Selective sanitation and racial health inequality*

Johan Fourie† Kelsey Lemon‡ Jan Pretorius§

Abstract

How selective sanitation provision shaped racial health inequality remains less well documented than its contribution to mortality decline. We examine this question in early twentieth-century Paarl, South Africa, using death notices linked to streets, district births reconstructed by reverse survival, and the staggered introduction of municipal stormwater drains. We find that investment was concentrated along central streets with relatively high shares of white households and commercial property. The corresponding health effects are less clear: estimates are sensitive to geolocation and sparse death counts, and neither all-cause nor cause-specific analyses provide robust evidence that white infants benefited disproportionately. Spatial associations also fail to distinguish proximity to the works from the route taken by contaminated water. Paarl extends the history of sanitation to a colonial African town where public investment followed existing social divisions. It documents selective provision while showing why historical administrative records do not necessarily reveal who gained from it.

Keywords: infant mortality; health inequality; sanitation; South Africa

JEL codes: I15; I18; N37; H51

*The authors would like to thank Rulof Burger, Johannes Norling, Carolyn Chisadza, Robert Garlick, Jonas Hjort, John Landon-Lane, Calumet Links, Neil Lloyd and participants from various workshops and conferences for valuable feedback on an earlier version of this paper. This paper was created with the help of Anthropic's Claude Code (Fable 5.1) and OpenAI's Codex (GPT-6 Astra). Cite this paper as: Fourie, Johan, Kelsey Lemon, and Jan Pretorius. 2026. "Selective sanitation and racial health inequality." Working Paper, Department of Economics, Stellenbosch University.

†Department of Economics, Stellenbosch University. Email: johanf@sun.ac.za.

‡Department of Economics, Stellenbosch University.

§Department of Economics, Stellenbosch University. Corresponding author: jhpretorius@sun.ac.za.

1 Introduction

The decline in infant and childhood mortality between the late nineteenth and early twentieth centuries was uneven. In colonial South Africa, infectious disease mortality remained high among much of the indigenous population even as white residents, descendants of European settlers, began to see gains. What enabled some groups to escape high mortality while others lagged behind? Public health infrastructure is part of the answer, but its distribution matters: investments could reduce mortality without benefiting all residents equally. We examine this question through early twentieth-century sanitation provision in a South African town, extending recent work on the *sanitary revolution* and the origins of persistent health inequalities (Aidt, 2022; Gallardo-Albarrán, 2025).

We study how a municipal stormwater drainage scheme related to the geography and racial distribution of infant mortality in Paarl, a small but rapidly developing town in the Cape region of South Africa. We link death notices to streets and compare mortality before and after drainage by race and cause, using reconstructed district births to estimate street exposure. The municipality's efforts to improve sanitation and its administration were slow and uneven, but stormwater drainage was among its earliest and most complete projects. The works coincided with broader changes in Cape public health practice and the early stages of South Africa's mortality transition (Van Heyningen, 1989). Decades before formal apartheid legislation, Paarl was already an unequal town. White and coloured residents lived close to one another, but wealthier, mostly white households occupied larger plots on the higher slopes of Paarl Mountain, while poorer, mostly coloured households lived in crowded areas nearer the Berg River (Beyers, 1933; Albertyn, 1992; Clift, 2012).[1]

The town council selected central thoroughfares where white households and commercial property were concentrated. The works can be dated and affected some streets rather than others, as in the infrastructure projects studied by the recent quasi-experimental literature. Their placement, however, was not random. Nor do pre-intervention mortality comparisons reliably explain the selection: their magnitude depends on reconstructed street exposure and on whether cells with no deaths are included. We therefore distinguish the historical evidence on municipal priorities from the statistical claim that the scheme responded specifically to white diarrhoeal deaths.

We ask how mortality changed across the town after the council prioritised these streets. The non-random placement of drainage prevents us from treating it as an exogenous shock and recovering an average causal effect on infant mortality. Instead, we bring together pre-intervention mortality, the timing and location of the works, topography and recorded causes of death. Triple-difference regressions and event studies summarise how mortality evolved by treatment status, race and disease category after a politically targeted intervention. The sparse estimating panel limits what those comparisons can establish.

The mortality evidence does not establish that the works achieved their intended health effect. Our primary analysis uses recorded residence, includes cells with no deaths and follows complete

[1] A.G. Oberholster and P. Van Breda (1987) note that substantial black African migration to the town only began after 1928; before then, black residents were a tiny minority.

birth cohorts through their first birthday. The joint cause-specific estimates do not show a decline in water- and food-borne mortality on treated streets. The all-cause estimates are too imprecise to distinguish a change for either racial group, and a difference between groups is harder still to establish with only four treated streets. The evidence for selective provision is stronger than the evidence for unequal health gains.

We also ask whether mortality changed with proximity to the works. We classify the surrounding street network geographically, but estimate mortality models only on streets represented in the death notices. These comparisons add no independent drainage interventions. Once we allow for pre-existing cause composition and common cause-specific changes, the spatial estimates do not establish benefits confined to white infants. Directed run-off exposure is also nearly indistinguishable from simple proximity to the treated streets.

Paarl slopes towards the Berg River, making the displacement of contaminated water a plausible explanation for geographic differences. The network and mortality records cannot distinguish that explanation from residential location, recording practices or other changes around the same streets. We therefore report the geographic and illness-duration comparisons as descriptive evidence, without claiming to identify a hydrological mechanism.

Research in development economics, health economics, economic history and historical demography has examined the decline in infectious disease mortality in the early 1900s. Nutrition, income and medical technology mattered, but evidence also points to clean water and sanitation as major contributors to lower infant mortality (Cutler and Miller, 2005; Beach et al., 2016; Kesztenbaum and Rosenthal, 2017; Peltola and Saaritsa, 2019). Studies of water purification credit chlorination and filtration with nearly half the decline in infant mortality in major American cities between 1900 and 1936 (Werner Troesken, 2004; David M. Cutler and Grant Miller, 2005; Joseph P. Ferrie and Werner Troesken, 2008; Dora L. Costa, 2015). The reliability of supply mattered too: Werner Troesken, Nicola Tynan and Yuanxiaoyue Artemis Yang (2021) show that constant, high-pressure water reduced waterborne disease mortality in London and accounted for a meaningful share of its late-nineteenth-century mortality transition.

Water and sewerage could reinforce one another. Marcella Alsan and Claudia Goldin (2019) find that their combined effect was decisive in saving infant lives at the start of the U.S. mortality transition. Effective systems removed pathogens responsible for diarrhoeal and enteric diseases, to which infants were particularly vulnerable. D. Mark Anderson, Daniel I. Rees and Tianyu Wang (2020) nevertheless caution against attributing the decline in U.S. childhood diarrhoeal deaths to public health interventions alone. Other industrialising countries also experienced gains as sanitation expanded. Infant mortality in England remained around 150 per 1,000 births until about 1900, then fell to nearly 100 by 1910 alongside the expansion of clean water and sewerage (Woods, Watterson and Woodward, 1988). In Paris, broader neighbourhood access to sewers increased the mortality benefits of improved water quality (Lionel Kesztenbaum and Jean-Laurent Rosenthal, 2017). Contemporary quasi-experimental evidence from lower-income settings likewise finds large health returns to urban water disinfection and piped water where networks function well (Geruso

and Spears, 2018; Bhalotra et al., 2021; Fan and He, 2023).

Who benefits depends on access and implementation. Widely shared services, such as city-wide water purification, could particularly help communities previously left exposed to disease. Werner Troesken (2004) proposed this explanation for black urban populations in the United States, where black and white residents often used the same municipal water systems. Consistent with that argument, chlorination between the 1910s and 1930s reduced black infant mortality with little measurable effect on white infant mortality, narrowing the black–white gap by roughly 13 per cent (Anderson et al., 2021). Yet access was unequal from the outset: segregation shaped where U.S. water mains were first installed (Beach, Parman and Saavedra, 2022). In colonial settings, race and imperial politics also influenced who received piped water and when (Broich, 2007). These studies connect institutions to infrastructure choices and their consequences (Ashraf, Donald and Lunt, 2016; Bogart, 2022; Aidt, 2022; Gallardo-Albarrán, 2025). Less well documented is how those choices affected people living in different parts of the same town.

South African history gives this question particular force. During the 1918 influenza pandemic, Kimberley's black mine residents died at roughly seven times the rate of its white town residents: 22.4 per 100 population compared with 3.1 (Phillips, 1984). In the 1930s, infant mortality among South Africa's coloured urban population[2] resembled England's half a century earlier, while white South African infants had mortality rates comparable to their European contemporaries (Rip, Bourne and Woods, 1988; Woods, Watterson and Woodward, 1988; Simkins and Van Heyningen, 1989). Historians have also shown how infrastructure concentrated in white neighbourhoods, often with the aim of shielding white residents from disease and far less regard for black communities. During the bubonic plague outbreaks of 1901–1904, Cape Town authorities invoked *sanitation* to remove African residents to Uitvlugt, a former sewage dumping site roughly ten kilometres away (Swanson, 1977; Bickford-Smith, 1995; White, 2018). Colonial investment in water, sewerage and housing generally favoured European areas. Houses in the racially mixed District Six lacked direct water connections while supply expanded elsewhere (Grant, 1992). Cape Town had a functioning water-borne sewerage system by 1904, but the Cape Flats, including Uitvlugt, received piped water and sewerage only in the mid-1940s (Buirski, 1983). Sanitary reform served both public health and social engineering purposes.

Paarl extends this largely Eurocentric sanitation literature to colonial Africa. Historians have written extensively on urban sanitation under colonial rule, but few studies examine how such projects affected mortality, or whose mortality (Swanson, 1977; Wright, 2001; Bin-Kasim, 2025). Paarl shares important features with the cities they describe: all but Europeans were excluded from infrastructure and sanitary planning, and officials were preoccupied with European health and living conditions. In one respect it was unusual. Cities such as Nairobi and Dakar separated European quarters with *cordons sanitaires* and buffer zones; Paarl remained residentially integrated

[2]We classify individuals into two broad racial categories: *white*, referring to those of European descent; and *coloured*, a distinct racial group historically defined by mixed ancestry, including significant lineage from the indigenous Khoi, Cape Malay slaves and other local populations. This follows the classifications in the historical records and contemporary South African terminology in common use.

well into the twentieth century.

Our question therefore differs from that of the European and American studies. They ask whether sanitation reduced mortality (Cutler and Miller, 2005; Alsan and Goldin, 2019); we ask who received it and whether those who received it gained. Paarl suits this question for three reasons. White and coloured households lived within a few streets of one another, so a scheme confined to selected streets drained some households' streets and not those of their near neighbours. The works can be dated by street. And the death notices record residence, race and cause, which allows mortality on the same streets to be compared before and after drainage. We make two contributions. We document that a colonial municipality directed one of its earliest sanitary investments towards white and commercial streets. We then show that detailed administrative records, even when linked to streets, do not settle who benefited: selective recording and four treated streets limit what the mortality comparisons can establish. For other colonial African towns, where provision was typically selective, the location of infrastructure does not by itself show who was protected.

That distinction remains relevant a century later, when infrastructure is central to public health policy in many low- and middle-income countries. Extensive networks can produce large health returns (Bhalotra et al., 2021; Fan and He, 2023), while partial coverage and poor siting can leave communities exposed (Geruso and Spears, 2018). Paarl supplies historical evidence on the choices behind provision. Its mortality estimates leave open how far the benefits of those choices extended.

The rest of the paper describes Paarl's demographic divisions and sanitary works, then introduces the death notices and district demographic sources. We explain birth reconstruction, street allocation and recording uncertainty before examining selection, mortality changes and geographic associations. The conclusion considers what this evidence adds to the history of health inequality.

2 Health and inequality in Paarl

Paarl lies roughly sixty kilometres northeast of Cape Town, between the Berg River and Paarl Mountain. It is the largest town in a predominantly agricultural district, historically also known as Paarl (Oberholster and Van Breda, 1987). In 1904, the town had 11,923 residents out of a district population of 30,423.[3] The district was relatively urbanised for the period, and the town's population was fairly evenly divided between white and coloured residents. Residential, industrial and agricultural properties surrounded a densely populated commercial centre along Main Street and the Lady Grey Bridge. This combination allows us to examine how infrastructure, topography and politics interacted within one town.

By the turn of the century, race and class shaped where people lived. Apartheid-era segregation had not yet hardened these divisions, but wealthier, predominantly white households tended to live in the south and west of Main Street, often on the higher slopes of Paarl Mountain. Poorer, predominantly coloured households were concentrated in the north and east, notably in the Jubilee and Ou Tuin neighbourhoods (Albertyn, 1992; Clift, 2012; Beyers, 1933). These were the ends

[3]Cape of Good Hope. (1905). *Census of the colony of the Cape of Good Hope 1904.* (G19-1905). Cape Town: Government Printer, p. 12, 14.

of a spectrum, rather than wholly separate worlds. As Beyers (1933) observed, in a small town "abundance and poverty, order and disorder, happiness and unhappiness lie side by side".[4] Wage data show white workers earning up to 2.5 times as much as coloured workers in 1904. White workers were more strongly represented in professional and commercial occupations; coloured residents were concentrated in agriculture, industry and domestic service.[5] These divisions shaped both exposure to unsanitary conditions and the economic and political means to seek reform.

Paarl faced a sanitary crisis by the late nineteenth century. Medical officers' reports from the 1890s and early 1900s describe overcrowding, poor housing and inadequate or absent sanitary infrastructure.[6] Medical officers, magistrates and wealthier residents repeatedly complained about roadside drains, or *sluits*. Without piped sewerage, households disposed of waste in poorly maintained cesspits or directly into these drains. In 1899, the district surgeon reported that "night soil, slop water, household and other refuse are for the greater part disposed of into the drains."[7] The drains often overflowed into the streets before emptying into the Berg River, from which some residents drew water.[8] Conditions were reportedly worst in neighbourhoods such as Ou Tuin, home to many coloured residents. The archives thus record a town-wide crisis alongside particular pressure to improve the central streets, where businesses and wealthier households were concentrated.

The town's water supply improved gradually from the 1880s, first with the construction of the Victoria Dam. Over the next three decades, the supply was augmented by the deepening of the Victoria Dam and the construction of two further dams, the Nantes and Bethel dams. Pipes supplied water to the main populated parts of the town, with households paying for individual connections. Though it remains unclear how many houses were connected, medical officers' reports show that coverage was not universal. Many residents still used wells, the mill stream or the Berg River, all of which were regarded as polluted.[9] Contemporary doctors linked the consumption of such water, along with the town's general sanitary conditions, to recurring typhoid and diarrhoeal illness. Although the piped supply met the town's water needs, water levels were insufficient to support either expansion or a water-borne sewerage system.[10]

Medical facilities did not keep pace with population growth. Paarl's first hospital opened in 1921; until then, residents needing hospital treatment had to travel to Cape Town (Oberholster and Van Breda, 1987). We do not know whether Paarl had maternity homes like those emerging in Cape

[4]Translated from the original: "Oorvloed en armoede, orde en wanorde, geluk en ongelukkigheid lê neffens mekaar" (Author's translation).

[5]Cape of Good Hope. (1905). *Statistical Register of the Colony of the Cape of Good Hope, 1904 (first six months).* Cape Town: Government Printer, p. 245.; Cape of Good Hope. (1905). *Census of the colony of the Cape of Good Hope 1904.* (G19-1905). Cape Town: Government Printer, p 420

[6]WCARS, CO 7651, file 1221, *Sanitary conditions in Paarl*; *Reports on the Public Health* between 1895 and 1908.

[7]Cape of Good Hope. (1900). *Reports on the Public Health, 1899.* (G.56–1900). Cape Town: Government Printer, p. 71

[8]WCARS, CO 7651, Report by Assistant MOH Gregory, 24 December 1897; WCARS, MOH 240 L72C, Report by Assistant MOH Mitchell, 23 June 1902.

[9]*Reports on the Public Health, 1895*, pp. 124–125; WCARS, MOH 240 L72C, Report by Assistant MOH Mitchell, 23 June 1902.

[10]A waterborne sewerage system was built in Paarl in the 1930s.

Town, or what support was available to its most vulnerable mothers.[11] The resident magistrate reported eight doctors in 1903, but the distribution of their services is unclear.[12] Research on the Colony as a whole suggests that access to medical care was probably unequal in Paarl too (Fourie and Jayes, 2021).

Public health administration was also changing across the Colony. The Public Health Act of 1883, amended in 1897, and the Births and Deaths Registration Act of 1894 gave authorities better mortality information and powers to enforce sanitary improvements. These reforms increased awareness of infant mortality and pressure for municipal action (Van Heyningen, 1989). In Cape Town, high infant mortality prompted the appointment of a "Lady Inspector" in 1900 to visit the homes of infants who had died and investigate the circumstances. Her work documented practices such as "artificial feeding" in place of breastmilk, whether because mothers could not breastfeed, possibly owing to malnutrition, or had to return to work early. She also found that most deceased infants were illegitimate, a finding that helped trigger the "Baby-farming scandal" and South Africa's child welfare movement. How far this early attention contributed to declining infant mortality remains unclear.

For Paarl, the Municipal Act No. 45 of 1882 was especially consequential. It allowed towns to raise funds and impose bye-laws for municipal improvements (Van Heyningen, 1989). After Paarl was proclaimed under the Act in 1898, its council introduced a "proper" disease notification system and appointed a sanitary inspector and town engineer the following year. The engineer then prepared a scheme of Public Works Improvements.[13]

The council announced the plan in 1902 after the mayor, S. P. H. de Villiers, secured a £20,000 loan from the colonial government. Construction began in 1905. Underground cement pipes replaced the open roadside *sluits* and discharged towards the Berg River. Medical Officer of Health (MOH) reports provide the only surviving construction chronology: Lady Grey Street and part of Main Street were drained in 1905, a northern section of Main Street followed in 1906, and Van der Lingen, Malherbe and Machinery Streets were completed in 1907.[14] No formal plans survive to explain the sequence or siting of the drainage scheme. The early works nevertheless show that the council prioritised the main commercial axis and higher-status residential streets.

Figure 1 locates the works. Engineering considerations and political priorities shaped a project that served the main commercial axis and prominent residential streets while many crowded coloured streets retained open *sluits*. By showing where officials directed their attention, the archives reveal the rationale behind the drainage scheme more clearly than patterns in cause-specific mortality do.

[11] See Burman and Naude (1991) for a discussion of vulnerable motherhood in Cape Town in this period.

[12] Western Cape Archives and Records Service (WCARS), MOH 240 L72C, Letter from Resident Magistrate to MOH Gregory, 28 July 1903.

[13] Proclamation no. 361, *Cape of Good Hope Government Gazette*, 30 December 1898; Cape of Good Hope. (1900). *Reports on the Public Health, 1899.* (G.56–1900). Cape Town: Government Printer, p. 167.

[14] Cape of Good Hope. (1906). *Reports on the Public Health, 1904-1905.* (G.39–1906). Cape Town: Government Printer, p. 242; Cape of Good Hope. (1907). *Reports on the Public Health, 1906.* (G.40–1907). Cape Town: Government Printer, p. 96; Cape of Good Hope. (1908). *Reports on the Public Health, 1907.* (G.33–1908). Cape Town: Government Printer, p. 69.

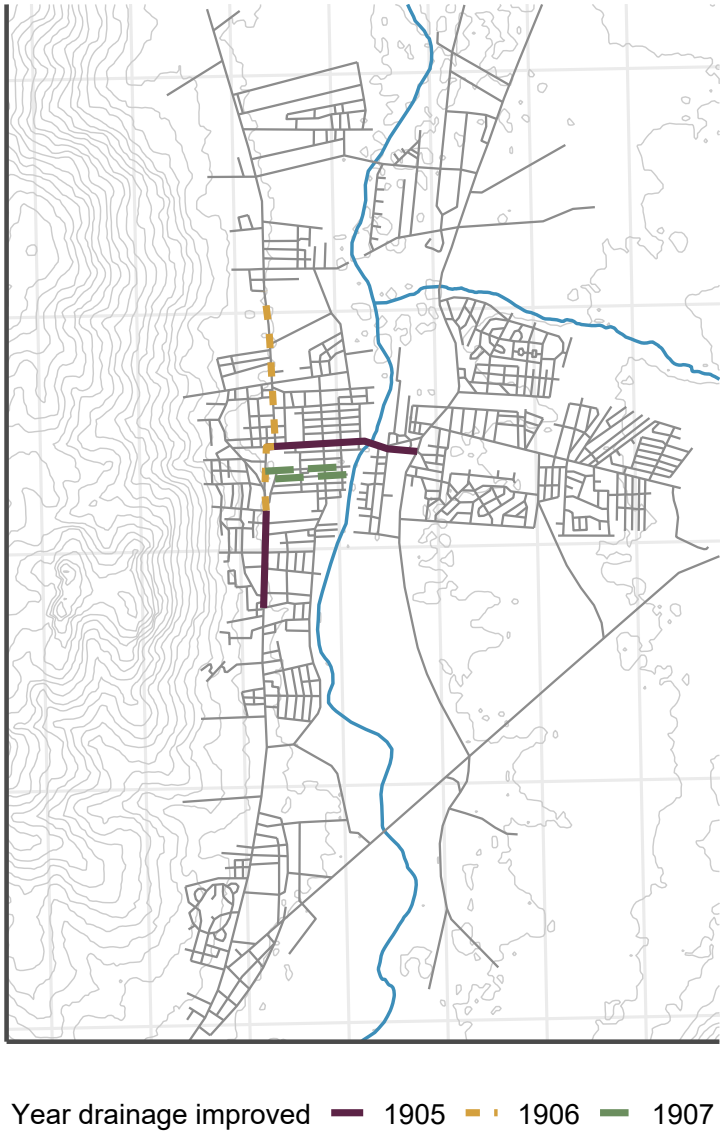


Figure 1: Map of Paarl showing streets with storm-water drainage improvements from 1905–1907

Notes: Lady Grey and a section of Main Street were drained in 1905; Main Street was extended in 1906, and Van der Lingen and Malherbe followed in 1907. The map depicts digitised phase segments. Machinery Street could not be located and is excluded. Main Street receives first-phase coding in the primary analysis and last-phase coding in the sensitivity.

Source: Based on a printed map sourced from SU Africana Library, E PAM 26 STR, *Street Plan of Paarl* (1958). The map was scanned, georeferenced, and matched to historical street names by the authors.

The scheme improved some streets while leaving others exposed. It removed surface waste and standing water locally, but other neighbourhoods continued to face contaminated run-off and may have experienced greater exposure. The health question therefore concerns both the size and the distribution of any gains. We link newly digitised death notices to the timing and location of the works to examine how infant mortality changed across streets and racial groups.

3 Data and death notices

We combine death notices from 1895–1931 with district demographic counts to examine sanitation and infant health in Paarl. Births are observed in seven years. For the remaining years, we reconstruct district birth and population series by race and sex from census age groups and the death register.[15] The primary comparisons use birth cohorts 1896–1930, which can be followed through their first birthday before the register's last complete year.

Death notices provide the individual records. Introduced under the Births and Deaths Registration Act of 1894, they record age, sex, race, residence, duration of final illness and stated cause of death, as well as the informant and any medical attendant. The cleaned file contains 27,512 notices across all ages. We use non-infant notices to reconstruct demographic series and estimate street composition, selecting the infant sample after checking age, stillbirth status and cohort eligibility.

Recorded residence determines which notices enter the street analysis. Of 8,226 white and coloured infant deaths in complete study cohorts, 4,377 record Paarl town and 1,158 identify a residential street. Requiring a classified cause leaves 1,135 deaths on 65 streets. We harmonise historical spellings against the municipal street plan.[16] We use place of death to fill missing residence only in a sensitivity check.

The birth and census counts cover the district, not just the town. To estimate street exposure, we first multiply district births by the smoothed share of district infant deaths recorded at town streets, including deaths without a classified cause. We then distribute that exposure across observed streets using estimated racial composition and one of three spatial weights: street length, line-segment counts or road density. Each allocation preserves the specified race–year total. We call these quantities conditional birth-equivalent exposures because they are estimates, rather than observed town or street births. Appendix A gives the assumptions and sensitivity checks.

We calculate district infant mortality rates as infant deaths per 1,000 reconstructed district live births. Figure 2 plots these cohort series by race. The descriptive differences motivate the street-level comparisons, but do not by themselves identify the contribution of drainage.

For each street, we divide deaths by allocated exposure and multiply by 1,000. The resulting level depends on the assumptions about coverage and residential allocation. Our count models include cells with no deaths. In the cause-relative specification, street–cohort–race fixed effects absorb the denominator.

[15] Appendix A describes transcription, geocoding, reverse survival and the conditional allocation of birth exposure to streets.

[16] SU Africana Library: E PAM 26 STR, Street Plan of Paarl, 1958.

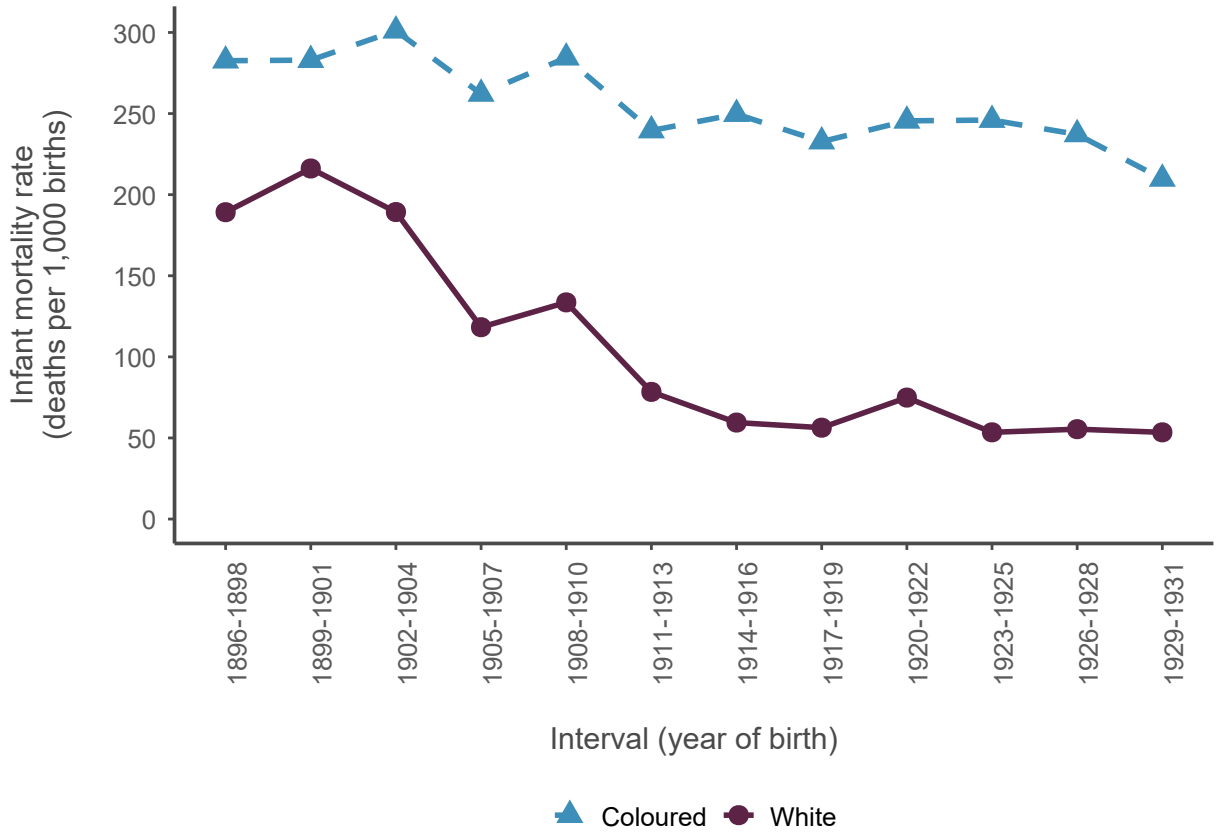


Figure 2: Infant mortality rate per 1,000 births by race, 1896–1931.

Notes: District infant deaths divided by reconstructed district births for the same cohort. The final plotted interval is truncated at 1930 to allow complete observation through the first birthday.
Source: Authors' calculations based on data described in the text.

We also examine what infants died from. The ICD10h system developed by Reid et al. (2024) groups the recorded causes into seven mutually exclusive categories: (1) water- and food-borne diseases, including diarrhoeal illnesses and enteric fevers; (2) broncho-pneumonia; (3) common infectious diseases, such as measles and whooping cough; (4) perinatal conditions and congenital weakness, including birth trauma, prematurity and debility; (5) convulsions, often associated with high fever or neurological problems; (6) other causes; and (7) ill-defined causes. These categories capture the period's main patterns of mortality. We refer to water- and food-borne diseases as *targeted* causes because they were particularly sensitive to sanitation and water quality. This distinction allows us to compare cause-specific changes across streets and racial groups.

Figure 3 shows the contribution of each classified cause to district infant mortality. Water- and food-borne disease, respiratory infections and neonatal conditions account for substantial mortality in both racial groups. These district trends describe a broader mortality transition; attributing any part of it to the drainage scheme requires the more local comparisons below.

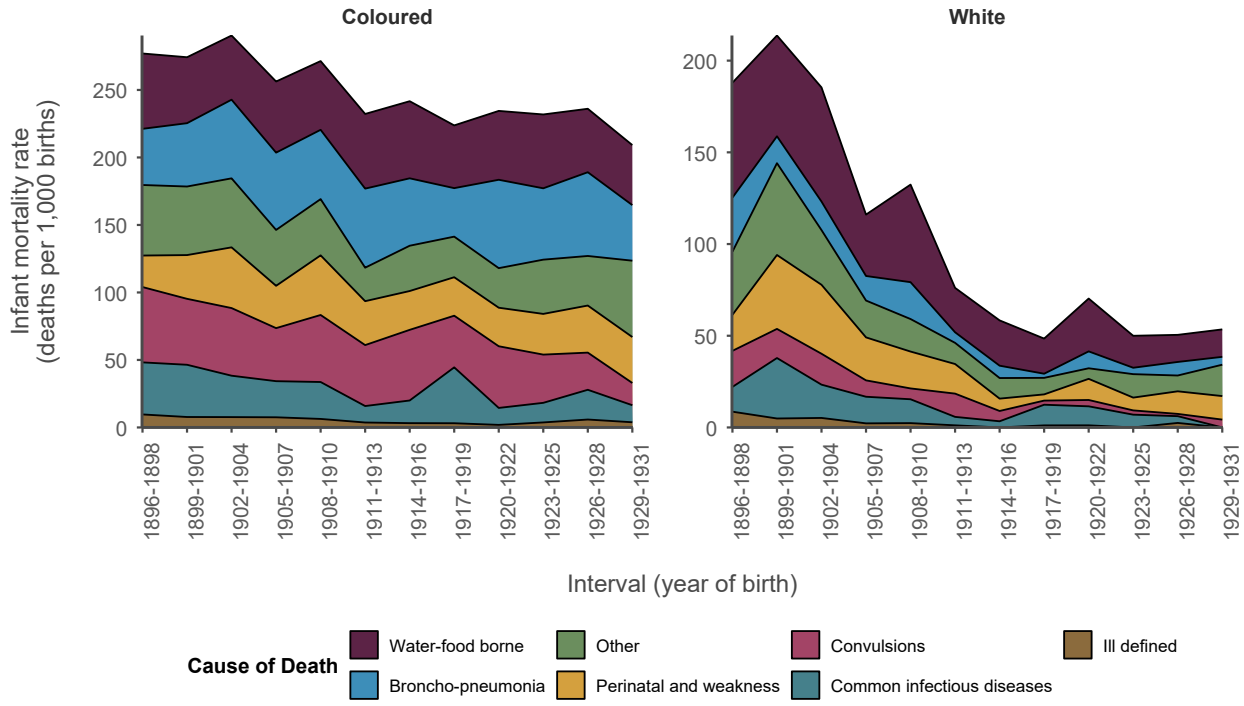


Figure 3: Cause-specific infant mortality rate per 1,000 births by race, 1896–1931.

Notes: Historical ICD10h cause groups for white and coloured infant deaths in the district, birth cohorts 1896–1930. Unclassified causes are excluded from this composition figure (257 notices). Numerator and denominator refer to the same cohort; the final plotted interval is truncated at 1930. Vertical scales differ by race.
Source: Authors' calculations based on data described in the text.

Table 1 describes the district population and the subset of infant notices used in the street regres-

sions. Panel A concerns district population; Panels B and C concern the observed residential-street sample. Panel D reports cause-specific deaths per 1,000 conditional birth-equivalent exposures, with zero-death years included.

Table 1: District population and observed street sample, 1896–1930

A: Population characteristics

	Population		Under 5 population			Under 1 population		
Race	Mean	SD	Mean	SD	Prop.*	Mean	SD	Prop.*
Coloured	18,846	660	3,013	173.0	16.0%	713	8.7	3.8%
White	12,697	833	1,257	69.9	9.9%	264	14.9	2.1%

B: Death-notice sample characteristics

Variable	Coloured	White
Number of death notices	961	174
District births	28,146	9,800
Count male/female	517/444	85/89
Medical attendant (%)	87.3%	94.8%
Unique doctors**	38	29
Informant signed (%)	41.1%	95.4%
Illness duration noted (%)	94.8%	96.6%
Unique street names	60	38
Father occupation noted (%)	32.0%	60.9%
Mother occupation noted (%)	16.0%	2.9%

C: Death notices by street

Race	Street	n	%
Coloured	Lady Grey	23	2.4%
Coloured	Main	59	6.1%
Coloured	Malherbe	21	2.2%
Coloured	Vd Lingen	33	3.4%
Coloured	All other	825	85.8%
White	Lady Grey	20	11.5%
White	Main	50	28.7%
White	Malherbe	4	2.3%
White	Vd Lingen	21	12.1%
White	All other	79	45.4%

D: Cause-specific infant mortality rates per 1,000 births

	Coloured				White			
Cause of death	Mean	SD	Min	Max	Mean	SD	Min	Max
Water-food borne	66.2	56.3	0	231.4	39.3	52.6	0	243.8
Broncho-pneumonia	48.8	44.6	0	172.4	6.2	11.5	0	41.6
Perinatal and weakness	45.7	43.9	0	172.4	17.8	28.3	0	110.8
Other	39.8	33.9	0	114.9	20.7	24.1	0	85.7
Convulsions	24.8	25.6	0	95.8	3.7	8.6	0	28.3
Common infectious diseases	24.8	28.3	0	124.6	12.4	21.1	0	92.6
Ill defined	6.4	14.8	0	76.6	2.8	8.1	0	30.9

Notes: Cohorts 1896–1930. Panel A concerns interpolated district population and Panel B lists district births alongside the primary residential-street death sample. Smooth-series standard deviations are descriptive, not sampling uncertainty. Panel D uses conditional birth-equivalent exposure with the single declared coverage adjustment; each cause-year with no deaths enters as zero. Medical attendance and informant signature are different fields. No additional level calibration is imposed.

*Prop. refers to the proportion the demographic contributes to the total population.

**Unique doctors refer to the names of the 'medical men' listed on the death notices.

Source: Authors' calculations based on data described in the text.

Panel B of Table 1 summarises the recorded characteristics of the primary street sample. Medical attendance, informant signature and recorded occupations are distinct fields. The district birth total provides context and should not be read as births observed on the sampled streets.

Panel C shows how sampled deaths are distributed across the four mapped treated streets and other observed streets. Machinery Street could not be located and is excluded. These counts describe the estimating sample, not the distribution of all households in Paarl.

Panel D reports cause-specific deaths per 1,000 conditional birth-equivalent exposures, averaged across all study years including zero-death years. These levels depend on the district-to-street coverage assumption and the allocation procedure. They are distinct from the district rates plotted in Figure 3.

The records have several limitations. Death registration was incomplete across South Africa, particularly among rural and black or coloured families, so some infant deaths never entered the register. Surviving notices also contain missing, incomplete or vague entries. Diagnoses such as *convulsions* or *debility* could conceal underlying conditions, a problem recognised at the time (Buirski, 1983). We exclude notices giving a farm or hamlet outside the town's street grid because they cannot be reliably matched for spatial analysis. Allocating births by street length also introduces measurement error into the mortality denominator.

Registration may have changed over time as well. We cannot observe completeness directly, but historical evidence points to particularly severe under-registration during crises. The 1918–1919 influenza pandemic, for example, led to substantial undercounting of deaths among South Africa's non-white population (Fourie and Jayes, 2021). Changes in reporting that differed by group could alter measured mortality rates. The initial drainage rollout did not coincide with the periods in which we most expect disruption, and comparisons within the town are less sensitive than aggregate rates to uniform reporting changes. Neither consideration rules out time-varying measurement error.

Linking individual deaths to streets and causes gives us a detailed view of recorded infant mortality before and after the works. The birth estimates make rate comparisons possible, but those comparisons remain conditional on recorded location and reconstructed exposure. We next explain how we use this variation and the pre-intervention patterns.

4 Empirical strategy and results

We first examine where the council built drains, comparing mortality on future-treated and never-treated streets before construction. These comparisons assess whether the historical focus on central streets coincided with unusually high white infant mortality from sanitation-sensitive causes; their sensitivity limits that interpretation. We then use triple differences and event studies to describe subsequent mortality changes by street, race and cause. Finally, we relate mortality and the duration of infants' final illnesses to geography, asking whether the patterns are consistent with changes in where people encountered contaminated water.

4.1 Empirical framework

Births are not observed at the street level. Let B_{tr} denote the conditional total obtained by scaling reconstructed district births by the share represented by recorded town streets, and let $w_s^{(m)}$ include the spatial weight and estimated racial share for street s. We allocate this exposure within the declared observed-street universe $\mathcal{S}$:

$$\widehat{B}_{trs}^{(m)} \;=\; B_{tr}\,\frac{w_s^{(m)}}{\sum_{j\in\mathcal{S}} w_j^{(m)}}, \text{ and } \sum_{s\in\mathcal{S}} \widehat{B}_{trs}^{(m)} \;=\; B_{tr}.$$

We preserve the specified exposure total once for each street–race–year before repeating it across causes. This accounting rule does not establish that the allocated exposure equals the births that actually occurred on each street.

The three weighting schemes make different assumptions about residential intensity. The first uses street length; the second counts the straight-line segments in each street's geometry; the third sums a grid-based road-density measure over the cells that the street crosses. We use them as alternative ways to measure exposure.[17]

Our main outcome is a street-level infant mortality rate (*IMR*) by race and cause of death. For each street s, birth year t, race $r \in \{\text{white}, \text{coloured}\}$, and cause-of-death category c, we define

$$IMR_{trsc} \;=\; \frac{deaths_{trsc}}{\widehat{B}_{trs}^{(m)}} \times 1{,}000,$$

where $deaths_{trsc}$ counts deaths before the first birthday and $\widehat{B}_{trs}^{(m)}$ is the conditional exposure defined above. For comparison, we retain weighted OLS estimates of log rates among cells containing deaths. Our principal count models use Poisson pseudo-maximum likelihood (PPML), retaining zero-death cells and including log exposure as an offset.

To structure the analysis, we estimate variants of the following fully saturated specification:

$$\begin{aligned}\log(IMR_{trsc}) = {} & \beta_0 + \beta_1\, Treatment_{st} + \beta_2\, Targeted_c + \beta_3\, Race_r \\ & + \beta_4\,(Treatment_{st}\cdot Targeted_c) + \beta_5\,(Treatment_{st}\cdot Race_r) \\ & + \beta_6\,(Targeted_c\cdot Race_r) + \beta_7\,(Treatment_{st}\cdot Targeted_c\cdot Race_r) \\ & + \alpha_s + \gamma_t + \varepsilon_{trsc}, \end{aligned} \tag{1}$$

Here, $Treatment_{st}$ equals one from the coded adoption year onwards, including that year: 1905 for Lady Grey and Main, and 1907 for Malherbe and Van der Lingen. We use Main Street's later phase in 1906 as a timing sensitivity check. Because notices locate infants only at street level, we cannot distinguish the drained and undrained sections of a partially treated street.

The indicator $Targeted_c$ equals one for water- and food-borne causes. The displayed OLS model includes street and year fixed effects, birth-exposure weights and standard errors clustered

[17] See Appendix A.

by street. Its exclusion of zero-death cells is a material limitation. PPML includes the complete grid for observed streets. A further, cause-relative model absorbs street × cohort × race and cause fixed effects. These remove any offset that is constant within street–cohort–race, but do not resolve selective observation.

Equation (1) makes three comparisons: treated with untreated streets, targeted with other causes, and white with coloured infants. Given the pre-intervention patterns discussed below, we use this triple difference to describe relative mortality changes rather than to recover an average treatment effect of drainage. The coefficients $(\beta_1, \ldots, \beta_7)$ jointly describe how the treated–untreated gap in log mortality changed after drainage for each racial and cause group.

To trace dynamics more flexibly we estimate pooled PPML indexed to each street's own adoption year. In the following expression, g indexes the four race–cause groups, with other causes pooled against targeted causes:

$$\log E[D_{stg} \mid X] = \log \widehat{B}_{str} + \sum_{k \neq [-3,-1]} \beta_{kg} D^k_{st} + \alpha_{sg} + \gamma_{tg}, \qquad k = t - a_s, \tag{2}$$

where a_s is the adoption year and D^k_{st} is the corresponding event-bin indicator for ever-treated streets. The bins are ≤ -10, -9 to -7, -6 to -4, -3 to -1 (reference), 0 to 2, 3 to 5, 6 to 8, and 9+ years. Never-treated streets supply the untreated benchmark. The model is jointly estimated with group-specific street and year effects, retaining cross-cause and cross-race covariance within street clusters.

Coding the two adoption cohorts separately distinguishes a 1907 birth on a street drained in 1905 from one on a street drained in 1907. We estimate pooled PPML rather than the interaction-weighted estimator of Sun and Abraham (2021), so differences in effects across adoption cohorts may affect interpretation. With only four treated streets, pointwise intervals and joint contrasts are approximate. Omitting one treated street at a time shows how much individual streets influence the estimates. Neither non-significant pre-period estimates nor a large control group establishes reliable causal inference in this small treated sample.

To examine possible mechanisms, we turn to individual notices and geographic differences across streets. Drains redirect surface run-off and waste, potentially changing where contaminated water collects or disperses. A street's physical features and position in the drainage network may therefore matter. We relate the duration of last illness ($DoLI$) for infant i on street s in year t to these features:

$$\begin{aligned} \log(DoLI_{ist} + 1) = {} & \delta_0 + \delta_1 \, Treatment_{st} + \delta_2 \, Geo_s + \delta_3 \, Race_i + \delta_4 \, (Treatment_{st} \cdot Geo_s) \\ & + \delta_5 \, (Geo_s \cdot Race_i) + \delta_6 \, (Treatment_{st} \cdot Geo_s \cdot Race_i) \\ & + \delta_7 (Treatment_{st} \cdot Race_i) + \theta \, Targeted_i + \mu \log(Age_i + 1) + \gamma_t + \varepsilon_{ist}, \end{aligned} \tag{3}$$

where Geo_s denotes standardised street slope, mean elevation, minimum sampled river distance or orientation. The models control for targeted cause and log age. We also estimate fully saturated

treatment–cause–geography–race interactions, with and without street fixed effects. We exclude doctor attendance as a control because treatment may itself affect attendance. Reported gradients sum the relevant terms using their joint covariance. These comparisons concern infants who died and whose illness duration was recorded; they identify neither survival changes nor the route of infection.[18]

Equations (1)–(3) organise the comparisons that follow. We begin with the streets the council selected and their mortality patterns before drainage.

4.2 Pre-intervention targeting and pre-trends

Why did the municipality choose these streets? No engineering plans survive, but petitions and medical reports drew attention to the foul open *sluits* along Main Street and neighbouring central streets, where white elites owned and operated properties. The death notices allow us to ask whether those concerns coincided with systematic mortality differences between future-treated and never-treated streets before drainage.

Table 2 compares pre-intervention mortality levels on streets that would and would not receive drains. These selection regressions distinguish water- and food-borne from other causes and white from coloured infants. They use births before drainage on treated streets and a matched pre-treatment window for never-treated streets, include birth-year fixed effects and weight observations by estimated street-level births. Since no treated street had yet received drainage in this sample, the coefficients describe pre-existing differences.

[18] Appendix A describes measurement and standardisation.

Table 2: Pre-intervention selection on infant mortality by race and cause of death

Street-level births measure	Street length		Street segments		Raster	
	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)
Ever treated	−0.212	0.560***	0.488	0.442**	−0.205	0.495**
	(0.406)	(0.194)	(0.411)	(0.186)	(0.428)	(0.193)
Targeted	0.033	0.097	−0.082	−0.035	0.033	0.094
	(0.134)	(0.147)	(0.130)	(0.103)	(0.134)	(0.142)
White	0.884**	1.708***	0.287	0.344	0.647	1.447***
	(0.406)	(0.224)	(0.528)	(0.455)	(0.462)	(0.265)
Ever Treated × Targeted	0.125	−0.168	0.157	−0.018	0.133	−0.150
	(0.236)	(0.141)	(0.231)	(0.111)	(0.239)	(0.135)
Ever Treated × White	−1.972***	−2.810***	−1.473***	−1.561***	−1.718***	−2.530***
	(0.440)	(0.230)	(0.531)	(0.457)	(0.492)	(0.272)
Targeted × White	0.196	−0.534	0.897***	0.433	0.374	−0.342
	(0.326)	(0.401)	(0.219)	(0.275)	(0.294)	(0.376)
Ever Treated × Targeted × White	−0.057	0.988**	−0.641*	0.057	−0.233	0.788*
	(0.448)	(0.427)	(0.375)	(0.300)	(0.427)	(0.403)
Mean street elevation		0.021**		−0.031***		0.021**
		(0.008)		(0.008)		(0.008)
Distance from river (m)		−0.001***		0.001***		−0.001***
		(0.000)		(0.000)		(0.000)
North-to-South street		−1.247***		−1.217***		−1.317***
		(0.145)		(0.138)		(0.141)
Num.Obs.	254	254	254	254	254	254
R^2 Adj.	0.499	0.728	0.290	0.634	0.480	0.738
R^2 Within Adj.	0.401	0.676	0.158	0.566	0.378	0.687
FE: Birth Year	Yes	Yes	Yes	Yes	Yes	Yes
Street clusters	44	44	44	44	44	44

Notes: *p<0.1, **p<0.05, ***p<0.01. OLS estimates for the pre-intervention sample. The outcome is log street-level infant mortality per 1,000 weighted births. The sample is limited to pre-drainage births on treated streets and the comparable pre-period on never-treated streets. *Ever treated* indicates streets ever drained; *Targeted* indicates water- and food-borne causes. Street-clustered standard errors are in brackets; observations are weighted by street-level births.
Source: Authors' calculations based on data described in the text.

The comparisons in Table 2 change with the allocation scheme and geographic controls. They also condition on a death having occurred in each cell. The interaction coefficients therefore cannot establish that the council selected streets because of a particular cause-specific mortality excess.

Including zero-death cells and street fixed effects changes the comparison. The replication tables report these count models and their samples, separating pre-existing differences from subsequent mortality changes. The historical evidence for selective provision does not depend on a significant pre-intervention mortality coefficient.

The pooled descriptive rates count exposure once per street–cohort–race, including when other causes are combined. Their levels still depend on reconstructed births and coverage assumptions. Similar point estimates do not establish equivalence, and these comparisons cannot recover the council's reasons for choosing particular streets.

The archives place the works on prominent central thoroughfares where white households and commercial property were concentrated. The mortality comparisons complement that evidence, but their instability prevents us from concluding that white diarrhoeal deaths specifically prompted the scheme.

We next examine mortality trends on the streets eventually drained. Table 3 uses only years before each street's adoption and allows the birth-year trend, centred on 1900, to vary by race and cause. These trends describe cells containing deaths. They do not test whether treated and control streets would otherwise have followed parallel paths.

The interaction of targeted cause, birth-year trend and coloured classification describes a racial difference in the pre-intervention cause gradient. Because the sample includes only treated streets, it cannot show whether mortality moved similarly elsewhere. The event study provides that broader comparison, although sparse outcomes and non-random treatment still limit its interpretation.

Only four treated streets can be mapped, so the table contains four clusters. We report standard errors for completeness, without significance stars or a claim that pre-trends are absent. The replication logs and sensitivity results document repairs to the variance matrix and the influence of individual streets.

The historical and mortality evidence gives us no basis for treating street selection as random. The pre-intervention comparisons examine differences in mortality levels and trends alongside petitions and official attention to the central *sluits*; their sensitivity limits how closely we can connect those differences to municipal priorities. The works were concentrated in a few prominent central streets. Four can be mapped and enter the analysis: Lady Grey, Main, Malherbe and Van der Lingen. That small treated group limits every street-level comparison. We now examine how the racial distribution of infant mortality changed after construction.

4.3 Consequences for infant mortality

Table 4 uses Equation (1) to describe street-level infant mortality over 1896–1930. It compares changes on treated and untreated streets across racial groups and causes of death. Because treatment was not randomly assigned, we do not interpret these estimates as causal effects.

Table 3: Pre-intervention trends in infant mortality by race and cause of death on treated streets

Street-level births measure	Street length	Street segments	Raster
	(1)	(2)	(3)
Intercept	4.709	5.241	4.788
	(0.377)	(0.333)	(0.379)
Targeted	0.199	0.170	0.199
	(0.218)	(0.235)	(0.217)
Coloured	1.217	1.251	1.193
	(0.173)	(0.166)	(0.172)
Birth year (centred on 1900)	0.075	0.069	0.075
	(0.036)	(0.024)	(0.036)
Targeted × Coloured	0.054	0.097	0.051
	(0.876)	(0.933)	(0.827)
Targeted × Birth year	0.135	0.134	0.135
	(0.024)	(0.021)	(0.025)
Birth year × Coloured	0.194	0.198	0.194
	(0.049)	(0.047)	(0.048)
Targeted × Birth year × Coloured	−0.407	−0.435	−0.403
	(0.876)	(0.764)	(0.836)
Num.Obs.	67	67	67
Streets (clusters)	4	4	4
R^2 Adj.	0.330	0.405	0.327
Street clusters	4	4	4

Notes: OLS pre-trend estimates for the streets that were later drained, restricted to years before drainage arrived. The outcome is log street-level infant mortality per 1,000 weighted births, allocated under the scheme named in the column header, and observations are weighted by the same measure. *Targeted* identifies water- and food-borne causes; *Coloured* is the race contrast, with white infants omitted. Year of birth is centred on 1900, so the level terms describe the middle of the pre-period. Standard errors are clustered on street of residence and reported in brackets. The sample contains four treated streets, so these are four clusters and no significance is claimed from them.
Source: Authors' calculations based on data described in the text.

Table 4: Street-level drainage and infant mortality

Street-level births measure	Street length			Street segments			Raster		
	(1.1)	(1.2)	(1.3)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)
Intercept	5.113***	5.070***		5.620***	5.578***		5.207***	5.168***	
	(0.465)	(0.505)		(0.333)	(0.342)		(0.463)	(0.500)	
Treatment	−0.819***	−0.795**	0.314**	−0.778***	−0.747***	0.204	−0.835***	−0.816**	0.303**
	(0.300)	(0.334)	(0.122)	(0.187)	(0.188)	(0.126)	(0.298)	(0.329)	(0.124)
Targeted		0.122	0.363***		0.137	0.287***		0.112	0.359***
		(0.139)	(0.052)		(0.082)	(0.077)		(0.134)	(0.052)
Coloured	0.299	0.328	0.480	−0.302	−0.268	0.261	0.253	0.279	0.439
	(0.536)	(0.574)	(0.359)	(0.511)	(0.527)	(0.413)	(0.548)	(0.584)	(0.360)
Treatment × Targeted		−0.072	−0.228***		−0.106	−0.168**		−0.060	−0.224***
		(0.116)	(0.044)		(0.075)	(0.066)		(0.114)	(0.045)
Treatment × Coloured	0.539	0.515	0.143	1.132*	1.097*	0.393	0.564	0.543	0.162
	(0.643)	(0.673)	(0.333)	(0.577)	(0.582)	(0.391)	(0.651)	(0.679)	(0.333)
Targeted × Coloured		−0.070	−0.363***		−0.106	−0.268***		−0.066	−0.356***
		(0.147)	(0.064)		(0.097)	(0.085)		(0.141)	(0.064)
Treatment × Targeted × Coloured		0.076	0.244***		0.126	0.167*		0.072	0.237***
		(0.129)	(0.064)		(0.098)	(0.087)		(0.125)	(0.064)
Num.Obs.	949	949	935	949	949	935	949	949	935
R^2 Adj.	0.139	0.136	0.853	0.034	0.030	0.855	0.128	0.125	0.850
R^2 Within Adj.			0.193			0.139			0.176
FE: Street residence			Yes			Yes			Yes
FE: Birth year			Yes			Yes			Yes
Street clusters	65	65	52	65	65	52	65	65	52

Notes: *p<0.1, **p<0.05, ***p<0.01. OLS estimates of Eq. 1. The sample includes infant deaths in Paarl, 1896–1930, with age below one year and a recorded residential street. The outcome is the logarithm of street-level infant mortality per 1,000 estimated births. *Treatment* indicates birth after drainage on the street; *Targeted* indicates a water- or food-borne cause; *Coloured* indicates racial classification. Street-clustered standard errors are in brackets; observations are weighted by street-level births. These specifications condition on street–year–race–cause cells containing at least one death, which is 3.0 per cent of the notional panel; Table 5 re-estimates the same specifications on the full grid with the zeros retained.
Source: Authors' calculations based on data described in the text.

The cause interactions describe relative changes, not total changes in mortality from targeted causes. For white infants, the targeted total is the sum of the treatment coefficient and its interaction with targeted cause. For coloured infants, the racial interactions must also be included. The uncertainty of each sum depends on the full covariance matrix.

Table 4 includes only street–cohort–race–cause cells with at least one death. Most cells have none, and many positive cells have exactly one. For a one-death cell, $\log(\text{IMR}) = \log(1000) - \log(\text{exposure})$. Much of the variation in these log-rate comparisons therefore comes from the constructed denominator.

Table 5 includes zero-death cells and estimates counts by PPML, with both exposure-offset and denominator-free specifications. We reconcile the input samples explicitly, distinguishing groups removed by the estimator because all outcomes are zero from observations excluded for missing data. These coefficients do not establish that mortality benefits from drainage differed by race.

We consequently do not claim an established racial difference in the response to drainage. A cause-relative interaction also cannot establish a fall in overall infant mortality. The predeclared all-cause auxiliary model in Appendix C includes unclassified causes and estimates a separate coefficient for each racial group, using each street's own adoption year.

Figure 4 shows the timing of mortality changes using joint cause–race covariance. The estimates are imprecise and do not establish a sanitation-specific decline among white infants after drainage. The pointwise intervals are conditional on reconstruction and measurement choices, and are approximate with four treated streets.

Table 5: Infant mortality dynamics retaining zero-death cells

	PPML (length)	PPML (segments)	PPML (raster)	PPML, no denominator	log(IMR + 1) OLS
	(1)	(2)	(3)	(4)	(5)
Treatment	−0.486***	−0.486***	−0.486***		−0.629***
	(0.163)	(0.163)	(0.163)		(0.062)
Targeted	1.322***	1.322***	1.322***		0.340
	(0.201)	(0.201)	(0.201)		(0.222)
Coloured	0.990***	1.015***	0.969***		0.207***
	(0.143)	(0.143)	(0.143)		(0.041)
Treatment × Targeted	0.007	0.007	0.007	0.007	0.514***
	(0.186)	(0.186)	(0.186)	(0.187)	(0.156)
Treatment × Coloured	0.252	0.252	0.252		0.292***
	(0.173)	(0.173)	(0.173)		(0.052)
Targeted × Coloured	−0.517**	−0.517**	−0.517**	−0.517**	0.192
	(0.228)	(0.228)	(0.228)	(0.228)	(0.268)
Treatment × Targeted × Coloured	−0.160	−0.160	−0.160	−0.160	−0.349
	(0.212)	(0.212)	(0.212)	(0.212)	(0.238)
Num.Obs.	30 940	30 940	30 940	4473	31 850
FE: Street residence	Yes	Yes	Yes		Yes
FE: Birth year	Yes	Yes	Yes		Yes
FE: Cause of death				Yes	
FE: Street × Birth year × Race				Yes	
Street clusters	65	65	65	65	65

Notes: PPML count models retain zero-death cells in the observed-street panel. Columns (1)–(3) use log conditional birth exposure as an offset and street/year fixed effects. Column (4) instead absorbs street–cohort–race and cause fixed effects; these absorb any exposure constant within the cell. The common-sample offset identity is checked separately and does not resolve selective recording. Column (5) is weighted log(rate + 1) OLS. Standard errors cluster by street. Stars use conventional approximate inference; only four streets were treated.

Source: Authors' calculations based on data described in the text.

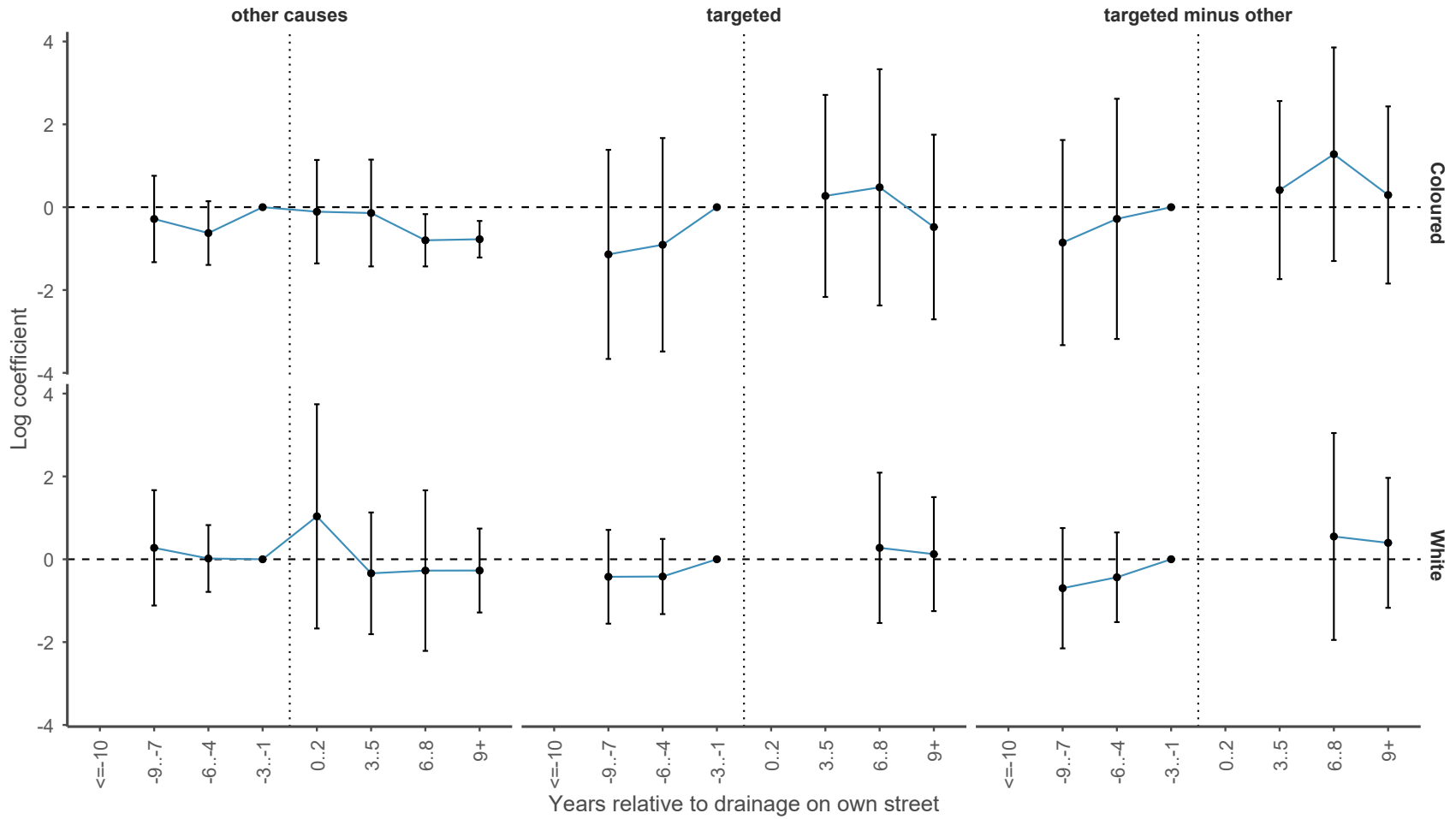


Figure 4: Event-study estimates of drainage improvements on infant mortality by race and disease category.

Notes: Pooled PPML event-time estimates with group-specific street and year effects and a log exposure offset. Event time follows own-street adoption; the omitted bin is three to one years before adoption and the end bins are open. Never-treated streets supply the untreated benchmark. A joint cause–race model supplies full street-cluster covariance for the difference column; disjoint cause samples are not statistically independent when they share streets. Intervals are pointwise estimate ±1.96 clustered standard errors and are approximate with four treated streets. Fixed street allocation factors are absorbed by the group-specific fixed effects. Blank estimates have no exposed deaths or fail a solver-tolerance check; the latter affects the white targeted coefficient in years zero to two and its contrasts. The dotted line separates pre- and post-adoption bins.

Source: Authors' calculations based on data described in the text.

Appendix C reports a post-period model that pools years after each street's adoption rather than averaging event bins. It separates changes in targeted and other causes within each racial group. The cause-relative contrasts do not establish a difference between white and coloured infants.

Changes in other causes could reflect municipal investment, residential composition, recording or the broader mortality transition. The event study cannot distinguish these explanations. Failure to reject zero also does not show that a group received no benefit.

The third column of Figure 4 subtracts the other-cause coefficient from the targeted-cause coefficient. Because both outcomes occur on the same streets, their covariance enters the standard error. The between-race difference is likewise a joint contrast. A significant coefficient for one group and an insignificant coefficient for another would not, by itself, establish unequal benefits.

A small sample does not justify dismissing pre-period departures as noise. The replication results report a joint pre-period diagnostic and estimates that omit each treated street in turn. Low power means that non-rejection cannot validate parallel trends.

The regressions and event study leave the mortality consequences of selective sanitation uncertain. We can date the works, but these comparisons firmly establish neither a mortality gain nor its racial distribution. The next step is to examine geographic associations.

4.4 Geography and possible redistribution of risk

Stormwater drains moved contaminated water without necessarily treating it. They could therefore change exposure on nearby streets, although risk displacement need not make health effects zero-sum. Given Paarl's slope towards the Berg River, drainage could plausibly clear filth from one area, providing highly localised relief, while redistributing that filth, and the risk it carried, elsewhere.

The data cannot identify this mechanism. Directed run-off exposure and simple network proximity are nearly collinear. Relative elevation and river proximity describe the same selected network, and their coefficients do not show which way pathogens travelled.

We can make a narrower comparison. Table 6 relates the duration of final illness to each geographic feature of the residential street, separately by race, for sanitation-sensitive deaths on treated streets after drainage.[19]

We analyse reported illness duration as $\log(\text{days}+1)$. Since duration is observed only for infants who died, the comparison cannot identify a survival effect or avoid selection into the death sample. Four street features enter the analysis: *Slope*, which may affect surface-water removal; *Elevation*, which locates the street within the drainage basin; *River distance*, which measures proximity to the Berg River; and *Road direction*, which describes street orientation. The estimates relate geography to recorded fatal illness.

Table 6 reports the post-drainage geographic gradients with pointwise intervals based on the joint covariance matrix. We standardise slope across unique streets in the primary observed sample. River distance here is the minimum distance across sampled street points.

[19] Appendix B plots water- and food-borne infant deaths against standardised street slope by race and period. That figure describes the raw spatial pattern; it is not offered as evidence for a hydrological mechanism.

Table 6: Geographic gradients in duration of last illness

Contrast	Estimate	Std. error	95% interval
Slope (street-standardised) White, targeted, post	-0.811	0.193	[-1.189, -0.433]
Slope (street-standardised) Coloured, targeted, post	0.035	0.069	[-0.100, 0.169]
Minimum river distance White, targeted, post	0.001	0.000	[0.001, 0.002]
Minimum river distance Coloured, targeted, post	0.000	0.000	[-0.000, 0.000]
Elevation White, targeted, post	0.032	0.007	[0.018, 0.047]
Elevation Coloured, targeted, post	0.001	0.004	[-0.007, 0.009]
North–south White, targeted, post	0.636	0.158	[0.327, 0.945]
North–south Coloured, targeted, post	0.039	0.087	[-0.131, 0.209]

Notes: Race-specific post-drainage gradients for targeted deaths in fully interacted OLS with birth-year fixed effects and log-age controls. Street-cluster covariance; approximate pointwise intervals. Slope is standardised across unique primary streets; elevation and minimum sampled river distance are in metres. The sample conditions on death and recorded illness duration. Estimation observations: 1078; street clusters: 65.

These gradients concern infants who died and whose notices provide usable information. Illness duration can vary with severity, diagnosis, care and recording. The estimates therefore cannot isolate a change in environmental exposure.

The analysis remains descriptive. An individual geographic association, especially among several related comparisons, cannot establish that drainage displaced risk or restricted benefits to one racial group.

4.5 Spillovers beyond the treated streets

Drainage could affect adjoining streets as well as those receiving pipes. We examine this possibility by relating mortality to the surrounding network. The comparisons retain the same limits arising from selective placement and the small number of independent interventions.

We divide the mapped network into treated streets, untreated streets downhill of them, streets near the river and downstream of the inferred outfall, and the remaining streets. These categories describe eventual network position. Mortality estimation uses only the 65 observed streets, not the entire map. A continuous exposure measure weights treated length by downhill network distance and river proximity, using primary decay and river scales of 400 and 600 ground metres. Appendix A explains the check using scales equivalent to the original coordinate settings.

We estimate PPML with street × cohort × race and cause fixed effects. The model includes cause–exposure and common post–cause terms, their race counterparts and the higher interactions.

The common post indicator begins in 1905, comparing scheme introduction across eventual network positions. It does not indicate when every outfall became active. A timing alternative excludes 1905–06; the available topology cannot support separate claims about annual source activation.

Table 7 compares each class with Unaffected streets and reports continuous exposure slopes by race, using the full covariance matrix. These estimates describe changes relative to other causes, not all-cause mortality effects. Allowing for pre-existing cause composition and common cause-specific changes materially alters the earlier spatial estimates.

The contrasts do not establish spatial mortality benefits confined to white infants. The coloured slope is the sum of the reference slope and racial interaction, and an imprecise sum does not imply an absent effect. The all-cause auxiliary model separately examines the broader mortality claim.

Directed run-off exposure and simple network proximity are almost perfectly correlated, making their joint coefficients unstable. Relative elevation provides another geographic association. Neither it nor proximity identifies a hydrological mechanism.

A placebo comparison treats broncho-pneumonia as the targeted cause. A non-significant placebo would not rule out differential recording, spatial confounding or correlated changes in other causes.

All exposure variation still comes from four treated streets. Adding nearby streets supplies no further independent works and does not justify precise causal inference. The spatial analysis therefore leaves the uncertainty in the treated-street comparison unresolved.

4.6 Robustness and limitations

The interpretation depends on assumptions about historical recording, spatial matching and reconstructed exposure. Appendix C and the replication files report checks on these assumptions, separating sample changes from estimator changes.

Usable residence information determines entry into the sample. We distinguish notices identifying town residence, those identifying a street within town, and those meeting the remaining analysis requirements. These checks cannot establish stable recording by unobserved treatment status: when the street is unknown, so is its treatment status.

Small denominators and sparse counts are a further concern. Winsorisation and exclusions of small-exposure cells check the sensitivity of positive-cell OLS, but cannot address the selection caused by excluding cells with no deaths. Estimates omitting one treated street at a time show how far the findings depend on individual streets.

The placebo and within-street medical-attendance comparisons do not validate the mortality design. A recorded maternal occupation is only a proxy for employment. Residence–death agreement is observable only when both streets were recorded independently. Illness duration and attendance concern infants who died, so they cannot explain a racial survival difference that the mortality models have not established.

The four treated streets were not randomly chosen. Street fixed effects cannot remove unobserved changes in housing, other municipal services, household behaviour or recording around them.

Table 7: Cause-relative spatial contrasts by race

Contrast	Estimate	Std. error	95% interval
White Downhill of treated minus Unaffected	1.232	1.209	[-1.138, 3.602]
Coloured Downhill of treated minus Unaffected	-0.138	0.464	[-1.048, 0.771]
White Downstream of outfall minus Unaffected	-0.377	1.338	[-2.999, 2.245]
Coloured Downstream of outfall minus Unaffected	-0.191	0.475	[-1.123, 0.741]
White Treated minus Unaffected	0.721	1.027	[-1.291, 2.733]
Coloured Treated minus Unaffected	0.694	0.436	[-0.161, 1.549]
White overland_z cause-relative marginal slope	0.029	0.039	[-0.048, 0.106]
Coloured overland_z cause-relative marginal slope	0.084	0.017	[0.051, 0.117]
White river_z cause-relative marginal slope	0.118	0.246	[-0.363, 0.600]
Coloured river_z cause-relative marginal slope	-0.013	0.124	[-0.257, 0.231]

Notes: PPML with street × cohort × race and cause fixed effects; all cause–exposure and common post–cause lower-order terms and race counterparts are included. Classes subtract Unaffected explicitly. Continuous slopes sum the relevant race terms. Post starts in 1905 against eventual network position. Full street-cluster covariance; approximate pointwise intervals. Four treated source streets. These are cause-relative changes, not all-cause effects. Estimation observations: 4473; street clusters: 65.

Conventional clustered intervals are approximate. We claim neither exact randomisation inference nor a bootstrap correction with validated coverage in this sparse sample.

Street exposure remains model-dependent. On a common sample, the denominator-free specification absorbs any street–cohort–race offset, an identity we check algebraically and numerically. This does not establish that the offset and denominator-free models estimate the same quantity or that selection is harmless. Partial-street treatment and errors in recorded residence also need not bias every contrast towards zero.

The evidence supports selective municipal provision, while leaving the direction, size and racial distribution of mortality changes uncertain. This distinction shapes what Paarl can contribute to the history of sanitation and health inequality.

5 Conclusion

We know more about sanitation's contribution to falling mortality than about how selective provision shaped health inequality within colonial towns. Paarl brings these questions together. Between 1905 and 1907, its council directed stormwater drainage towards central streets where white households and commercial property were concentrated, while many crowded coloured neighbourhoods retained open drains. The scheme thus extended public investment along existing social divisions. Whether it also widened the racial mortality gap is a separate empirical question, which the surviving evidence leaves unresolved.

We address that question by linking individual death notices to the location and timing of drainage works, reconstructing district births, and comparing infant mortality across streets, racial groups and causes of death. The analysis follows birth cohorts from 1896 to 1930 and combines triple-difference comparisons with event studies and count models that retain cells with no deaths. The results do not consistently establish a decline in water- and food-borne mortality after drainage, and the all-cause estimates are too imprecise to determine whether either racial group benefited more. Geographic comparisons likewise cannot separate exposure to contaminated run-off from proximity to the works. These findings are insufficient to determine whether drainage had a health effect, because selective recording, reconstructed street exposure and four non-randomly treated streets limit what the comparisons can establish.

The contribution is to connect the political geography of sanitation provision to an explicit assessment of its distributional consequences. Studies of American water purification show that shared systems could narrow racial mortality gaps (Troesken, 2004; Anderson et al., 2021); research on segregation shows that early investment could coexist with delayed universal access (Beach, Parman and Saavedra, 2022). Paarl extends this discussion to a colonial African town where reform was confined to selected streets. The result was an incomplete sanitary system, an instance of what Wright (2001) termed the "ambiguous modernisms" of colonial cities. Paarl's experience shows that the presence of infrastructure does not translate uniformly into protection from disease. That finding complements the evidence on the joint benefits of effective water and sewerage systems

(Alsan and Goldin, 2019). The location of new drains records a municipal allocation decision; establishing who gained requires evidence on mortality and exposure beyond that location.

The paper also shows why the way historical deaths are observed and counted matters for this distributional question. Among street–cohort–race–cause cells containing a single death, variation in log mortality rates comes from reconstructed exposure. Excluding cells with no deaths can therefore make comparisons depend heavily on the denominator rather than on differences in recorded deaths. Retaining those cells and estimating racial contrasts jointly changes the assessment of apparent health gains. Linking detailed notices to infrastructure is valuable, but the resulting spatial detail does not by itself resolve selective observation or supply independent policy variation. This is a practical qualification to the use of historical administrative records to explain health inequality.

The remaining gap is to establish when selective sanitation provision translated into unequal protection. Household addresses and independent birth records would help distinguish residence, exposure and recording; engineering records on connections and outfalls would help separate local improvements from downstream risks. Comparisons across towns or later network extensions could provide more independent interventions and show whether broader coverage changed the distribution of benefits. Such evidence would connect the municipal choices documented here to the racial mortality trajectories they may have influenced. Paarl establishes the unequal reach of an early sanitary reform and specifies the evidence needed to understand its place in South Africa's health transition.

References

Aidt, Toke S. 2022. "New perspectives on the contribution of sanitary investments to mortality decline." *The Economic History Review*, 75(2): 647–686.

Albertyn, Elizabeth. 1992. "Die argitektuur van die Paarl tussen die twee wêreldoorloë: 'n Kultuurhistoriese ondersoek." PhD diss. Stellenbosch University.

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): 586–638.

Anderson, D. Mark, Daniel I. Rees, and Tianyu Wang. 2020. "The phenomenon of "summer diarrhea" and its waning, 1910–1930." *Explorations in Economic History*, 76: 101320.

Anderson, D. Mark, Kerwin Kofi Charles, Daniel I. Rees, and Tianyi Wang. 2021. "Water purification efforts and the black–white infant mortality gap, 1906–1938." *Journal of Urban Economics*, 122: 103345.

Ashraf, Nava, Stephen G. Donald, and Tarek Lunt. 2016. "Infrastructure, incentives, and institutions." *American Economic Review: Papers & Proceedings*, 106(5): 77–82.

Beach, Brian, John Parman, and Martin Saavedra. 2022. "Segregation and the initial provision of water in the United States." *AEA Papers and Proceedings*, 112: 193–198.

Beach, Brian, Joseph P. Ferrie, Martin H. Saavedra, and Werner Troesken. 2016. "Typhoid Fever, Water Quality, and Human Capital Formation." *The Journal of Economic History*, 76(1): 41–75.

Beyers, M.S.M. 1933. "Die verbreiding van die menslike woonplekke in die distrikte van die Pêrel en Wellington." Master's diss. Stellenbosch University.

Bhalotra, Sonia, Damian Clarke, Jose Gomes, and Atheendar Venkataramani. 2021. "Urban water disinfection and mortality decline in lower-income countries." *American Economic Journal: Economic Policy*, 13(1): 400–434.

Bickford-Smith, Vivian. 1995. "South African urban history, racial segregation and the unique case of Cape Town?" *Journal of Southern African Studies*, 21(1): 63–78.

Bin-Kasim, Waseem-Ahmed. 2025. *Urban Saniscapes: Slums, Housing and Everyday Sanitation in Accra and Nairobi, 1908–1963.* East Lansing:Michigan State University Press.

Bogart, Dan. 2022. "Infrastructure and institutions: Lessons from history." *Regional Science and Urban Economics*, 94: 103740.

Broich, John. 2007. "Engineering the empire: British water supply systems and colonial societies, 1850–1900." *Journal of British Studies*, 46(2): 346–365.

Buirski, Peter. 1983. "Mortality rates in Cape Town, 1895-1980: A broad outline." In *Studies in the History of Cape Town.* Vol. 5, , ed. Christopher Saunders, 125–166. University of Cape Town.

Burman, Sandra, and Margaret Naude. 1991. "Bearing a bastard: the social consequences of illegitimacy in Cape Town, 1896–1939." *Journal of Southern African Studies*, 17(3): 373–413.

Clift, H. 2012. "Historical Overview: Drakenstein Municipal Area." Drakenstein Heritage Survey Unpublished report.

Costa, Dora L. 2015. "Health and the economy in the United States from 1750 to the present." *Journal of Economic Literature*, 53(3): 503–570.

Cutler, David M., and Grant Miller. 2005. "The role of public health improvements in health advances: The twentieth-century United States." *Demography*, 42(1): 1–22.

Fan, Mingjie, and Guojun He. 2023. "Clean water and infant health: Evidence from piped water provision in China." *Journal of the Association of Environmental and Resource Economists*, 10(6): 1129–1166.

Ferrie, Joseph P., and Werner Troesken. 2008. "Water and Chicago's mortality transition, 1850–1925." *Explorations in economic history*, 45(1): 1–16.

Fourie, Johan, and Jonathan Jayes. 2021. "Health Inequality and the 1918 Influenza in South Africa." *World Development*, 141: 105407.

Gallardo-Albarrán, Daniel. 2025. "The global sanitary revolution in historical perspective." *Journal of Economic Surveys*, 39(2): 567–598.

Geruso, Michael, and Dean Spears. 2018. "Sanitation and health externalities: Resolving the Muslim mortality paradox." *American Economic Journal: Applied Economics*, 10(2): 125–162.

Grant, Duncan. 1992. "The politics of water supply: The history of Cape Town's water supply, 1840-1920." Master's diss. University of Cape Town.

Kesztenbaum, Lionel, and Jean-Laurent Rosenthal. 2017. "Sewers' diffusion and the decline of mortality: The case of Paris, 1880–1914." *The Journal of Economic History*, 77(1): 172–211.

Oberholster, A.G., and P. Van Breda. 1987. *Paarl Valley, 1687-1987.* Pretoria:HSRC.

Peltola, Jari, and Sakari Saaritsa. 2019. "Portfolios of public health: Sanitary investments, water and infant mortality in early twentieth-century Finland." *The Economic History Review*, 72(2): 605–637.

Phillips, Howard. 1984. "'Black October': The impact of the Spanish Influenza Epidemic of 1918 on South Africa." PhD diss. University of Cape Town.

Reid, Alice, Eilidh Garrett, Maria Hiltunen Maltesdotter, and Angelique Janssens. 2024. "Historic cause of death coding and classification scheme for individual-level causes of death – Infant Categorisations." *Dataset, Apollo - University of Cambridge Repository*, DOI: 10.17863/CAM.109961.

Rip, M.R., D.E. Bourne, and D.L. Woods. 1988. "Characteristics of infant mortality in the RSA 1929-1983." *South African Medical Journal*, 73(1): 227–229.

Simkins, Charles, and Elizabeth Van Heyningen. 1989. "Fertility, Mortality, and Migration in the Cape Colony, 1891–1904." *International Journal of African Historical Studies*, 22(1): 79–111.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Swanson, Maynard W. 1977. "The Sanitation Syndrome: Bubonic Plague and Urban Native Policy in the Cape Colony, 1900–1909." *The Journal of African History*, 18(3): 387–410.

Troesken, Werner. 2004. *Water, Race, and Disease.* Cambridge, MA:MIT Press.

Troesken, Werner, Nicola Tynan, and Yuanxiaoyue Artemis Yang. 2021. "What are the health benefits of a constant water supply? Evidence from London, 1860–1910." *Explorations in economic history*, 81: 101402–.

Van Heyningen, Elizabeth. 1989. "Public health and society in Cape Town, 1880-1910." PhD diss. University of Cape Town.

White, Alexandre IR. 2018. "Global risks, divergent pandemics: contrasting responses to Bubonic Plague and smallpox in 1901 Cape Town." *Social Science History*, 42(1): 135–158.

Woods, Robert, Patricia Watterson, and John Woodward. 1988. "The causes of rapid infant mortality decline in England and Wales, 1861–1921." *Population Studies*, 42(3): 343–366.

Wright, Gwendolyn. 2001. "The Ambiguous Modernisms of African Cities." In *The Short Century: Independence and Liberation Movements in Africa, 1945–1994*, ed. Okwui Enwezor, 225–233. Munich:Prestel.

A Data construction and measurement

This appendix explains how we construct the main variables. We describe the death notices and their geographic links, reconstruct annual demographic series for Paarl, allocate conditional birth exposure across streets and derive the street-level geographic measures.

A.1 Death notices and sample construction

Official death notices provide our individual records. Introduced under the Births and Deaths Registration Act of 1894, these forms were completed after a death by a relative, acquaintance or public official. A typical infant notice gives name, age, sex, race, residence, final illness duration, stated cause of death, informant and, where applicable, the attending doctor or certifying official. Their combination of demographic information, residential detail and written cause descriptions is unusually rich for a historical African source.

The example in Figure A.1 shows both the detail available and the work needed to use it. Handwriting, spelling variation and uneven diagnostic precision require careful standardisation during transcription.

We transcribed notices for all ages and standardised ages, racial classifications, causes and residential references. After removing duplicates supported by the available evidence, the cleaned file contains 27,512 notices. We link names to the historical street plan at street rather than house level.

We remove exact repeated source rows and supported same-name matches within register groups. Ambiguous fuzzy matches, including possible twins, remain in the primary file and are flagged for a sensitivity check using the earlier collapse rule. The private adjudication ledger records source fields and reasons for each decision. Without confirmation from the original register, ambiguous matches remain unresolved.

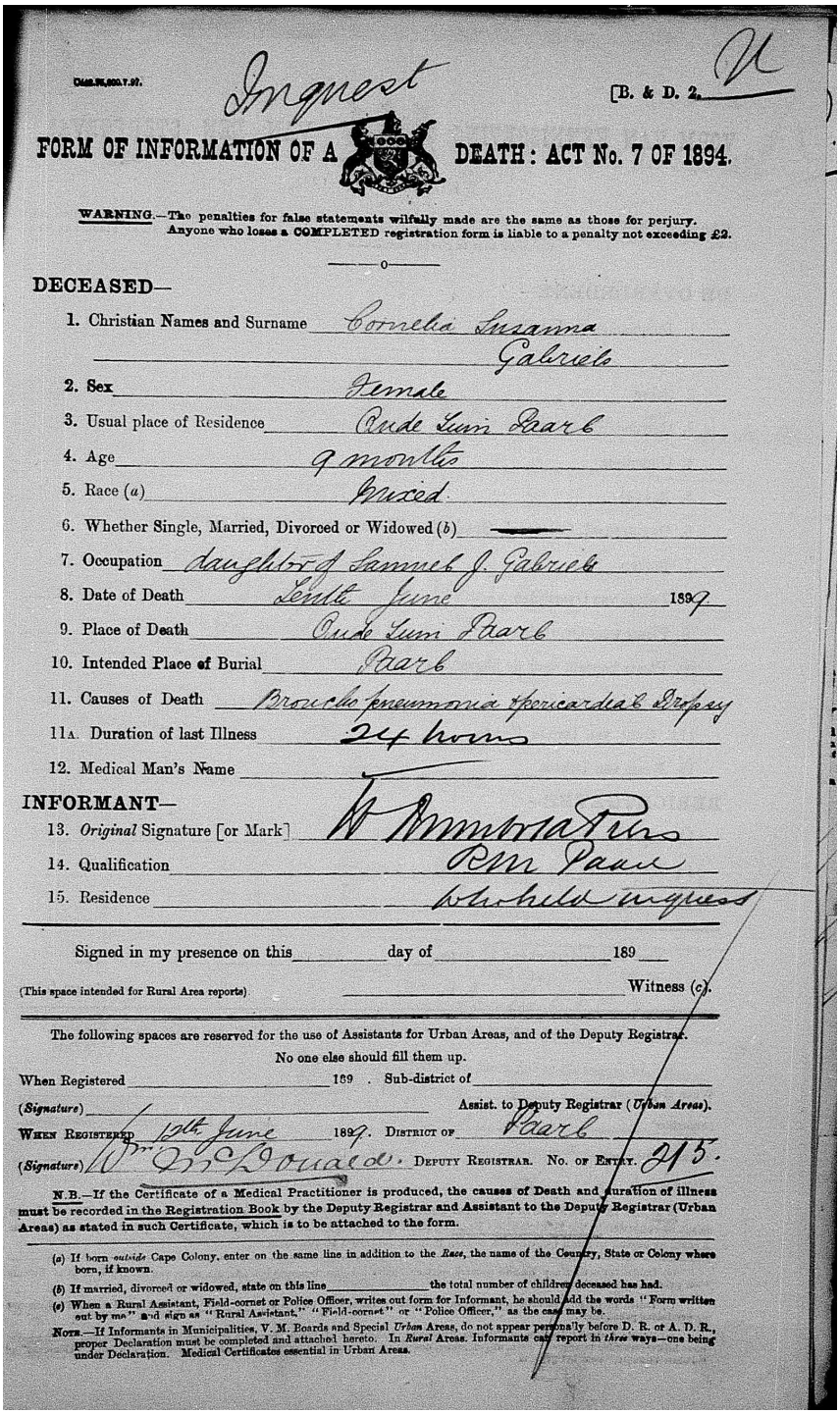
We retain age, stillbirth, residence and cause-classification flags for all non-duplicate notices, including those ineligible for the main analysis. We do not infer an age where it is ambiguous. Explicit stillbirths are excluded from live-born mortality and demographic reconstruction; uncertain day-zero cases remain, with a separate check excluding them. We infer birth year from age and death timing and flag incomplete dates. Table A.1 reconciles these steps with the final sample.

The sample flow distinguishes a known town address from an identifiable street within town. Farm and estate addresses help explain district-to-town inclusion, but cannot explain further losses among notices already identified as town residents. Appendix B.2 examines these stages separately.

District descriptive figures use the eligible full register; within-town regressions use recorded residential streets and complete birth cohorts. These are different geographic populations.

A.2 Construction of annual birth and population series

Infant mortality rates require births by race and year, but surviving records give births directly for only a few years. We reconstruct annual series for 1895–1936. To make the uncertainty clear, we



Ingnest

[B. & D. 2.

FORM OF INFORMATION OF A DEATH: ACT No. 7 OF 1894.

WARNING.—The penalties for false statements wilfully made are the same as those for perjury. Anyone who loses a COMPLETED registration form is liable to a penalty not exceeding £2.

DECEASED—

1. Christian Names and Surname: Cornelia Susanna Gabriels
2. Sex: Female
3. Usual place of Residence: Oude Zuin Paarl
4. Age: 9 months
5. Race (a): Mixed
6. Whether Single, Married, Divorced or Widowed (b): —
7. Occupation: daughter of Samuel J. Gabriels
8. Date of Death: Tenth June 1897
9. Place of Death: Oude Zuin Paarl
10. Intended Place of Burial: Paarl
11. Causes of Death: Broncho pneumonia Pericardial Dropsy
11A. Duration of last Illness: 24 hours
12. Medical Man's Name: —

INFORMANT—

13. *Original* Signature [or Mark]: W Immelman Rev
14. Qualification: Dutch Reformed Minister
15. Residence: Wisheld Ingress

Signed in my presence on this ____ day of ____ 189__

(This space intended for Rural Area reports) ____ Witness (c).

The following spaces are reserved for the use of Assistants for Urban Areas, and of the Deputy Registrar. No one else should fill them up.

When Registered ____ 189 . Sub-district of ____

(*Signature*) ____ Assist. to Deputy Registrar (*Urban Areas*).

When Registered 12th June 1897. District of Paarl

(*Signature*) W. Mc Donald. Deputy Registrar. No. of Entry. 215.

N.B.—If the Certificate of a Medical Practitioner is produced, the causes of Death and duration of illness must be recorded in the Registration Book by the Deputy Registrar and Assistant to the Deputy Registrar (Urban Areas) as stated in such Certificate, which is to be attached to the form.

(a) If born *outside* Cape Colony, enter on the same line in addition to the *Race*, the name of the Country, State or Colony where born, if known.

(b) If married, divorced or widowed, state on this line ____ the total number of children deceased has had.

(c) When a Rural Assistant, Field-cornet or Police Officer, writes out form for Informant, he should add the words "Form written out by me" and sign as "Rural Assistant," "Field-cornet" or "Police Officer," as the case may be.

Note.—If Informants in Municipalities, V. M. Boards and Special *Urban* Areas, do not appear personally before D. R. or A. D. R., proper Declaration must be completed and attached hereto. In *Rural* Areas, Informants can report in *three* ways—one being under Declaration. Medical Certificates essential in Urban Areas.

Figure A.1: Scan of a handwritten death notice

Notes: Notice of 9-month-old Cornelia Susanna Gabriels, who died from broncho-pneumonia and pericardial effusion secondary to her underlying condition.

Source: FamilySearch, "Paarl, Cape Province, South Africa records," image 221 of 605, citing National Archives of South Africa, Image Group Number 004532680.

Table A.1: Construction of the estimating sample

Restriction	Notices
Cleaned notices (all ages/races)	27,512
Study racial categories	26,790
Known plausible age; explicit stillbirth excluded	26,569
Live-born infants	8,678
Complete birth cohorts 1896-1930	8,226
Recorded Paarl town residence	4,377
Resolved residential street	1,158
Classified cause	1,135

Notes: Nested restrictions from the corrected cleaned file. Recorded residence is primary; the place-of-death fallback is a sensitivity. Cohorts end in 1930 to permit observation through age one. Explicit stillbirths and ambiguous ages are excluded from live-born infant analyses.

first distinguish observed quantities from those we must estimate.

A.2.1 What the sources observe

Let y_{trsg} denote the observed value of demographic variable g (for example, births or population aged 0–4, 5–14, and so forth) for race $r \in \{white, coloured\}$, sex $s \in \{male, female\}$, and year t.

Annual births are recorded for seven years: 1895–1898 and 1904–1906. Population counts by age, sex and race are available only in census years $t \in T_c = \{1904, 1911, 1921, 1926, 1931, 1936\}$. The sources overlap in just one year, 1904. Births for the remaining thirty-five years must therefore be estimated.

The available census tables impose a further limit. They supply only European age-specific counts for 1921, 1926 and 1931, leaving a long gap in the coloured series between 1911 and 1936. This limitation concerns the tables available to us, not the population enumerated by the entire 1921 census. The death register's last complete year is 1931, making later coloured birth estimates especially dependent on reconstruction assumptions.

A.2.2 Reverse survival

We reconstruct births by adding deaths before the census to survivors in each census age interval. Let I be that interval's implied birth-date window and a_{It} the fraction of calendar year t it spans. In a closed population with complete recording, the accounting restriction is:

$$\sum_t a_{It} B_{trs} = N_{Irs}(T) + D_{Irs}(T).$$

Survivors and prior deaths must refer to the same birth interval, which may include parts of several calendar years. Broad age groups therefore constrain sums of births, not annual observations

assigned to a midpoint. We allocate incomplete dates and imprecise ages uniformly within their intervals. Registered births are preserved exactly by sex; we fit the remaining census constraints and report the residuals. Matching these constraints does not independently test registration completeness.

A.2.3 From anchors to an annual series

Reverse survival supplies overlapping cohort constraints. To obtain an annual series we model births as a smooth rate multiplied by women aged 15–49:

$$B(t) \;=\; \mathrm{GFR}(t) \times W(t),$$

where $W(t)$ is interpolated between observed population counts. Separately by race and sex, we fit a natural spline to the log rate by minimising weighted squared log discrepancies in census cohort sums and registered-year rates. Registered sex-specific births replace the fitted values exactly. The annual path between observed years remains an assumption.

A.2.4 Uncertainty

We report specified reconstruction scenarios, not probability bands. The baseline extrapolates age-specific deaths beyond 1931 using averages from 1927–1931. One alternative omits those deaths; another shifts census dates by thirty days. These scenarios show sensitivity to particular assumptions rather than a confidence interval or a complete uncertainty distribution.

Child migration and incomplete birth or death registration can change the reconstructed series. Matching registered years by construction cannot validate the assumptions about either. The evidence provides no defensible numerical migration or registration adjustment, so we acknowledge this uncertainty without assigning an unsupported range. Regression intervals are conditional on the chosen reconstruction.

Figure A.2 plots the annual district series alongside registered births. Exact agreement in observed years follows from the anchoring rule; it does not validate the reconstructed years between them.

A.3 Allocation of conditional birth exposure to streets

The source data give district births. In what follows, B_{tr} is the conditional birth-equivalent total after the coverage adjustment, and $\mathcal{S}$ is the set of streets represented by eligible observed infant deaths.

For each allocation method, a non-negative spatial weight represents the assumed residential importance of each street. We first allocate exposure as:

$$\widehat{B}^{(m)}_{trs} = B_{tr} \frac{w^{(m)}_s}{\sum_{j \in \mathcal{S}} w^{(m)}_j}.$$

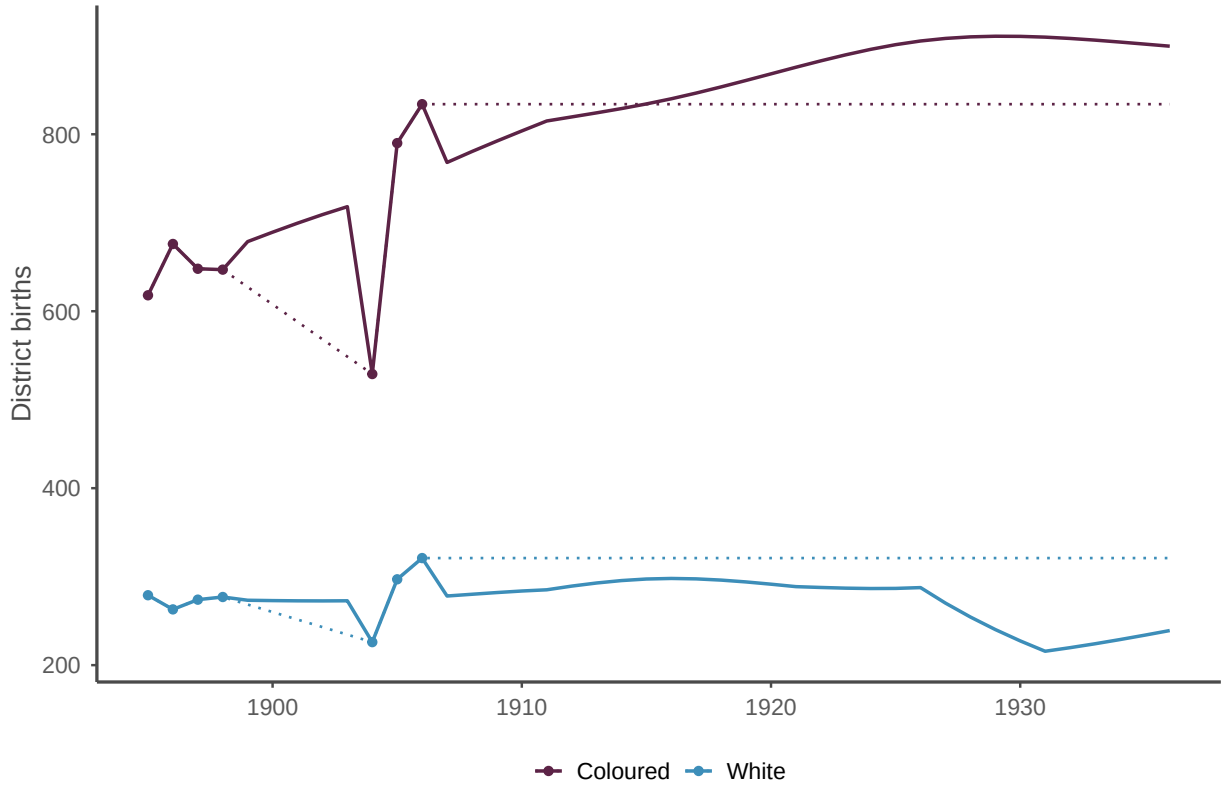


Figure A.2: Reconstructed annual births by race, Paarl, 1895–1936

Notes: The solid line is the reconstructed district birth series; the dotted line interpolates between registered-birth observations. Points mark registered births only. Census age-group sums constrain the fit but are not annual observed births. The named scenarios in Appendix C are sensitivity checks, not confidence bands.

Source: Authors' calculations based on data described in the text.

The normalised weights sum to one across the specified streets. The allocation preserves conditional exposure, not an independently observed town birth total.

We use three alternative weighting schemes to proxy residential intensity along the street network, each based on a different feature of the historical road geometry:

1. *Street-length weights.* We set $w_s^{(1)}$ to the length in metres of street s's primary polyline. Where several road features share a street name, we retain the longest. The method assumes that residential density is proportional to total street length.

2. *Segment-based weights.* Each street polyline consists of straight segments joining consecutive vertices. We set $w_s^{(2)}$ to the number of vertices minus one, with a minimum of one. This geometric complexity proxies street frontage and subdivision intensity.

3. *Raster-based density weights.* We overlay a grid of 50×50 metre cells on the historical network. Cell density is the number of road features crossing a cell. We obtain $w_s^{(3)}$ by summing these densities across all cells intersected by street s. This permits residential intensity to vary along a street rather than treating it as uniform.

We report all three allocations as alternative measurement strategies, rather than competing structural models.

A.3.1 Racial composition of streets

Road geometry alone would allocate white and coloured births in identical proportions across streets. In a residentially segregated town, this would assign a geometric share of white births even to a street that was effectively all-coloured, creating a denominator beneath a nearly empty death count. The resulting error need not cancel in the racial comparisons central to the paper.

We estimate street racial composition from non-infant deaths with recorded residence during 1896–1931, adjusting for race-specific district non-infant mortality rates. Both deaths and population person-years used in those rates cover the district and the same period. Adjusted town counts provide a prior of strength eight, also assigned to streets with no supporting non-infant deaths. Sensitivity checks use strengths four and sixteen and a composition window ending before 1905.

Writing π_s^r for the estimated share of race r on street s, the allocation becomes

$$\widehat{B}_{trs}^{(m)} \;=\; B_{tr}\,\frac{w_s^{(m)}\pi_s^r}{\sum_{j\in\mathcal{S}} w_j^{(m)}\pi_j^r},$$

so that the spatial allocation incorporates estimated racial shares. Shrinkage makes the shares positive; it does not establish how many births occurred on each street. Agreement with the composition of infant deaths would not independently validate an allocation itself derived from mortality.

A.3.2 Coverage and level

District birth totals and recorded town-street deaths concern different populations. We define p_{tr} as the smoothed share of district infant deaths recorded at town streets, irrespective of cause classification, and use $p_{tr}B_{tr}^{district}$ as the conditional street exposure total.

This adjustment assumes that recorded town streets' share of district deaths proxies their share of district births. Town–rural mortality differences and unequal recording can violate that assumption. Street rate levels are therefore conditional, while the district descriptive rates use numerators and denominators for the same district population.

We impose no further race-level calibration. We check conservation once per street–cohort–race before repeating exposure across causes, and report exclusions separately from all-zero groups removed by the estimator. The allocation gives finite positive exposure to supported streets, including five that previously lacked denominators because no non-infant deaths supported their composition estimates. A further check excludes those five from both comparison models without reallocating their exposure.

On a common sample, the denominator-free model absorbs exposure by construction. This reduces one source of measurement dependence, but neither verifies the coverage assumption nor recovers deaths at unknown treatment locations.

A.4 Construction of street-level geographic indices

We derive geographic measures from each street's physical features and location to examine possible spatial redistribution of disease exposure.

We harmonise street names and retain the longest polyline where names repeat. Horizontal distances use WGS84/UTM zone 34S (EPSG:32734), checked against geodesic distances in twenty comparisons. The primary grid, decay and river scales are 50, 400 and 600 ground metres. A sensitivity check converts the earlier Mercator-coordinate settings to their local ground equivalents.

Street slope measures elevation change along a street. We sample a fixed number of evenly spaced points along its polyline and obtain their elevations from a digital elevation model. We then divide the sampled elevation range by total street length:

$$\text{Slope}_s = \frac{\max(\text{elev}_s) - \min(\text{elev}_s)}{\text{length}_s}.$$

This measure approximates elevation range per metre of street length. We standardise using the mean and standard deviation across unique streets in the primary observed panel, fixed across samples:

$$\text{Slope}_s^z = \frac{\text{Slope}_s - \overline{\text{Slope}}}{\text{sd}(\text{Slope})}.$$

Mean elevation is the average across sampled street points, measured in metres above sea level. It describes the street's vertical position in the town and proxies baseline drainage conditions separately from slope.

Distances from fifteen sampled street points to the nearest river segment are aggregated as mean and minimum distance. Descriptive geographic controls use mean distance; the illness-duration comparison uses minimum sampled distance. Neither quantity measures an infant's exact household location.

We classify orientation from the polyline endpoints. A street is north–south if its net displacement in that direction exceeds its east–west displacement; otherwise it is east–west. This indicator captures broad alignment with the town's slope and water flows rather than detailed curvature.

These measures describe different aspects of the geography through which drainage might have changed waterborne disease exposure. They are consistently constructed proxies for spatial differences, not precise measurements of hydrological processes.

B Supplemental figures and robustness tables

This appendix reports geographic descriptions, sample selection, robustness checks among cells containing deaths and within-street comparisons. Appendix C gives further measurement sensitivities and the all-cause auxiliary model.

B.1 Descriptive evidence on geographic redistribution

Figure B.1 plots targeted infant death counts against standardised street slope by race and period. Counts are not divided by births, so differences can reflect residence and recording as well as mortality risk. The figure is descriptive.

The pattern is consistent with drains moving contaminated water downhill, reducing exposure above while concentrating it below. It is also consistent with other explanations. Explicitly measuring run-off exposure does not distinguish that mechanism from the alternatives.

Run-off exposure and proximity are nearly collinear, so the geographic comparisons cannot separate water routing from nearness to the works. Relative elevation and river proximity add associations without identifying the hydrological mechanism.

We therefore do not advance a hydrological mechanism or treat this descriptive figure as evidence of a benefit confined to one racial group. The explicit spatial contrasts are reported in Section 4.5.

B.2 Selection into the street-identified sample

We next examine changes in the usable street sample. Figure B.2 reports the share reaching the analysis sample by race and cohort interval. Visible continuity or a non-significant interaction cannot establish stable recording by treatment status.

Table B.1 separates district-to-town inclusion, street recording within town and subsequent eligibility. Its universe is 8,226 white and coloured infant deaths in birth cohorts 1896–1930; 4,377 record town residence, 1,158 a resolved residential street, and 1,135 a classified cause. Unknown street location is never assigned an observed treatment status.

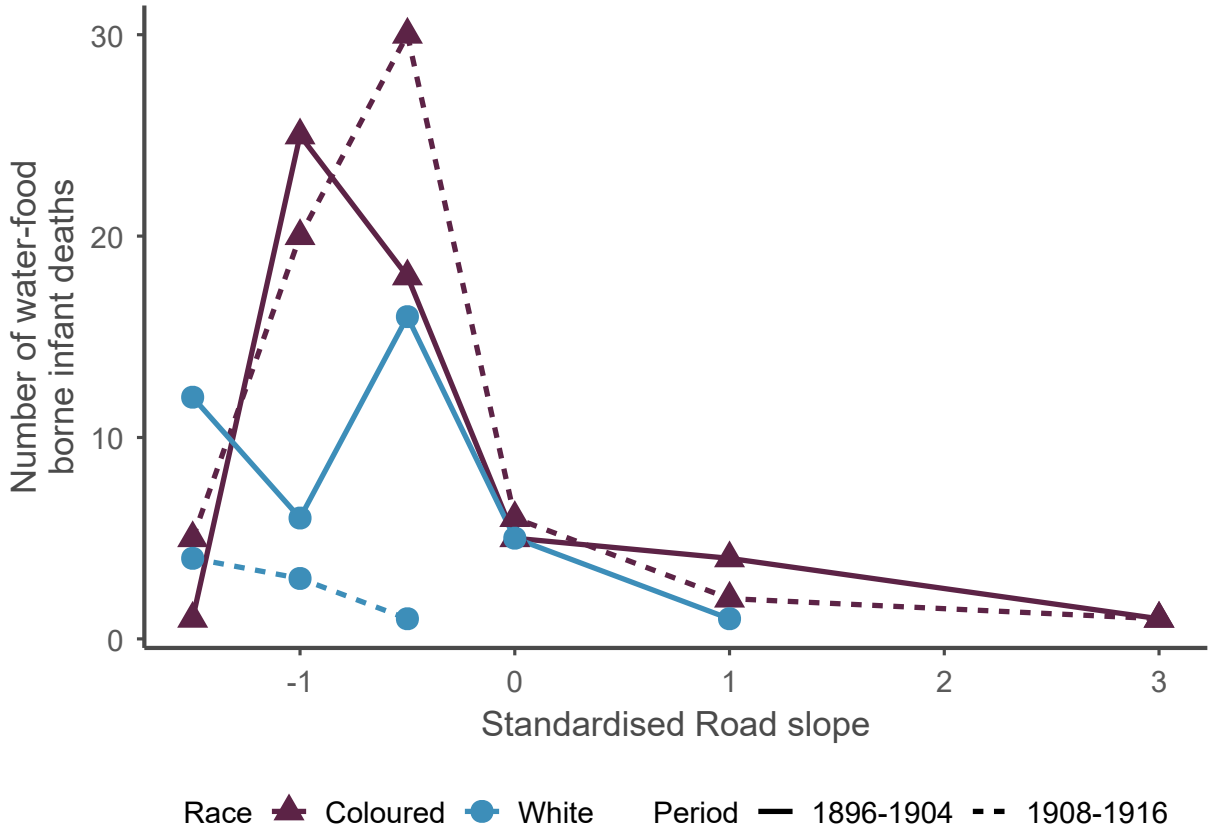


Figure B.1: Water- and food-borne infant deaths by road slope, race, and period.

Notes: The figure plots the number of infant deaths from water- and food-borne causes against standardised street slope, separately for white and coloured infants and for the pre-intervention (1896–1904) and post-intervention (1908–1916) periods.

Source: Authors' calculations based on data described in the text.

The stages of sample selection behave differently and must be examined separately. The post-period indicator equals one for infants born in 1905 or later, when the drainage rollout began.

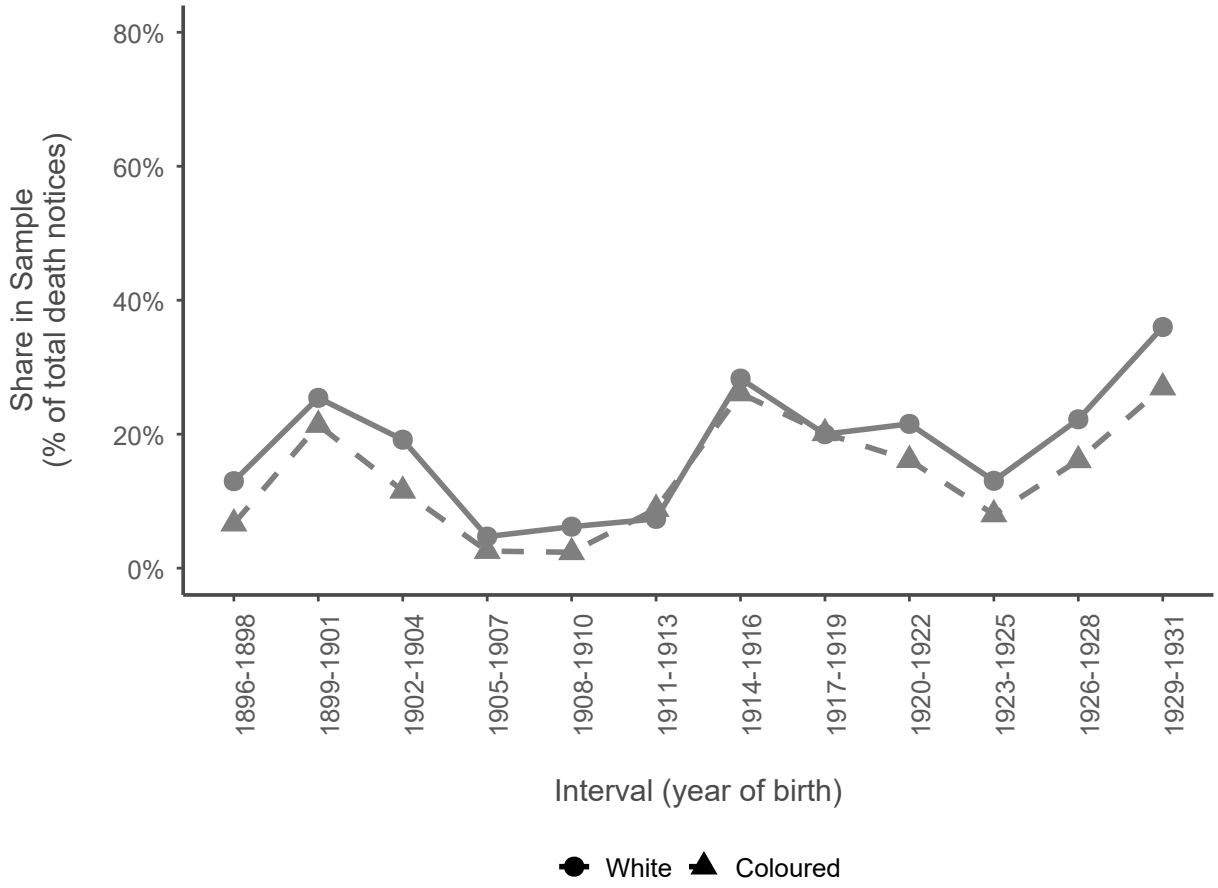


Figure B.2: Within-race share of infant death notices reaching the analysis sample, by birth cohort.

Notes: Plot based on all death notices in Paarl for infants under one year.

Source: Authors' calculations based on data described in the text.

Column (1) describes whether a notice records Paarl town rather than another location. This is distinct from street recording within town, and a district rural share cannot explain the latter margin.

Column (2) conditions on town residence. The street sample is selective, and the coefficient describes recording among observed deaths. Heteroskedasticity-robust standard errors are used because no independent register-level dependence unit has been established for notices without streets.

Columns (3) and (5) allow the post-period association with street recording and final eligibility to differ by race. Their uncertainty concerns the observed recording margins. Neither a small point estimate nor a non-significant interaction establishes that recording by treatment geography was

stable.

Column (4) allows post-period recording changes to vary by cause. Reduced recording of targeted deaths can manufacture an apparent mortality decline. Its bias for the drainage comparison depends on differential recording by treatment geography, which unlocated notices cannot reveal. The recording-factor scenarios therefore state assumptions explicitly and report the factor needed to move a cause-relative contrast to zero.

A final margin is cause classifiability among located town notices. Its regression conditions on known street and cohort; non-significance cannot establish that classification was stable in the unlocated population.

Selection continues to limit the mortality comparisons. We report the observable stages of selection and sensitivity to recording assumptions without claiming that weighting identifies missing outcomes.

Table B.1: Selection into the street-identified sample

	(1) In Paarl town	(2) Has street	(3) Has street	(4) Has street	(5) In analysis
Constant	0.543***	0.258***	0.315***	0.203***	0.156***
	(0.028)	(0.034)	(0.040)	(0.039)	(0.024)
Post			−0.100***	0.019	−0.051**
			(0.037)	(0.036)	(0.023)
Coloured	0.002	−0.046**	−0.114***		−0.046**
	(0.017)	(0.020)	(0.031)		(0.020)
Log(Age + 1)	−0.010**	0.005	0.005	0.005	0.001
	(0.004)	(0.005)	(0.005)	(0.005)	(0.003)
Post × Coloured			0.116***		0.051**
			(0.040)		(0.025)
Common infectious diseases	0.061***	0.019	0.019	−0.007	0.025*
	(0.022)	(0.026)	(0.026)	(0.047)	(0.015)
Convulsions	−0.124***	−0.030	−0.031	−0.041	−0.039***
	(0.020)	(0.027)	(0.027)	(0.052)	(0.012)
Ill-defined	0.039	0.032	0.033	0.186**	0.025
	(0.040)	(0.049)	(0.049)	(0.087)	(0.028)
Other	0.046**	0.028	0.029	0.025	0.027**
	(0.019)	(0.023)	(0.023)	(0.043)	(0.013)
Perinatal and weakness	0.092***	0.059**	0.058**	0.129***	0.060***
	(0.022)	(0.026)	(0.026)	(0.046)	(0.016)
Water- and food-borne	0.096***	0.065***	0.065***	0.118***	0.062***
	(0.017)	(0.021)	(0.021)	(0.041)	(0.012)
Post × Common infectious diseases				0.046	
				(0.057)	
Post × Convulsions				0.013	
				(0.059)	
Post × Ill-defined				−0.221**	
				(0.105)	
Post × Other				0.011	
				(0.051)	

Continued on next page

Table B.1: Selection into the street-identified sample (Continued)

	(1) In Paarl town	(2) Has street	(3) Has street	(4) Has street	(5) In analysis
Post × Perinatal and weakness				−0.094*	
				(0.052)	
Post × Water- and food-borne				−0.065	
				(0.047)	
Num.Obs.	8226	4377	4377	4377	8226
R^2 Adj.	0.020	0.007	0.008	0.007	0.016

Notes: *p<0.1, **p<0.05, ***p<0.01. OLS selection estimates. The dependent variable is an indicator for surviving the margin named in the column header. The sample is every European or coloured infant death in the district between 1896 and 1930, not the modelling frame. Column (1) asks whether the notice records a Paarl-town address at all rather than a farm or another town. Columns (2)–(4) condition on Paarl town and therefore isolate recording practice. Column (5) combines all three margins, including whether the cause of death could be classified. Post indicates birth years from 1905 onward. Broncho-pneumonia is the omitted cause. Heteroskedasticity-robust standard errors are in brackets; dependence between notices may remain.
Source: Authors' calculations based on data described in the text.

B.3 Additional checks on infant mortality patterns

Table B.2 reports three checks based on Equation (1) in the main text. The first two address sensitivity to extreme cells; the third is a placebo on the cause of death.

Small denominators make a single death produce a large log rate, so sensitivity to extreme cells matters. But selecting cells by the outcome would introduce another problem: discarding the highest measured rates would mechanically yield a larger triple interaction and smaller standard error. That apparent reassurance would reflect the selection rule rather than the estimate's reliability. The two checks below therefore leave the outcome ranking untouched.

Columns (1.1)–(1.3) winsorise the log outcome at the first and ninety-ninth percentiles. Columns (2.1)–(2.3) exclude the bottom decile of estimated exposure. Both still condition on cells containing deaths. We report their estimates as sensitivity checks, not as validation of the full mortality design.

Columns (3.1)–(3.3) redefine targeted cause as broncho-pneumonia. These placebo comparisons test a different cause contrast; neither a significant nor a non-significant coefficient by itself validates the drainage interpretation.

Cause coding, general mortality changes and selective recording may affect these outcomes together. The placebo is therefore read alongside the sample analysis and joint cause-specific contrasts, rather than as evidence that all alternative explanations have been ruled out.

Table B.2: Additional checks on infant mortality patterns

	Winsorised			Small cells dropped			Placebo Targeted		
Street-level births measure	Str length	Str segments	Raster	Str length	Str segments	Raster	Str length	Str segments	Raster
	(1.1)	(1.2)	(1.3)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)
Treatment	0.394***	0.183	0.363***	0.329***	0.231*	0.324***	0.247**	0.162	0.239**
	(0.130)	(0.126)	(0.130)	(0.101)	(0.120)	(0.097)	(0.110)	(0.103)	(0.110)
Targeted	0.360***	0.283***	0.356***	0.348***	0.259***	0.359***	−0.153	−0.147	−0.133
	(0.054)	(0.077)	(0.054)	(0.066)	(0.097)	(0.059)	(0.167)	(0.219)	(0.174)
Coloured	0.484	0.263	0.444	0.688**	0.417	0.692**	0.328	0.164	0.293
	(0.356)	(0.412)	(0.357)	(0.278)	(0.413)	(0.274)	(0.335)	(0.379)	(0.335)
Treatment × Targeted	−0.220***	−0.157**	−0.213***	−0.209***	−0.144*	−0.217***	−0.152	−0.220	−0.176
	(0.047)	(0.065)	(0.048)	(0.057)	(0.082)	(0.052)	(0.200)	(0.261)	(0.211)
Treatment × Coloured	0.058	0.399	0.096	−0.059	0.267	−0.078	0.254	0.454	0.267
	(0.338)	(0.389)	(0.337)	(0.254)	(0.409)	(0.255)	(0.307)	(0.353)	(0.306)
Targeted × Coloured	−0.358***	−0.260***	−0.351***	−0.345***	−0.238**	−0.352***	0.131	0.122	0.114
	(0.066)	(0.088)	(0.066)	(0.073)	(0.103)	(0.066)	(0.161)	(0.217)	(0.167)
Treatment × Targeted × Coloured	0.238***	0.151*	0.226***	0.220***	0.148	0.224***	0.051	0.130	0.073
	(0.070)	(0.088)	(0.070)	(0.072)	(0.095)	(0.067)	(0.202)	(0.281)	(0.214)
Num.Obs.	935	935	935	845	841	844	935	935	935
R^2 Adj.	0.853	0.856	0.850	0.857	0.853	0.856	0.851	0.854	0.847
FE: Street residence	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE: Birth year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Street clusters	52	52	52	47	49	49	52	52	52

Notes: *p<0.1, **p<0.05, ***p<0.01. Columns (1.1)–(1.3) winsorise the outcome at the 1st and 99th percentiles, retaining every observation. Columns (2.1)–(2.3) instead drop street–year–race–cause cells in the bottom decile of estimated births, trimming on the denominator rather than on the outcome. Columns (3.1)–(3.3) redefine the Targeted category as broncho-pneumonia, a placebo cause that surface drainage should not directly affect. The outcome is log street-level infant mortality per 1,000 estimated births, and the birth-allocation scheme varies by column as indicated in the header. All columns include street and birth-year fixed effects. Street-clustered standard errors are in brackets; observations are weighted by street-level births.
Source: Authors' calculations based on data described in the text.

B.4 Racial disparities within treated streets

A final comparison examines medical attendance and residential exposure among recorded deaths on the same streets. These may be associated with unequal health conditions, but they cannot explain a racial mortality response that the main estimates do not establish.

Every column clusters by recorded street. Residence–death matching requires both streets to have been recorded independently, and its smaller sample is reported explicitly. A mother's recorded occupation is a proxy for employment, not a validated labour-force measure.

We also exclude doctor signature as a control. Medical attendance could respond to treatment, household resources and illness severity, making it a possible post-treatment variable. In the illness-duration models, it could absorb much of the variation those comparisons are intended to describe.

Table B.3: Mechanism checks on medical access and residential exposure by race

Dependent Variable	Doctor Signature		SoR = SoD				Log(Dur. Illness + 1)	
	(1.1)	(1.2)	(2.1)	(2.2)	(2.3)	(2.4)	(3.1)	(3.2)
Constant	0.946***							
	(0.022)							
Treatment	0.008	0.010					−1.609	−1.610
	(0.024)	(0.042)					(1.163)	(1.157)
Coloured	−0.075***	−0.067**	−0.006	−0.001	−0.001	−0.001		−0.120
	(0.025)	(0.028)	(0.005)	(0.003)	(0.003)	(0.003)		(0.152)
Present			0.001		0.008	0.008		
			(0.012)		(0.017)	(0.017)		
Working				−0.002	0.001	0.003		
				(0.004)	(0.007)	(0.007)		
SoR = SoD							−0.442	−0.440
							(1.141)	(1.142)
Treatment × Coloured	0.009	0.035						0.059
	(0.031)	(0.027)						(0.165)
Present × Coloured			0.005		−0.003	−0.002		
			(0.016)		(0.014)	(0.014)		
Working × Coloured				−0.016**	−0.033***	−0.036***		
				(0.007)	(0.012)	(0.013)		
Present × Working					−0.011	−0.010		
					(0.019)	(0.020)		
Present × Working × Coloured					0.033**	0.033**		
					(0.015)	(0.015)		
Treatment × SoR = SoD							1.172	1.150
							(1.166)	(1.171)
Coloured × SoR = SoD								–
								(–)
Treatment × Coloured × SoR = SoD								–
								(–)

Continued on next page

Table B.3: Mechanism checks on medical access and residential exposure by race (Continued)

Dependent Variable	Doctor Signature		SoR = SoD				Log(Dur. Illness + 1)	
	(1.1)	(1.2)	(2.1)	(2.2)	(2.3)	(2.4)	(3.1)	(3.2)
Log(Age + 1)						−0.002	0.397***	0.397***
						(0.001)	(0.018)	(0.018)
Num.Obs.	1135	1121	813	813	813	813	759	759
R^2 Adj.	0.005	0.059	−0.043	−0.035	−0.031	−0.031	0.339	0.338
Street clusters	65	52	50	50	50	50	47	47

Notes: *p<0.1, **p<0.05, ***p<0.01. OLS mechanism checks. Columns (1.1)–(1.2) test whether treatment affected doctor-signed notices by race. Columns (2.1)–(2.4) examine whether place of death matches place of residence by maternal presence, maternal employment, and race. Columns (3.1)–(3.2) test whether treatment effects on illness duration vary with place-of-death matching and race. Columns (1.2), (2.1)–(2.4), and (3.1)–(3.2) include street-of-residence and birth-year fixed effects; columns (2.4), (3.1), and (3.2) also control for log age at death. Standard errors are clustered by street of residence in every column and reported in brackets. Doctor signature is not included as a control, being post-treatment.
Source: Authors' calculations based on data described in the text.

Columns (1.1) and (1.2) describe racial differences in recorded medical attendance, with and without street and cohort fixed effects. Attendance is a proxy for access to formal care among deaths; it does not measure all medical care received by surviving infants.

The treatment and racial interaction estimates describe whether recorded attendance changed around adoption. An imprecise interaction does not establish that drainage left a gap unchanged, nor does it identify the role of medical care in survival.

Columns (2.1)–(2.4) examine recorded residential exposure. The indicator $SoR = SoD$ equals one only if independently recorded residential and death streets match. *Present* identifies informant descriptions referring to a mother or grandmother; it does not directly observe maternal attendance at an inquest. *Working* indicates a recorded maternal occupation. The interactions describe associations among recorded deaths, not infants' daily movements or exposure away from home.

The maternal and residential interactions are reported with street-clustered uncertainty. Missing independently recorded locations reduce the usable sample; these coefficients cannot establish daily mobility or explain a differential sanitation benefit.

Columns (3.1) and (3.2) examine duration of final illness. They condition on death, usable duration and independently observed place information. Doctor attendance remains excluded as a control because it may be post-treatment.

Table B.3 documents medical-attendance and residential associations among recorded deaths. It does not establish a mobility mechanism or explain an identified racial divergence in the drainage response.

C Joint contrasts and measurement checks

The all-cause auxiliary model counts each infant once, including unclassified causes on the primary observed streets. It uses adoption indicators by race, log exposure as an offset, and street–race and cohort–race fixed effects. The static cause model and the more flexible joint model use different fixed effects; their disagreement is a substantive specification limitation. None of these comparisons provides experimental assignment.

We assess differential recording through $RR_{observed} = RR_{true}q$, with $q = 0.5, 0.75, 1, 1.25, 1.5, 2$. Notices with unknown streets receive no treatment assignment. For each stated cause-relative contrast, the factor that would move the estimate to the null is $q = \exp(\widehat{\beta})$. The replication files report the numerical scenarios, estimates omitting one treated street at a time, zero-group removals, event-bin support and spatial-scale checks. Bins with no exposed deaths have boundary coefficients and no finite reported interval. Because the network does not date outfall activation, the spillover timing check excludes 1905–06.

Table C.4: All-cause infant mortality contrasts

Contrast	Estimate	Std. error	95% interval
White: all causes	-0.011	0.292	[-0.583, 0.561]
Coloured: all causes	-0.164	0.262	[-0.677, 0.350]
Coloured minus White: all causes	-0.153	0.441	[-1.018, 0.712]

Notes: PPML; full street-cluster covariance and approximate pointwise intervals. Four treated streets. Unknown-cause deaths are included; this is a separate estimand from the cause-relative models. Estimation observations: 3210; street clusters: 65.

Table C.5: Joint post-adoption contrasts by race and cause

Contrast	Estimate	Std. error	95% interval
White: other causes	-0.245	0.387	[-1.004, 0.513]
White: targeted	0.431	0.454	[-0.460, 1.321]
Coloured: other causes	-0.326	0.270	[-0.855, 0.202]
Coloured: targeted	0.523	0.339	[-0.141, 1.188]
White: targeted minus other	0.676	0.592	[-0.484, 1.836]
Coloured: targeted minus other	0.850	0.302	[0.258, 1.441]
Coloured minus White: cause-relative	0.174	0.706	[-1.209, 1.556]

Notes: Pooled post-adoption PPML with separate street and year effects for each race–cause group, and full street-cluster covariance. The post coefficient pools years after own adoption; it does not average event-bin coefficients. Estimation observations: 18471; street clusters: 65.

Table C.6: Within-race totals in the static offset model

Contrast	Estimate	Std. error	95% interval
White: targeted total	-0.479	0.191	[-0.854, -0.104]
Coloured: targeted total	-0.386	0.240	[-0.857, 0.084]
Coloured minus White: targeted total	0.092	0.214	[-0.327, 0.512]
White: cause-relative	0.007	0.186	[-0.358, 0.373]
Coloured: cause-relative	-0.152	0.161	[-0.468, 0.163]
Coloured minus White: cause-relative	-0.160	0.212	[-0.575, 0.256]

Notes: Linear combinations of the static offset PPML. A targeted total includes treatment and its cause/race interactions; a cause-relative contrast subtracts the other-cause response. Full street-cluster covariance. Estimation observations: 30940; street clusters: 65.

Table C.7: Named one-at-a-time sensitivities

Variant	White cause-relative	Coloured minus White
primary	0.007 (0.187)	-0.160 (0.212)
fallback address	-0.320 (0.199)	0.266 (0.199)
legacy duplicates	0.007 (0.187)	-0.160 (0.212)
drop uncertain zero	0.007 (0.187)	-0.160 (0.212)
strict365	0.007 (0.187)	-0.160 (0.212)
drop boundary	0.007 (0.187)	-0.160 (0.212)
drop imputed date	0.007 (0.187)	-0.160 (0.212)
include 1931	0.044 (0.186)	-0.201 (0.208)
main last phase	0.058 (0.182)	-0.210 (0.207)
prior4	0.007 (0.187)	-0.160 (0.212)
prior16	0.007 (0.187)	-0.160 (0.212)
preperiod composition	0.007 (0.187)	-0.160 (0.212)
late zero	0.007 (0.187)	-0.160 (0.212)
census plus30	0.007 (0.187)	-0.160 (0.212)
exclude five streets	0.007 (0.187)	-0.168 (0.212)

Notes: Denominator-free PPML, with street–cohort–race and cause effects. Entries are log coefficients (street-clustered standard errors). Duplicate/status and date-boundary alternatives that retain the same primary deaths can return the same estimate. Reconstruction and fixed allocation factors are absorbed here; their exposure-offset counterparts and sample changes are supplied in the replication files. These checks are separate alternatives, not an exhaustive combination.

Table C.8: Complete-cohort age windows and Main Street timing

Age limit / phase	Cohorts	Deaths	White cause-relative	Coloured minus White
Under 1 / first	1896-1930	1135	0.007 (0.187)	-0.160 (0.212)
Under 2 / first	1896-1929	1498	-0.033 (0.187)	-0.404 (0.308)
Under 5 / first	1896-1926	1505	0.055 (0.183)	-0.440 (0.327)
Under 1 / last	1896-1930	1135	0.058 (0.182)	-0.210 (0.207)
Under 2 / last	1896-1929	1498	0.010 (0.194)	-0.447 (0.320)
Under 5 / last	1896-1926	1505	0.099 (0.193)	-0.483 (0.340)

Notes: Denominator-free PPML. Birth cohorts end in 1931 minus the age limit; numbers in parentheses are approximate street-clustered standard errors. First/last phase changes Main Street adoption from 1905 to 1906.

Table C.9: Differential-recording tipping factors

Joint cause-relative contrast	Observed ratio	Null tipping q
White: targeted minus other	1.966	1.966
Coloured: targeted minus other	2.339	2.339

Notes: Observed targeted/other post-change ratio relative to controls equals the corresponding true ratio multiplied by q. The displayed factor would move the point contrast to zero under that assumption. It is not an estimate of recording completeness, and no unknown-street death is assigned treatment.

Table C.10: District birth totals under reconstruction scenarios

Scenario	White	Coloured
baseline	9800	28146
census plus30	9812	28134
late zero	9779	27251

Notes: Totals cover cohorts 1896–1930. Late zero removes extrapolated post-1931 deaths; census plus30 shifts census dates by thirty days. Registered sex-specific counts remain exact. Scenarios are not confidence intervals. Migration and registration uncertainty lack an empirically supported numerical range.