The welfare effects of industrial pollution in the 19th century Netherlands

Auke Rijpma and Robin Philips

September 2, 2023

Utrecht University and UCLouvain

Introduction

Introduction

- Impact of pollution commonly thought to be high, e.g on health, mortality, and even cognitive ability and productivity (Currie et al. 2014; Graff Zivin and Neidell 2013)
- In a context of growing industry, using polluting technology, concentrated in cities, and little regulation, we would expect air pollution to have similar outcomes in the past
- Literature has found negative impacts on e.g. mortality and height (Hanlon 2020, 2018; Beach and Hanlon 2018; Franke 2022),
- Could be important part of our understanding of changing living standards during industrialisation.
- Little work at individual level (cf. Bailey, Hatton, and Inwood 2018).

Today

- A historical micro-data approaches to tackle this subject
- Case of Leiden, 1890.
- Combining fine grained housing and factory locations, with civil registry deaths to explore mortality impact.
- Find that people living close to (heavier) factories in 1890 are predicted to live shorter lives.
- Work in progress!

Context

- Leiden not a key industrialising centre in the Netherlands, but by c. 1900 substantial industry had built up, especially in metal, printing, food processing, and textiles.
- General health issues: urban west of the Netherlands had high mortality compared with rest until early twentieth century; Leiden was no exception.
- Regulation of factories and pollution is the so-called “hinderwet” (1810/1824/1874):
    - In practice a repressive rather than preventive law.
    - Thought to have been a relatively weak law.

Context

Figure 1: Leiden gasworks on the Langegracht, 1923

Context

Figure 2: Leiden skyline, 1940, with the flour mill and gas works (again)

Data

Demographic data

- Starting point are Leiden's population registers
- Registers of all inhabitants in a municipality and their mutations, esp. moving house.
- Constructed data from this for Leiden, 1890–1923, though we will focus on 1890 when new registers were started on basis of the 1889 census.
- Caveat: digitised version exceptionally rich, but includes only streets A-H!

Demographic data

Links and enrichment needed.

1. coordinates – openstreetmap addresses (Haklay and Weber 2008), in turn based on Dutch BAG.
2. sex – bride/groom first names from marriage certificates.
3. life tables – mortality.org.
4. (occupations) – occupation lists from HSN (Mandemakers, Mourits, and Muurling 2019).
5. death records – civil registry deaths, covering 90% of Netherlands.

1-4 are easy, done using 90% similarity of flattened strings; or even hard links.

Data linkage

- Linking population registers to death certificates more challenging, because:
    1. spelling varied more;
    2. names are not unique;
    3. there are a lots of death certificates (11m).
- Linking done using a supervised machine learning approach (Rijpma, Cilliers, and Fourie 2020),

Data linkage

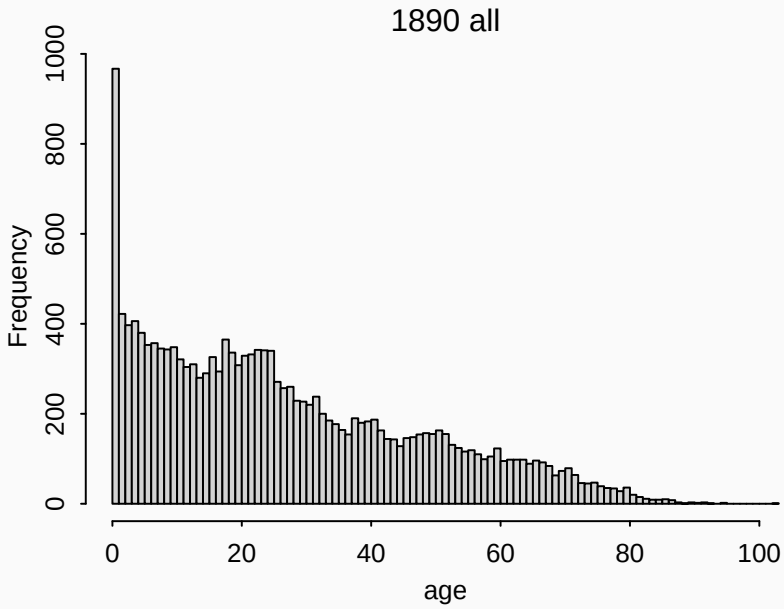
- Label 282 population register entries, blocked on year of birth (and a 1-year buffer) and first letter of the surname, and manually linked to 190 death certificates.
- Manual link found for 2/3 of the data. Imperfect because
  - Some provinces do not provide age at death, so we can't get year of birth.
  - Deaths after 1970 are not in the data.
  - Death reporting has inaccuracies.
- Gradient boosting classifier trained (95 % precision; 92 % recall) and used to predict links in rest of data (Chen and Guestrin 2016).

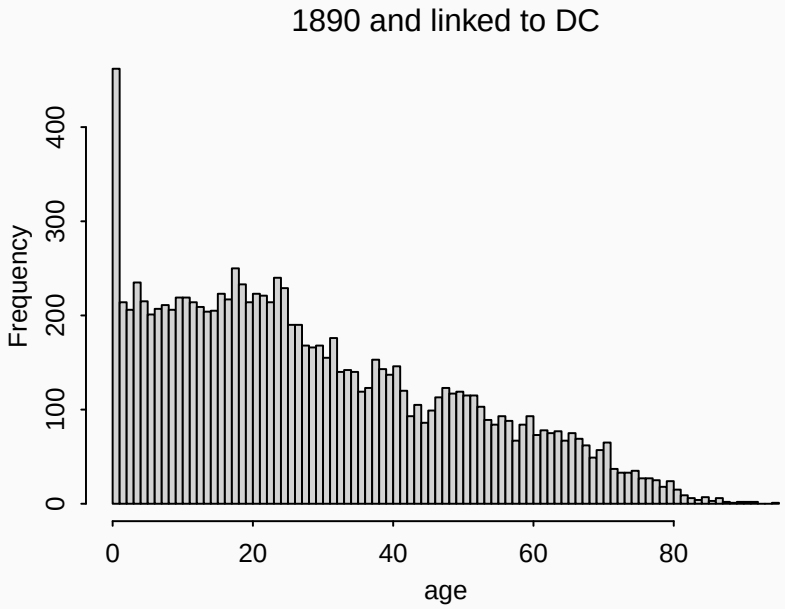
	FALSE	TRUE
FALSE	338642	2
TRUE	3	35

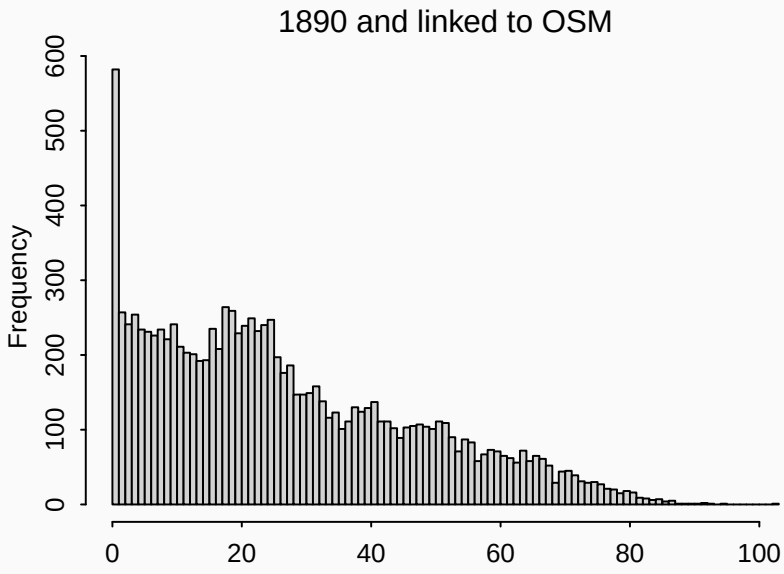
Data attrition and biases

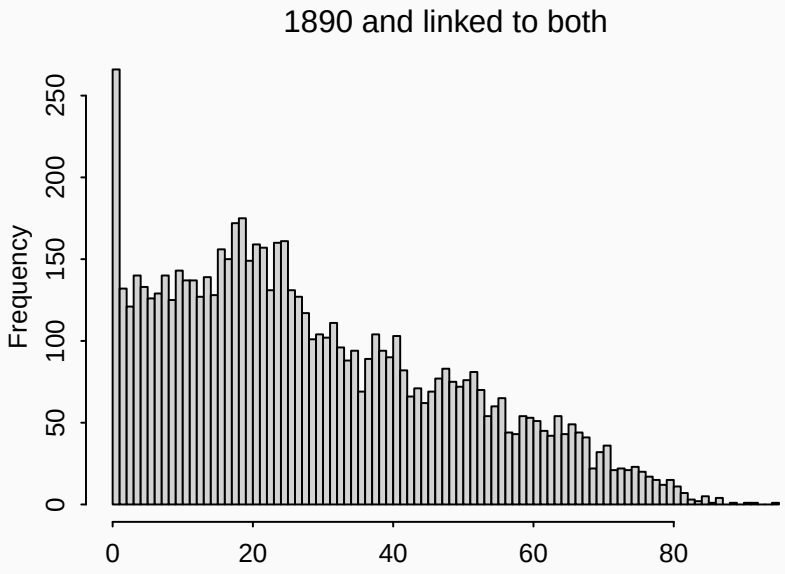
step	N	share	loss
1890 all	14332	1.00	NA
+ linked to DC	9627	0.67	0.33
+ linked to OSM	6267	0.44	0.35
+ sex determined	6218	0.43	0.01
+ life table	5965	0.42	0.04
+ has occupation	2485	0.17	0.58
+ occupation linked	2271	0.16	0.09

Selection and descriptives: age distribution

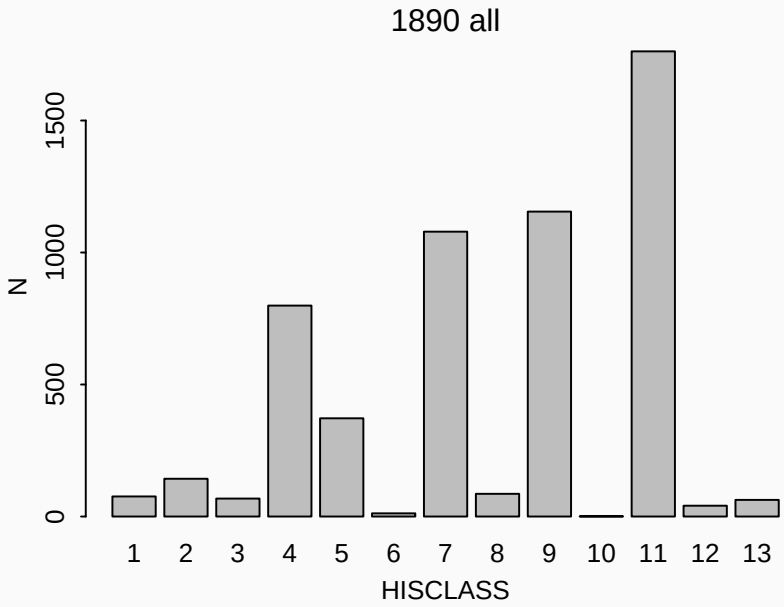


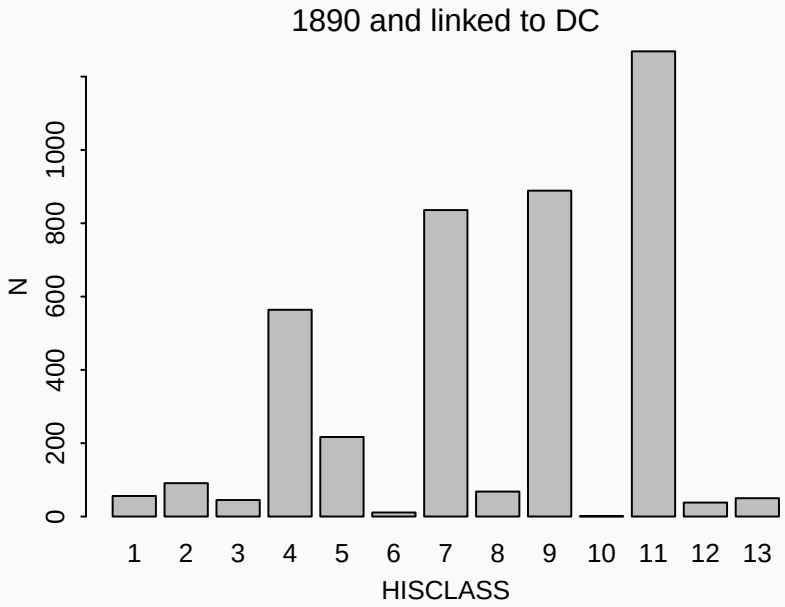


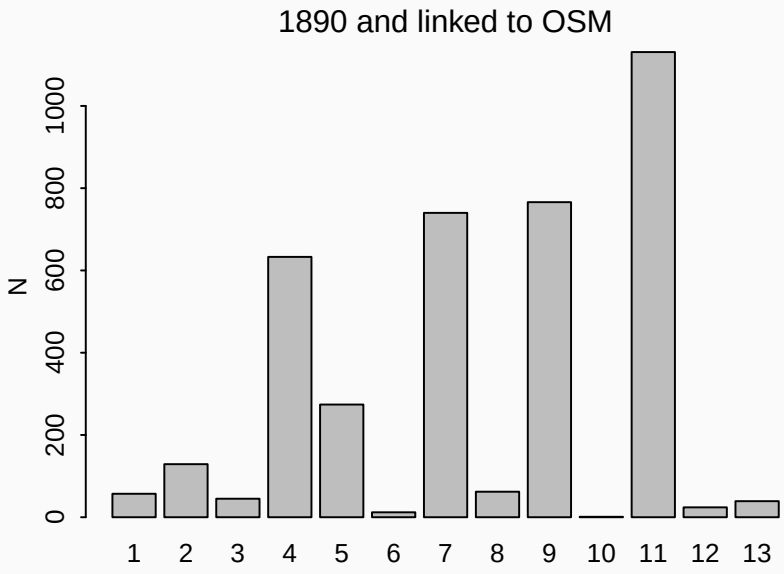


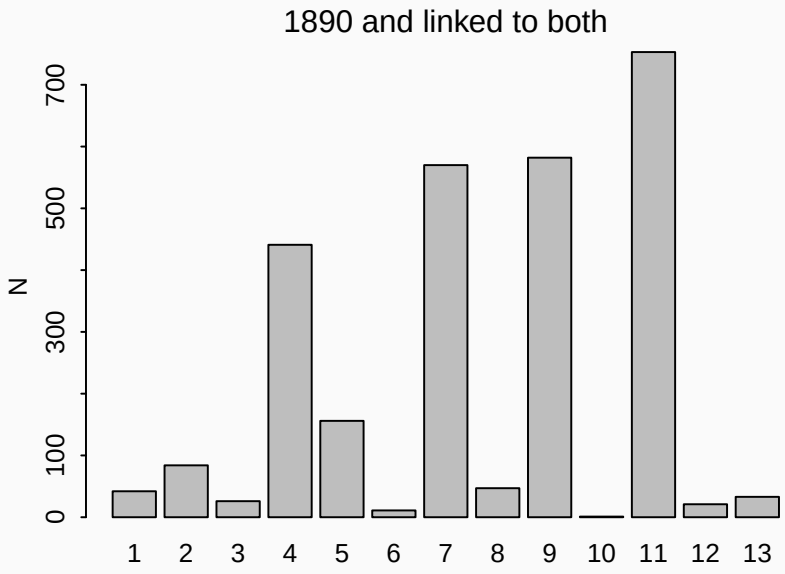


Selection and descriptives: occupational stratification

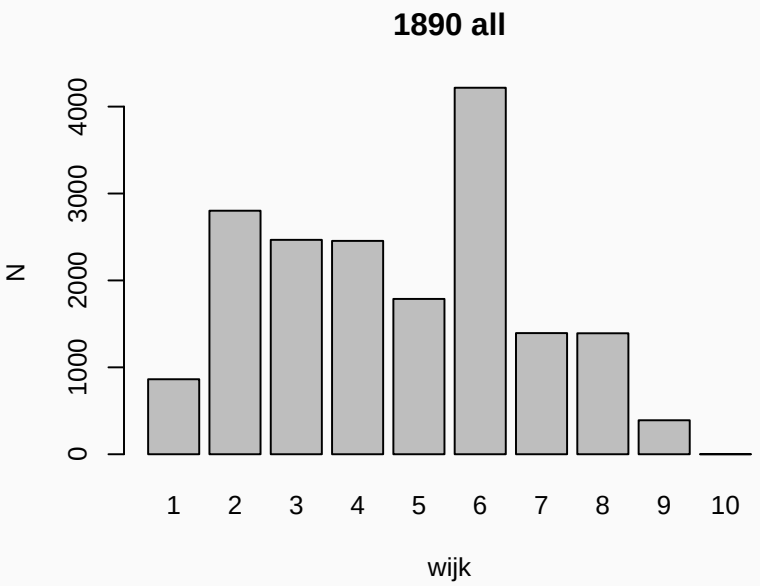


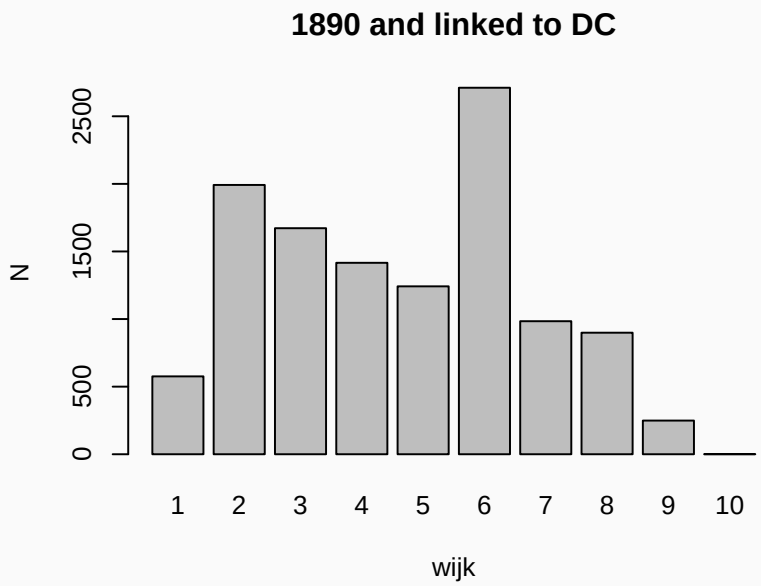


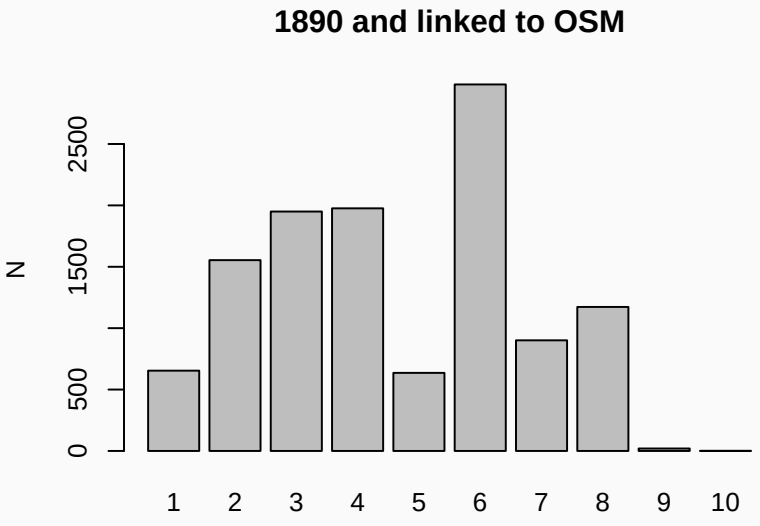


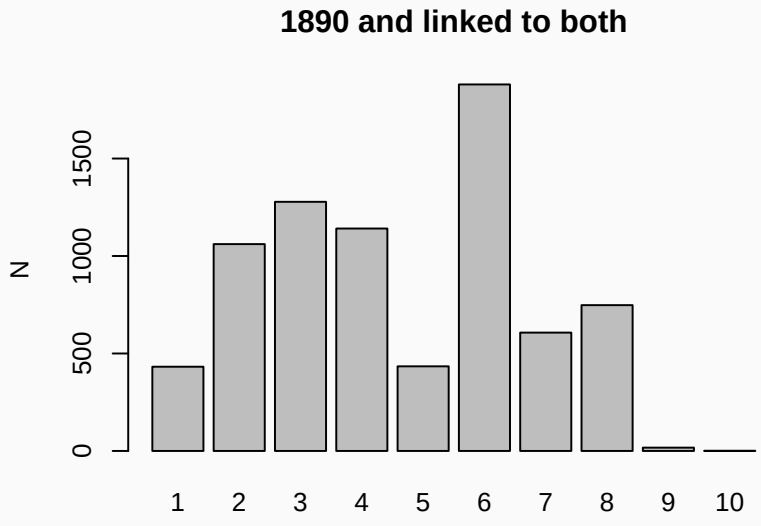


Selection and descriptives: geographic distribution (neighbourhood)









Factories

- List of factories (establishments with at least 10 employees and/or an operating steam engines) obtained from the yearly municipal reports (Philips 2019).
- Contains ISIC group, number of employees, steam engines.
- Supplemented with Struve and Bekaar inquiry of 1887-1889 (with data about i.a. working hours and working conditions).
- Geocoded through addresses: available from the Struve and Bekaar inquiry, advertisements in local newspapers, telephone list of 1915 and secondary literature

Factories: location

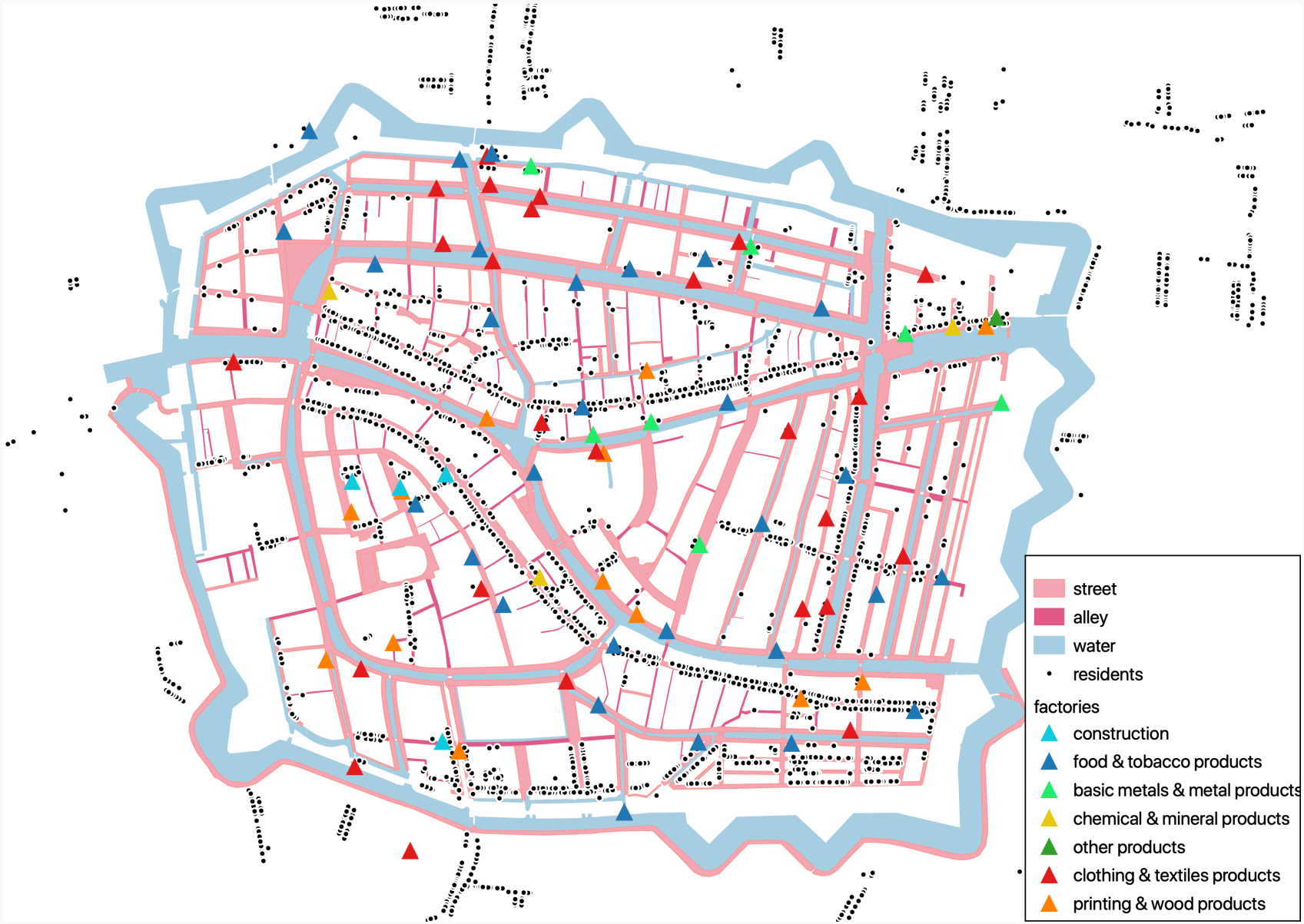
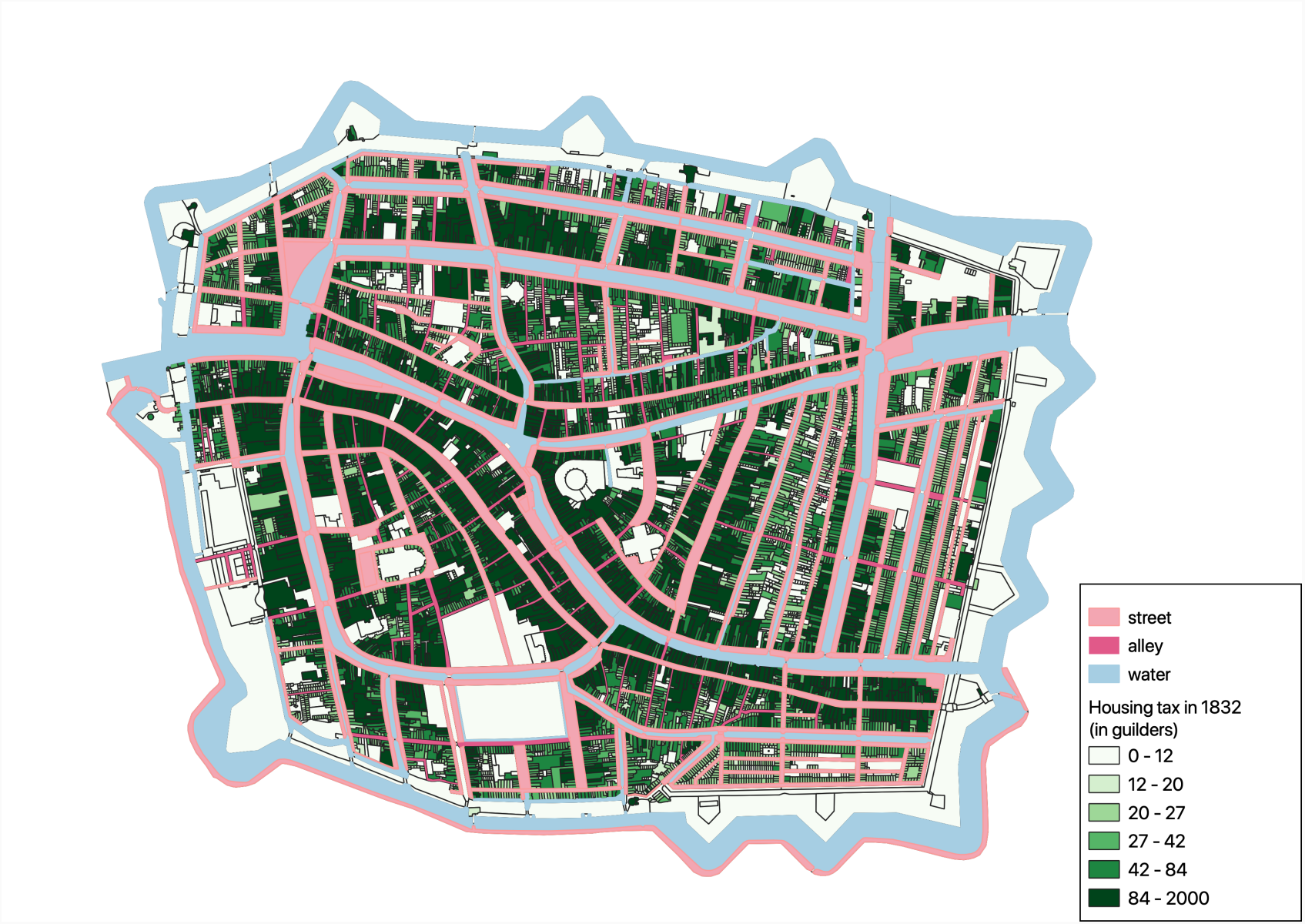


Figure 6: Locations of industry

Factories: location

- Not biased towards periphery: outside city walls or near railways and major waterways.
- Spread throughout the city, in residential areas, and close to canals.
- Some tendency towards areas with lower-value housing, but not extreme.

1832 house values

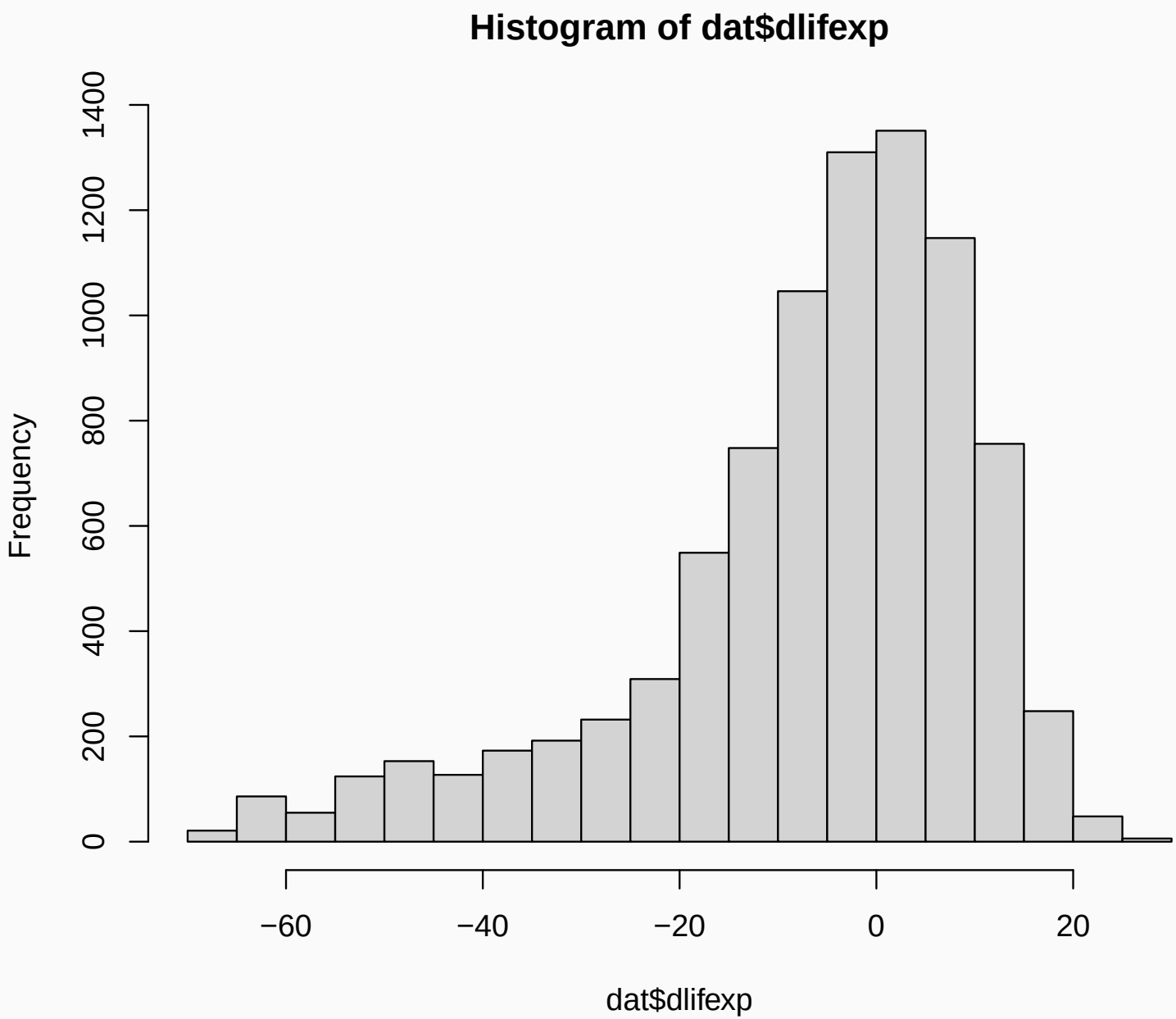


Methods

Methods: loss of life expectancy

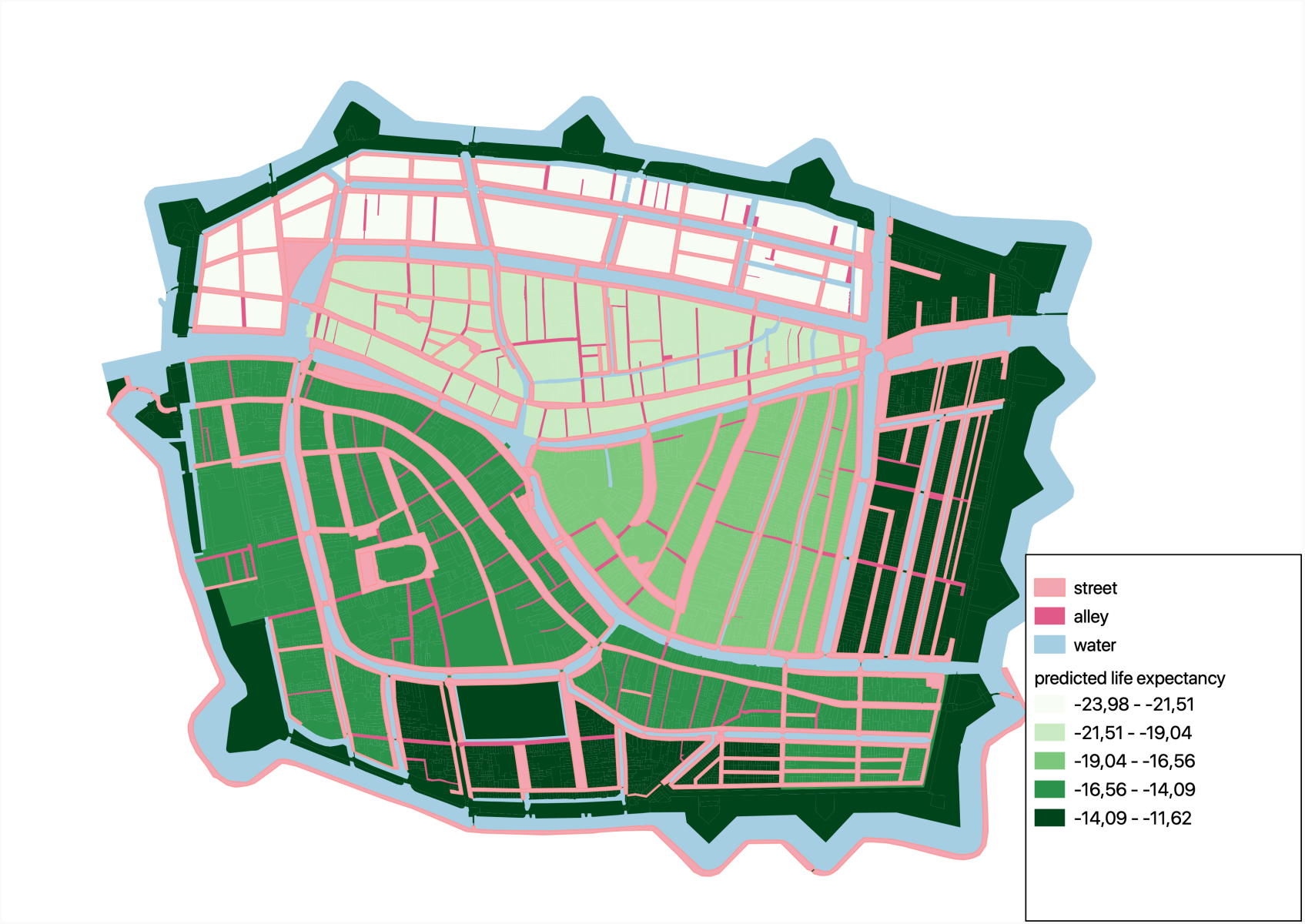
- OLS: loss of life expectancy on factory exposure (Rabl 2003).
- Data mixes cohorts, so simple longevity has issues.
- Take cohort life tables to calculate remaining life expectancy at age, and calculate difference with actual achieved age at death.
- Include basic demographic controls, 1832 house value, neighbourhood FE.

Methods: loss of life expectancy



Method: loss of life expectancy

Method: loss of life expectancy



Methods: factory impact

- Distance-to-house weighted sum of factories, broken down by ISIC groups.
- Weights are not linear, because effect should decrease more than linearly
- Rather, use $\frac{1}{d_{i,f}} X_f$

Methods: factory impact (metal)

Results

Results

Dependent Variable:			dlifexp		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
food.and.tobacco.exp	-2.199**				
	(0.9234)				
metal.exp		-8.830***			
		(1.527)			
chemical.and.mineral.exp			-4.240		
			(50.13)		
printing.and.wood.exp				-5.994**	
				(2.003)	
clothing.and.textile.exp					3.091
					(1.806)
Fixed-effects					
wijk	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	5,593	5,593	5,593	5,593	5,593
R^2	0.02425	0.02465	0.02422	0.02469	0.02431
Within R^2	0.02158	0.02198	0.02154	0.02202	0.02163

Clustered (wijk) standard-errors in parentheses

Conclusions and remaining issues

- Association of proximity of (heavier) industries with reduced life expectancy.
- Many remaining issues though!
- Causality:
    - Work in factory
    - Low incomes, education at poor locations
- Measurement
    - Large parts of *Wijk* 5 not linked
    - Model pollution rather than proximity to factory
    - Geographic careers for longer time path

Additional tables

Years to death

Dependent Variable:	year_to_death				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
male	-2.596***	-2.568***	-2.589***	-2.594***	-2.586***
	(0.4045)	(0.4114)	(0.4158)	(0.4055)	(0.4048)
log1p(belast_ink_gebouw)	0.0054	0.0036	0.0045	0.0040	0.0046
	(0.0298)	(0.0294)	(0.0295)	(0.0297)	(0.0299)
food.and.tobacco.exp	-2.100				
	(1.353)				
metal.exp		-9.634***			
		(1.352)			
chemical.and.mineral.exp			-10.02		
			(47.42)		
printing.and.wood.exp				-6.544**	
				(2.320)	
clothing.and.textile.exp					3.628*
					(1.852)
Fixed-effects					
wijk	Yes	Yes	Yes	Yes	Yes
age	Yes	Yes	Yes	Yes	Yes
Fit statistics					

House clustered errors

Dependent Variable:			dlifexp		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
food.and.tobacco.exp	-2.199				
	(2.556)				
metal.exp		-8.830***			
		(1.553)			
chemical.and.mineral.exp			-4.240		
			(35.15)		
printing.and.wood.exp				-5.994**	
				(2.503)	
clothing.and.textile.exp					3.091*
					(1.690)
Fixed-effects					
wijk	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	5,593	5,593	5,593	5,593	5,593
R^2	0.02425	0.02465	0.02422	0.02469	0.02431
Within R^2	0.02158	0.02198	0.02154	0.02202	0.02163

Clustered (id) standard-errors in parentheses

Conley errors

Dependent Variable:			dlifexp		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
food.and.tobacco.exp	-2.199				
	(1.652)				
metal.exp		-8.830***			
		(1.444)			
chemical.and.mineral.exp			-4.240		
			(45.53)		
printing.and.wood.exp				-5.994***	
				(2.016)	
clothing.and.textile.exp					3.091
					(1.934)
Fixed-effects					
wijk	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	5,593	5,593	5,593	5,593	5,593
R^2	0.02425	0.02465	0.02422	0.02469	0.02431
Within R^2	0.02158	0.02198	0.02154	0.02202	0.02163

Conley (0.053km) standard-errors in parentheses

Inverse square dissipation

Dependent Variable:	dlifexp				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
food.and.tobacco.exp	-2.199**				
	(0.9234)				
metal.exp		-8.830***			
		(1.527)			
chemical.and.mineral.exp			-4.240		
			(50.13)		
printing.and.wood.exp				-5.994**	
				(2.003)	
clothing.and.textile.exp					3.091
					(1.806)
Fixed-effects					
wijk	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	5,593	5,593	5,593	5,593	5,593
R^2	0.02425	0.02465	0.02422	0.02469	0.02431
Within R^2	0.02158	0.02198	0.02154	0.02202	0.02163

Clustered (wijk) standard-errors in parentheses

References

References i

Bailey, Roy E., Timothy J. Hatton, and Kris Inwood. 2018. "Atmospheric Pollution, Health, and Height in Late Nineteenth Century Britain." *The Journal of Economic History* 78 (4): 1210–47. https://doi.org/10.1017/S0022050718000578.

Beach, Brian, and W. Walker Hanlon. 2018. "Coal Smoke and Mortality in an Early Industrial Economy." *The Economic Journal* 128 (615): 2652–75. https://doi.org/10.1111/ecoj.12522.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. KDD '16. New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939785.

References ii

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. "What Do We Know about Short- and Long-Term Effects of Early-Life Exposure to Pollution?" *Annual Review of Resource Economics* 6 (1): 217–47. https://doi.org/10.1146/annurev-resource-100913-012610.

Franke, Richard. 2022. "Poverty, Pollution, and Mortality: The 1918 Influenza Pandemic in a Developing German Economy." *The Economic History Review* n/a. https://doi.org/10.1111/ehr.13165.

Graff Zivin, Joshua, and Matthew Neidell. 2013. "Environment, Health, and Human Capital." *Journal of Economic Literature* 51 (3): 689–730. https://doi.org/10.1257/jel.51.3.689.

Haklay, Mordechai, and Patrick Weber. 2008. "Openstreetmap: User-Generated Street Maps." *IEEE Pervasive Computing* 7 (4): 12–18. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4653466.

References iii

Hanlon, W. Walker. 2018. “London Fog: A Century of Pollution and Mortality, 1866-1965.” w24488. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w24488.

———. 2020. “Coal Smoke, City Growth, and the Costs of the Industrial Revolution.” *The Economic Journal* 130 (626): 462–88. https://doi.org/10.1093/ej/uez055.

Mandemakers, Kees, Rick Mourits, and Sanne Muurling. 2019. “HSN_HISCO_release_2018_01.” https://datasets.iisg.amsterdam/dataset.xhtml?persistentId=hdl:10622/MUZMAL.

Philips, Robin C. M. 2019. “Construction of a Census of Companies for the Netherlands in 1896.” *TSEG - The Low Countries Journal of Social and Economic History* 16 (1). https://doi.org/10.18352/tseg.1064.

References iv

Rabl, Ari. 2003. “Interpretation of Air Pollution Mortality: Number of Deaths or Years of Life Lost?” *Journal of the Air & Waste Management Association* 53 (1): 41–50. https://doi.org/10.1080/10473289.2003.10466118.

Rijpma, Auke, Jeanne Cilliers, and Johan Fourie. 2020. “Record Linkage in the Cape of Good Hope Panel.” *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 53 (2): 112–29. https://doi.org/10.1080/01615440.2018.1517030.