17% for firms, although the voter count estimate is now only significant at the 10% level.

4.2.2 Compositional changes: Industry, social class, migration status

In this section we evaluate whether the response to sewers was different for certain types of firms or residents, exploring whether sewers induced shifts in the firm composition of a

Figure 5: Sectoral heterogeneity

(a) *Treatment effects by sector* **(b)** *Treatment effects by type*



Notes: PPML difference-in-differences estimates of the effect of sewer access on firm counts, comparing 1872 with 1856, with hexagon and quartier-year fixed effects. The sample is restricted to hexagons that had no sewer access in 1856 but were connected by 1896. Panel A reports separate estimates for each sector. Dark green sectors are customer-facing businesses; light green sectors are non-customer-facing. Panel B pools firms into customer-facing and non-customer-facing businesses and estimates both effects jointly, with group-specific fixed effects; Δ is the difference between the two. Coefficients are in log points. Standard errors are clustered at the hexagon level; bars show 90% confidence intervals.

location and whether they generated gentrification. To examine sectoral heterogeneity in firm responses to sewer expansion, we estimate specification (2) by PPML separately for each of the NAICS industry categories observed in our dataset.

Sectoral heterogeneity. Figure 5, Panel A reports the sector-specific estimates. Estimates highlighted in dark green are for customer-facing industries, such as health practitioners, administrative and legal services, clothing, cosmetics, food stores and restaurants. Firms in these sectors tend to locate close to their potential customers, as they interact with them daily. In light green we highlight the estimates for the non-customer-facing sectors, such as engineering services, publishing, construction and manufacturing. Panel A shows that the strongest proportional responses occur among customer-facing businesses. Health practitioners, administrative and legal services, cosmetics and clothing all increase by about 30–40 log points. Manufacturing and most other non-customer-facing sectors, by contrast, do not appear to sort into locations with improved sanitary conditions. Panel B formally

Table 2: Sewers and the composition of voters

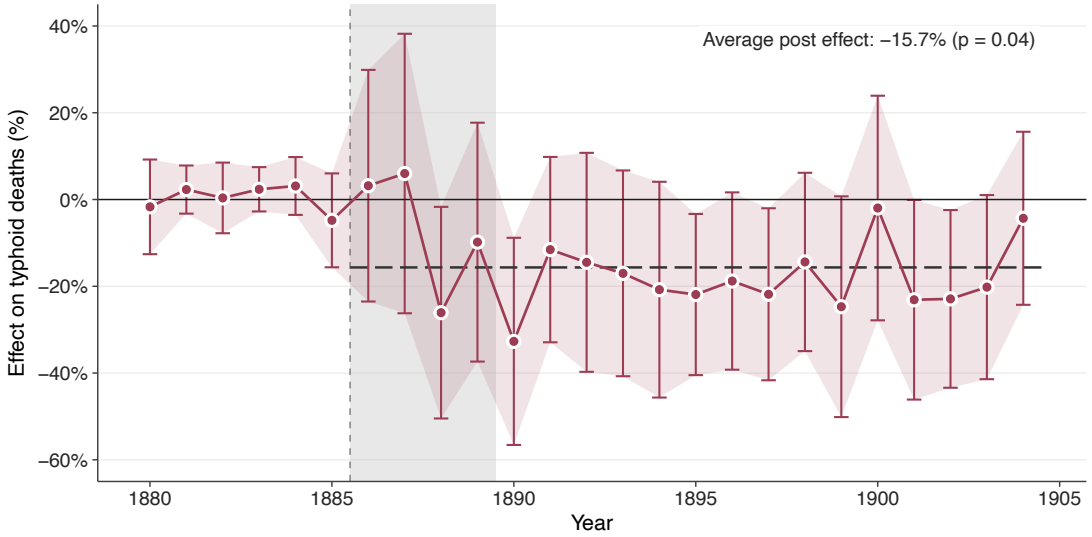
		Share in HISCLASS-5 class				Migration
	HISCAM	Elite	Lower middle	Skilled	Unskilled	Born outside Paris
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	-0.326	-0.003	0.015	-0.014	-0.000	0.006
	(0.427)	(0.013)	(0.017)	(0.016)	(0.008)	(0.015)
Control mean (pre)	62.8	0.177	0.378	0.375	0.066	0.744
Observations	1,997	2,045	2,045	2,045	2,045	2,014
Hexagon FE	Yes	Yes	Yes	Yes	Yes	Yes
Quartier-year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Difference-in-differences estimates of specification (2), comparing 1871 with the 1859/60 baseline (1st, 2nd, 3rd, 7th, 9th and 12th arrondissements). The outcome in column (1) is the mean HISCAM occupational status score of voters in a hexagon; in columns (2)–(5) the share of voters in each HISCLASS-5 social class; in column (6) the share of voters born outside Paris. Shares require at least one voter with a classified occupation (birthplace) in both periods, which restricts the sample to around 2,000 hexagons. HISCAM is undefined for voters without a stated occupation, such as rentiers. Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

tests for the heterogeneity of treatment effects across these two business types. Customer-facing businesses increased by about 17 log points in treated hexagons, while the number of non-customer-facing businesses did not change, and the difference between the two is statistically significant. The results point to heterogeneous responses to sanitation improvements across sectors and suggest that customer-facing services in particular benefit from locating in urban environments with improved public health conditions.

Social class. We next turn to analyse whether sewers created compositional changes on the household side. One natural hypothesis is that richer households disproportionally sorted into the newly sewered neighbourhoods. If richer households also consume more floorspace and are able to pay higher prices, this would be important to capture. In the most extreme case, even the absence of an aggregate population response to sewers could hide a gentrification-driven sizeable effect, with richer households replacing poorer ones and paying higher rents. Such sorting would also be an important dimension of heterogeneity to account for in our quantitative model. To test for it, we use the occupations recorded in the electoral lists to classify voters into the five HISCLASS social classes and to assign them an occupational status score (HISCAM). Table 2 shows no such effects. Neither the average occupational status of voters (column 1) nor the shares of voters in the different social classes (columns 2–5) changed differentially in treated hexagons.

Figure 6: Typhoid deaths



Notes: Synthetic difference-in-differences event study estimates for deaths per hexagon, expressed as a percentage of the pre-period control mean. The grey rectangle marks the 1886–1889 treatment window; the dashed vertical line marks the 1885 reference year. N = 5,717 treated and 16,087 control hexagons. The shaded band and the capped vertical bars show 90% confidence intervals from a unit-level block bootstrap (200 replications).

Migration status. Another interesting dimension is whether the increase in voters is driven by migrants moving to Paris or by people already living in the city. A disproportionate role of migrants would suggest that sewered areas grew rapidly as new neighbourhoods absorbing inflows from outside the city. Using information on voters' place of birth, we find no change in the share of voters born outside Paris (column 6). This suggests that the population increase we document is driven by people born in Paris and by those who moved to Paris alike.

4.2.3 Typhoid deaths

We close our reduced-form analysis of local effects by examining whether sewers improved public health, the channel through which we argue they made locations more attractive. For this, we use the annual maps of typhoid deaths published by the city of Paris. Since these maps start only in 1880, we cannot study the 1856–1872 expansion wave. Instead, we exploit a later wave of the network expansion. Among hexagons on the street network without sewer access in 1885, we compare those connected between 1886 and 1889 with those not yet connected by 1889. Treated and control hexagons in this later wave are less comparable than in our main design. Furthermore, our outcome variable is significantly sparser. For these reasons we use a synthetic difference-in-differences estimator following Arkhangelsky et al. (2021), which allows us to give higher weights to more comparable control hexagons, evolving in parallel to treated hexagons in terms of typhoid death counts.

Figure 6 reports the event study estimates for typhoid deaths. Before 1886, typhoid deaths in treated hexagons closely track their synthetic control. After sewer arrival, they fall below it, and remain lower throughout the 1890s. Averaged over the post-treatment period, typhoid deaths decline by about 15.7% of the control-group mean, with a p-value of 0.04. This result supports our conclusion that the effects on economic outcomes outlined above are driven by improved sanitary conditions arising from sewerage arrival.

4.3 Effect of nearby sanitation improvements

The evidence presented so far shows that small parcels connected to the sewer network during the 1856–1872 expansion evolved differently from parcels that were connected only later. There are, however, reasons to expect the effects of sewerage investments to extend beyond the locations in which they were built. Sewers remove wastewater from the streets and reduce ambient pollution, such as unpleasant odours. They also reduce the likelihood of disease, which travels with people. People are in fact at the core of these spillover effects. A given share of neighbours gaining access to sewers means something different in sparsely and in densely populated areas. If half of the residents of a neighbourhood gain access to sewers, the resulting reduction in local pollution is likely to be much larger in a neighbourhood of several thousand people than in one of a few hundred. If sewers relieve a congestion force of this kind, the benefit of nearby sewers should therefore increase with density.

We test this with a triple-difference design. For each hexagon h, NearbySewers_h is the share of addresses within 250 metres (own hexagon excluded) that were newly connected to the sewer network between 1856 and 1872, and Density_h is the number of addresses per hectare within the same 250-metre radius. We use the addresses listed in the Vasserot during the 1840s. Figure 7 illustrates both measures for one hexagon. We estimate the following specification:

$$\begin{aligned} y_{ht} = {} & \beta_1 \,\text{NearbySewers}_h \times \text{Density}_h \times \text{Post}_t + \beta_2 \,\text{NearbySewers}_h \times \text{Post}_t + \beta_3 \,\text{Density}_h \times \text{Post}_t \\ & + \rho \,\text{OwnSewer}_h \times \text{Post}_t + X'_{ht}\gamma + \delta_h + \delta_{qt} + \varepsilon_{ht} \end{aligned} \tag{3}$$

where β_1 is our coefficient of interest. It measures how the spillover from nearby sewers varies with density. A positive β_1 means that the same share of neighbours gaining access to sewers makes a location more attractive in a dense neighbourhood than in a sparse one, consistent with sewers relieving a dispersion force. β_3 allows denser neighbourhoods to evolve differently from sparser ones irrespective of sewers. OwnSewer_h indicates whether the hexagon itself was connected between 1856 and 1872, so that ρ is the direct effect of a hexagon's own sewer, and β_1 and β_2 capture the effect of nearby sewers over and above a hexagon's own connection. The controls X_{ht} include the share of addresses within 250 metres already sewered by 1856, interacted with Post_t and with $\text{Density}_h \times \text{Post}_t$, and an indicator for the hexagon itself being

Figure 7: Nearby sewers and density



Notes: The map shows the 250 m neighbourhood (dashed circle) of one 25 m hexagon (blue outline in the centre). Each dot is an address. Red dots are addresses newly connected to the sewer network between 1856 and 1872; grey dots are all other addresses. NearbySewers_h is the share of red dots among all addresses in the circle, excluding the hexagon itself. Density_h is the (log) number of addresses per hectare in the circle.

sewered by 1856, interacted with Post_t. X_{ht} also includes the controls for contemporaneous changes of the Haussmann renovation used in Section 4, each interacted with Post_t. δ_h and δ_{qt} are hexagon and quartier-year fixed effects. For residents, we compare 1871 to 1859/60; for firms, 1872 to 1856.

The variation we exploit comes from the timing of the network expansion. All neighbouring hexagons that enter our measure were eventually connected to the sewer network by 1896. What differs across locations is how many of their neighbours were connected during the 1856–1872 expansion rather than before or after it. We add hexagon FE and quartier-year fixed effects as in our main specification. Our identifying assumption is that, conditional on the fixed effects (and the rest of the controls), the timing with which a location's surroundings were connected is unrelated to other changes that affected dense and sparse neighbourhoods differently. Put differently, the specification allows locations whose neighbours were connected early to evolve differently from those whose neighbours were connected later (β_2), and denser neighbourhoods to evolve differently from sparser ones (β_3). What we assume is that, absent the sewer expansion, the difference between locations with many and with few neighbours connected in 1856–1872 would have evolved in the same way in dense and in sparse neighbourhoods. Because we include quartier-year fixed effects, this comparison is made within quartiers, where the timing of connection was governed by the

Table 3: Effects of own sewer and nearby sewers

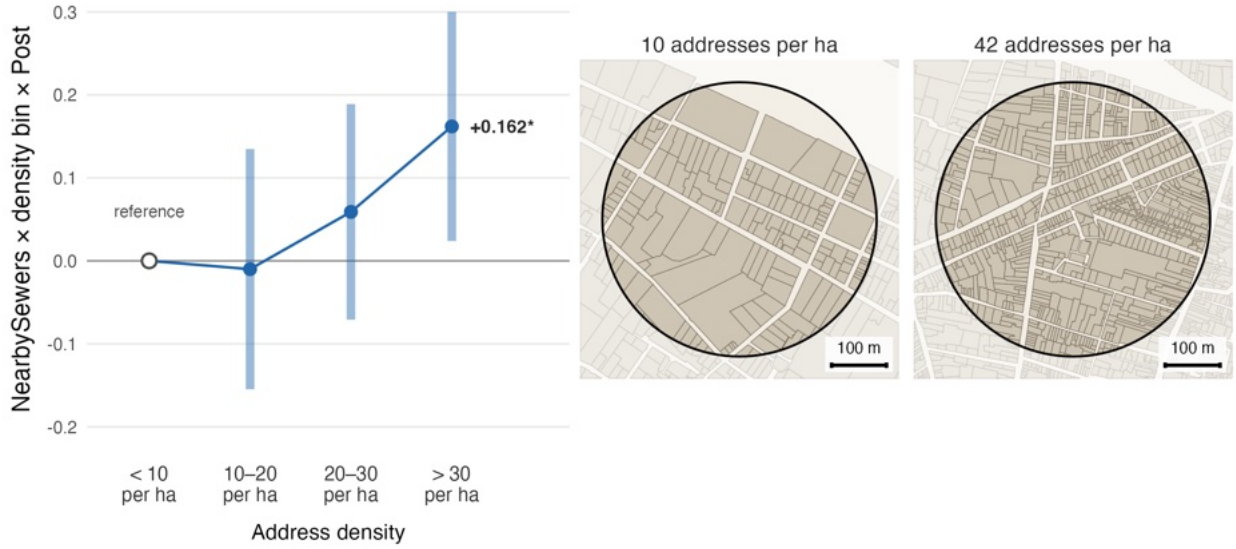
	Residents		Firms	
	Main	Placebo	Main	Placebo
	1859/60–1871	1853–1859/60	1861–1872	1850–1856
	(1)	(2)	(3)	(4)
NearbySewers × Density × Post (β_1)	0.074**	−0.034	0.015	−0.011
	(0.030)	(0.065)	(0.027)	(0.028)
NearbySewers × Post (β_2)	−0.034	−0.013	0.018	−0.009
	(0.045)	(0.069)	(0.035)	(0.028)
OwnSewer × Post (ρ)	0.054	−0.111	0.132*	0.004
	(0.063)	(0.079)	(0.080)	(0.050)
Density × Post	Yes	Yes	Yes	Yes
Already-sewered controls × Post	Yes	Yes	Yes	Yes
Haussmann controls × Post	Yes	Yes	Yes	Yes
Hexagon FE	Yes	Yes	Yes	Yes
Quartier × year FE	Yes	Yes	Yes	Yes
Hexagons	5,745	1,173	12,290	8,360
Mean outcome, pre	11.4	7.3	5.9	4.4

Notes: PPML estimates of specification (3). Residents are voter counts from the digitised electoral lists. Columns (1) and (3) compare 1871 with 1859/60 for residents and 1872 with 1861 for firms. Columns (2) and (4) are placebo tests on the pre-expansion periods 1853 to 1859/60 (residents in quartiers 4 and 25–28) and 1850 to 1856. NearbySewers is in units of 10 percentage points and Density is centred at its sample mean. Already-sewered controls are the share of addresses within 250 m already sewered by 1856, its interaction with Density, and an indicator for the hexagon itself being sewered by 1856. Haussmann controls are as in Table 1. The samples exclude a 250 m band at the edge of the area covered by the Vasserot addresses. Conley standard errors with a 250 m cutoff are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

incremental, street-by-street progress of construction. The assumption is therefore weaker than a standard parallel trends assumption. It is only violated by shocks that coincided with the 1856–1872 expansion and specifically affected locations whose dense surroundings were connected during it. Furthermore, our measure of density precedes the expansion by more than a decade. Table 3 reports the estimates.

For residents, the effect of nearby sewers rises with density. The coefficient β_1 in column (1) is 0.074 and significant at the 5% level. A 10 percentage point increase in the share of nearby addresses newly connected raises the number of residents by about 7 log points more for each additional log point of density. At average density, nearby sewers have no detectable effect ($\beta_2 = -0.034$, s.e. 0.045). By contrast, at the 90th percentile of density, a 10 percentage point increase in the share of newly connected nearby addresses raises residents by about 0.069 log points. The placebo in column (2) shows no such pattern before the expansion. The

Figure 8: Effect of nearby sewers on residents, by density

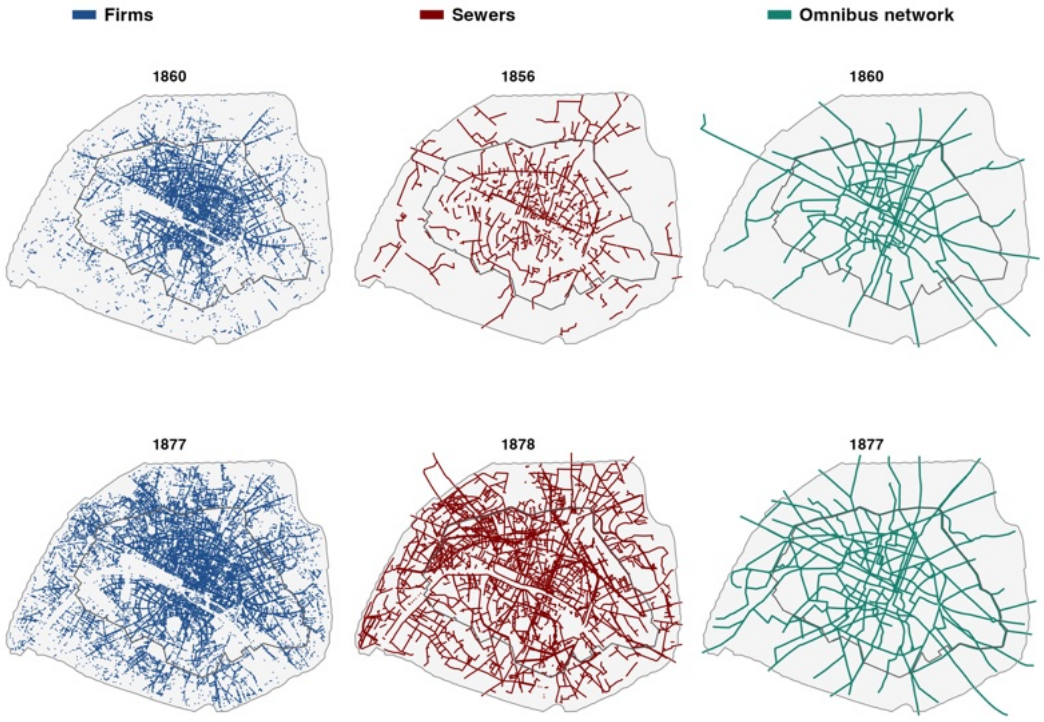


Notes: The left panel reports PPML estimates of a version of specification (3) for residents (voter counts, 1859/60 to 1871) in which log density is replaced by four bins of address density within 250 m (below 10, 10–20, 20–30 and above 30 addresses per hectare). NearbySewers × Post and the share of nearby addresses already sewered by 1856 × Post enter separately for each bin, together with bin × Post, the own-sewer indicators × Post, the Haussmann controls × Post, hexagon fixed effects and quartier-year fixed effects. Each point is the effect of NearbySewers (in units of 10 percentage points) in that bin minus its effect in the sparsest bin. Outcomes are winsorised at the 99th percentile in each year. Bars show 90% confidence intervals based on Conley standard errors with a 250 m cutoff. $^*p < 0.10$. The maps on the right show the 250 m neighbourhood of two hexagons in the sample, one in a sparse bin (10 addresses per hectare) and one in the densest bin (42 addresses per hectare), with the parcels of the Vasserot cadastre.

coefficient on the triple interaction is small, negative and imprecise (−0.034, s.e. 0.065). For firms, the interaction with density is close to zero in both periods (columns 3 and 4). Firms instead respond to sewer access in their own location. A hexagon's own connection raises firm counts by 0.132 log points, significant at the 10% level, while the placebo estimate is zero (0.004). For residents, the own-sewer effect is positive but imprecise once nearby sewers are controlled for.

The linear interaction masks where the effect comes from. In Figure 8 we replace log density with four bins of address density within 250 m and estimate the effect of nearby sewers separately in each bin (see also Appendix Table A11). The effect is flat across the three sparser bins and appears only in the densest neighbourhoods, above 30 addresses per hectare. Relative to the sparsest bin, the effect of nearby sewers on residents is −0.010 for 10 to 20 addresses per hectare and 0.059 for 20 to 30. It rises to 0.162 above 30, significant at the 10% level. Compared with the average of the three sparser bins, the effect in the densest bin is 0.146 log points higher (s.e. 0.066). The maps in Figure 8 show what these densities look like on the ground. Nearby sewers raise the attractiveness of a location only where

Figure 9: Firms, sewers, and the omnibus network over time



Notes: Firms: 25 m hexagons with at least one geocoded firm in the business directories, 1860 and 1877. Sewers: digitised sewer network, 1856 and 1878 (the digitised snapshots closest to the omnibus years). Omnibus network: digitised omnibus lines, 1860 and 1877. Dark grey line: boundary of the pre-1860 city; light grey line: boundary of Paris after the 1860 annexation.

people live packed closely together. This is what we expect if sewers relieve a congestion force that rises with the number of people nearby.

4.4 Sewers and market access

In this section, we examine whether access to sanitation affects the impact of market access on firm growth. Locations better connected to the rest of the city are more attractive places to live and to set up a business. Congestion forces, such as higher land rents, floorspace constraints, or pollution, push in the other direction and can partially offset these gains. If sewers alleviate congestion, they should yield differential gains in firm growth from market access increases. We test this using the joint expansion of the omnibus and sewer networks in the annexed communes after 1860. The omnibus expansion reduced travel times and raised market access, while the sewer network potentially helped these gains materialise into increases in firm activity. In Figure 9, we overlay the distribution of firm counts, the sewer and omnibus networks at the start and end of the main sewerage expansion wave.

Over the period of the main wave of sewerage expansion (1856–1872), public transport in Paris was primarily reliant on the horse-drawn omnibus. Figure A19 shows the number of passengers carried each year by every public transport service, from the statistics compiled in the *Étude historique et statistique sur les moyens de transport dans Paris*, from 1894. Between 1856 and 1872 the omnibus carried between three quarters and five sixths of all passengers (82% in 1860, 76% in 1872), far ahead of cabs, the Ceinture railway, or any other form of transport. It remained the largest single mode until 1877, when the new tramways started taking over. We are therefore capturing the main transport-related determinant of a location's value.

To measure how the omnibus expansion changed each location's access to the rest of the city, we construct market access using an exponential decay in travel time (Ahlfeldt et al. 2015; Heblich et al. 2020). The access of 25 m hexagon h in year t is

$$\mathrm{MA}_{ht} = \sum_{g \neq h} L_g \, e^{-\theta \, \tau_{hg,t}}, \tag{4}$$

where L_g is the number of firms in hexagon g in 1860, held fixed so that access changes only through the transport network, and $\tau_{hg,t}$ is the travel time in minutes on the fastest route combining walking (3 mph) with the omnibus (6 mph, 3 minutes to board) on the 1860 or 1877 network. As in Heblich et al. (2020), we build travel times from the historical transport network itself, and we take the speed of the horse-drawn omnibus from their evidence for London. We set the decay rate to $\theta = 0.07$ per minute. A destination 30 minutes away then receives about 12% of the weight of one next door. Appendix A.2.3 provides further context and descriptives on our market access measure.

4.4.1 Results

We compare the growth in firm activity between 1856 and 1877 across two groups of hexagons in the annexed communes. The first group was newly sewered by 1872. The second was not yet sewered in 1872 but was sewered by 1896. Restricting the sample to hexagons that eventually received a sewer keeps the two groups comparable in the type of street they are on. We then ask whether the growth advantage of the sewered hexagons depends on how well connected they are through the omnibus network. We estimate the following specification

$$y_{ht} = \alpha_h + \alpha_{q(h)t} + \beta \, \mathrm{Access}_{ht} + \gamma \, \mathrm{Sewer}_{ht} + \delta \, \mathrm{Access}_{ht} \times \mathrm{Sewer}_{ht} + \lambda \, \mathrm{Access}_{h,1860} \times \mathrm{Post}_t + \varepsilon_{ht} \tag{5}$$

where y_{ht} is either the inverse hyperbolic sine of the number of firms in hexagon h in year t or an indicator for the presence of at least one firm, α_h are hexagon fixed effects and $\alpha_{q(h)t}$ quartier $\times$ year fixed effects. Sewer_{ht} indicates that the hexagon has a sewer. It is zero for

Table 4: Annexed communes: omnibus access, sewers, and firm growth

	All firms		By sector (market access)		
	Market access	Travel time (min. saved)	Customer-facing	Manu-facturing	Other business
	(1)	(2)	(3)	(4)	(5)
Panel A: ihs(firms)					
Access × Sewer	0.173***	0.012***	0.147***	0.011	0.033***
	(0.039)	(0.003)	(0.030)	(0.009)	(0.012)
Access	0.055	0.002	0.036	0.037***	0.016
	(0.068)	(0.004)	(0.053)	(0.014)	(0.018)
Sewer	0.073***	0.072***	0.049***	0.004	0.017***
	(0.016)	(0.016)	(0.012)	(0.004)	(0.005)
Mean outcome, 1877	0.375	0.375	0.240	0.044	0.077
Panel B: firm presence					
Access × Sewer	0.079***	0.006***	0.090***	0.013*	0.030***
	(0.022)	(0.001)	(0.019)	(0.008)	(0.010)
Access	0.039	0.002	0.032	0.030**	0.006
	(0.040)	(0.002)	(0.036)	(0.012)	(0.016)
Sewer	0.044***	0.043***	0.033***	0.004	0.016***
	(0.010)	(0.010)	(0.008)	(0.003)	(0.004)
Mean outcome, 1877	0.251	0.251	0.176	0.041	0.068
Baseline access × Post	Yes	Yes	Yes	Yes	Yes
Already-sewered share 1856 × Post	Yes	Yes	Yes	Yes	Yes
Hexagon, quartier × year FE	Yes	Yes	Yes	Yes	Yes
Observations	38,924	38,924	38,924	38,924	38,924

Notes: OLS estimates of specification (5). The sample is 25 m hexagons in the annexed communes that were not sewered in 1856 and were sewered by 1896, observed in 1856 and 1877. The dependent variable is ihs(firms) in Panel A and an indicator for at least one firm in Panel B. Sewer indicates that the hexagon is sewered by 1872. Access is log market access in columns (1) and (3)–(5) and minus the average travel time in minutes in column (2), both centred at their 1860 mean. Columns (3)–(5) split firms into customer-facing (retail, food, restaurants and personal services), manufacturing and other business. All regressions include hexagon and quartier × year fixed effects, baseline access × Post and the share of the neighbourhood already sewered in 1856 × Post. Results are robust to starting the sample in 1860. Standard errors clustered by 250 m grid cell in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

every hexagon in 1856. Access_{ht} is one of two measures of omnibus access, log market access or minus the average travel time to the city's firms, centred at its 1860 mean. We control for baseline access interacted with Post_t, so that δ is not driven by well-connected places being on a different growth path, and for the share of the neighbourhood already sewered in 1856 interacted with Post_t. The coefficient of interest is δ. It measures whether hexagons with higher access grow more when sewered, relative to not-yet-sewered ones. Because access is centred, γ is the effect of a sewer at average access, and β captures how growth varies with access among hexagons without a sewer. Standard errors are clustered by 250 m grid cell, each containing about 22 hexagons, to account for spatial correlation of treatment between neighbouring hexagons.

Table 4 reports the results. Sewered hexagons grow more than not-yet-sewered ones. At average access, a sewer raises ihs(firms) by 0.07 (columns 1 and 2). It raises the probability that a hexagon has at least one firm by 4 percentage points, against a mean of 25% in 1877. This advantage increases with access. The coefficient on Access × Sewer is 0.173 for market access and 0.012 per minute of travel time saved, both significant at the 1% level. A hexagon with market access one standard deviation above the mean (0.5 log points) has a sewer effect of about 0.16. This is more than twice the effect at the mean. One standard deviation of travel time (7.4 minutes) gives a similar figure. For firm presence the pattern is the same (0.079 and 0.006). The coefficient on access itself (for unsewered locations) is positive but smaller and statistically insignificant for both measures. Among hexagons without a sewer, better access is not clearly associated with more firm growth. Better-connected locations grow faster once they are sewered. This follows our prediction that congestion holds back growth in accessible but unsanitary places.

Columns 3 to 5 split firms by sector, using market access. The interaction is driven by customer-facing firms such as shops, restaurants and food stores (0.147 for ihs and 0.090 for presence). It is smaller for other business (0.033 and 0.030) and close to zero for manufacturing (0.011 and 0.013). Manufacturing is the one sector whose growth rises with access regardless of sewers. The sectors that respond depend on people being willing to live in and visit a location. This again points to sewers improving local amenities and reducing congestion. The main results are robust to starting the sample in 1860 instead of 1856.

Because no hexagon in the sample is sewered at the start, the Access × Sewer term in Table 4 compares sewered and not-yet-sewered hexagons by their 1877 level of access, which is the sum of their 1860 access and the gain from the omnibus expansion. Table A14 separates the two parts. Per unit of access, the effects are of similar size. With market access, the coefficient is 0.11 for the gain and 0.17 for the 1860 level. With travel time, they are 0.010 and 0.012 per minute. In no column can we reject that the two are equal (p-values between 0.25 and 0.91). Sewers pay off more where access is higher, whether that access was already there in 1860 or came with the omnibus expansion. The term on the gain is less precisely estimated because access gains vary much less across hexagons than baseline access. Figure A22 shows the same data non-parametrically. Hexagons sewered by 1872 grow more than not-yet-sewered ones at almost every level of the access gain, and the not-yet-sewered hexagons catch up only where the gains were largest. Figure A23 groups hexagons by their 1860 access instead. In the lowest-access decile, sewered and not-yet-sewered hexagons grow alike, and the gap widens steadily with baseline access, to 0.26 in ihs(firms) and 0.10 in firm presence in the top decile.

5 Quantitative model

To quantify the effects of sewers, we develop an urban model that builds on the contributions of Ahlfeldt et al. (2015) and Heblich et al. (2020). The city consists of a discrete set of locations $\mathbb{N}$. Households maximise utility by jointly choosing residential and workplace locations. Firms produce a homogeneous final good in each location, which is costlessly traded across the city, and maximise profits using labour and floorspace as inputs. Sewers affect residential amenities and location-specific firm productivity through two channels. First, they may generate private benefits for residents and firms in locations with sewer access. Second, they may reduce the local externalities generated by surrounding residents and workers through pollution, externalities that increase with local density. We further allow for endogenous entry, to quantify the extent to which sewers attract migrants to the city. Households have the option of living elsewhere in the country and the city's total population responds endogenously to changes in the city's attractiveness.

5.1 Households

Households receive utility from consumption goods (c_{ij}) and residential floorspace (h_{ij}). Their wages w_j and rents q_i are determined by where they live i and work j. We assume that the utility function takes a Cobb–Douglas form so that the utility maximisation problem of household o is

$$\begin{aligned} \max_{ij,c_{ij},h_{ij}} \quad & U_{ijo} = \frac{B_i}{d_{ij}} \left(\frac{c_{ij}}{\beta} \right)^{\beta} \left(\frac{h_{ij}}{1-\beta} \right)^{1-\beta} z_{ijo} \\ \text{s.t.} \quad & w_j = P_i c_{ij} + q_i h_{ij} \end{aligned} \tag{6}$$

where B_i is the residential amenity value of location i. Travelling between the residential and workplace locations is costly and commuting costs increase with travel time,

$$d_{ij} = e^{\kappa \tau_{ij}} \tag{7}$$

where τ_{ij} is the travel time in minutes and κ governs how quickly commuting costs rise. Households have an idiosyncratic utility component z_{ijo} following an i.i.d. Fréchet distribution

$$F(z_{ijo}) = \exp\left[-z_{ijo}^{-\epsilon} \right] \tag{8}$$

Utility maximisation leads to the following indirect utility function for households

$$u_{ijo} = \frac{B_i w_j z_{ijo}}{d_{ij} P_i^{\beta} q_i^{1-\beta}} \tag{9}$$

Amenity effect of sewers. We model the effect of sewers on residential amenities through two channels, a private benefit of sewer access and a reduction in the negative

externalities that nearby residents and workers without sewer access impose on a location in the absence of sanitation.

$$B_i = \bar{b}_i \, e^{\eta_1 \, \text{Sewer}_i} \, e^{-\eta_2 \, (1 - \text{ShareSewered}_i) \, \text{Density}_i^{\gamma_B}} \left(\Omega_i^R\right)^{\eta_3} \tag{10}$$

Following the standard parameterisation in the literature, $\bar{b}_i$ denotes the fundamental amenity value of location i, capturing geographical features as well as the availability of infrastructure such as street lighting. Sewer_i indicates whether location i has access to a sewer, capturing the private benefits of sewer access. These include easier wastewater disposal and improved sanitary conditions in the immediate surroundings of buildings. The parameter η_1 therefore captures the increase in amenities associated with sewer access.

The negative externalities are captured by the second exponential term. Density_i measures the number of residents and workers within 250 metres of location i, while ShareSewered_i measures the share of these residents and workers located in places that themselves have access to sewers. The parameter η_2 captures the negative effect of having more people in nearby unsewered locations, conditional on overall local density. The parameter γ_B determines how this effect varies with density. Overall, the expression captures the costs of local population density arising from the increasing volume of polluting wastewater generated in the surrounding area, with the magnitude of these costs depending on the extent of sewer access.

Finally, we allow for positive residential externalities arising from proximity to other residents. Following the standard parameterisation in the literature, we define

$$\Omega_i^R = \sum_k e^{-\rho \, \tau_{ik}} R_k, \tag{11}$$

where R_k denotes the number of residents in location k, weighted by the walking time τ_{ik} between i and k, with decay parameter ρ. The parameter η_3 is the elasticity of residential amenities with respect to this measure of nearby residential population. We expect η_3 to be positive, reflecting the benefits of proximity to other residents, for example through a more vibrant social life.

Choices. The properties of the Fréchet distribution give rise to three results regarding the choices made by utility-maximising households. First, the probability that a household chooses to live in location i and work in location j is given by

$$\pi_{ij} = \frac{\left(\frac{B_i w_j}{d_{ij} P_i^{\beta} q_i^{1-\beta}}\right)^{\epsilon}}{\sum_k \sum_l \left(\frac{B_k w_l}{d_{kl} P_k^{\beta} q_k^{1-\beta}}\right)^{\epsilon}}. \tag{12}$$

Second, conditional on living in location i, the probability of working in location j is given by

$$\pi_{ij|i} = \frac{\left(\frac{w_j}{d_{ij}}\right)^{\epsilon}}{\sum_s \left(\frac{w_s}{d_{is}}\right)^{\epsilon}}. \tag{13}$$

Intuitively, the higher the wage offered in a workplace and the closer it is to a worker's residential location, the more attractive it becomes as a place of work.

Third, the expected utility of living in Paris, before the idiosyncratic preference draw z_{ijo} is realised, is given by

$$u_P = \Gamma\left(1 - \frac{1}{\epsilon}\right)\left[\sum_k \sum_l \left(\frac{B_k w_l}{d_{kl} P_k^{\beta} q_k^{1-\beta}}\right)^{\epsilon}\right]^{1/\epsilon}, \tag{14}$$

where $\Gamma(\cdot)$ denotes the gamma function. This expression captures the expected utility available in Paris under optimal within-city location choices and enters households' decisions about whether to live in Paris, as described in Section 5.5.

5.2 Firms

Production takes place in each location under perfect competition. Firms produce a homogeneous final good, which is costlessly traded across the city and which we choose as the numéraire. Output in location j is

$$y_j = A_j L_j^{\alpha} \left(H_j^c\right)^{1-\alpha}, \tag{15}$$

where A_j is the productivity of location j, L_j is the number of workers and H_j^c is commercial floorspace. Firms choose the number of workers and floorspace to maximise profits, taking the wage w_j and the price of floorspace q_j as given. The first-order conditions imply that expenditure shares on both input factors are constant,

$$q_j H_j^c = \frac{1-\alpha}{\alpha} w_j L_j. \tag{16}$$

Perfect competition drives profits to zero, which pins down the price that firms can pay for floorspace in location j given the local wage and productivity,

$$q_j = (1-\alpha)\left(\frac{\alpha}{w_j}\right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}}. \tag{17}$$

Locations with higher productivity therefore support higher floorspace prices and, for a given price, higher wages. Sewers enter the model through their effect on A_j.

Productivity effect of sewers. We model the effect of sewers on local productivity analogously to their effect on residential amenities, distinguishing between a private benefit of sewer access and a potential reduction in negative externalities.

$$A_j = \bar{a}_j \, e^{\lambda_1 \, \text{Sewer}_j} \, e^{-\lambda_2 \, (1 - \text{ShareSewered}_j) \, \text{Density}_j^{\gamma_A}} \left(\Omega_j^L\right)^{\lambda_3} \tag{18}$$

The fundamental productivity of location $\bar{a}_j$ captures features, such as proximity to the riverfront, that affect productivity independently of the density of surrounding economic activity. The parameter λ_1 captures the extent to which sewer access increases location-specific productivity. Sewers may allow firms to dispose of wastewater more easily and improve the sanitary conditions of their premises. They may also make a location a more attractive marketplace, encouraging customers to visit more frequently and allowing firms to sell more output for a given set of inputs. The potential negative sanitary externalities are captured by the second exponential term, which uses the same measures of local density and sewer access as the residential amenity specification. This term captures how productivity may be adversely affected by high levels of waste pollution in the surrounding area. Our empirical results suggest that firms respond to sewer access at their own location rather than to sewer access in the surrounding area, consistent with $\lambda_2 = 0$. One possible explanation is that firms follow workers and customers, whose location choices are themselves affected by local sanitary conditions.

Finally, we allow for positive production externalities arising from proximity to other workers. Following the standard parameterisation in the literature, we define

$$\Omega_j^L = \sum_k e^{-\delta \, \tau_{jk}} L_k, \tag{19}$$

where L_k denotes the number of workers in location k, weighted by the walking time τ_{jk} between j and k, with decay parameter δ. The parameter λ_3 is the elasticity of productivity with respect to this measure of nearby employment. It reflects the productivity benefits of proximity to other workers and firms, for example through knowledge spillovers or demand spillovers.

5.3 Housing supply

Floorspace in each location is supplied by a competitive construction sector that combines land with capital. Following Ahlfeldt et al. (2015), we model the supply of floorspace H_i as increasing in the available land K_i and in the price of floorspace,

$$H_i = h \, q_i^{\mu} K_i, \tag{20}$$

where h is a constant and $\mu \geq 0$ is the elasticity of floorspace supply with respect to its price. Floorspace is used both by households for housing and by firms for production, and a single price q_i clears the market for floorspace in each location. In the limiting case $\mu = 0$, floorspace is fixed by land alone and prices absorb the entire adjustment. Since our hexagons are of equal area, K_i is constant.

5.4 Market clearing

The model has two markets that clear in each location, the labour market and the market for floorspace. The market for the final good clears by Walras' law.

Labour market. Employment in location j equals the number of residents throughout the city who choose to commute to j. Using the conditional commuting probabilities in (13), labour market clearing requires

$$L_j = \sum_s \pi_{sj|s} R_s = \sum_s \frac{\left(\frac{w_j}{d_{sj}}\right)^{\epsilon}}{\sum_k \left(\frac{w_k}{d_{sk}}\right)^{\epsilon}} R_s. \tag{21}$$

Wages adjust to clear this market. A location attracts more workers the higher its wage relative to other workplaces and the shorter the commute from where workers live. Through the zero-profit condition (17), locations with higher productivity A_j, including those with sewer access, can pay higher wages and therefore employ more workers.

Housing market. A single price q_i clears the market for floorspace in each location, which is used both by residents and by firms. Residents spend a share $1 - \beta$ of their income on housing, where the expected income of a resident of i is the commuting-probability-weighted average of the wages they can earn, $\bar{v}_i = \sum_j \pi_{ij|i} w_j$. Firms spend a share of their costs on floorspace given by the first-order condition (16). Floorspace market clearing therefore requires that the value of floorspace supplied in (20) equals total expenditure on floorspace by residents and firms,

$$q_i H_i = (1 - \beta)\, \bar{v}_i R_i + \frac{1 - \alpha}{\alpha}\, w_i L_i \equiv E_i. \tag{22}$$

Combining (22) with the supply of floorspace (20) gives the equilibrium price and quantity of floorspace in each location,

$$q_i = \left(\frac{E_i}{h K_i}\right)^{\frac{1}{1+\mu}}, \qquad H_i = h^{\frac{1}{1+\mu}} K_i^{\frac{1}{1+\mu}} E_i^{\frac{\mu}{1+\mu}}. \tag{23}$$

An increase in local demand for floorspace, for instance because sewers make a location more attractive to residents or firms, raises its price with elasticity $1/(1 + \mu)$ and the quantity of

floorspace with elasticity $\mu/(1+\mu)$. The elasticity of floorspace supply μ therefore determines how far sewer-induced increases in amenities and productivity are capitalised into floorspace prices rather than accommodated through higher densities.

5.5 Entry into the city

The total population of Paris is endogenous and we model entry into the city as a sequential choice. Households first decide whether to live in Paris or elsewhere in France. Conditional on living in Paris, they then choose their optimal pair of residential and workplace locations within the city. Solving the problem backwards, the value of living in Paris is the expected utility from the within-city choice.

Household o obtains utility from living in Paris (P) or elsewhere in France (F)

$$V_{mo} = u_m \, \zeta_{mo}, \qquad m \in \{P, F\}, \tag{24}$$

where u_P is the expected utility of living in Paris and $u_F \equiv \bar{u}$ is the deterministic part of the utility of living elsewhere. The idiosyncratic preferences ζ_{mo} are drawn independently from a Fréchet distribution,

$$F(\zeta) = \exp\left[-\zeta^{-\theta}\right], \tag{25}$$

where the shape parameter θ governs the dispersion of preferences for Paris across households. Each household chooses the option that gives it the higher utility, so that the share of the French population that lives in Paris is

$$\frac{R_P}{R_P + R_F} = \frac{u_P^\theta}{u_P^\theta + \bar{u}^\theta}, \tag{26}$$

where R_P and R_F denote the number of residents of Paris and of the rest of France, respectively. Inverting (26) and using the expected utility of living in Paris from (14) gives the population mobility condition,

$$\bar{u} \left(\frac{R_P}{R_F}\right)^{1/\theta} = \Gamma\left(1 - \frac{1}{\epsilon}\right) \left[\sum_k \sum_l \left(\frac{B_k w_l}{d_{kl} P_k^\beta q_k^{1-\beta}}\right)^\epsilon\right]^{1/\epsilon}. \tag{27}$$

Households move into Paris until the expected utility of living in the city equals the left-hand side, which acts as an outside option that rises with the size of Paris. The outside option rises because the marginal migrant has an increasingly weak idiosyncratic preference for living in the city. The parameter θ governs how strongly the population of Paris responds to changes in its attractiveness. We assume that migration to and from Paris leaves the utility of living elsewhere in France, $\bar{u}$, unaffected. Migration does, however, affect the expected utility

of living in Paris, u_P, through floorspace prices, wages, agglomeration and the sanitation externality.

Sewer investments raise the population of Paris through this extensive margin. By increasing amenities and productivity within the city, they raise u_P and attract households until the mobility condition (27) holds again.

5.6 General equilibrium

We now characterise the general equilibrium of the model. The spatial allocation of households and firms is determined by the parameters $\{\beta, \alpha, \epsilon, \kappa, \mu, \theta, \boldsymbol{\eta}, \boldsymbol{\lambda}, \gamma_B, \gamma_A, \rho, \delta\}$, geography $\{\mathbb{N}, \boldsymbol{\tau}, \mathbf{K}\}$, the total population of France $\{R_P + R_F\}$ and the utility of living elsewhere in France $\{\bar{u}\}$, and location characteristics $\{\text{Sewer}_i, \bar{a}_i, \bar{b}_i\}$ that determine residential amenities and firm productivity.

Given these primitives, the endogenous variables $\{\mathbf{R}, \mathbf{L}, \mathbf{w}, \mathbf{q}, \mathbf{H}\}$ adjust such that workers choose residential–workplace pairs according to (12), firms maximise profits according to (17) and (16), labour markets clear according to (21), and floorspace markets clear according to (22) and (20). Entry into Paris occurs until expected utility is equalised between Paris and the rest of France according to (27).

6 Quantitative Analysis

In historical settings, linking the model to the data often comes with challenges. In our case, the challenge is that the electoral lists provide citywide population records only in selected years. We overcome this challenge by estimating the model in two steps. First, we invert the model to recover the amenity and productivity of each location as structural residuals that rationalise the observed distribution of residents and workers in 1872, the year for which we observe both throughout the city. Second, we estimate the parameters that govern the effects of sewers by Simulated Method of Moments (SMM). We choose them such that the model reproduces the two empirical results of Section 4, the effect of a location's own sewer access and the way in which the effect of nearby sewer access varies with local density. For residents, we compute these moments on the same subset of arrondissements as in the empirical analysis. Table 5 summarises the data inputs and the calibrated and estimated parameters.

Table 5: Overview of model quantification

Parameter	Description	Value	Detail/Source
Data inputs			
R_i	Residents	$N \times 1$	Electoral lists
L_j	Workers	$N \times 1$	Business directories
Sewer_i	Sewer access	$N \times 1$	Historical sewer maps
τ_{ij}	Travel times	$N \times N$	Historical transport network
K_i	Land	$N \times 1$	Hexagon area
Externally calibrated parameters			
α	Labour share in production	0.70	Piketty and Zucman (2014)
$1-\beta$	Housing expenditure share	0.24	Kesztenbaum and Rosenthal (2017)
ϵ	Fréchet shape, within-city choice	5.25	Heblich et al. (2020)
κ	Commuting cost per minute	0.013	Ahlfeldt et al. (2015): $\epsilon\kappa = 0.07$
μ	Floorspace supply elasticity	1.83	Heblich et al. (2020)
η_3	Residential agglomeration	0.15	Ahlfeldt et al. (2015)
λ_3	Workplace agglomeration	0.08	Ahlfeldt et al. (2015)
ρ	Decay of residential agglomeration	0.91	Ahlfeldt et al. (2015)
δ	Decay of workplace agglomeration	0.36	Ahlfeldt et al. (2015)
θ	Fréchet shape, entry into Paris	3	Bryan and Morten (2019)
Estimated parameters			
η_1	Private amenity effect of sewers	0.025	SMM
η_2	Sanitation externality, amenities	30%†	SMM
γ_B	Curvature of the sanitation externality	5	SMM
λ_1	Private productivity effect of sewers	0.05	SMM
λ_2	Sanitation externality, productivity	0	SMM
Recovered by model inversion			
$\bar{b}_i$	Fundamental amenities	$N \times 1$	Model inversion
$\bar{a}_j$	Fundamental productivity	$N \times 1$	Model inversion
$\bar{u}_t$	Utility of living elsewhere in France	by year	Mobility condition (27)

† Amenity loss in a fully unsewered location at the 95th percentile of Density_i in 1872. With $\lambda_2 = 0$, γ_A is not identified.

6.1 Model inversion

We first recover wages and floorspace prices from the market-clearing conditions, and then recover amenities and productivity as structural residuals.

Wages. Given travel times $\boldsymbol{\tau}$, the commuting parameters $\{\epsilon, \kappa\}$, and the observed numbers of residents R_i and workers L_j, labour market clearing (21) defines a system of equations that identifies wages w_j up to a normalisation. We compute travel times between all pairs of locations using the fastest combination of walking and public transport, which at the time consisted of horse-drawn omnibuses and tramways. To do so, we digitise the omnibus and tram network operating in each year from historical maps and combine it with the street network (Appendix A.1.6). Since bilateral commuting flows are not observed for nineteenth-century Paris, we take estimates from the literature for the commuting parameters. The earliest historical setting for which these parameters have been estimated is London, where the 1921 population census first recorded where residents worked. We therefore set the Fréchet shape parameter to $\epsilon = 5.25$, as estimated by Heblich et al. (2020) using these data. For commuting costs, we use the semi-elasticity of commuting flows with respect to travel time of $\epsilon\kappa = 0.07$ per minute estimated by Ahlfeldt et al. (2015) using block-level data for Berlin, which matches the fine spatial scale of our analysis most closely, implying $\kappa \approx 0.013$. With these values, a commute of 10 minutes halves the attractiveness of a residence–workplace pair relative to working in the same location, and a commute of 30 minutes reduces it by almost 90%.

Floorspace prices. Given wages, the expenditure shares $\{\alpha, \beta\}$ and the distribution of residents and workers, floorspace market clearing (22), together with floorspace supply (20), identifies floorspace prices q_i up to a normalisation. We assume that residents and firms pay the same price for floorspace in a given location. Land K_i is the area of each hexagon, which is the same across locations, and we set the elasticity of floorspace supply to $\mu = 1.83$, following Heblich et al. (2020). Kesztenbaum and Rosenthal (2017) construct measures of average income and rent at the arrondissement level for 1891. A population-weighted average across arrondissements gives a housing expenditure share of $1 - \beta = 0.24$, which we assume to be stable over our period, at the upper end of the range considered by Redding and Sturm (2024) for London in the late 1930s. We set the labour share to $\alpha = 0.7$, its 1880 value for France in Piketty and Zucman (2014). We infer floorspace prices from the model rather than using the transaction prices recorded in the *sommier foncier*, as transactions are observed only for a sample of properties and do not record floorspace. Using the observed prices instead leaves the recovered amenities and productivity largely unchanged, since prices enter them with weights of only $1 - \beta = 0.24$ and $1 - \alpha = 0.30$ (Appendix Figure A24).

Amenities and productivity. Given wages and floorspace prices, we recover amenities and productivity as the values that rationalise the observed distribution of residents and workers in 1872 as an equilibrium of the model. Summing the choice probabilities (12) over workplaces gives the probability $\pi_i = R_i/R_P$ of living in location i. Taking logs and rearranging yields

$$\ln B_i = \frac{1}{\epsilon} \ln \Phi + \frac{1}{\epsilon} \ln \pi_i + (1 - \beta) \ln q_i - \frac{1}{\epsilon} \ln \left[\sum_k \left(\frac{w_k}{d_{ik}} \right)^{\epsilon} \right], \tag{28}$$

where Φ denotes the denominator of (12), which is common to all locations, and the last term measures the market access of location i to employment opportunities across the city. Locations that, all else equal, have more residents, higher floorspace prices or worse access to jobs must offer higher amenities to compensate residents. Similarly, the zero-profit condition (17) implies

$$\ln A_j = \alpha \ln \left(\frac{w_j}{\alpha} \right) + (1 - \alpha) \ln \left(\frac{q_j}{1 - \alpha} \right), \tag{29}$$

so that locations in which firms pay higher wages or floorspace prices must be more productive for firms to nevertheless hire many workers in these locations. For given parameters, these conditions determine amenities and productivity uniquely up to a normalisation (Ahlfeldt et al. 2015).

The amenities and productivity recovered in this way are those of the city in 1872. They therefore include the effects of the sewer network in place in 1872 and of the residents and workers in that year, through agglomeration and the sanitation externality. At this stage, we do not yet separate them into the fundamentals $\bar{b}_i$ and $\bar{a}_j$ and these endogenous components. This is the purpose of the next step.

6.2 Estimating the sewerage-related parameters

We estimate the sewerage-related parameters by Simulated Method of Moments (SMM), an indirect inference approach that builds on our empirical results. We place the parameters $\{\boldsymbol{\eta}, \boldsymbol{\lambda}, \boldsymbol{\gamma}\}$ on a grid and, for each candidate, solve for the model-implied distribution of residents and workers before the expansion of the sewer network. We then replicate our empirical analysis on these model-implied outcomes. The SMM estimates are the parameter values for which the model most closely reproduces the empirical effects of sewers.

Simulation. For a candidate set of parameters, we use (10) and (18) to split the amenities and productivity recovered for 1872 into fundamentals and the components that depend on sewers and on the distribution of residents and workers. Holding the fundamentals fixed, we remove the sewers built between 1856 and 1872 and solve for the equilibrium of the

city under the 1856 sewer network, holding total population fixed at its 1856 level. This yields a model-implied distribution of residents and workers before the expansion of the network.

Moments. We estimate the regressions of Section 4 on the simulated pre-period and the observed post-period, using the same specifications, including hexagon and quartier-by-year fixed effects. We target two sets of moments. The first is the difference-in-differences estimate of the effect of a location's sewer access on residents and workers. The second is the effect of nearby sewer access on residents by bins of local address density, relative to the least dense bin, which captures how the value of sewer access in the surroundings increases with density. The SMM estimates minimise the distance between the model-implied and the empirical moments, weighted by the inverse of their variance.

Identification. The two sets of moments contain different information. The effect of a location's sewer access combines the private benefit of sewer access, spillovers from the reduction in externalities as nearby locations also gain access, and general equilibrium adjustments such as the relocation of residents and firms across the city. On its own, it therefore cannot separate the private benefit from the externality. The density gradient in the effect of nearby sewer access provides the additional variation needed to do so. The private benefit raises the attractiveness of a sewered location regardless of how dense its surroundings are, and therefore cannot generate this gradient. The gradient thus identifies the strength and curvature of the sanitation externality, η_2 and γ_B, and, given the externality, the sewer access effects identify the private benefits η_1 and λ_1.

Estimates. We estimate the model under two scenarios. In the first scenario (S1), sewers affect amenities and productivity only through the private benefit of sewer access, as in a standard treatment of infrastructure as a local amenity ($\eta_2 = \lambda_2 = 0$). In the second scenario (S2), we additionally allow sewers to reduce the sanitation externality. The estimates reveal that sewer access substantially raises the attractiveness of locations for both residents and firms. In S1, sewer access raises the amenity value of a location by 3% and its productivity by 5%. Since amenities enter utility multiplicatively, the amenity effect is equivalent to a 3% increase in consumption for residents of sewered locations, while the productivity effect means that firms produce 5% more output with the same inputs. Allowing for the sanitation externality in S2 reduces the private amenity effect to 2.5%, as part of the effect that S1 attributes to the private benefit of sewer access in fact reflects the reduction in pollution from nearby locations. The productivity effect remains unchanged at 5%. The estimated externality is substantial but concentrated in the densest parts of the city. In a fully unsewered location, it is negligible at the median density of 1872 but lowers amenities by 30% at the 95th percentile. Consistent with the empirical evidence that firms respond to

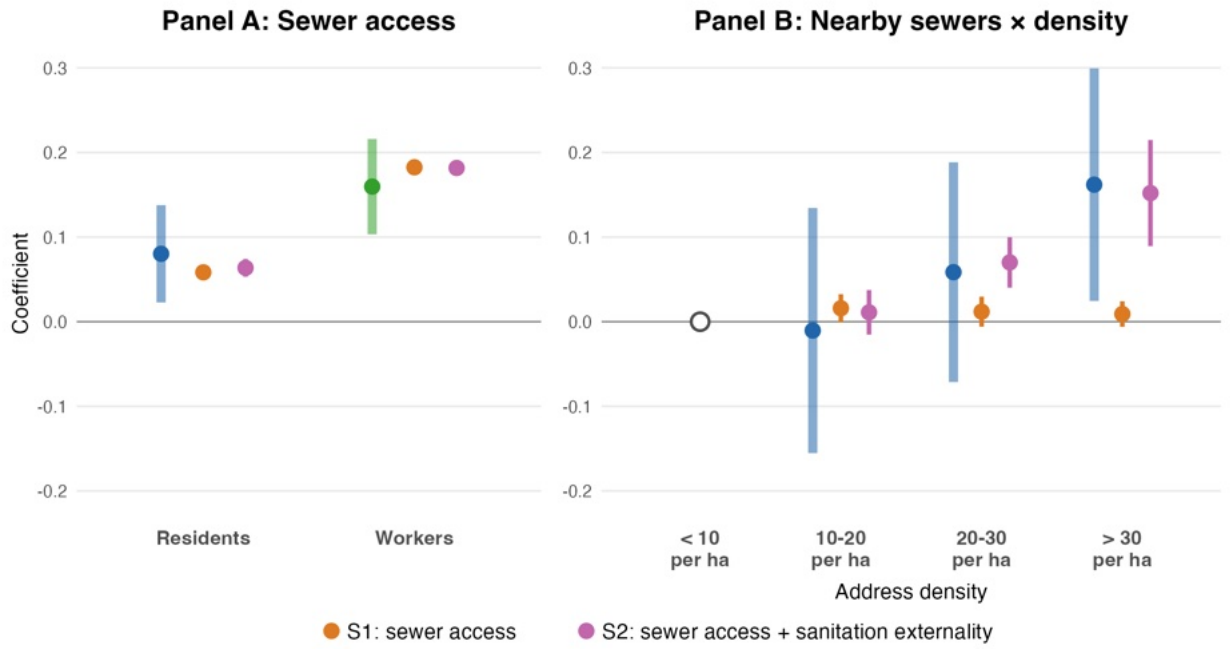
their own sewer access rather than to that of their surroundings, we find no externality on the production side ($\lambda_2 = 0$).

6.3 Model fit

Targeted moments. Figure 10 compares the model-implied effects of sewers at the SMM estimates of the two scenarios with their empirical counterparts. Both scenarios match the effect of a location's sewer access on residents and workers (Panel A). They differ in their ability to match the density gradient in the effect of nearby sewer access (Panel B). In the data, the effect of nearby sewer access on residents in the densest locations exceeds that in the sparsest locations by 0.16 log points. Scenario S2 implies a gradient of 0.15, whereas S1 implies a gradient close to zero in all density bins. In S1, sewers in nearby locations attract residents and workers to the surrounding area, which in turn generates positive agglomeration spillovers. These spillovers, however, appear to be too weak and to vary too little with density to account for the pattern in the data. Once we allow sewers to remove the negative externality that residents and workers impose on each other in the absence of sanitation, the model matches both sets of moments. Our interpretation is that in initially denser locations, there is more wastewater pollution to remove, so that nearby sewers make these locations considerably more attractive and draw in further residents. The density gradient in the data is therefore consistent with the removal of the sanitation externality, rather than with agglomeration forces triggered by nearby sewer investments.

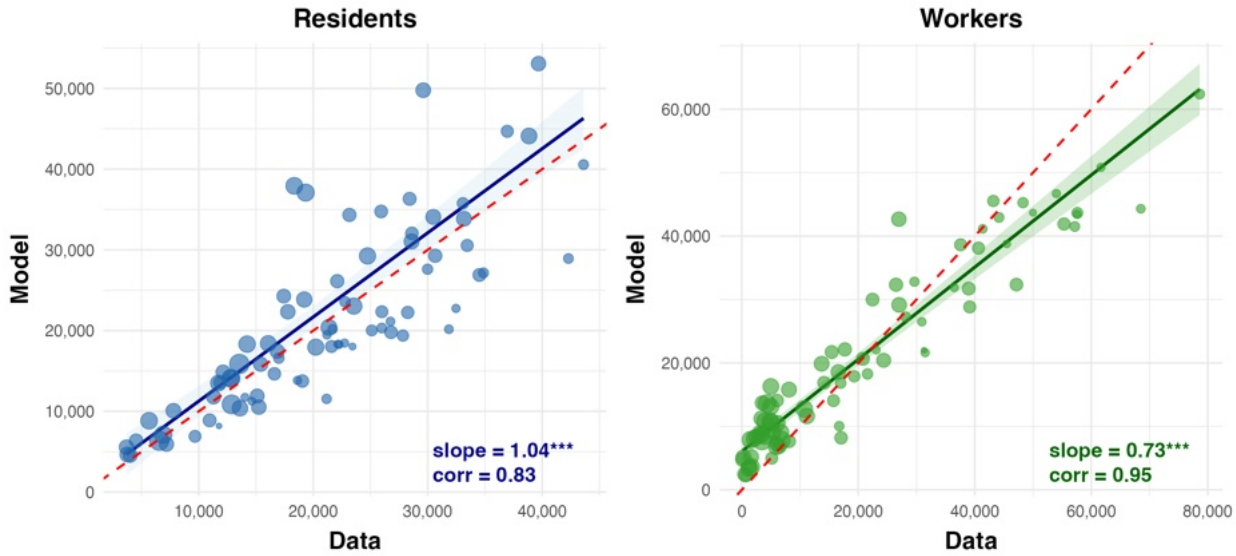
Untargeted moments. Since only the second scenario reproduces the density gradient in the data, we use it in the remainder of the paper, and now evaluate its fit on moments outside the estimation. The estimation targets only relative changes between locations, and the resident moments draw on variation from the six arrondissements for which electoral lists survive in the pre-period. Figure 11 therefore assesses the model on two dimensions that the estimation does not use. First, it compares levels rather than changes, namely the number of residents and workers in each quartier in the model-implied equilibrium before the expansion of the sewer network and in the data in 1861, the earliest year for which the census reports population at the quartier level. Second, for residents, the comparison covers all 80 quartiers of the city, including those outside the six arrondissements used in estimation. The model reproduces the distribution of residents across quartiers closely, with a slope of 1.04 and a correlation of 0.83. For workers, the correlation is 0.95, while the slope of 0.73 indicates that the model somewhat understates differences in employment across quartiers. The model thus captures the spatial distribution of economic activity in the baseline period, including in parts of the city whose variation it was not used to estimate.

Figure 10: Model fit on targeted moments



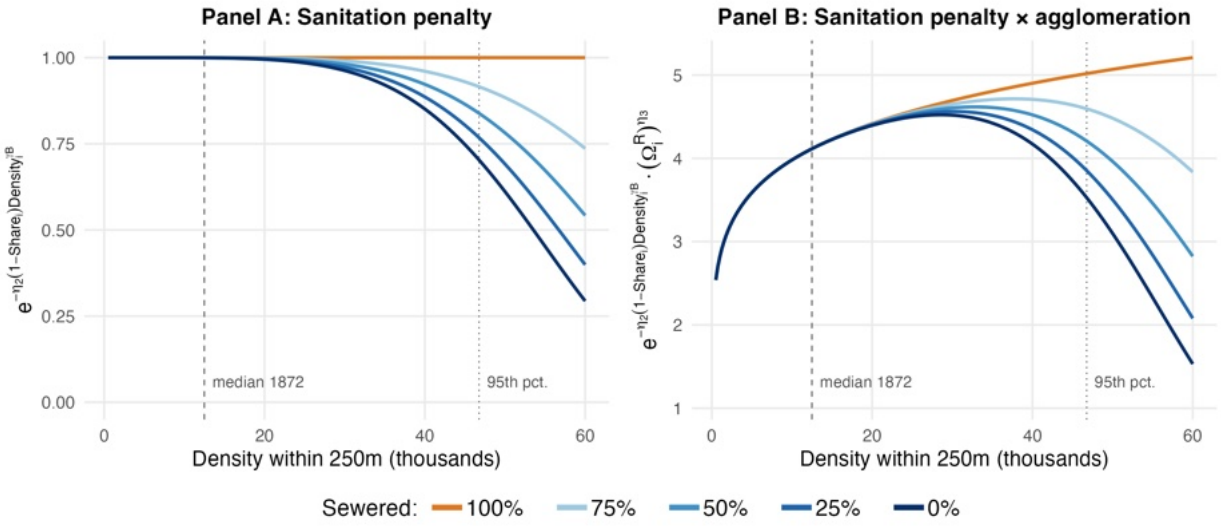
Notes: Panel A: difference-in-differences estimates of the effect of sewer access on residents and workers. Panel B: effect of nearby sewer access on residents by bin of address density, relative to the least dense bin. Data estimates with 90% confidence intervals in blue (residents) and green (workers); model-implied estimates for scenario S1 (sewer access only) in orange and S2 (sewer access and sanitation externality) in purple, with 90% confidence intervals from estimating the same regressions on the model output.

Figure 11: Untargeted fit on residents and workers by quartier



Notes: Each dot is one of the 80 quartiers of Paris, sized by its number of hexagons. Vertical axis: number of residents and workers in the model-implied equilibrium before the expansion of the sewer network. Horizontal axis: residents from the 1861 census and workers from the 1861 business directories. Solid lines are weighted OLS fits with 95% confidence bands; the dashed line is the 45-degree line.

Figure 12: Channeling density



Notes: Panel A: sanitation penalty on amenities, $e^{-\eta_2(1-\text{ShareSewered}_i)\text{Density}_i^{\gamma B}}$, as a function of the number of residents and workers within 250 metres, for shares from 100% to 0% of the surroundings with sewer access. Panel B: sanitation penalty multiplied by the agglomeration term $(\Omega_i^R)^{\eta_3}$, assuming that nearby residents scale with local density. Dashed and dotted vertical lines mark the median and the 95th percentile of local density in 1872. Parameters at the estimates of the second scenario.

6.4 Channeling density

The estimated sanitation externality already offers insights into the costs of density in the absence of sanitation, and into how sewers allow cities to benefit more fully from the advantages of density. Figure 12 illustrates how the sanitation externality varies with local density and sewer coverage.

Panel A plots the sanitation penalty in the amenity function (10) against the number of residents and workers within 250 metres, for different shares of the surroundings with sewer access. In a fully sewered neighbourhood, there is no penalty. Without sewers, the penalty is negligible at low and moderate densities, including at the median density in Paris in 1872, but becomes substantial in the densest parts of the city, where the implied amenity loss reaches 30% at the 95th percentile. At low densities, sewers therefore mainly provide a private benefit to the residents of the newly sewered locations. Residents appear willing to tolerate urban density and the neighbourhood-level pollution that comes with it up to a point. Beyond that point, the volume of waste becomes too large, living conditions deteriorate sharply, and each additional resident lowers the amenity value of the location for everyone else.

Panel B shows the trade-off between this negative force and the positive forces of density. It plots the product of the sanitation penalty and the residential agglomeration term in (10), which follows the specification of Ahlfeldt et al. (2015). We view these agglomeration forces

as present in cities in general, but counteracted by other forces, among which our focus is the sanitation externality. The figure reveals an inverse U-shaped pattern. At low densities, the positive forces dominate, as the sanitation externality has not yet reached harmful levels. At higher densities, the balance shifts, and each additional resident reduces the attractiveness of the location. For a fully unsewered neighbourhood, the turning point lies at about the 80th percentile of local density in 1872. Sewers remove this downturn, so that additional residents continue to make the location more attractive.

Sewers are therefore a technology that raises the benefit of having more residents in a location. We believe that this insight helps us understand the development of cities and of neighbourhoods within them. Cities often emerge before sanitation is provided, likely because positive forces draw people together. To sustain high densities, however, cities eventually need to improve sanitary conditions.

7 Counterfactuals

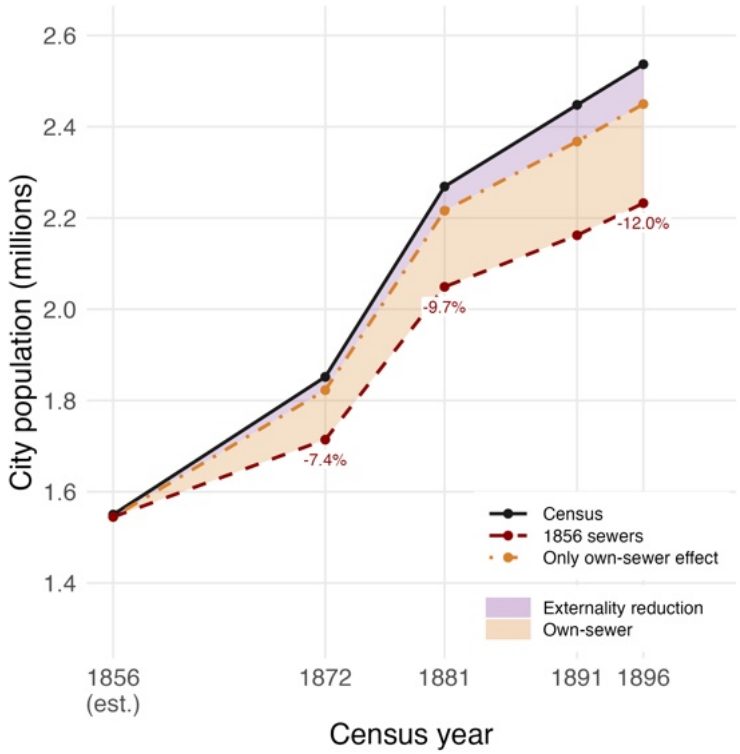
We now use the estimated model to quantify the aggregate implications of the expansion of the Parisian sewer network. In contrast to the estimation, in which we hold the population of Paris fixed, the counterfactuals allow households to move between Paris and the rest of France, as described in Section 5.5. The population of the city is therefore endogenous and responds to changes in its attractiveness.

7.1 Sewers and the growth of Paris

We first examine to what extent the growth of Paris in the second half of the nineteenth century can be attributed to the expansion of the sewer network. Between 1856 and 1896, the population of Paris increased from about 1.55 million to 2.54 million. For each census year, we solve for the implied utility of the outside option $\bar{u}_t$ in (27) such that the model, with the sewer network of that year, matches the population recorded in the census. The outside option thereby absorbs all other changes that made Paris more or less attractive relative to the rest of France in that year. Holding $\bar{u}_t$ and all other model primitives fixed, we then solve for the equilibrium of the city under the sewer network of 1856. The difference between the two equilibria measures the contribution of the sewer expansion to the population of Paris in each year.

Results. Figure 13 shows the results. Without the expansion of the sewer network after 1856, Paris would have had 7.4% fewer residents in 1872, 9.7% fewer in 1881 and 12.0% fewer in 1896. The effect grows over time as the network expands and as the growing city makes

Figure 13: Sewers and the growth of Paris



Notes: Population of Paris in the census (black) and in two counterfactual equilibria. The dashed red line keeps the sewer network at its 1856 extent. The dash-dotted orange line grants each location the private benefit of sewer access under the network of the respective year, while the sanitation externality is determined by the 1856 network. The orange area is the contribution of the private benefit of sewer access, the purple area that of the reduction in the sanitation externality. Labels give the population shortfall without the sewer expansion relative to the census. In each year, the utility of the outside option is chosen such that the model with that year's sewer network reproduces the census population. The 1856 population is an estimate for the post-1860 boundaries of the city.

the absence of sewers increasingly costly. Relative to the growth of the census population between 1856 and 1896, the sewer expansion accounts for about 30% of the growth of Paris by the end of the 19th century.

To separate the two channels through which sewers affect the city, we solve for a third equilibrium in which each location receives the private benefit of sewer access under the sewer network of the respective year, while the sanitation externality is still determined by the sewer network of 1856. The gap between this equilibrium and the census population measures the contribution of the reduction in the sanitation externality. The reduction in the externality accounts for about 21% of the effect of sewers in 1872 and 29% in 1896, or about a quarter on average (Appendix Table A15, Panel C). Its share rises over time, consistent with the convexity of the externality in density. As the city grows, the costs of density in unsewered neighbourhoods rise more than proportionally, and so do the gains from removing them. The two channels also operate in different parts of the city, as documented in

Appendix Figure A26. The private benefit of sewer access accounts for almost all of the gains of the locations sewered between 1856 and 1872. The reduction in the sanitation externality instead mainly benefits the dense centre of the city, which had already been sewered by 1856 but whose surroundings only received sewers afterwards.

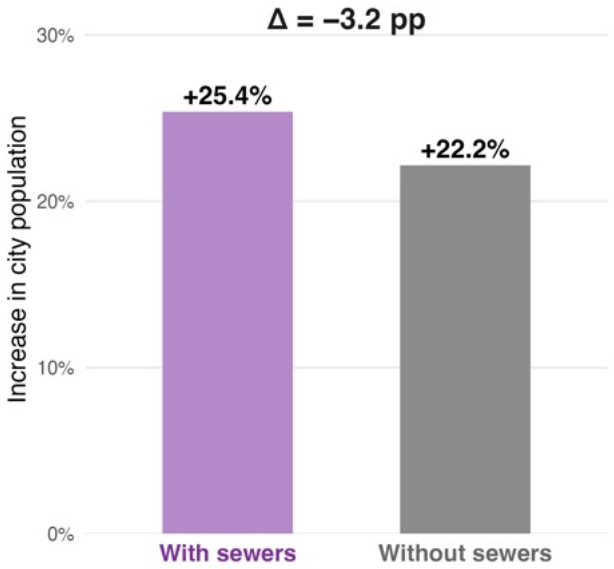
Sensitivity. The magnitude of these effects depends on parameters that we take from the literature. Most importantly, the entry elasticity θ governs how strongly the population of Paris responds to changes in its attractiveness relative to the rest of France, and therefore scales the population effects directly. In our baseline, we set $\theta = 3$, in line with the estimates of Bryan and Morten (2019) for the United States (2.7) and Indonesia (3.2). Appendix Table A15 reports the counterfactual for values of θ between 1 and 5. The population effect of the sewer expansion scales almost proportionally with θ. In 1896, Paris would have had between 4.4% fewer residents for $\theta = 1$ and 18.4% fewer for $\theta = 5$ without the sewer expansion, so that the sewers account for between 11% and 47% of the growth of Paris since 1856. The split between the two channels, by contrast, is almost unaffected. For all values of θ, the reduction in the sanitation externality accounts for about 20% of the effect of sewers in 1872 and about 28% in 1896. The magnitude of the aggregate effect therefore depends on how mobile households are between Paris and the rest of France, while the importance of the externality relative to the private benefit of sewer access does not.

7.2 Complementarities with other investments

Because sewers reduce the disutility that residents impose on each other, their benefits may extend beyond the locations they serve. Any investment that makes the city more attractive draws in additional residents and raises local densities. In the absence of sanitation, these higher densities increase the costs of the sanitation externality and thereby dampen the gains from the investment. Sanitation investments may therefore be complementary to a wide range of other public investments in cities. Such complementarities have direct implications for policy. Under a budget constraint, they suggest that sanitation should be provided before other investments. Moreover, since the gains from complementarities accrue to the city as a whole rather than to the locations that receive sewers, they provide an argument for public provision or subsidies.

We examine this complementarity for public transport, a form of urban infrastructure that has received considerable attention in the literature. The first line of the Paris metro opened in 1900, and by 1910 the metro network had already reached a substantial extent. At this point, most of the city was already connected to the sewer network. We take the model calibrated to 1896, the last year for which we observe the sewer network and the year closest to the opening of the metro. Since the model features a commuting structure, we can introduce

Figure 14: Population gain from the metro with and without sewers



Notes: Increase in the population of Paris from adding the metro network of 1910 to the omnibus network in the city of 1896, relative to the equilibrium without the metro. With sewers: the city of 1896 with its sewer network. Without sewers: a hypothetical city of 1896 without any sewers, in which the sanitation externality is active in every location, with fundamentals adjusted such that the model reproduces the same equilibrium of 1896 before the metro opens. The utility of the outside option is held fixed.

the metro by reducing travel times between locations to the levels implied by the omnibus network combined with the metro network of 1910. We solve for the resulting equilibrium, holding the utility of the outside option fixed, and measure the increase in the population of Paris relative to the equilibrium without the metro. We then repeat this exercise in a hypothetical city of 1896 without any sewers, in which the sanitation externality is active in every location. In this city, we adjust the fundamentals such that the model reproduces the same distribution of residents and workers in 1896 before the metro opens. Both settings therefore start from the same equilibrium and differ only in the extent to which the sanitation externality responds to changes in density.

Figure 14 shows that the metro raises the population of Paris by 25.4% when the city has its sewer network, but only by 22.2% when the sanitation externality is present everywhere. Without sewers, the population gain from the metro would therefore have been about one-eighth smaller. As the metro improves access to jobs across the city, it draws additional residents into Paris and raises local densities. In the absence of sewers, these higher densities increase the costs of the sanitation externality and dampen the gain. Sewers and transport infrastructure are therefore complements.

This example illustrates how sanitation amplifies the effectiveness of other public investments through its effect on the externalities that residents impose on each other. More

generally, it suggests that improving sanitation is a precondition for cities to sustain further growth.

7.3 Costs and benefits of sewerage investments

We conclude by comparing the benefits of the sewer expansion with its costs. We measure the benefits as the increase in the total value of the building stock implied by the model. To express this value in francs, we scale the value of the building stock in the equilibrium of 1872 to the total value of Parisian property in 1872, which we estimate at 8.8 billion francs.[3] Comparing the equilibrium of 1872 with the counterfactual equilibrium under the sewer network of 1856, the sewer expansion raised the value of the building stock by 8.6%, or about 760 million francs. Of this increase, 83% is due to the private benefit of sewer access and 17% to the reduction in the sanitation externality. The municipal expenditure on the sewer network between 1853 and 1877 amounted to 45.5 million francs, which implies a benefit–cost ratio of about 17.

For comparison, the construction of the metro network between 1899 and 1910 was financed through bonds of about 480 million francs. In the counterfactual of Section 7.2, the metro raised the value of the building stock of 1896 by 24.8%, or about 3.0 billion francs, which implies a benefit–cost ratio of about 6. Per franc spent, the sewer expansion therefore generated about 2.7 times the benefits of the metro. This comparison likely understates the relative returns to sewers, since part of the gains from the metro itself relied on the sewer network already being in place. Sanitation was thus not only a highly profitable investment in its own right, but also laid the foundation for the returns to the public investments that followed.

8 Conclusion

Haussmann famously described sewers as the "internal organs of the great city", functioning like arteries that sustain urban health and vitality. Our analysis provides an economic interpretation of this insight. We show that sewerage investments were not merely a public health intervention but a fundamental determinant of urban development. By reducing the costs associated with density, sewerage increased the attractiveness of locations for both households and firms, enabling Paris to sustain larger concentrations of population and eco-

[3] This estimate combines the average transaction price in the *sommier foncier*, adjusted for the selection of streets with recorded transactions, with a conservative count of 75,000 buildings. The franc values below are therefore lower bounds, while the comparison between the two investments does not depend on this scaling.

nomic activity. Through these general equilibrium effects, sanitation infrastructure helped cities capture the benefits of agglomeration while mitigating the congestion forces that historically constrained urban growth.

Our empirical analysis documents these effects at a fine spatial scale. Locations that received sewers during the first expansion wave gained about 10% more voters and 18% more firms than locations connected later. The increase in firms was driven by customer-facing businesses, while the social class and origin of residents did not change. Nearby sewers increased the number of residents only in the most densely populated neighbourhoods, consistent with sewers reducing the costs that dense populations impose on one another. In a later wave of the expansion, typhoid deaths declined by about 16% in locations that received sewers.

The quantitative importance of these effects is substantial. Our counterfactuals imply that, without the sewer expansion, Paris would have had 12% fewer residents in 1896, which accounts for about 30% of the city's population growth between 1856 and 1896. About one quarter of this effect operates through reduced sanitation externalities, while the remainder reflects the private benefits of sewer access in connected locations. These findings suggest that a significant share of the city's expansion can be attributed to improvements in sanitation infrastructure rather than to broader economic or demographic trends alone.

More broadly, our results highlight the central role of basic urban infrastructure in shaping long-run economic development. While the benefits of sewerage are often evaluated through the lens of public health, our findings demonstrate that sanitation investments also generate large economic returns by facilitating urban density and supporting agglomeration economies. This lesson remains highly relevant today. As rapidly urbanising cities across the developing world continue to face severe sanitation deficits, investments in sewerage may be among the most effective policies for promoting both public health and economic development. The experience of nineteenth-century Paris suggests that the capacity to sustain dense and productive cities depends critically on the infrastructure that allows urban residents and firms to coexist at scale.

References

Ahlfeldt, Gabriel M., Stephen J. **Redding**, Daniel M. **Sturm**, and Nikolaus **Wolf** (2015): "The Economics of Density: Evidence From the Berlin Wall". *Econometrica* 83.6, pp. 2127–2189.

Allen, Treb and Costas **Arkolakis** (2014): "Trade and the Topography of the Spatial Economy". *Quarterly Journal of Economics* 129.3, pp. 1085–1140.

Allen, Treb and Dave **Donaldson** (2020): *Persistence and Path Dependence in the Spatial Economy*. NBER Working Paper 28059. National Bureau of Economic Research.

Alsan, Marcella and Claudia **Goldin** (2019): "Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880–1920". *Journal of Political Economy* 127.2, pp. 586–638.

Ambrus, Attila, Erica **Field**, and Robert **Gonzalez** (2020): "Loss in the Time of Cholera: Long-Run Impact of a Disease Epidemic on the Urban Landscape". *American Economic Review* 110.2, pp. 475–525.

Anderson, D. Mark, Kerwin Kofi **Charles**, and Daniel I. **Rees** (2022): "Reexamining the Contribution of Public Health Efforts to the Decline in Urban Mortality". *American Economic Journal: Applied Economics* 14.2, pp. 126–157.

Arkhangelsky, Dmitry, Susan **Athey**, David A. **Hirshberg**, Guido W. **Imbens**, and Stefan **Wager** (2021): "Synthetic Difference-in-Differences". *American Economic Review* 111.12, pp. 4088–4118.

Bryan, Gharad, Edward **Glaeser**, and Nick **Tsivanidis** (2020): "Cities in the Developing World". *Annual Review of Economics* 12, pp. 273–297.

Bryan, Gharad and Melanie **Morten** (2019): "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia". *Journal of Political Economy* 127.5, pp. 2229–2268.

Chay, Kenneth Y. and Michael **Greenstone** (2005): "Does Air Quality Matter? Evidence from the Housing Market". *Journal of Political Economy* 113.2, pp. 376–424.

Costa, Dora L. and Matthew E. **Kahn** (2015): "Declining Mortality Inequality within Cities during the Health Transition". *American Economic Review* 105.5, pp. 564–569.

Coury, Michael, Toru **Kitagawa**, Allison **Shertzer**, and Matthew A. **Turner** (2024): "The Value of Piped Water and Sewers: Evidence from 19th Century Chicago". *Review of Economics and Statistics*, pp. 1–47.

Currie, Janet, Lucas **Davis**, Michael **Greenstone**, and Reed **Walker** (2015): "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings". *American Economic Review* 105.2, pp. 678–709.

Cutler, David and Grant **Miller** (2005): "The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States". *Demography* 42.1, pp. 1–22.

Donaldson, Dave (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure". *American Economic Review* 108.4-5, pp. 899–934.

Eichholtz, Piet, Matthijs **Korevaar**, Thies **Lindenthal**, and Ronan **Tallec** (2021): "The Total Return and Risk to Residential Real Estate". *Review of Financial Studies* 34.8, pp. 3608–3646.

Faber, Benjamin (2014): "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System". *Review of Economic Studies* 81.3, pp. 1046–1070.

Galiani, Sebastian, Paul **Gertler**, and Ernesto **Schargrodsky** (2005): "Water for Life: The Impact of the Privatization of Water Services on Child Mortality". *Journal of Political Economy* 113.1, pp. 83–120.

Glaeser, Edward (2011): *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier*. New York: Penguin Press.

Graff Zivin, Joshua and Matthew **Neidell** (2012): "The Impact of Pollution on Worker Productivity". *American Economic Review* 102.7, pp. 3652–3673.

Greenstone, Michael and Rema **Hanna** (2014): "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India". *American Economic Review* 104.10, pp. 3038–3072.

Hanlon, W. Walker (2020): "Coal Smoke, City Growth, and the Costs of the Industrial Revolution". *Economic Journal* 130.626, pp. 462–488.

Heblich, Stephan, Stephen J. **Redding**, and Daniel M. **Sturm** (2020): "The Making of the Modern Metropolis: Evidence from London". *Quarterly Journal of Economics* 135.4, pp. 2059–2133.

Heblich, Stephan, Alex **Trew**, and Yanos **Zylberberg** (2021): "East-Side Story: Historical Pollution and Persistent Neighborhood Sorting". *Journal of Political Economy* 129.5, pp. 1508–1552.

Henderson, J. Vernon and Matthew A. **Turner** (2020): "Urbanization in the Developing World: Too Early or Too Slow?" *Journal of Economic Perspectives* 34.3, pp. 150–173.

Keiser, David A. and Joseph S. **Shapiro** (2019): "Consequences of the Clean Water Act and the Demand for Water Quality". *Quarterly Journal of Economics* 134.1, pp. 349–396.

Kesztenbaum, Lionel and Jean-Laurent **Rosenthal** (2017): "Sewers' Diffusion and the Decline of Mortality: The Case of Paris, 1880–1914". *Journal of Urban Economics* 98, pp. 174–186.

Kremer, Michael, Jessica **Leino**, Edward **Miguel**, and Alix Peterson **Zwane** (2011): "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions". *Quarterly Journal of Economics* 126.1, pp. 145–205.

McCulloch, Sean E., Matthew P. **Schaelling**, Matthew **Turner**, and Toru **Kitagawa** (2025): *Sewers and Urbanization in the Developing World*. NBER Working Paper 33597. National Bureau of Economic Research.

Papayanis, Nicholas (1996): *Horse-Drawn Cabs and Omnibuses in Paris: The Idea of Circulation and the Business of Public Transit*. Baton Rouge: Louisiana State University Press.

Piketty, Thomas and Gabriel **Zucman** (2014): "Capital Is Back: Wealth-Income Ratios in Rich Countries 1700–2010". *Quarterly Journal of Economics* 129.3, pp. 1255–1310.

Pinkney, David H. (1958): *Napoleon III and the Rebuilding of Paris*. Princeton, NJ: Princeton University Press.

Redding, Stephen J. and Daniel M. **Sturm** (2024): *Neighborhood Effects: Evidence from Wartime Destruction in London*. NBER Working Paper 32333. National Bureau of Economic Research.

Robert, Jean (1959): *Les Tramways Parisiens*. Paris: Published by the author.

Steinwender, Claudia (2018): "Real Effects of Information Frictions: When the States and the Kingdom Became United". *American Economic Review* 108.3, pp. 657–696.

van Leeuwen, Marco H. D. and Ineke **Maas** (2011): *HISCLASS: A Historical International Social Class Scheme*. Leuven: Leuven University Press.

van Leeuwen, Marco H. D., Ineke **Maas**, and Andrew **Miles** (2002): *HISCO: Historical International Standard Classification of Occupations*. Leuven: Leuven University Press.

WHO and UNICEF (2025): *Progress on Household Drinking Water, Sanitation and Hygiene 2000–2024: Special Focus on Inequalities*. Tech. rep. Geneva: World Health Organization (WHO) and the United Nations Children's Fund (UNICEF).

A Appendix

Contents

A.1 Data Appendix . . . 55
A.1.1 Historical sewer maps . . . 55
A.1.2 Historical addresses . . . 60
A.1.3 Electoral lists . . . 63
A.1.4 Sommier Foncier . . . 73
A.1.5 Typhoid deaths . . . 79
A.1.6 Commuting times . . . 80
A.2 Empirical Appendix . . . 85
A.2.1 Haussmann's renovation . . . 85
A.2.2 Nearby sewers by density bin . . . 86
A.2.3 Sewers and market access . . . 87
A.3 Quantitative Appendix . . . 92
A.4 Counterfactual Appendix . . . 93
A.4.1 Sensitivity to the entry elasticity . . . 93
A.4.2 Population changes by location type in 1872 . . . 93

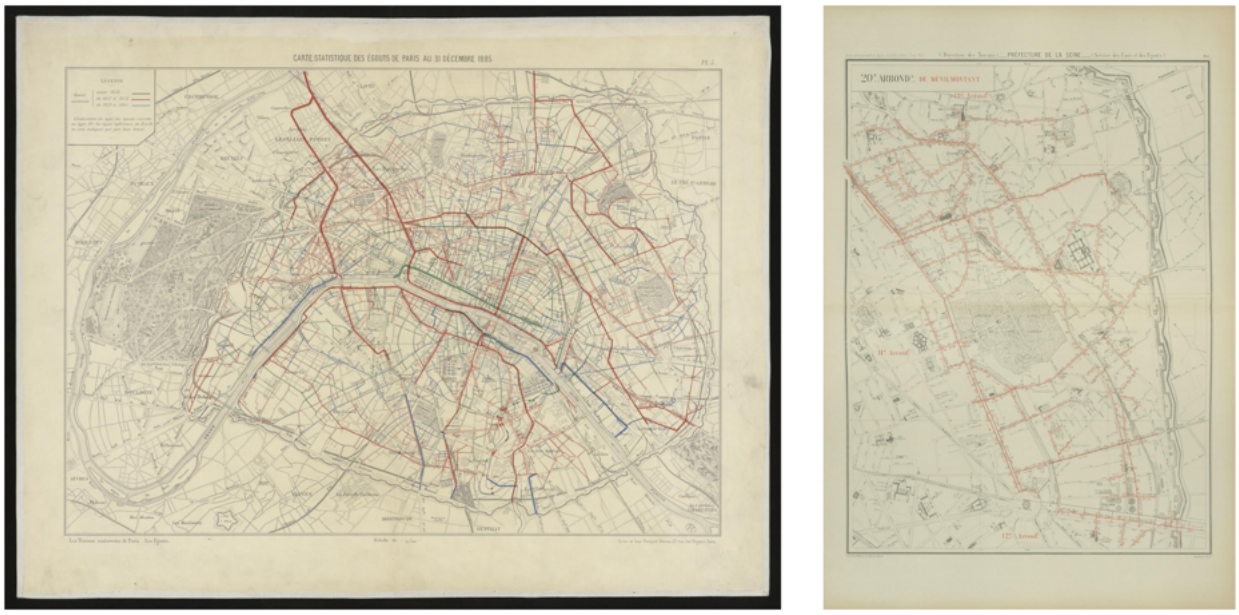
A.1 Data Appendix

A.1.1 Historical sewer maps

We digitise historical sewer maps of the Parisian sewerage network across eight cross-sections spanning 1856 to 1896. The maps derive from three archival sources. The *Carte statistique des égouts de Paris* is a single city-wide map displaying the sewerage network in three periods (1856, 1878, and 1885) simultaneously using different ink colours (green, red, and blue respectively). The *Atlas administratif des égouts de la ville de Paris* and *Plan général des égouts de la Ville de Paris* are each organised into sub-maps, with each map depicting a different section of Paris and its sewer network for that year. Table A1 lists the year and source for each cross-section, and Figure A1 shows an example of each map type.

Digitisation. For each map, we first manually sample the RGB values of more than 100 pixels corresponding to sewer lines and use the 5th to 95th percentile of each colour channel to define the target colour range. An automated detection routine then classifies every pixel in the image as sewer or non-sewer based on whether its RGB values fall within this range. The raw detection output contains noise arising from annotations printed in a similar colour

Figure A1: Examples of historical sewer maps



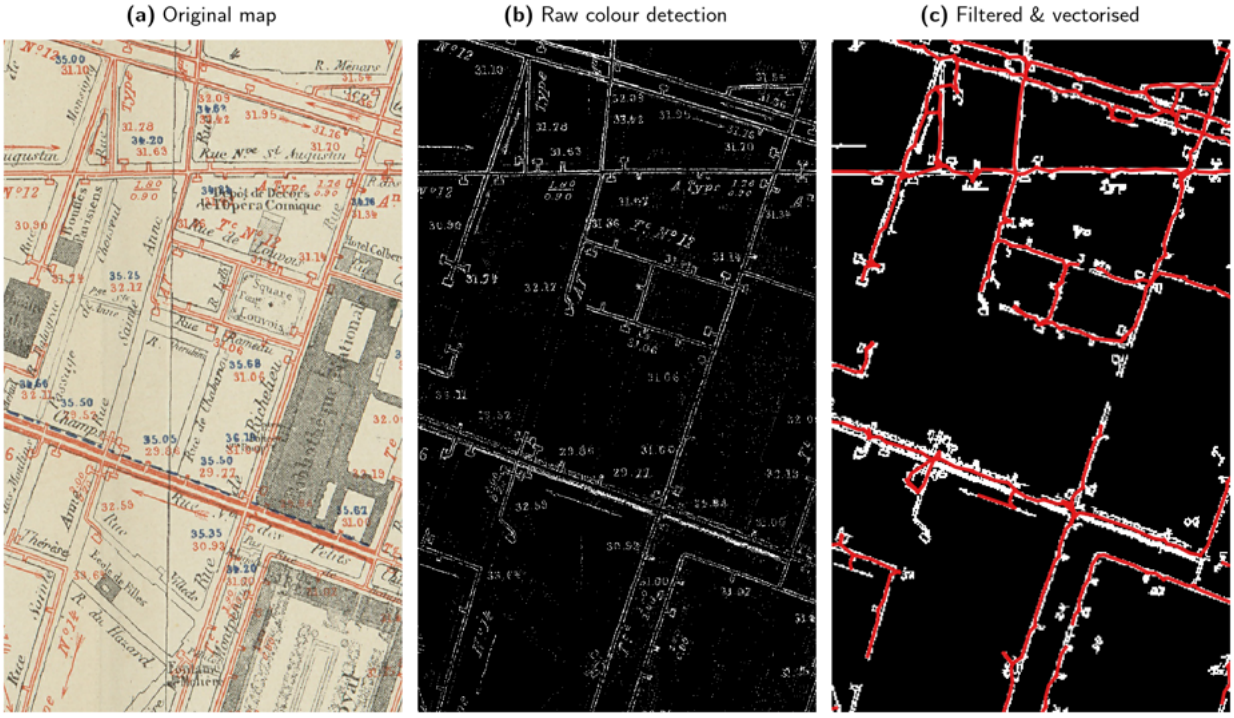
Notes: The left panel shows the *Carte statistique des égouts de Paris*, a city-wide map displaying the sewer network expansion across three periods in different ink colours: green for lines constructed before 1856, red for 1857–1878, and blue for 1879–1885. The right panel shows a map from the *Atlas administratif des égouts de la ville de Paris* (1880), depicting the 20th arrondissement.

Table A1: Data sources and years of sewer maps

Year	Source	Maps
1856	*Carte statistique des égouts de Paris* (displays 1856, 1878, and 1885)	1
1872	*Plan général des égouts de la Ville de Paris*	12
1875	*Plan général des égouts de la Ville de Paris*	12
1878	*Carte statistique des égouts de Paris* (displays 1856, 1878, and 1885)	1
1880	*Atlas administratif des égouts de la ville de Paris*	16
1885	*Carte statistique des égouts de Paris* (displays 1856, 1878, and 1885)	1
1889	*Atlas administratif des égouts de la ville de Paris*	16
1896	*Atlas administratif des égouts de la ville de Paris*	16

Notes: The *Carte statistique* encodes three cross-sections in a single map through the use of differently coloured inks. The other two series provide maps for different areas of Paris for the respective years.

Figure A2: Illustration of digitisation pipeline



Notes: The three panels show successive stages of the digitisation pipeline applied to a single 300×500 m tile from the 1880 *Atlas administratif*. Panel (a) shows the original scanned map tile. Panel (b) shows the raw binary output of the RGB colour-detection step, where white pixels are classified as sewer. Panel (c) overlays the final vectorised line network (red) on the filtered binary mask.

to the sewer lines. We remove this noise by applying a minimum connected-component size filter that discards isolated pixel clusters. The filtered binary raster is subsequently vectorised using graph-based tracking. Figure A2 illustrates the three stages of this pipeline for a representative 300×500 m tile from the 1880 *Atlas administratif*.

Validation. To assess the quality of the automated classification, we compare it against a manually digitised reference network for the three years covered by the *Carte statistique des égouts de Paris*. We digitised this map by hand as part of a separate effort, tracing each ink layer individually. For each of the three years, we classify all 146,184 hexagons in our 25 m grid as sewered (if crossed by a sewer line) or not sewered under both procedures, and compare the two classifications across four mutually exclusive categories. Table A2 reports the results. Agreement rates are 91.8% in 1856, 82.7% in 1878, and 79.8% in 1885. The automated procedure classifies more hexagons as sewered than the manual tracing in all three years, consistent with false positives from map labels, road markings or image artefacts that fall within the target colour range, which the manual digitisation avoids. The agreement

rate declines across years as the share of sewered hexagons grows and opportunities for disagreement increase.

Table A2: Validation – manual vs. automated classification

	1856	1878	1885
Both sewered	5,920 (4.0%)	18,589 (12.7%)	23,034 (15.8%)
Only manual	3,453 (2.4%)	9,192 (6.3%)	11,956 (8.2%)
Only automated	8,494 (5.8%)	16,089 (11.0%)	17,642 (12.1%)
Neither sewered	128,317 (87.8%)	102,314 (70.0%)	93,552 (64.0%)
Agreement rate	134,237 (91.8%)	120,903 (82.7%)	116,586 (79.8%)

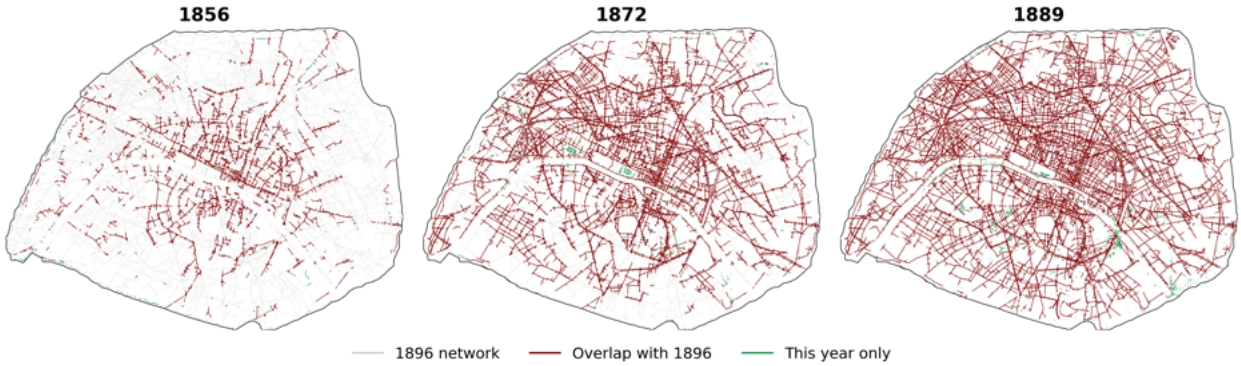
Notes: N = 146,184 hexagons. Agreement is defined as the sum of hexagons classified as sewered under both procedures and hexagons classified as not-sewered under both procedures.

Harmonisation to the 1896 map. Historical maps were produced at different times, printed at varying scales, and subsequently digitised from scanned images. Minor differences in georeferencing across map editions can shift the detected network by several metres relative to its true location. Left uncorrected, these shifts cause the same physical sewer to appear as distinct lines in different cross-sections, confounding year-to-year comparisons in network coverage.

To address this, we use the 1896 map as a spatial anchor and harmonise all earlier networks against it. For each year, every detected line is classified either as an *overlap* line, which lies within a calibrated distance of the nearest line in the 1896 network, or as a *year-only* line, which has no nearby 1896 counterpart. The distance threshold is calibrated separately for each year as the 95th percentile of the nearest-neighbour distance distribution between the two networks, bounded between 5 and 20 metres. This range accommodates the typical scale of georeferencing imprecision while preserving genuine differences in network extent. In practice, thresholds range from approximately 10 to 18 metres across years, reflecting variation in scan quality and original map precision.

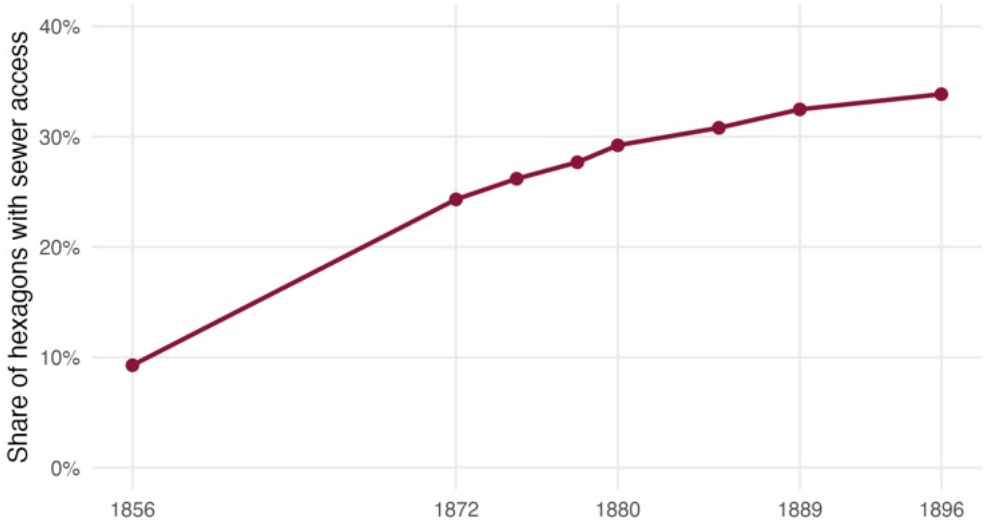
Figure A3 illustrates the resulting classification for 1856, 1872, and 1889. Year-only lines (green) are sparse in all three cross-sections. Moreover, lines detected in a given year but absent from the 1896 network do not consistently reappear in later cross-sections: lines classified as year-only in 1872, for instance, are largely absent in 1889. Since physical sewers, once constructed, remained in place for decades, such intermittent detections are more plausibly explained by noise, arising from variation in ink colour, scan quality, or printing artefacts, than by genuine infrastructure that was subsequently demolished.

Figure A3: Harmonised sewer networks in 1856, 1872 and 1889



Notes: The three panels show the harmonised sewer network for 1856, 1872, and 1889 relative to the 1896 anchor year. Maroon lines indicate overlap: lines present in the respective year that align within the calibrated distance threshold of the 1896 network. Green lines indicate year-only: lines present in the respective year with no nearby 1896 counterpart. Faint grey shows the full 1896 reference network.

Figure A4: Sewer access over time



Notes: Share of 25-metre hexagons in Paris with access to sewers, by year of the sewer maps. A hexagon has access to sewers if it is crossed by a sewer line that is matched to the 1896 network, and access is cumulative.

Sewer access. We define a hexagon as having access to sewers in a given year if it is crossed by a line of that year that is matched to the 1896 network. Since sewers, once constructed, remained in place, we treat sewer access as permanent, so that a hexagon that is classified as sewered in one cross-section remains sewered in all later ones. This cumulative definition also reduces sensitivity to differences in detection across map editions. Figure A4 shows the resulting expansion of sewer access. The share of hexagons with access to sewers rises from 9% in 1856 to 24% in 1872 and 34% in 1896.

A.1.2 Historical addresses

To geo-locate the addresses recorded in historical documents, we construct a reference register of the Paris street network as it existed in 1880. The challenge is that while the physical layout of Parisian roads has remained largely stable over the past 150 years, street names occasionally changed. Our approach is to take the contemporary street network as a starting point and systematically verify, for each street, whether its name and physical extent match those recorded in 1880 to obtain a reference register for 1880, which falls in the middle of our study period.

Street network. We take the contemporary Paris street network and manually inspect each of the more than 6,000 streets against the arrondissement-level maps of the sewerage network from 1880 that we georeferenced, which also show the street network and street names (*Atlas administratif des égouts de la ville de Paris*). We classify each street into one of six edit types, summarised in Table A3. In 78% of cases, the street has neither changed its name nor its layout since 1880 (Type 1). In a further 12%, the street layout is unchanged but the name has been updated (Type 2). The remaining cases require more substantive adjustments: 7% of streets today correspond to a single 1880 road that has since been split into several segments (Type 3), while 2% are the result of several 1880 roads being merged into a single contemporary road (Type 4). A further 1% of streets present in the 1880 network no longer exist today (Type 5), while a small residual of contemporary roads were created after 1880 and have no 1880 equivalent (Type 6). Figure A5 maps the spatial distribution of edit types across Paris and reports the share of the total street network in each category.

House numbers. We assign house numbers to the reconstructed 1880 street network using contemporary address data from the Base Adresse Nationale, the official French national address register. Paris has operated a stable house-numbering convention since the early nineteenth century, under which even numbers run along the right-hand side of a street and odd numbers along the left, in ascending order. The direction of travel, and thus the starting point of the numbering, is determined by the Seine. On streets running parallel to the river, numbers increase in the direction of the river's flow; on streets running perpendicular to it, numbers increase with distance from the river. This convention has remained unchanged since 1880, meaning that contemporary house numbers provide a reliable proxy for the location of historical addresses along a given street.

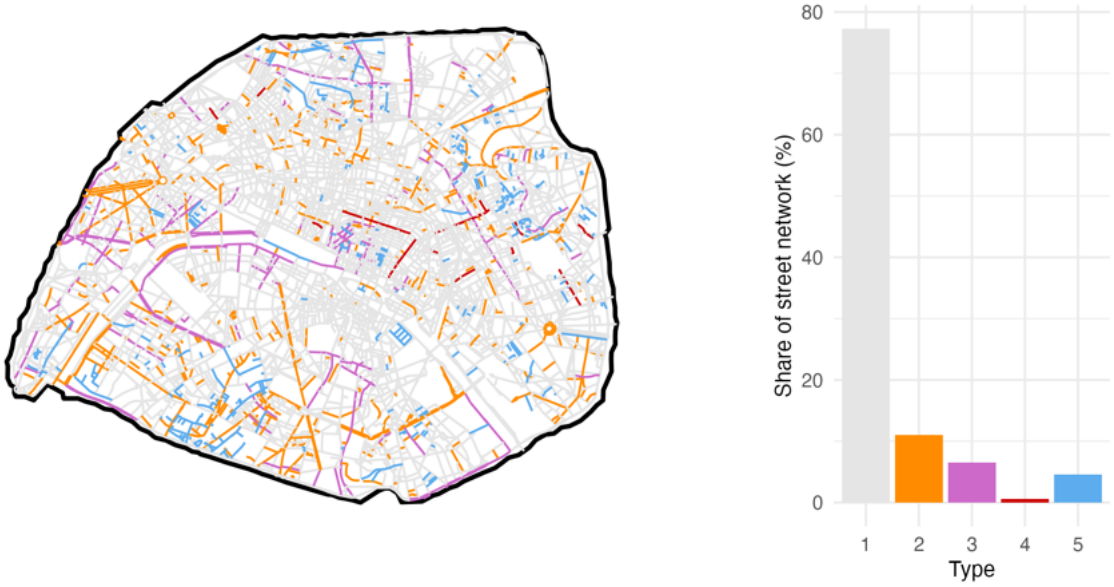
For streets whose name and layout have not changed (Types 1 and 2), we assign contemporary house numbers directly. For streets where a single 1880 road has since been split into several segments (Type 3), we pool all house numbers from the corresponding modern

Table A3: Street edit types

Type	Scenario and adjustments
1	*Road is still at the same place and name has not changed.* **Road name:** No changes required. **House number:** Today's house numbers used directly.
2	*Road is still at the same place but name has changed.* **Road name:** Road name updated to 1880 name. **House number:** Today's house numbers used directly.
3	*A single road in 1880 has since been split into several roads today.* **Road name:** All resulting segments assigned the 1880 road name. **House number:** House numbers from all today-segments pooled and renumbered sequentially along the 1880 road.
4	*Multiple roads in 1880 have since been combined into one road today.* **Road name:** Today's road replaced by the original 1880 road segments. **House number:** Each today house number spatially assigned to its nearest 1880 segment and renumbered within that segment.
5	*A road present in 1880 no longer exists today.* **Road name:** Road added back to today's network with its 1880 name. **House number:** Assigned ex-post using available historical records.
6	*A road exists today but had no equivalent in 1880.* **Road name:** Road excluded from the 1880 reference register. **House number:** Addresses on this road are not assigned an 1880 location.

Notes: Classification of the more than 6,000 streets in the contemporary Paris street network according to how they compare to the 1880 road network. Each street was manually inspected against arrondissement-level maps from 1880.

Figure A5: Distribution of street edit types



Notes: The left panel maps the spatial distribution of street edit types across Paris. The right panel shows the share of the total street network length in 1880 accounted for by each type. Type 1 (grey): road name and layout unchanged. Type 2 (orange): layout unchanged, name updated. Type 3 (purple): one 1880 road split into multiple today-roads. Type 4 (red): multiple 1880 roads merged into one today-road. Type 5 (blue): roads present in 1880 but absent today.

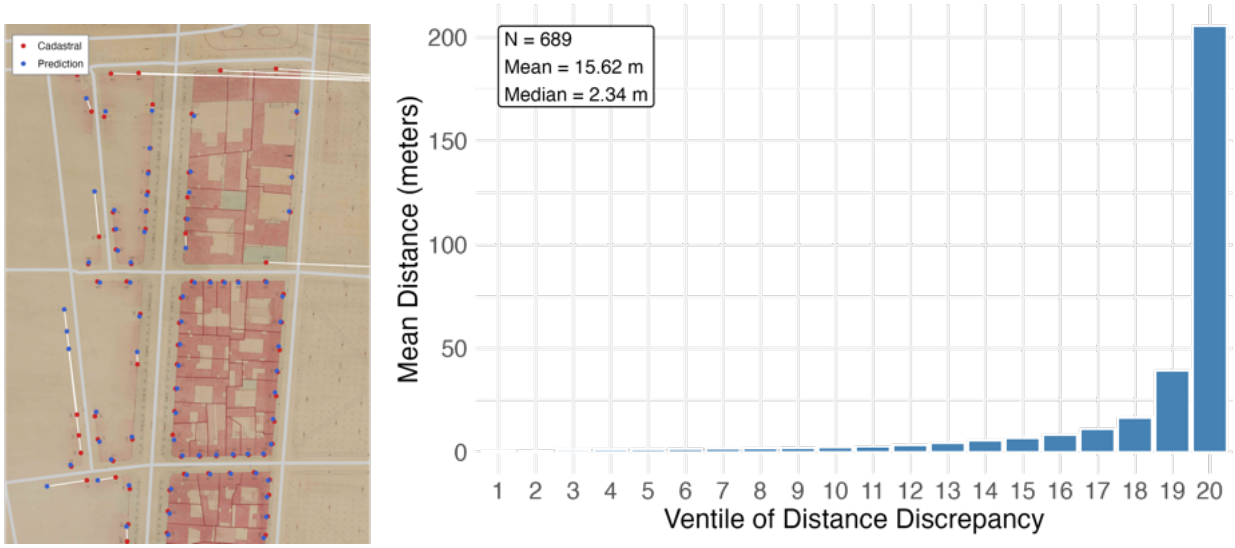
segments and renumber them sequentially along the original 1880 road. For streets where several 1880 roads have been merged into a single contemporary road (Type 4), we assign each contemporary house number to its nearest original 1880 segment, and then renumber addresses within each segment.

Because the sewer expansion was determined primarily at the street level, small errors in the precise location of individual house numbers do not materially affect whether an address is classified as sewered or unsewered. Nevertheless, we formally validate the accuracy of the geo-location procedure, as described below.

Validation. To assess how accurately contemporary house numbers reproduce the locations of historical addresses, we manually digitise house numbers from four selected cadastral maps dating from the second half of the nineteenth century. These maps display house numbers and serve as ground truth. We randomly select four maps, two from inner arrondissements and two from outer arrondissements, yielding a total of 689 digitised addresses. For each cadastral address, we identify the closest match in our reference register on the basis of street name and house number, and compute the spatial distance between the two locations. Figure A6 presents the results. Panel (a) shows an example overlay for one of the four sample areas, with cadastral house-number locations in red, our reconstructed locations in blue,

Figure A6: Validation of the address geo-location procedure

(a) *Cadastral map* **(b)** *Distribution of spatial discrepancies*



Notes: Panel (a) shows an example overlay for one of the four cadastral sample areas (7th arrondissement). Red points indicate house-number locations digitised from the cadastral map; blue points show the corresponding locations in our reference register; white lines connect matched pairs. Panel (b) reports the mean spatial discrepancy within each ventile of the discrepancy distribution, pooling all four sample areas. The validation sample comprises 689 addresses across two inner and two outer arrondissement sample areas.

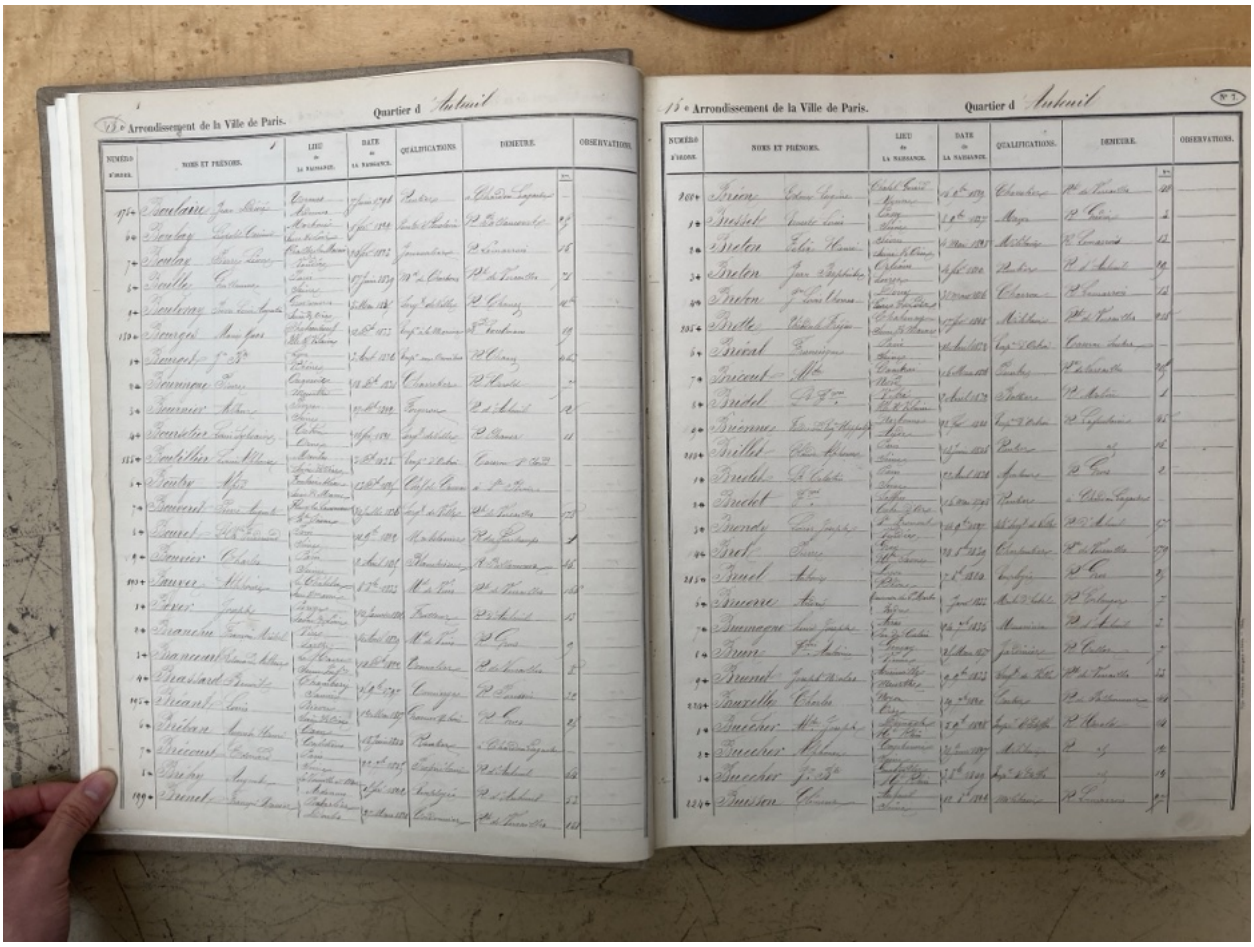
and connecting lines indicating matched pairs. Panel (b) reports the distribution of spatial discrepancies across the full validation sample, expressed as mean distance within each ventile. The median discrepancy is 2.34 metres and the mean is 15.62 metres. Up to the 90th percentile, distances remain well below 20 metres. The small number of larger discrepancies arises predominantly at street intersections, where house numbers have occasionally shifted from one side of a junction to the other over time.

A.1.3 Electoral lists

Electoral lists (*listes électorales politiques*) are administrative registers maintained by the Paris municipal government and preserved at the Archives de Paris. Following the introduction of universal male suffrage in 1848, these registers recorded all eligible voters and provide individual-level information including name, date and place of birth, occupation, and residential address. After 1852, recording a residential address became mandatory, allowing individuals to be located at the street and house-number level.

The electoral lists have been preserved from 1848 onwards, although coverage is incomplete during the nineteenth century. The years of 1871 and 1892, are the only ones during the 19th century for which lists survive for all twenty Paris arrondissements. For most other years,

Figure A7: Electoral list sample page



Notes: Sample image from an 1880 electoral list. The columns record, the running number, name, place of birth, date of birth, occupation (*qualifications*), address and observations.

coverage is limited to a subset of arrondissements for which the corresponding registers have survived. The registers are classified in the D1M2 series of the Archives de Paris. Each volume corresponds to a single arrondissement and year, with voters listed alphabetically by surname. Figure A7 shows a sample page.

Because the records contain individual residential addresses, they provide a unique source for constructing spatially disaggregated population data and complement census statistics, which are available only at the quartier level. The population recorded in the electoral lists represents adult men aged 21 or older. Despite this restriction, voter counts provide the best available proxy for the spatial distribution of population density within the city.

Photographing protocol. We visited the Archives de Paris on multiple occasions to photograph electoral list registers. Images were captured following a standardised protocol

Table A4: Electoral list documents digitised

Register year	Arrondissements	Volumes	Pages	Entries	Geolocated (%)
1853	7, 8	3	2,122	44,607	71.0
1859/60	1, 2, 3, 7, 9, 12	10	5,254	117,631	73.2
1871	1–20	20	15,264	343,760	74.4
1880	1, 3, 6, 9, 12, 15	10	5,040	115,719	80.7
1892	1–20	44	27,170	620,319	79.4
Total		**87**	**54,850**	**1,242,036**	**77.2**

Notes: Documents digitised by year. Pages are counts of single pages (left and right side of each double-page). Entries are transcribed rows carrying individual level data. Geolocated is the share of entries matched to a house-number coordinate in the 1880 address register.

whereby registers were photographed using a camera mounted on a tripod. In total, we photographed 87 register volumes spanning five electoral years (1853, 1859–1860, 1871, 1880 and 1892), yielding 54,850 page images. Table A4 summarises the coverage of our digitisation effort.

Transcription with a vision–language model. We transcribe the registers with a vision–language model, Claude Sonnet 5 (Anthropic), accessed through the Anthropic API. The model reads each page image and returns one record per row of the register. Before transcription, each photographed two-page spread is deskewed, and split into its two pages. Each page is sent as one request, which contains the page image and the fixed prompt shown below. For each field, the model returns its transcription, and a confidence score between 0 and 1. Abbreviations are expanded in the transcription, so that "R gravilliers" becomes "Rue des Gravilliers". The instruction asks the model to leave a field empty rather than guess when it cannot read it. The model detects a total of 1,285,244 rows, of which 43,208 are empty lines in the tables, while the remaining 1,242,036 rows are the true entries we use for our analysis.

Prompt

```
You transcribe 19th-century Parisian electoral registers.

You are given one page of a ruled table. Each page has roughly 25 numbered rows
and these columns, left to right:

  numéro d'ordre | noms et prénoms | lieu de naissance | date de naissance |
  qualifications | demeure (street name, then house number) | observations

The hand is French cursive in iron gall ink, legible but heavily abbreviated.

CONVENTIONS IN THE SOURCE
  7bre / 8bre / 9bre / Xbre    septembre / octobre / novembre / décembre
```

```
  Jer, Jr / Jllet, Jllt        janvier / juillet
  1er                          premier (day 1)
  "id."                        same as the row above
  o.                           ouvrier
  f.                           fabricant
  Md                           marchand
  g., gn, go.                  garçon de
  Empé Comal, Employé Comal    employé de commerce
  Cis, Comis                   commis
  voygr, Vojgr                 voyageur
  Nde Nazareth                 Notre-Dame-de-Nazareth
  St, Ste                      Saint, Sainte
  Birthplaces                  usually "commune, département"

WHAT TO RETURN

One object per ruled row, in top-to-bottom order, including rows that are blank
or struck through | return those with null fields rather than skipping them.
Row numbering starts at 1 at the top of the page and counts every ruled line.

Every field is an object with up to three keys:

  "v"  the normalised reading. Expand abbreviations here: "o. bijoutier"
       becomes "Ouvrier bijoutier", "Empé Comal" becomes "Employé de commerce",
       "St Malo" becomes "Saint-Malo". Use null if you cannot read it.

  "w"  the literal glyphs as they appear on the page, before any expansion or
       correction. OMIT THIS KEY ENTIRELY when it would be identical to "v".

  "c"  a number from 0 to 1 | how confident you are in the READING, not in the
       expansion. 1 means certain, 0 means a pure guess. OMIT THIS KEY ENTIRELY
       when it would be 0.95 or above; anything you are sure of needs no number.

       Use the range, not three favourite values. 0.62 and 0.78 are both useful
       and mean different things; rounding everything to 0.9 / 0.7 / 0.4 throws
       the information away. Anchors:

         0.95+  certain | clean glyphs, no competing reading (omit the key)
         0.85   confident, but the hand is untidy or a letter is ambiguous
         0.70   probably right; one plausible alternative reading exists
         0.50   a coin flip between two readings
         0.30   barely legible, mostly inference from context
         0.10   a guess you would not defend

So a clean, obvious entry is just {"v": "Paris"}. A hard one is
{"v": "Toulouse", "w": "Coulouer", "c": 0.35}. A legible-looking entry you are
still unsure of is {"v": "Abric", "c": 0.72} | confidence is about doubt, not
about how much you changed.

RULES

1. Never invent a plausible French surname, commune or trade to fill a gap. If
   you cannot read it: "v" is null, "c" is low (0.2 or less), and "w" holds
   are visible. A null is far more useful than a confident guess | downstream
   code cannot tell a guess from a reading, but it handles nulls correctly.

2. Confidence is about legibility. A heavily abbreviated but perfectly clear
```

```
   "Empé Comal" scores 0.95+ and the key is omitted. A cleanly written word that
   could be one of two names does not, however neat it looks.

3. Resolve "id." against the row above and set "w" to "id." so the substitution
   stays visible.

4. Do not compute the ISO date. Return the date exactly as written in "dt.raw",
   preserving the register's own abbreviations ("3 7bre 1844", "13 Jer 1818") |
   Stage 3 parses them. Read the day and month digits carefully: month
   abbreviations 7bre/8bre/9bre/Xbre are easily confused with each other and
   with ordinary numerals.

5. demeure has two parts, "rue" and "num".

   "rue" MUST be the COMPLETE street designation including its type and any
   article, exactly as it would be written on a modern envelope:
       page says "R gravilliers"   -> v: "Rue des Gravilliers"   w: "R gravilliers"
       page says "R Nde Nazareth"  -> v: "Rue Notre-Dame-de-Nazareth"
       page says "Bd Sebastopol"   -> v: "Boulevard de Sébastopol"
       page says "R St Martin"     -> v: "Rue Saint-Martin"
   Never return the bare name ("Gravilliers"). Expand R. to Rue, Bd to
   Boulevard, and restore the article (des, de la, du, de) that the clerk
   omitted. If you are unsure which article belongs, give your best reading and
   mark "c" accordingly | a downstream matcher handles small differences, but a
   missing street type breaks it.

   "num" is the house number alone. If it is struck through or illegible, "v"
   is null. House numbers are the least legible field in these registers and
   the one the project depends on most | read them carefully, and prefer null
   with "c": "low" over a confident guess.

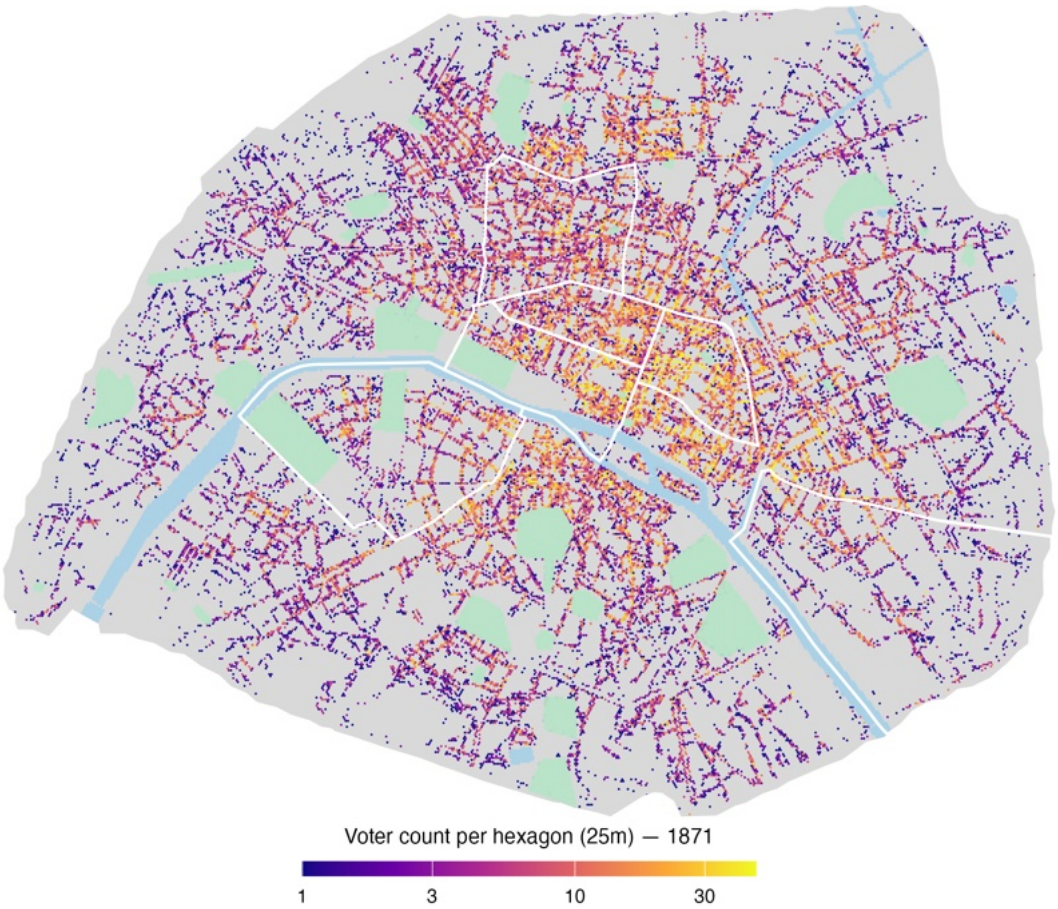
6. observations is usually empty. Return {"v": null} when it is.

7. "n" is the numero d'ordre from the leftmost column | the clerk's own running
   number. It is a plain integer, not a field object: no "w", no "c". It runs
   continuously through the volume and does NOT restart at 1 on each page, so a
   page may well begin at 476. Copy what is written; never renumber to match "r"
   and never infer it from position. If a row has no number, use null.

Return only rows you can see on this page. Do not carry over from another page.
```

Geolocation. We geolocate each entry by matching its address to our reference register of streets and house numbers in 1880 (Appendix A.1.2). We first match the street name to the streets of the same arrondissement, after translating street names that changed before 1880 into their 1880 names. We then look up the house number on the matched street. Each matched entry receives the coordinates of its address and is assigned to its quartier and to the 25 m hexagon used throughout the paper. Table A5 reports the results for each volume. Across all 87 volumes, we geolocate 77.2% of entries. Rates range from 71% in 1853 to 81% in 1880 and are similar in the two years of our main analysis, at 73.2% in 1859/60 and 74.4% in

Figure A8: Resident density across Paris (1871)



Notes: Number of geolocated voters in each 25 m hexagon in 1871, across all twenty arrondissements of Paris. White lines mark the six arrondissements of the main difference-in-differences sample.

1871. Figure A8 shows the result for 1871, when lists survive for all twenty arrondissements, as the number of geolocated voters in each 25 m hexagon across Paris.

Table A5: Electoral list digitisation output by volume

Volume	Year	Arr.	Pages	Entries	% Street	% Geo.	Volume	Year	Arr.	Pages	Entries	% Street	% Geo.
D1M2 31	1853	8	1,024	21,301	80.1	71.8	D1M2 454	1892	1	512	11,559	86.6	76.9
D1M2 32	1853	7	578	13,044	76.0	58.9	D1M2 455	1892	2	462	10,289	87.4	85.1
D1M2 33	1853	7	520	10,262	90.0	84.8	D1M2 456	1892	2	424	9,269	93.9	84.7
D1M2 52	1859	1	522	11,690	82.0	70.2	D1M2 457	1892	3	624	13,765	87.4	79.5
D1M2 53	1859	1	600	13,172	84.5	74.1	D1M2 458	1892	3	610	13,704	85.7	79.3
D1M2 54	1859	1	464	8,372	77.8	69.4	D1M2 459	1892	4	654	14,726	79.8	70.4
D1M2 58	1860	2	400	9,353	80.8	67.3	D1M2 460	1892	4	642	14,644	82.7	72.3
D1M2 59	1860	2	340	8,000	89.0	83.1	D1M2 461	1892	5	412	9,079	92.0	79.2
D1M2 60	1860	3	846	20,149	83.3	76.9	D1M2 462	1892	5	724	16,711	82.5	79.3
D1M2 61	1860	7	828	17,298	83.7	61.0	D1M2 463	1892	6	612	14,477	87.6	86.3
D1M2 62	1860	9	440	10,446	79.9	76.2	D1M2 464	1892	6	608	13,495	91.8	87.3
D1M2 63	1860	9	388	9,058	82.5	81.9	D1M2 465	1892	7	506	11,375	85.2	86.9
D1M2 64	1860	12	426	10,093	79.0	79.2	D1M2 466	1892	7	390	9,110	85.2	86.2
D1M2 166	1871	1	868	20,053	84.6	70.5	D1M2 467	1892	8	484	10,247	90.7	87.7
D1M2 168	1871	2	842	19,261	78.0	66.3	D1M2 468	1892	8	490	11,212	92.3	89.9
D1M2 170	1871	3	1,020	23,099	89.9	84.5	D1M2 469	1892	9	640	14,247	87.8	88.2
D1M2 172	1871	4	994	22,832	70.5	60.2	D1M2 470	1892	9	640	14,336	87.9	88.5
D1M2 174	1871	5	882	19,772	78.2	66.6	D1M2 471	1892	10	506	11,970	71.5	79.3
D1M2 176	1871	6	948	21,495	83.9	74.1	D1M2 472	1892	10	804	18,440	88.9	85.0
D1M2 179	1871	7	764	16,685	73.1	72.7	D1M2 473	1892	10	452	10,571	88.0	84.7
D1M2 180	1871	8	710	15,418	81.8	77.0	D1M2 474	1892	11	514	12,018	79.7	77.4
D1M2 182	1871	9	904	20,959	84.7	84.2	D1M2 475	1892	11	760	18,145	81.0	80.7
D1M2 184	1871	10	968	22,756	81.6	80.9	D1M2 476	1892	11	690	16,146	79.3	67.9
D1M2 186	1871	11	1,328	31,353	81.3	75.0	D1M2 477	1892	11	762	17,885	78.5	82.4
D1M2 188	1871	12	520	10,831	79.0	78.2	D1M2 478	1892	12	716	16,228	87.1	84.0
D1M2 190	1871	13	502	10,910	80.0	71.9	D1M2 479	1892	12	714	16,058	75.4	77.3
D1M2 192	1871	14	476	10,070	71.3	71.9	D1M2 480	1892	13	686	15,589	74.0	77.1
D1M2 194	1871	15	640	14,620	73.9	72.6	D1M2 481	1892	13	654	14,826	82.1	73.5
D1M2 195	1871	16	386	8,344	73.4	72.1	D1M2 482	1892	14	710	16,410	81.9	64.6
D1M2 198	1871	17	472	10,626	77.0	81.5	D1M2 483	1892	14	712	16,462	81.9	71.0
D1M2 200	1871	18	996	23,087	83.4	82.9	D1M2 484	1892	15	654	14,832	81.2	76.3
D1M2 202	1871	19	490	9,554	82.7	68.7	D1M2 485	1892	15	704	16,789	79.2	73.9
D1M2 204	1871	20	554	12,035	74.6	69.4	D1M2 486	1892	16	466	10,537	81.3	80.8
D1M2 297	1880	1	626	14,567	83.2	73.0	D1M2 487	1892	16	440	9,769	74.7	77.5
D1M2 298	1880	1	416	6,841	82.6	69.6	D1M2 488	1892	17	604	14,065	80.3	81.6
D1M2 299	1880	3	840	20,384	92.7	83.9	D1M2 489	1892	17	674	15,624	87.4	86.5
D1M2 300	1880	6	414	9,586	89.8	83.6	D1M2 490	1892	17	720	16,218	80.3	83.2
D1M2 301	1880	6	418	9,667	89.0	83.2	D1M2 491	1892	18	646	14,795	86.1	81.4
D1M2 302	1880	9	484	11,695	84.3	85.8	D1M2 492	1892	18	772	17,673	80.7	80.0
D1M2 303	1880	9	416	9,828	85.9	86.1	D1M2 493	1892	19	790	17,930	85.6	71.6
D1M2 304	1880	12	390	9,257	86.8	83.8	D1M2 494	1892	19	472	10,850	78.4	72.5
D1M2 305	1880	12	326	6,817	85.4	82.5	D1M2 495	1892	20	860	19,902	82.7	75.7
D1M2 306	1880	15	710	17,077	80.3	76.2	D1M2 496	1892	20	906	21,266	84.5	77.2
D1M2 453	1892	1	348	7,076	86.3	77.5							

Notes: Digitisation output for each volume of the electoral lists. % Street is the share of entries whose street name we match to a street in the 1880 register. % Geo. is the share of entries we match to the coordinates of a house number.

Table A6: Social class of residents

		1859/60		1871	
	Class	n	%	n	%
1	Elite (higher managers and professionals)	18,692	16.0	13,422	12.2
2	Lower middle class	28,098	24.0	27,846	25.3
3	Skilled workers (medium and lower skilled)	45,712	39.0	42,285	38.4
4	Farmers and fishermen	148	0.1	142	0.1
5	Unskilled workers and farm workers	5,717	4.9	4,581	4.2
	No class assigned	18,754	16.0	21,965	19.9
	All with a qualification	**117,121**	100.0	**110,241**	100.0

Notes: Summary statistics for the sixteen volumes of the main difference-in-differences sample. HISCLASS collapsed into five groups. Percentages are of all entries with an occupation. Entries without a class mostly record no trade, such as "employé", or a trade that could not be coded.

Social class. The registers record the occupation of each elector in a column headed *qualifications*. We use it to assign each elector to a social class in three steps. First, we separate the trade from the position. An entry such as "ouvrier bijoutier" records both a trade (jeweller) and a position within it (journeyman). We separate the two because clerks differed in how often they wrote down the position. Second, we assign each trade a code from the Historical International Standard Classification of Occupations (HISCO; van Leeuwen et al. 2002). We do this with a language model, Claude Opus 5, which receives the trade together with the full list of HISCO codes and must return one of them. Each distinct trade is coded only once, so the same trade always receives the same code. In total, about 13,600 distinct trades are mapped to 881 codes. Third, we combine the HISCO code with the position to obtain the social class of the elector in the HISCLASS scheme (van Leeuwen and Maas 2011). This scheme separates, for example, a master from a journeyman in the same trade. We group its thirteen classes into five broad groups, ranging from an elite of higher managers and professionals to unskilled workers. Table A6 shows that the class composition of electors is similar in 1859/60 and 1871. The main difference is a fall in the share of the elite from 16% to 12%. This is mostly because fewer electors in 1871 are recorded as rentiers or property owners, whom HISCLASS counts as elite.

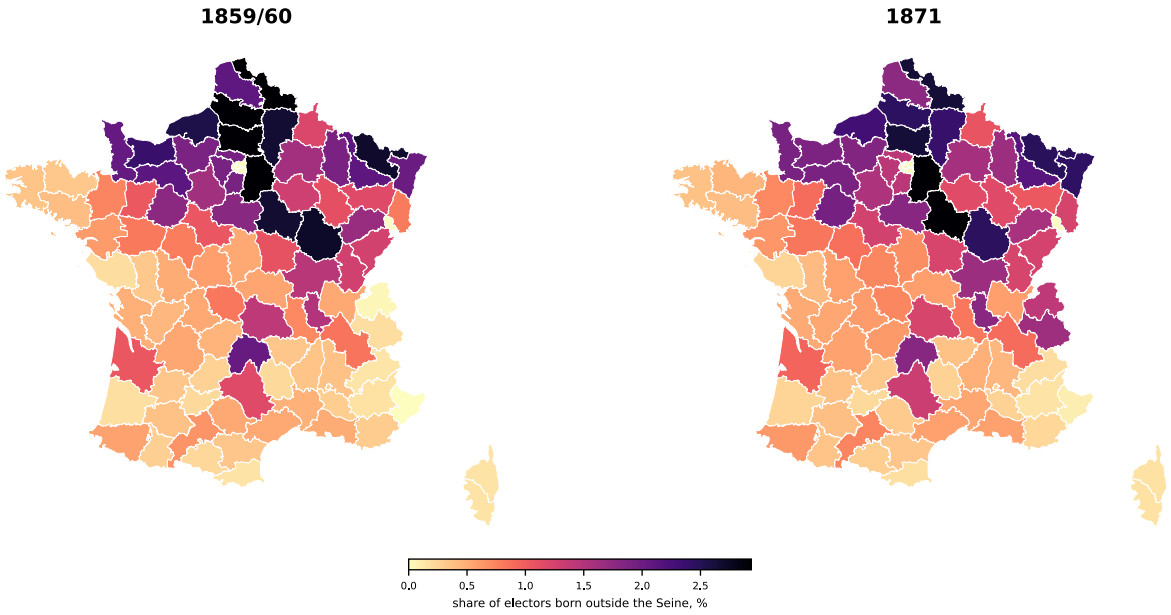
Migrants. The registers also record where each elector was born. We classify electors as born in Paris, born elsewhere in the Seine département, or born outside the Seine. Table A7 shows that about three-quarters of electors were born outside the Seine, and that this share is the same in 1859/60 and 1871. Figure A9 shows the départements from which these migrants moved to Paris. Most came from northern and eastern France, in particular from the départements around Paris.

Table A7: Birthplace of residents

	1859/60		1871	
Birthplace	n	%	n	%
Paris	24,963	24.5	27,082	24.6
Seine, outside Paris	2,057	2.0	2,619	2.4
Outside the Seine	74,826	73.5	80,552	73.1
With a recorded birthplace	**101,846**	100.0	**110,253**	100.0
No birthplace recorded	16,131		787	

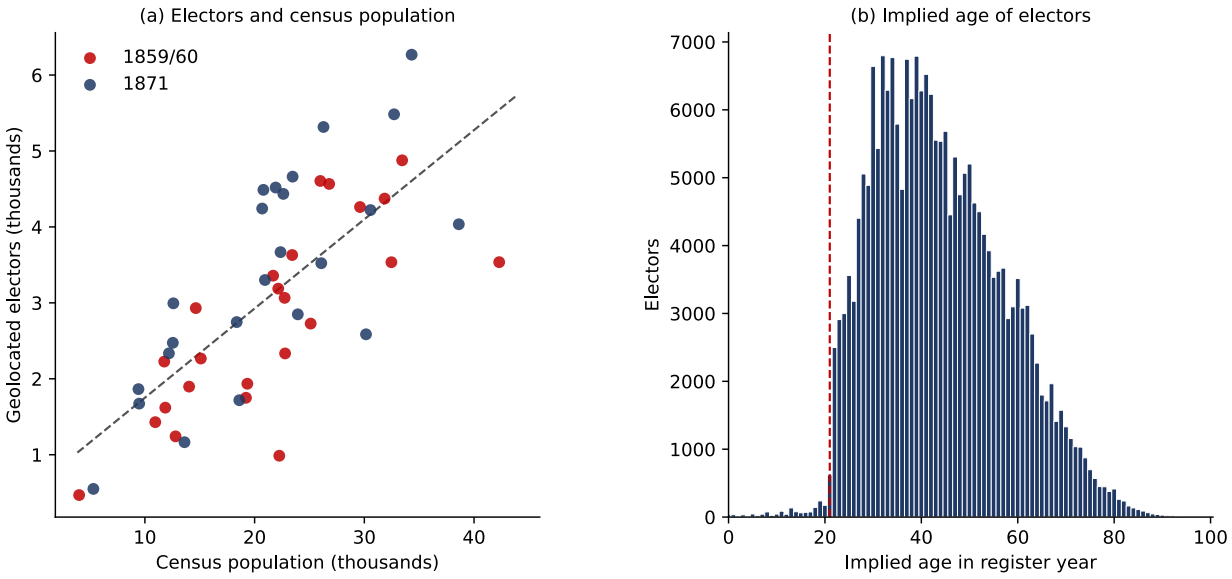
Notes: Summary statistics for the sixteen volumes of the main difference-in-differences sample. Birthplace is classified from the commune and département of birth. Percentages are of electors with a recorded birthplace.

Figure A9: Department of birth of residents born outside of Paris



Notes: Share of electors born outside the Seine department by département of birth, 1859/60 (left) and 1871 (right).

Figure A10: Validation of the electoral lists digitisation



Notes: The sixteen volumes of the main difference-in-differences sample. Panel (a) plots geolocated electors against the census population for each quartier and register year; the 1859/60 registers are compared with the 1861 census and the 1871 registers with the 1872 census. The dashed line is a linear fit over all points. Panel (b) shows the implied age of electors, the register year minus the transcribed year of birth. The dashed line marks age 21, the minimum age for registering as a voter.

Validation. We validate the transcription and geolocation in two ways. First, we compare the number of geolocated electors with the census population at the quartier level. The census is available in 1861 and 1872, so we compare the 1859/60 registers with the 1861 census and the 1871 registers with the 1872 census. Panel (a) of Figure A10 shows a close relationship across the 48 quartier-years of our main sample, with an R^2 of 0.54. Geolocated electors correspond to 13.2% of the census population in 1859/60 and 16.3% in 1871. This is somewhat below the share of adult male citizens in the population, as expected given that not all entries can be geolocated.

Second, we use the transcribed date of birth as check for the overall quality of the transcription. Only men aged 21 or over could register, but nothing in our procedure constrains the transcription to be limited to certain age ranges. Panel (b) of Figure A10 shows the transcribed age of each elector in the year of the register. The distribution starts sharply at age 21, and only 0.67% of electors in our main sample fall below it.

Table A8: Sommier Foncier digitisation coverage

Series	Period	Inventory range	Eichholtz et al. (2021)	Own digitisation
1	1809–1859	1–365	110	—
2	1859–1880	366–1,169	87	34
3	1880–1945	1,170–2,012	130	—

Notes: The Sommier Foncier consists of three series of register books. Counts reflect the number of unique register books for which at least one transaction appears in the final dataset. Eichholtz et al. (2021) draw on all three series; our own digitisation covers a subset of Series 2 registers.

A.1.4 Sommier Foncier

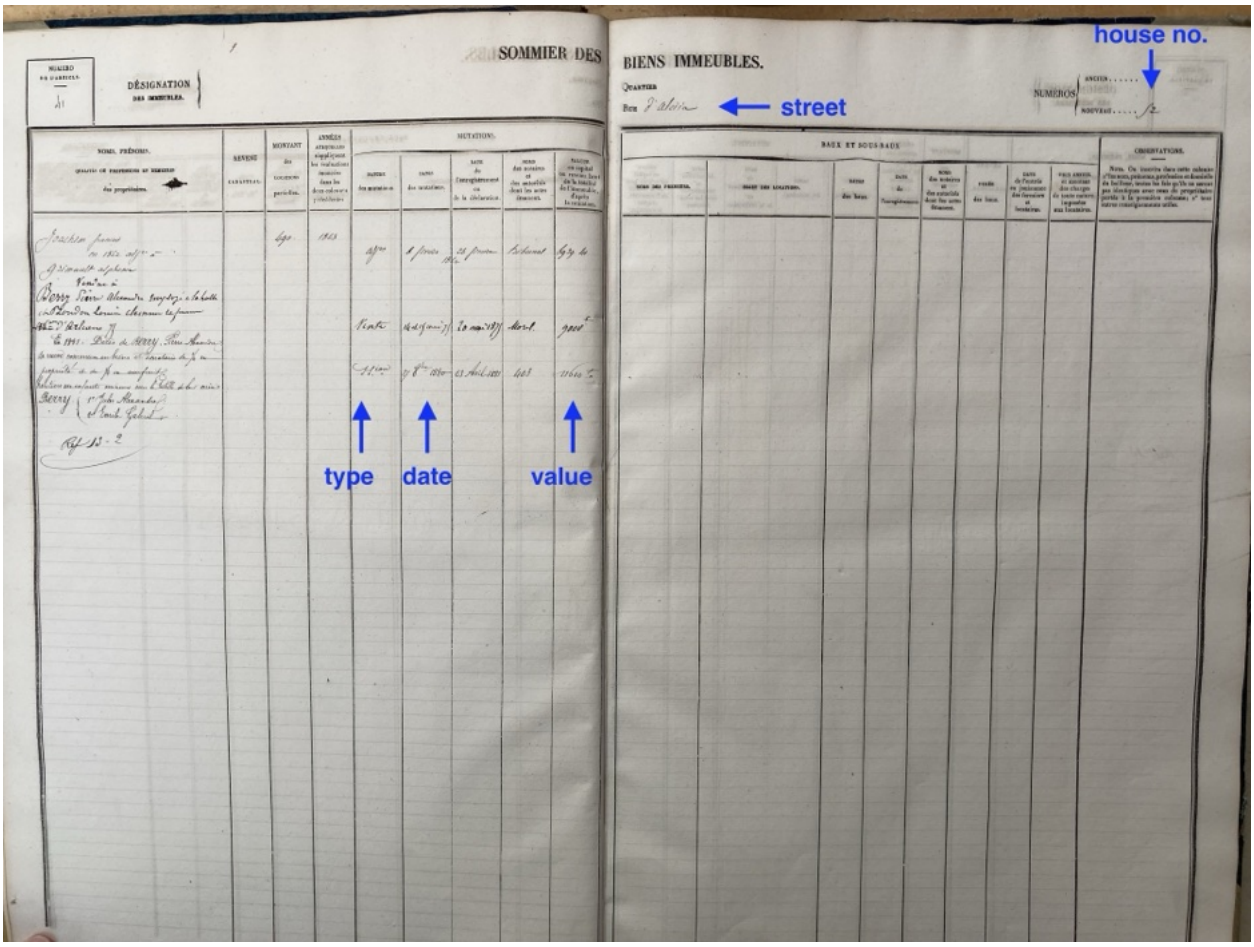
The *Sommier Foncier* is the historical land registry of Paris, maintained by the French administration as an instrument for verifying the accuracy of tax declarations on property transfers during the period from 1809 to 1943. The registers are held at the *Archives de Paris* and are organised into three chronological series: Series 1 covers 1809–1859 (DQ18 volumes 1–365), Series 2 covers 1859–1880 (DQ18 volumes 366–1,169), and Series 3 covers 1880–1945 (DQ18 volumes 1,170–2,012). Each volume is a physical register book in which each double-page records the transaction history of a single property for the respective period.

Our digitisation efforts focus on expanding a data set constructed by Eichholtz et al. (2021) who digitised approximately a 15% random sample of the complete registry, drawing on all three series and targeting properties in the historical centre of Paris. Table A8 documents the number of distinct register volumes digitised by Eichholtz et al. (2021). Our own digitisation covers 34 additional Series 2 volumes, extending coverage to the outer arrondissements for the 1859–1880 period, when the most substantial sewerage network expansion took place.

Figure A11 shows an example double-page from a Series 2 register book. The street name and house number are written at the top of the right page. The left page contains a five-column table (*mutations*), recording the type of transfer (Col. 1), the date of the transaction (Col. 2), and the transaction value in francs (Col. 5). The right page records information on lease arrangements, which we do not use in this paper.

Three transaction types account for the large majority of entries. *Ventes* are market sales in which seller and buyer agree a price. *Successions* record the transfer of property at death; the registered value is a taxable estate valuation rather than a negotiated market price. By law, this valuation equalled twenty times the annual rental value of the property for the period before 1918 (Eichholtz et al. 2021). *Adjudications* are auction sales conducted before a notary: the property is offered at a reserve price and sold to the highest bidder.

Figure A11: Example page from the Sommier Foncier (Series 2)



Notes: Example of a double-page from a Series 2 volume (DQ18 806). The left page records property transfers: each row corresponds to one transaction, with columns for the transfer type, date, and value in francs. The street name and house number of the property are recorded at the top of the right page. Blue annotations indicate the fields extracted during our digitisation process.

Auction sales were a regular and predominantly voluntary mechanism for selling investment property in Paris. Eichholtz et al. (2021) find that auction sales account for approximately 37% of all property sales in their sample, of which only around 13% were involuntary judicial sales arising from foreclosure or bankruptcy. Other transaction types encompass donations, income assignments (*revenu*), and partition deeds (*partage*).

Transaction prices differ systematically across types and between the two datasets. In the Eichholtz et al. (2021) Series 2 sample, the median transaction price is 60,000 francs for ventes, 82,000 francs for adjudications, and 8,000 francs for successions. The pattern is similar in our own digitisation, with medians of 16,800 francs, 29,600 francs, and 2,600 francs for ventes, adjudications, and successions respectively. The sizeable difference in price levels between the two datasets across all transaction types reflects the geographic composition of

Figure A12: Digitisation pipeline



Notes: Overview of the digitisation pipeline applied to 34 Series 2 register volumes (3,503 page images) of the *sommier foncier*. Workstream A (top) extracts and georeferences street addresses through a manual transcription tool and fuzzy matching against an 1880 house-number shapefile. Workstream B (bottom) detects table rows and column cells using two YOLOv8 models and transcribes their content using the Claude API. Both workstreams are merged in Part 4 to produce the master dataset.

each sample. While Eichholtz et al. (2021) draw primarily on inner arrondissements of the historical centre of Paris, our digitisation covers the outer arrondissements, where property values were considerably lower.

Digitisation approach. To acquire images of the register books, we visited the *Archives de Paris* on several occasions and photographed a total of 34 volumes. Images were captured following a standardised protocol, with a camera mounted on a tripod in a bird's-eye perspective to photograph each double-page in a volume (e.g. Figure A11). Pages were only skipped when they contained no pencil markings at all, while pages containing annotations but no recorded transactions were still photographed. The documents were subsequently digitised using a combination of automated and manual procedures. Figure A12 provides an overview of the digitisation pipeline.

Address transcription and georeferencing. Due to occasional inconsistencies in the placement of addresses on the page, we manually transcribed the address recorded on each page image. Because each double-page spread in the register corresponds to a single property, the address remains constant across all transactions recorded on that page, allowing us to transcribe one address per image. In addition, each volume typically covers only a small number of streets, albeit many properties, which substantially facilitated the manual

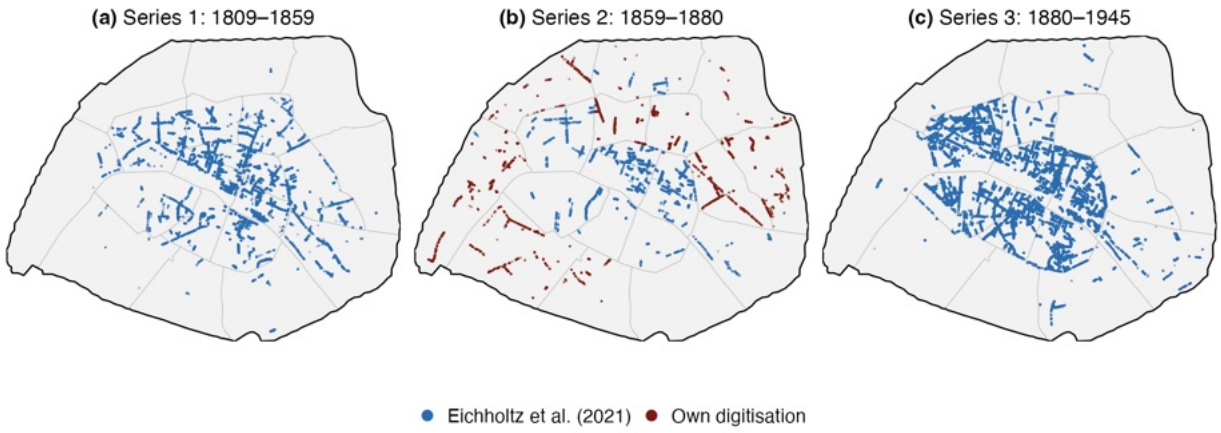
transcription process. The transcribed addresses were subsequently georeferenced using the 1880 address register we constructed (see Appendix A.1.2). Across the 3,503 double-page images, 53% were successfully geolocated to both a street name and a specific house number, while a further 32% were matched to a street without an exact house-number coordinate. The remaining 15% could not be matched geographically.

Table parsing and content transcription. To extract the transaction information, we trained two YOLOv8 oriented bounding-box (OBB) models in sequence. The first model detects individual table rows containing transactions within each page image; the second detects the five column cells within each detected row crop. This yields approximately 6,000 row crops across the 3,503 double-page images. We then used Anthropic's API to transcribe the content of three column cells from each row crop: transaction type (Col. 1), transaction year (Col. 2), and transaction value in francs (Col. 5). The API was instructed to return its best-guess value together with a confidence assessment, and to flag any crop it considered ambiguous or partially illegible rather than guessing. Flagged entries were subsequently reviewed manually. Across the three columns, the share of entries flagged for review was 27% for transaction type (1,484 entries reviewed), 33% for transaction year (1,674 entries reviewed), and 40% for transaction value (2,162 entries reviewed). Entries that remained unreadable after manual review were coded as missing.

Final merge. All outputs were merged using the image and row-crop identifiers to construct a single georeferenced transaction-level dataset. The resulting dataset contains 2,680 geolocated transactions, of which 1,919 could be matched to both a street and a specific house number. An additional 791 transactions could be matched to a street, but not to an exact house-number location (see Figure A12).

Data description. For our analysis, we use the combined dataset of Eichholtz et al. (2021) and our own digitisation efforts. The Eichholtz et al. (2021) data provide information on street and house number, but are only matched to quartiers; we therefore geolocate these observations using the same procedure as for our own data. Figure A13 documents the geographic coverage of the geolocated transactions across the three *Sommier Foncier* series. Panel (a) shows Series 1 transactions from Eichholtz et al. (2021), concentrated in the historical core of Paris. Panel (b) shows that the two datasets complement each other for Series 2: the Eichholtz et al. (2021) observations cluster in the inner arrondissements, while our digitisation extends coverage to the outer arrondissements incorporated into Paris following the 1860 annexation. Panel (c) shows Series 3 coverage from Eichholtz et al. (2021), which again focuses on the inner city.

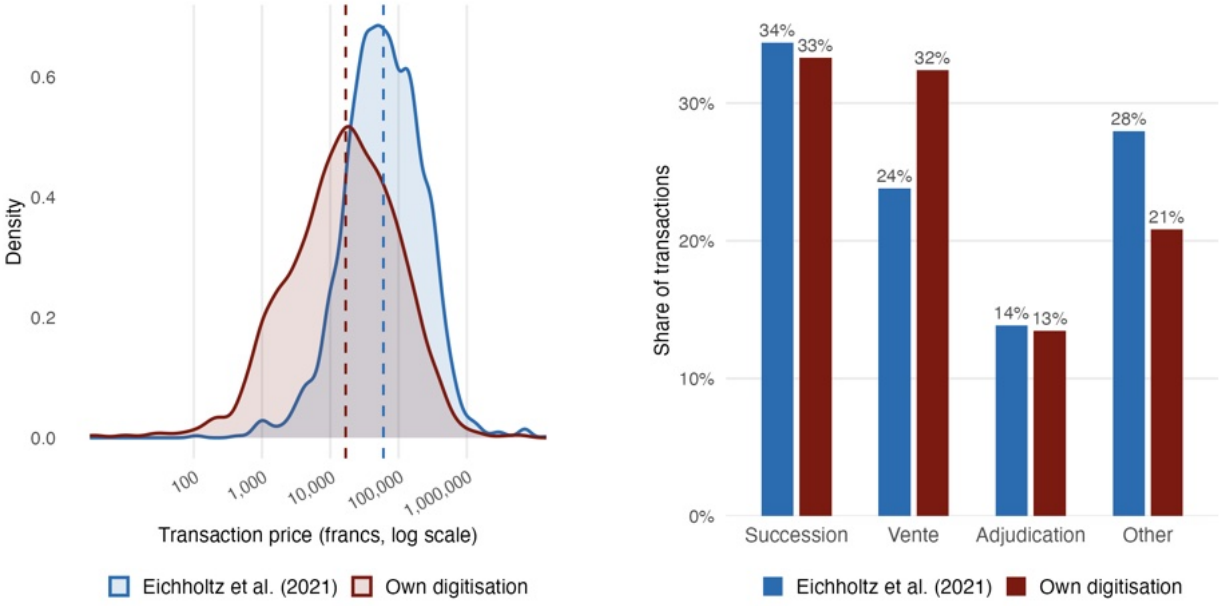
Figure A13: Geographic coverage of the Sommier Foncier datasets



Notes: Each point represents a geolocated property transaction. Blue: transactions from Eichholtz et al. (2021). Maroon: own digitisation.

Figure A14 characterises the two Series 2 datasets along two further dimensions. Panel (a) shows the distributions of the log transaction price, restricting both datasets to ventes to ensure comparability. The two distributions have a similar shape but the median in our digitisation (16,800 francs) is substantially below that of Eichholtz et al. (2021) (60,000 francs). As noted above, this gap is driven by differences in the spatial coverage. Panel (b) shows the composition of transaction types in each dataset. Ventes account for the largest share in both sources, followed by successions and adjudications. The composition is broadly comparable across the two datasets, suggesting that the type mix does not differ markedly between the inner and outer arrondissements during this period.

Figure A14: Price distribution and transaction type composition



(a) *Price distribution (ventes only)* **(b)** *Transaction type composition*

Notes: Both panels compare Series 2 data from Eichholtz et al. (2021) (blue) with our own digitisation (maroon). Panel (a) shows kernel density estimates of the log transaction price, restricting to ventes; dashed vertical lines mark the respective medians (Eichholtz: 60,000 francs; own digitisation: 16,800 francs). Panel (b) shows the share of transactions in each category, excluding entries with missing or unreadable transaction type.

A.1.5 Typhoid deaths

The *Annuaire statistique de la ville de Paris* published, for each year between 1880 and 1904, a map of Paris on which every death from typhoid fever recorded in that year is marked by a black dot at the address of the deceased. Appendix Figure A15 reproduces the 1890 map (Panel (a)) together with our digitised version (Panel (b)).

We digitise all 25 annual maps in three steps. First, we georeference each scanned sheet using the arrondissement and quartier boundaries and the fortification wall as control points. Second, we isolate the death markers by converting the sheet to greyscale and applying Otsu's method, which separates the dark dots from the lighter base map. Third, because the base map itself contains black elements (street names, the Seine's banks), we manually inspect and correct the extracted point set for every year, adding markers the thresholding misses and deleting spurious ones.

We then overlay the resulting point layer on our 25 m hexagon grid and count deaths within each hexagon. This yields a balanced hexagon–year panel of typhoid deaths for 1880–1904, containing 22,962 deaths in total. The annual count falls steadily over the period, from 1,777 deaths in 1880 and a peak of 3,203 in the 1882 epidemic year to 335 in 1904.

Figure A15: Typhoid death map 1890

(a) *Original map* **(b)** *Digitised deaths*



Notes: Panel (a) shows the 1890 map of typhoid deaths from the *Annuaire statistique de la ville de Paris*, on which each death is marked by a black dot at the address of the deceased. Panel (b) shows the deaths extracted from this map in our digitisation.

A.1.6 Commuting times

We construct historical measures of within-city travel times for Paris by digitising the horse-drawn omnibus and tram network operating in each cross-section. From scanned maps, we manually extract the network of each year as line shapefiles referenced to a common modern OpenStreetMap (OSM) road baseline. The underlying maps were produced by different publishers, at different scales and with different degrees of cartographic accuracy. Manual digitisation provides a single, internally consistent network.

Omnibuses and tramways in Paris. In 1826, Stanislas Baudry began running a horse-drawn coach between a public bathhouse he owned on the outskirts of Nantes and the city centre; the vehicle picked up passengers along the whole route rather than only at the two endpoints, and stopped in front of a shop bearing the sign *Omnes Omnibus* (Latin for "for all"), from which the service took its popular name. Baudry brought the idea to Paris in 1828, and several competing entrepreneurs quickly followed with their own lines. Competition among a large number of small operators persisted through the following decades until 1855, when the imperial administration consolidated the various omnibus companies into a single concessionaire, the *Compagnie Générale des Omnibus* (CGO), which was granted an exclusive right to operate fixed-route omnibus services in Paris (Papayanis 1996).

The same year, 1855, also saw the introduction of the tramway to Paris: Alphonse Loubat, drawing on the grooved-rail design he had observed in New York, built the first French tramway line along the Cours-la-Reine. Rail-borne tramways remained a marginal complement to the omnibus for the next two decades, but expanded rapidly from the 1870s once the wide, straight boulevards built under Haussmann's Second Empire renovation of Paris provided suitable rights-of-way for fixed track. Unlike the omnibus network, the tramway was never brought under a single concessionaire: our 1895 source map (Table A9) distinguishes lines operated by the CGO itself from those of at least two further major concessionaires, the *Compagnie des Tramways de Paris et du Département de la Seine* and the *Compagnie Générale Parisienne de Tramways*, together with a handful of smaller specialised operators running steam-hauled and funicular lines (Robert 1959).

The geographic extent of the network also tracks the administrative geography of Paris itself. The annexation of 1860, which absorbed the surrounding communes into the current 20 arrondissements, roughly doubled the city's official area; our c.1860 source map already reflects the CGO's extension of its routes into these newly incorporated outer districts, and later cross-sections show continued densification of coverage as the city grew (Pinkney 1958).

Figure A16: The Parisian Omnibus

(a) *1898* **(b)** *1899* **(c)** *1910*



Notes: Parisian omnibus carriages photographed by Eugène Atget in 1898, 1899 and 1910. All three vehicles are horse-drawn (two horses in 1898 and 1899, three in 1910) and share the double-deck body of the *Compagnie Générale des Omnibus* fleet, with an enclosed lower compartment and an open-air upper deck (*l'impériale*) reached by a rear staircase.

Horse traction remained the dominant technology for both omnibuses and trams for the rest of the century. Experiments with mechanical and electric traction began only in the 1890s, and the CGO did not begin operating motor omnibuses at scale until 1905; conventional accounts date the last day of horse-drawn omnibus service in Paris to January 1913, after which the network was fully motorised (Papayanis 1996; Robert 1959). This long persistence of horse traction is visible directly in Figure A16, which reproduces three photographs from a series of street scenes of Parisian public transport taken by Eugène Atget and held by the Bibliothèque historique de la Ville de Paris.

Because the vehicles were horse-drawn and ran in mixed traffic on cobblestone streets, their speed was substantially lower than that of the electric trams and motor omnibuses that later replaced them. Heblich et al. (2020) document reported travel speeds of around 6 miles per hour (roughly 9.7 km/h) for horse omnibuses and horse trams in London during the second half of the nineteenth century, rising to around 7–8 mph only once motor omnibuses and electric trams became widespread after 1900; given the similarity in vehicle technology, road surface and legally enforced pace (a trot, not a canter, within the city) between London and Paris in this period, we adopt the same single average travel-speed assumption of 6 mph for the combined horse omnibus and tram network in constructing route-level travel times.

Data sources. Table A9 lists the historical maps used to construct the digitised network for each of the four periods we use.

Digitisation approach. For each year, we first georeference the scanned historical map in QGIS using a small number of ground-control points tied to stable, identifiable features (i.e. parks and monuments) matched against their present-day location. Rather than tracing

Table A9: Data sources for the digitised omnibus/tram network

Year	Title	Cartographer / publisher	Network digitised
1860	*Nouveau plan de Paris en 20 arrondissements. Plan des omnibus* [c. 1860]	Dessiné et gravé par J. Sédille (rue Montmartre 78)	Omnibus (CGO)
1865	*Carte des tramways du département de la Seine* [c. 1865] (éch. 1:50,000)	Gravé chez L. Wuhrer (Louis Wuhrer, 1843–1925), Paris [undated]	Tram ("lignes en exploitation" only)
	Les chemins de fer souterrains et tramways dans Paris	Édité par P. Méa	Underground railway
1877	*Paris. Plan nouveau : 1877*	E. Andriveau-Goujon, éditeur (4 rue du Bac, Paris)	Omnibus + Tram (*Plan des Tramways* inset)
1895	*Paris. Réseau des tramways : 1895*	E. Andriveau-Goujon, éditeur	Omnibus + Tram (5 operators digitised separately)

Notes: For 1865, no named author or editor is recorded for the *Carte des tramways du département de la Seine*; the map's only identified contributor is its engraver-printer, Louis Wuhrer (1843–1925), and its date of publication is not recorded on the sheet itself (catalogued as c. 1865). The map distinguishes lines already "en exploitation" (operating, digitised here) from "lignes et prolongements en projet" (planned/under construction, excluded); the underground-railway (*chemin de fer souterrain*) lines are digitised separately from the P. Méa map. For 1895, the five sub-networks digitised separately and merged into "Tram" correspond to the operators listed in the source map's legend (Compagnie Générale des Omnibus tramways, Compagnie des Tramways de Paris et du Département de la Seine, Compagnie Générale Parisienne de Tramways, a steam-hauled line, and a funicular line).

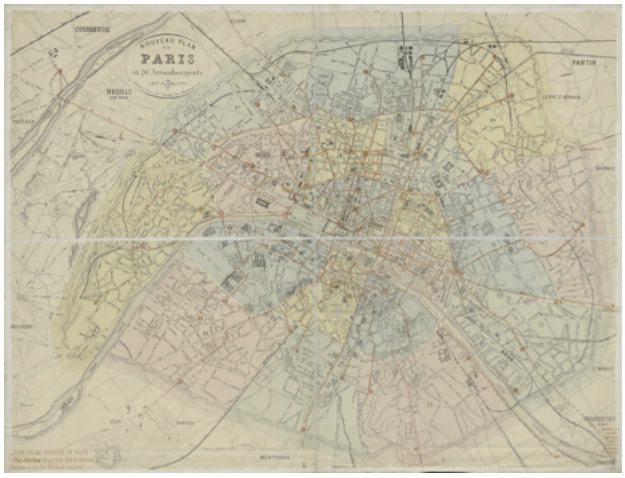
routes directly onto the historical street grid – whose scale and positional accuracy vary across publishers and editions – we overlay a modern OSM road-network extract for Paris as a common geometric baseline, and trace each digitised route onto this shared baseline. This ensures that differences in network coverage across cross-sections reflect genuine change in transit provision, in the same spirit as the 1896-anchor harmonisation procedure used for the historical sewer network. Within each year, we digitise separately by the operating company or vehicle technology distinguished on the source map (e.g. the five tram sub-networks for 1895 listed in Table A9).

Computing travel times. We use the digitised networks to compute bilateral travel times between the hexagonal grid cells that form the spatial units of our analysis. For any origin cell i and destination cell j, the travel time is the faster of two options: walking the whole way, or combining walking with a ride on the omnibus/tram network. Walking follows the straight-line distance between the two cell centroids at an assumed pace of 3 mph. A network trip consists of (i) a straight-line walk from the origin to the nearest point of the network, (ii) a ride along the network at 6 mph (the speed assumption discussed above), and (iii) a straight-line walk from the network to the destination, plus a fixed boarding penalty of three minutes. The traveller is assumed to take whichever option is quicker, so that

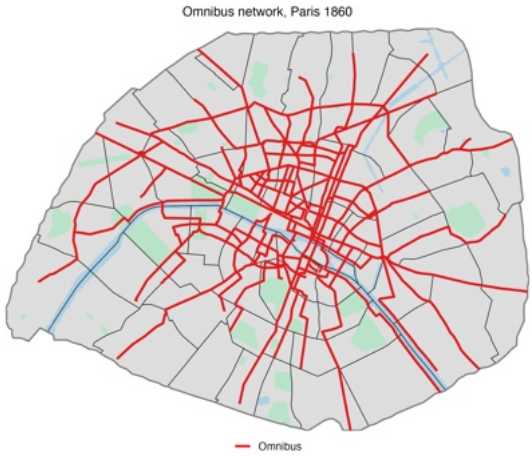
$$\tau_{ij} = \min \left\{ \frac{\ell_{ij}}{v_{\text{walk}}}, \quad \frac{\ell_i}{v_{\text{walk}}} + p + \frac{r_{ij}}{v_{\text{bus}}} + \frac{\ell_j}{v_{\text{walk}}} \right\},$$

Figure A17: Omnibus maps

(a) *1860 (map)*



(b) *1860 (omnibus routes)*



(c) *1877 (map)*



(d) *1877 (omnibus routes)*

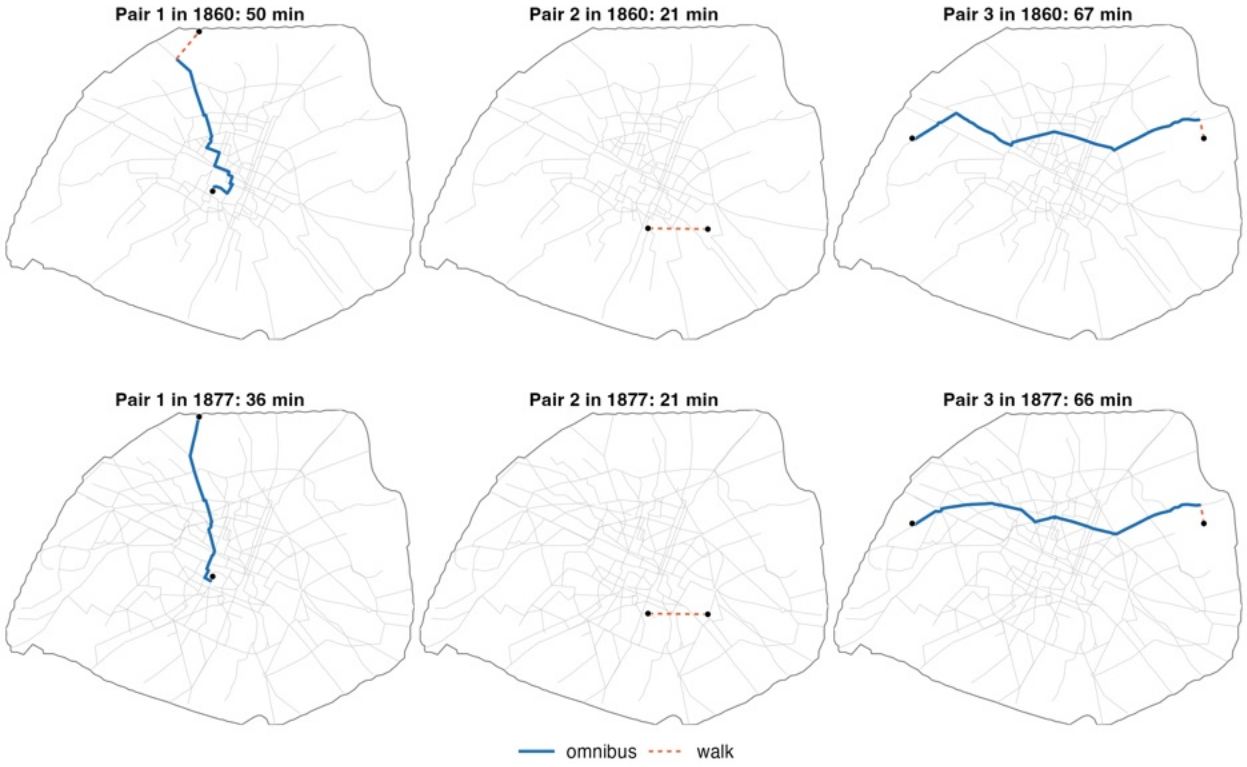


Notes: Left panels: source historical maps, georeferenced to a modern coordinate system. Right panels: digitised omnibus network traced from the source map onto the modern OSM road baseline (see *Digitisation approach* above and Table A9 for sources).

where the first term is the walk-only time and the second is the omnibus/tram option: ℓ_{ij} is the straight-line origin–destination distance, ℓ_i and ℓ_j are the walking distances from each cell to its nearest boarding point on the network, r_{ij} is the distance travelled along the network, $p = 3$ minutes is the boarding penalty, and $v_{\text{walk}} = 3$ and $v_{\text{bus}} = 6$ mph.

Because we do not observe individual stops, we allow passengers to board and alight at any point along a line. We build the ride network directly from the digitised line geometry: the lines are split at every intersection so that a passenger may transfer between lines at no cost wherever two lines cross (lines passing within 25 m of one another are treated

Figure A18: Computing travel times: example routes



Notes: Illustrative example of the route and travel-time computation for three randomly chosen origin–destination pairs (columns) in 1860 (top row) and 1877 (bottom row); the fastest travel time is reported in each panel title. Solid blue segments denote travel on the omnibus/tram network, orange dashed segments denote walking, and black dots mark the origin and destination. Grey lines show the digitised network for the corresponding year, clipped to the municipal boundary; passengers may board anywhere along a line and transfer between lines at no cost where lines cross. In these examples Pair 2 is faster on foot in both years, whereas Pairs 1 and 3 use the network.

as crossing, to absorb small differences in digitisation). The ride distance r_{ij} is then the shortest path through this connected network between the boarding point nearest the origin and the alighting point nearest the destination. Figure A18 illustrates the procedure for three example origin–destination pairs in 1860 and 1877.

A.2 Empirical Appendix

A.2.1 Haussmann's renovation

Table A10 compares treated and control hexagons within quartiers in terms of the elements of the Haussmann renovation that we use as controls in Table 1: access to public water fountains, parks and railway stations. Treated hexagons started from slightly better access to parks and railway stations in 1856. Over the expansion period, however, access to parks and railway stations improved slightly *more* for control hexagons, and access to water fountains is balanced in terms of distance.

Table A10: Balance of Haussmann renovation elements between treated and control hexagons

	Treated	Control	Difference
Water fountains			
Meter to nearest	453.6	456.3	-2.7
			(2.2)
log MA	-6.267	-6.480	0.213***
			(0.032)
Parks			
Baseline meter to nearest (1856)	2,380.4	2,423.3	-43.0***
			(3.8)
Δ meter to nearest	-691.4	-701.7	10.3***
			(3.1)
Baseline log MA (1856)	-15.538	-15.621	0.082***
			(0.010)
Δ log MA	1.793	1.782	0.010
			(0.009)
Stations			
Baseline meter to nearest (1856)	1,031.2	1,072.9	-41.7***
			(3.3)
Δ meter to nearest	-239.6	-296.2	56.6***
			(3.4)
Baseline log MA (1856)	-10.794	-11.079	0.284***
			(0.023)
Δ log MA	1.370	1.689	-0.318***
			(0.024)
Hexagons	21,973	13,935	

Notes: The sample is the main difference-in-differences sample: hexagons without sewer access in 1856 that were connected by 1896. Treated hexagons were connected by 1872. All columns are within quartier: the treated and control columns report quartier-adjusted means, and the difference is the coefficient on the treatment indicator in a regression with quartier fixed effects. MA: market access. Changes (Δ) refer to 1856–1872. Standard errors, clustered at the hexagon level, are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.2.2 Nearby sewers by density bin

Table A11 reports the estimates behind Figure 8. Log density in specification (3) is replaced by four bins of address density within 250 m, and the effect of nearby sewers is estimated separately in each bin. For residents, the effect is flat across the three sparser bins and rises above 30 addresses per hectare. For firms, it does not vary with density.

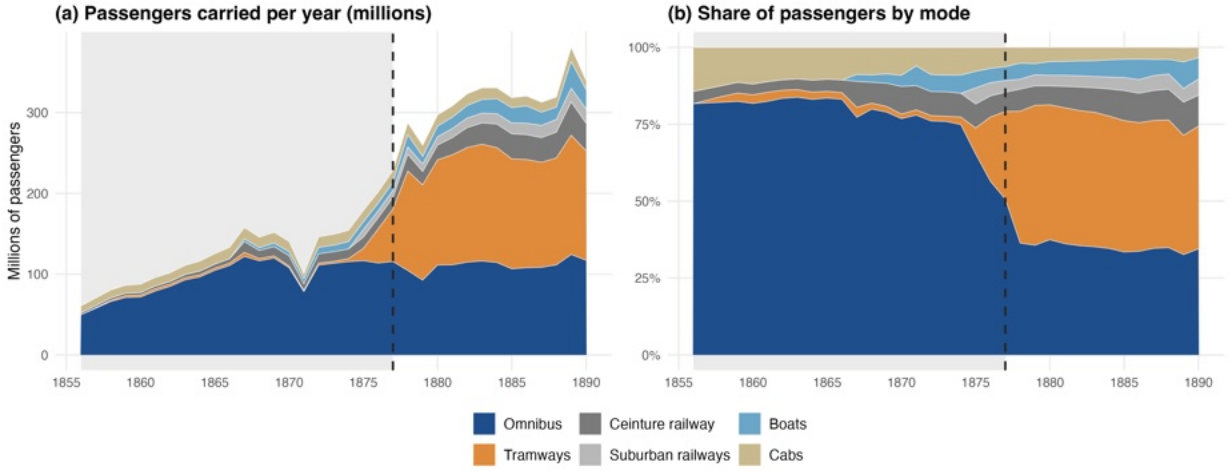
Table A11: Effect of nearby sewers by density bin

	Residents 1859/60–1871 (1)	Firms 1861–1872 (2)
NearbySewers × Post, sparsest bin (< 10)	−0.072	−0.005
	(0.065)	(0.037)
Relative to the sparsest bin:		
NearbySewers × Post × (10–20)	−0.010	−0.044
	(0.088)	(0.049)
NearbySewers × Post × (20–30)	0.059	−0.057
	(0.079)	(0.070)
NearbySewers × Post × (> 30)	0.162*	−0.002
	(0.084)	(0.054)
(> 30) − mean(< 30)	0.146**	0.032
	(0.066)	(0.038)
OwnSewer × Post	0.050	0.118*
	(0.060)	(0.066)
Density bin × Post	Yes	Yes
Already-sewered controls × Post	Yes	Yes
Haussmann controls × Post	Yes	Yes
Hexagons	5,745	12,290
Hexagons by bin	2,549 / 1,625 / 631 / 940	5,297 / 3,389 / 1,543 / 2,061
Mean outcome, pre	11.1	5.4

Notes: PPML estimates of a version of specification (3) in which log density is replaced by four bins of address density within 250 m (below 10, 10–20, 20–30 and above 30 addresses per hectare). NearbySewers × Post and the share of nearby addresses already sewered by 1856 × Post enter separately for each bin, together with bin × Post. The first row is the effect of NearbySewers (in units of 10 percentage points) in the sparsest bin; the next three rows are the effect in each bin minus the effect in the sparsest bin. (> 30) − mean(< 30) is the effect above 30 addresses per hectare minus the average effect in the three sparser bins. Residents are voter counts from the digitised electoral lists. Haussmann controls and samples are as in Table 3. All specifications include hexagon and quartier-year fixed effects. Outcomes are winsorised at the 99th percentile in each year. Conley standard errors with a 250 m cutoff are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A.2.3 Sewers and market access

Figure A19: Passengers carried by public transport mode (1856–1890)



Notes: Passengers carried per year, transcribed from Alfred Martin, *Étude historique et statistique sur les moyens de transport dans Paris* (Paris: Imprimerie nationale, 1894), Tableau n° 3, "Nombre total de voyageurs transportés à toute distance, par année et par compagnie (période 1855–1890)", pp. 274–277. Omnibus: lines and suburban services of the Compagnie générale des Omnibus and railway-station omnibuses. Tramways: the Compagnie générale des Omnibus' horse railway and tramways and the Tramways Nord and Sud companies (from 1874). Ceinture railway: all lines of the Chemin de fer de ceinture. Suburban railways: trips between the Paris termini and nearby localities (from 1875). Boats: passenger boats on the Seine (from 1867). Cabs: Compagnie générale des voitures à Paris only, estimated in the source at 12 passengers per cab and day. Shaded: 1856–1877, the period of the main sewerage infrastructure expansion. The dashed line marks its end.

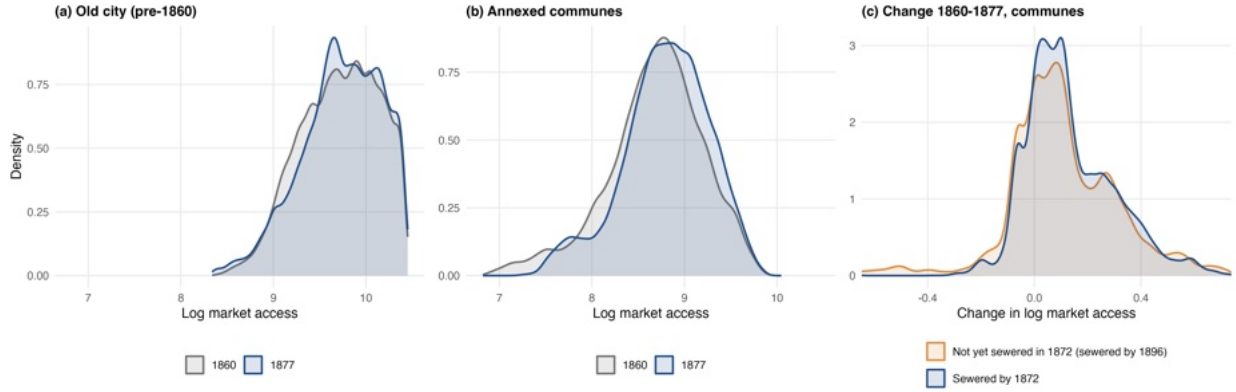
Table A12 puts the communes in context. In 1860 they already had some omnibus service, mostly radial lines running out from the old city. 43% of commune hexagons were within 300 m of a line (88% in the old city), the median distance to a line was 352 m (96 m), and average travel time to the city's firms was 42 minutes (27). By 1877 the share within 300 m of a line had risen to 68% and the median distance had halved to 187 m, while access in the old city barely changed. Table A13, Figure A21(b) and Figure A20(c) describe the resulting changes in the annexed communes. Log market access rises by a median of 0.09 log points between 1860 and 1877 (0.37 at the 90th percentile), it improves for 77% of hexagons, and the gains are concentrated along the new lines. Within quartiers, the 1860–1877 change accounts for 39% of the variation in 1877 access. Figure A20 shows the distribution of market access and changes in market access across hexagons. The distribution of gains is very similar for hexagons that were sewered by 1872 and those that were not yet sewered.

Table A12: Omnibus access in the old city and the annexed communes, 1860 and 1877

	Old city (pre-1860)		Annexed communes	
	1860	1877	1860	1877
Mean log market access	9.71	9.73	8.68	8.80
Market access relative to old city 1860	1.00	1.02	0.36	0.40
Mean travel time to 1860 firms (min)	27.3	27.0	42.4	40.6
Share within 300 m of an omnibus line	0.88	0.96	0.43	0.68
Median distance to nearest line (m)	96	70	352	187
Hexagons	62,233	62,233	83,951	83,951

Notes: All 25 m hexagons of the post-1860 city. Market access is $\log \sum_{g \neq h} L_g \, e^{-0.07 \, \tau_{hg,t}}$, computed on the 1860 and 1877 omnibus networks. The second row expresses it relative to the old city's 1860 mean. Travel time is the average over all 1860 firms in the city.

Figure A20: Distribution of market access (1860 and 1877)



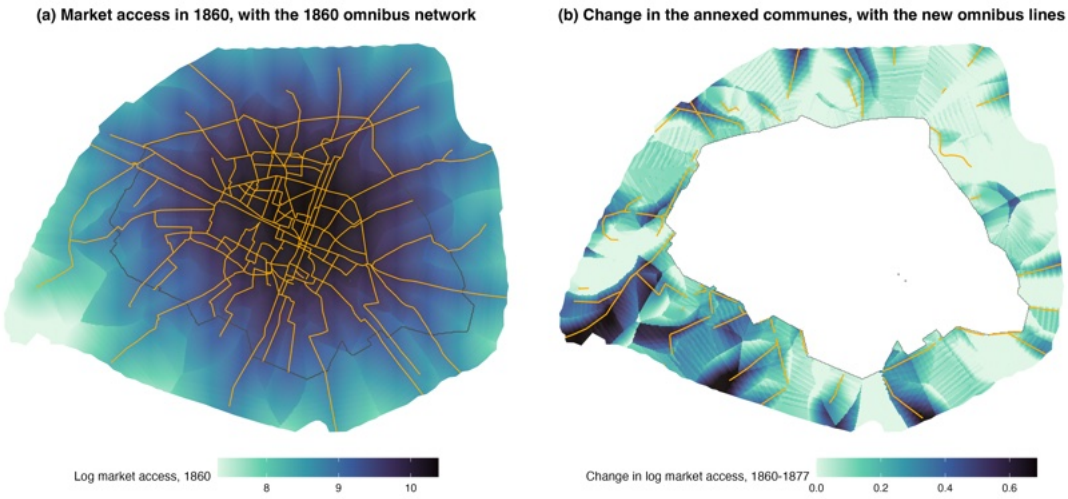
Notes: Log market access ($e^{-0.07\tau}$, 1860 firm weights). Panels (a) and (b): kernel densities across all 25 m hexagons of the old city and of the annexed communes, on the 1860 and 1877 omnibus networks. Panel (c): kernel density of the 1860–1877 change for the sample of Table 4, hexagons sewered by 1872 and hexagons not yet sewered in 1872 but sewered by 1896.

Table A13: Annexed communes: omnibus access in 1860 and its change to 1877

	Level, 1860		Change, 1860–1877				Share of within-
	Mean	SD	SD	p10	p50	p90	quartier variation
Log market access	8.807	0.502	0.190	-0.067	0.094	0.366	39%
Travel time to 1860 firms (minus minutes)	-40.629	7.424	3.284	-0.994	1.363	5.891	49%
Distance to nearest line (minus log m)	-5.741	0.793	0.751	-0.035	0.208	1.673	98%

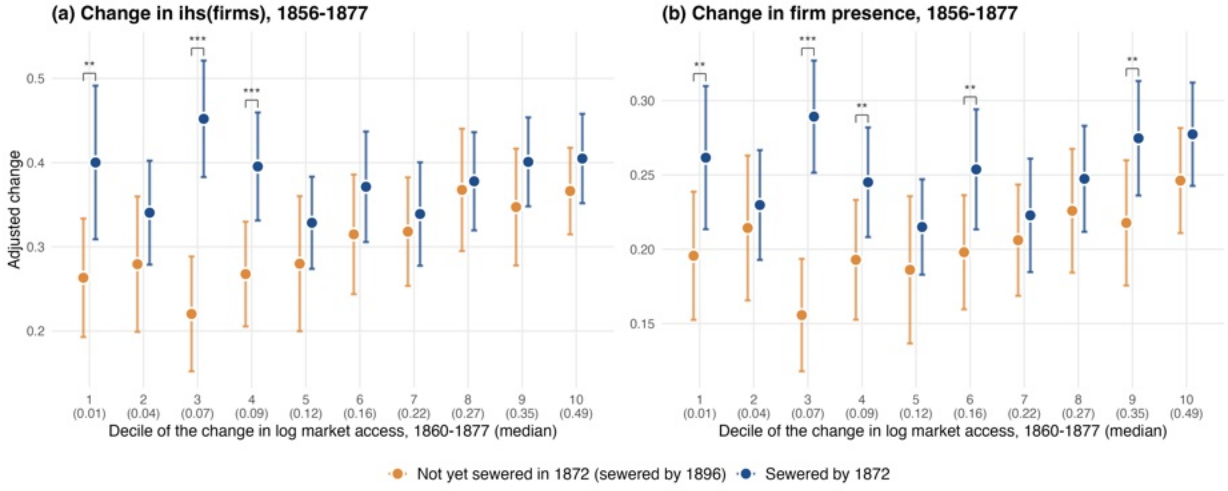
Notes: Sample of Table 4. Changes are between the 1860 and 1877 omnibus networks. Travel time and distance to the nearest omnibus line enter with a minus sign, so that an increase is an improvement. The last column reports the share of the within-quartier variance of 1877 access that is due to the 1860–1877 change.

Figure A21: Market access in 1860 and its change to 1877



Notes: Log market access ($e^{-0.07\tau}$, τ travel time in minutes, 1860 firm weights) of every 25 m hexagon. Panel (a): level in 1860 on the 1860 omnibus network (amber lines), all hexagons, colour scale from the 1st to the 99th percentile. Panel (b): change from 1860 to 1877 in the annexed communes, with decreases (23% of hexagons, mostly route changes) shown as zero and the scale capped at the 99th percentile; amber lines are omnibus lines on the 1877 network more than 50 m from any 1860 line. Grey line: boundary of the pre-1860 city.

Figure A22: Annexed communes: firm growth by gain in market access



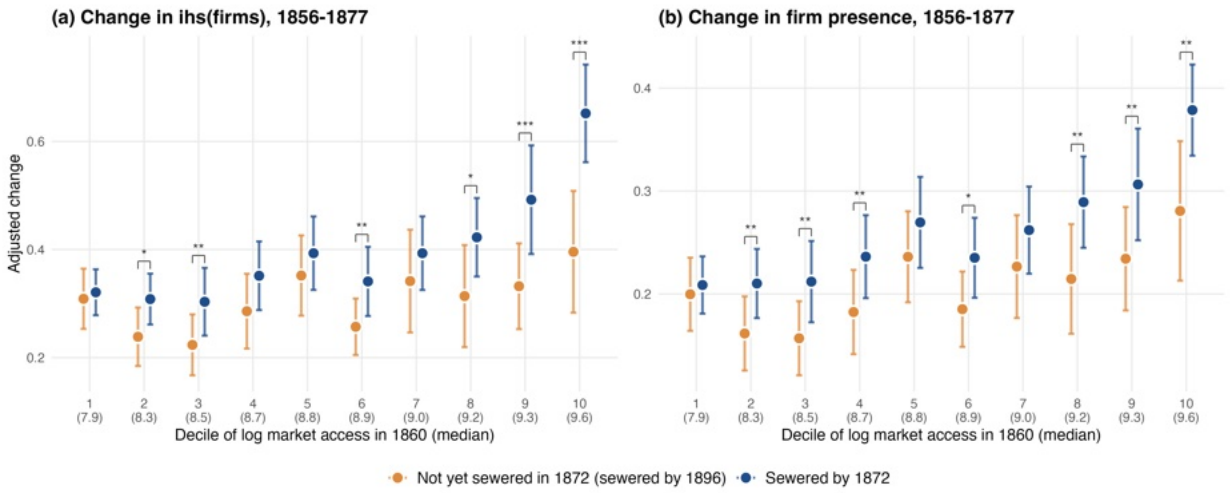
Notes: Sample of Table 4, 1856–1877, excluding hexagons whose market access fell. Hexagons are grouped into deciles of the 1860–1877 change in log market access ($e^{-0.07\tau}$; median change in parentheses). Each point is the mean change in ihs(firms) (panel a) or in firm presence (panel b) in that decile, after residualising on quartier fixed effects and a cubic in baseline log market access (pooled over both groups; sample mean added back), separately for hexagons sewered by 1872 and hexagons not yet sewered in 1872 but sewered by 1896. Bars: 95% confidence intervals, standard errors clustered by 250 m grid cell. Brackets mark deciles where the difference between the two groups, estimated jointly, is significant: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A14: Decomposition of Access × Sewer

	All firms		By sector (market access)		
	Market access	Travel time (min. saved)	Customer-facing	Manu-facturing	Other business
	(1)	(2)	(3)	(4)	(5)
Panel A: ihs(firms)					
ΔAccess × Sewer × Post	0.107	0.010*	0.104	0.009	0.003
	(0.100)	(0.006)	(0.077)	(0.020)	(0.029)
Baseline access × Sewer × Post	0.173***	0.012***	0.147***	0.012	0.033***
	(0.039)	(0.003)	(0.030)	(0.009)	(0.012)
Access	0.084	0.003	0.055	0.038**	0.029
	(0.086)	(0.005)	(0.067)	(0.018)	(0.023)
Sewer	0.082***	0.076***	0.055***	0.004	0.021***
	(0.020)	(0.020)	(0.016)	(0.005)	(0.006)
Equal effects, p-value	0.45	0.69	0.52	0.91	0.25
Mean outcome, 1877	0.375	0.375	0.240	0.044	0.077
Panel B: firm presence					
ΔAccess × Sewer × Post	0.086	0.007**	0.082	0.010	0.010
	(0.060)	(0.004)	(0.052)	(0.017)	(0.026)
Baseline access × Sewer × Post	0.080***	0.006***	0.090***	0.014*	0.030***
	(0.022)	(0.001)	(0.020)	(0.008)	(0.010)
Access	0.036	0.001	0.035	0.031**	0.015
	(0.050)	(0.003)	(0.045)	(0.015)	(0.020)
Sewer	0.043***	0.040***	0.034***	0.005	0.018***
	(0.013)	(0.012)	(0.011)	(0.004)	(0.005)
Equal effects, p-value	0.91	0.65	0.86	0.82	0.39
Mean outcome, 1877	0.251	0.251	0.176	0.041	0.068
Baseline access × Post (common)	Yes	Yes	Yes	Yes	Yes
Already-sewered share 1856 × Post	Yes	Yes	Yes	Yes	Yes
Hexagon, quartier × year FE	Yes	Yes	Yes	Yes	Yes
Observations	38,924	38,924	38,924	38,924	38,924

Notes: Specification of Table 4, with Access × Sewer split into the 1860–1877 change in access and its 1860 level, each interacted with Sewer and Post. Table 4 is the special case in which the two coefficients are equal. The row "Equal effects, p-value" tests this restriction. Sample, controls, fixed effects and clustering as in Table 4. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

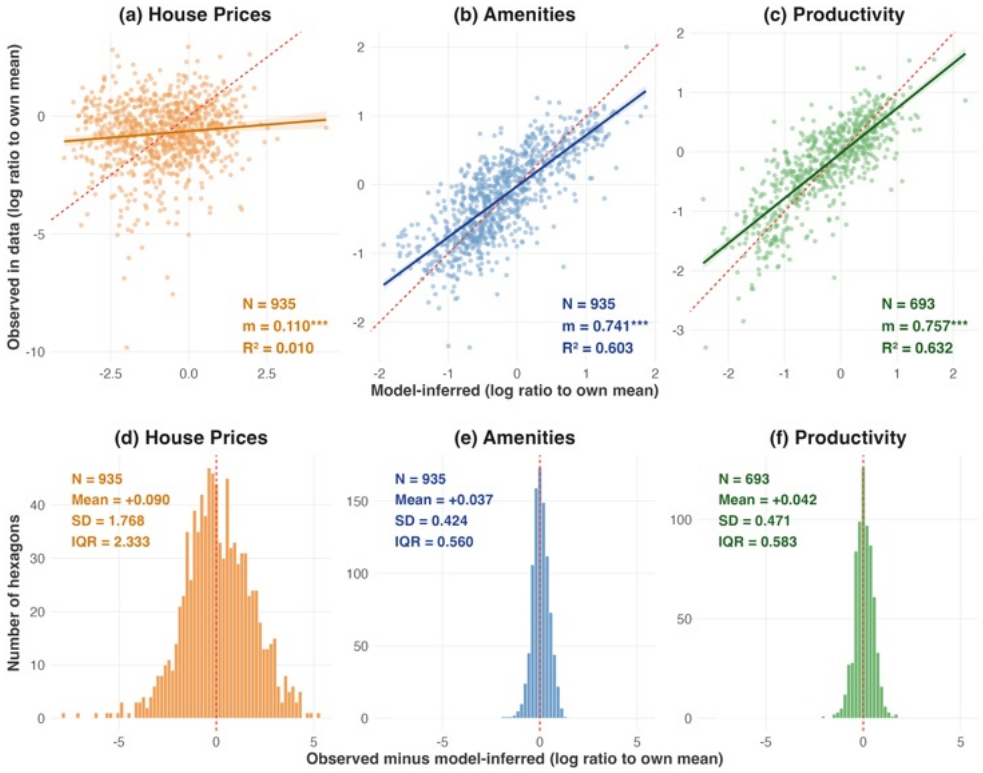
Figure A23: Annexed communes: firm growth by baseline market access



Notes: Sample of Table 4, 1856–1877. Hexagons are grouped into deciles of log market access in 1860 ($e^{-0.07\tau}$, 1860 firm weights; median in parentheses). Each point is the mean change in ihs(firms) (panel a) or in firm presence (panel b) in that decile, after residualising on quartier fixed effects and a cubic in the 1860–1877 change in log market access (pooled over both groups; sample mean added back), separately for hexagons sewered by 1872 and hexagons not yet sewered in 1872 but sewered by 1896. Bars: 95% confidence intervals, standard errors clustered by 250 m grid cell. Brackets mark deciles where the difference between the two groups, estimated jointly, is significant: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

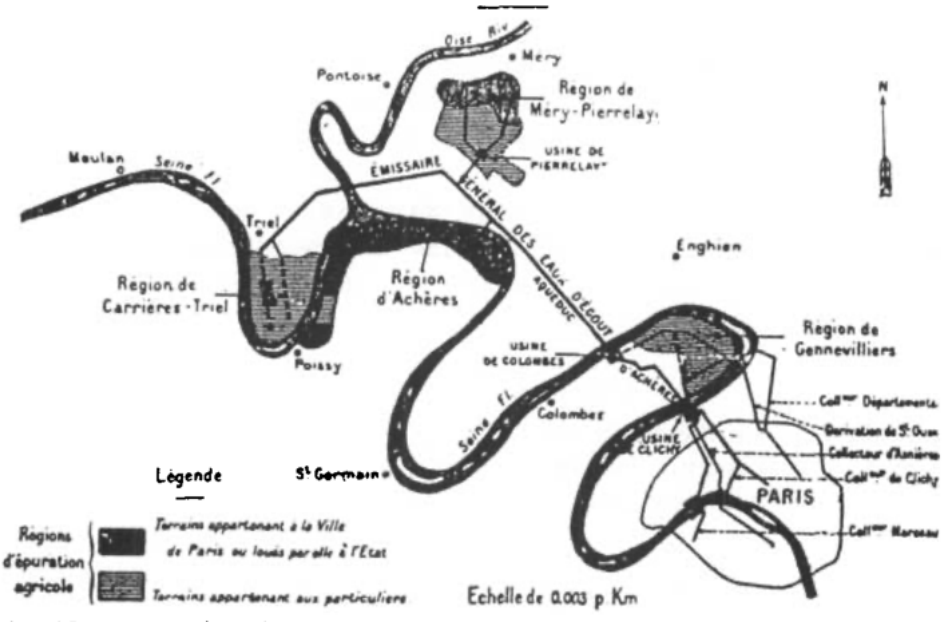
A.3 Quantitative Appendix

Figure A24: Sensitivity of the inversion to the price measure



Notes: The unit of observation is the hexagon. The sample is hexagons with a recorded transaction price in the ten-year window centred on 1871: 935 residential hexagons in panels (a), (b), (d) and (e), and 693 workplace hexagons in panels (c) and (f). Panels (a)–(c) plot the objects recovered using the inferred land price q (horizontal axis) against the same objects recovered when the price $\tilde{q}$ observed in the *sommier foncier* is substituted for it (vertical axis); panel (a) is the price itself. Because the price enters each structural residual through a single additive term, this substitution is exact and requires no re-solution of the model: $\ln \tilde{B}_i = \ln B_i + (1-\beta)(\ln \tilde{q}_i - \ln q_i)$ and $\ln \tilde{A}_j = \ln A_j + (1-\alpha)(\ln \tilde{q}_j - \ln q_j)$. The axis labels refer to the price series used to recover each object rather than to the object plotted, which is given by the panel title. Since the inferred rent is identified only up to scale and is denominated in model units while the observed price is in francs, each series is divided by its own mean in levels and then logged. Both axes are therefore unit-free log ratios to their own mean, and on the red dashed 45-degree line a hexagon x% above its own series mean on one measure is x% above it on the other. Solid lines are OLS fits with 95% confidence bands; N is the number of hexagons, m the OLS slope and R^2 the coefficient of determination, and *** denotes significance at the 1% level. Panels (d)–(f) show the distribution of the vertical minus horizontal difference from the panel above, on a common horizontal scale and bin width. The mean difference is small in each case and is not statistically distinguishable from zero once standard errors are clustered by quartier. Amenities and productivity stay close to the 45-degree line and their differences are tightly concentrated, whereas the inferred and observed prices are only weakly related; the contrast follows from the weights $1-\beta = 0.24$ and $1-\alpha = 0.30$ that land prices carry in the two residuals, so most of the variation in the recovered fundamentals comes from populations, wages and market access rather than from prices. Note finally that the two axes of panel (a) measure related but distinct objects: the inversion assigns every location the same floorspace, so q is closer to an expenditure density than to a price per unit of housing, while the *sommier foncier* records the value of an entire property without building characteristics.

Figure A25: Belgrand Plan



A.4 Counterfactual Appendix

A.4.1 Sensitivity to the entry elasticity

The entry elasticity θ in (27) governs how strongly the population of Paris responds to a change in its attractiveness relative to the rest of France. Table A15 reports the counterfactual of Section 7.1 for values of θ between 1 and 5. For each value, we recalibrate the utility of the outside option $\bar{u}_t$ such that the model with the sewer network of each census year reproduces the census population, and then solve for the equilibrium under the sewer network of 1856. Panel A reports the resulting population difference relative to the census, and Panel B the share of the growth of Paris since 1856 that is attributable to the sewer expansion. Panel C reports the share of this effect that is due to the reduction in the sanitation externality, with the remainder due to the private benefit of sewer access.

A.4.2 Population changes by location type in 1872

Figure A26 shows how strongly different types of locations are affected by the reduction of the sewer network to its 1856 extent in the counterfactual of Section 7.1. It compares the city of 1872 with two counterfactuals, one under the sewer network of 1856 and one in which locations keep the private benefit of their sewer access in 1872 while the sanitation externality is determined by the 1856 network. The second counterfactual therefore isolates the contribution of the externality reduction. The locations sewered between 1856 and 1872

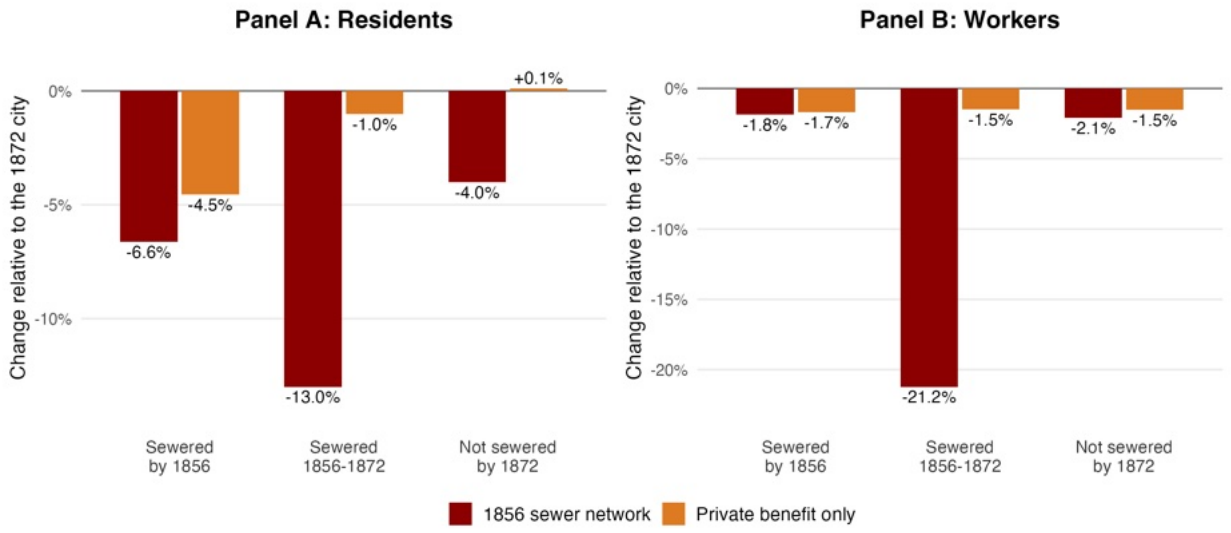
Table A15: Sensitivity of sewer-driven growth to the entry elasticity

	$\theta = 1$	$\theta = 2$	$\theta = 3$	$\theta = 4$	$\theta = 5$
Panel A: Population difference without the sewer expansion (%)					
1872	−2.6	−5.1	−7.4	−9.8	−11.9
1881	−3.5	−6.7	−9.7	−12.5	−15.1
1891	−4.2	−8.1	−11.7	−15.0	−18.0
1896	−4.4	−8.3	−12.0	−15.3	−18.4
Panel B: Share of population growth since 1856 due to the sewer expansion (%)					
1872	16.1	31.3	45.6	60.1	72.7
1881	11.0	21.1	30.6	39.4	47.7
1891	11.6	22.1	31.8	40.8	49.1
1896	11.3	21.4	30.8	39.4	47.4
Panel C: Share of the sewer effect due to the externality reduction (%)					
1872	19.9	19.8	21.1	20.7	20.2
1881	24.5	24.2	24.1	24.4	25.0
1891	27.8	27.9	28.1	28.2	28.4
1896	28.2	28.4	28.6	28.8	29.2

Notes: Counterfactual population of Paris with the sewer network held at its 1856 extent, for alternative values of the entry elasticity θ. Panel A: population difference relative to the census population of each year. Panel B: this difference as a share of the increase in the census population between 1856 and the respective year. Panel C: share of the total difference in Panel A that is due to the reduction in the sanitation externality, computed as the gap between the census population and an equilibrium in which each location receives the private benefit of sewer access under the network of the respective year, while the sanitation externality is determined by the 1856 network. The remainder is due to the private benefit of sewer access. In each year and for each θ, the utility of the outside option is chosen such that the model with that year's sewer network reproduces the census population. The baseline is $\theta = 3$, following Bryan and Morten (2019).

lose 13.0% of their residents without the sewer expansion, but only 1.0% once they keep the private benefit, so that their losses mainly reflect the loss of the private benefit. Locations that did not receive sewers over this period also lose residents. For those already sewered by 1856, most of this loss remains when the private benefit is kept, at 4.5% compared with 6.6%, as it stems from the positive spillovers of the sewers built in their surroundings. These are the densest parts of the city and are therefore most exposed to the sanitation externality. For locations not yet sewered by 1872, the loss instead reflects the smaller size of the city as a whole. For workers, the changes are driven by the private benefit and general equilibrium effects, since the externality only affects amenities.

Figure A26: Population changes by location type, 1872



Notes: Change in the number of residents (Panel A) and workers (Panel B) relative to the city of 1872, by sewer status of locations, in two counterfactual equilibria of 1872. 1856 sewer network: the sewer network is held at its 1856 extent. Private benefit only: each location receives the private benefit of its sewer access in 1872, while the sanitation externality is determined by the 1856 network. In both, the utility of the outside option is held at its 1872 level. Parameters at the estimates of the second scenario, $\theta = 3$.