

Supplementary Appendix

Note on the Digital Appendix Animation

Appendix for: Peltola, Jarmo, Sakari Saaritsa, and Henri Mikkola (2023). “Can’t Boil, Won’t Boil: Material Inequality, Information and Disease Avoidance during a Typhoid Epidemic in Tampere, Finland, in 1916.” *Social Science History*.

Digital Appendix: <https://www.dropbox.com/scl/fi/tcz2k4kp22z2oxz7j2ypr/Peltola-Saaritsa-Mikkola-Tampere-1916.gif?rlkey=mx6zjl36qpa7dym941fa2prow&dl=0>

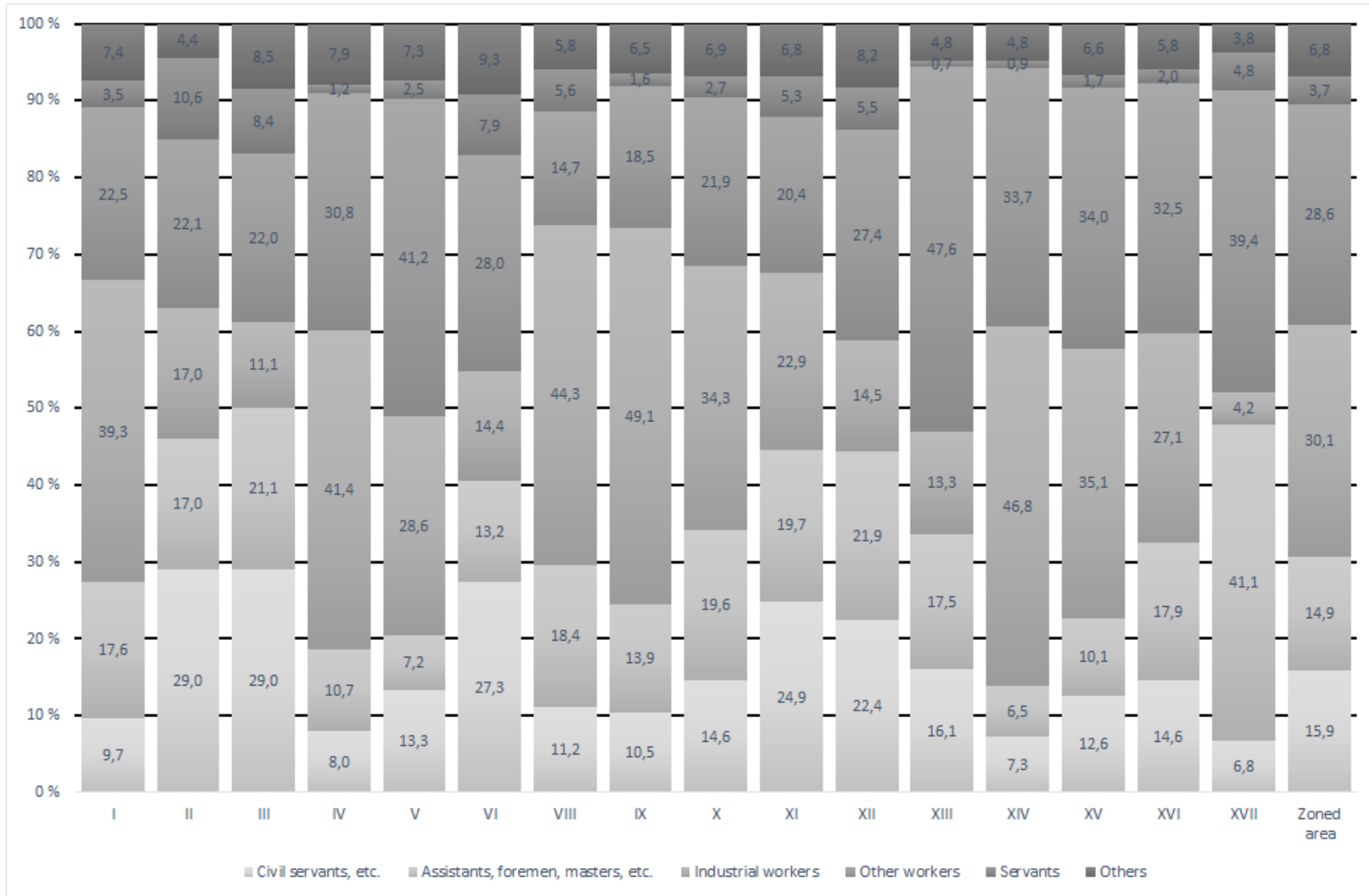
Animation created by Henri Mikkola.

Brown spots indicate patients. Patients are shown at their addresses in Tampere over a map from 1916, whereas in actuality they were hospitalized. Patients exit through recovery (disappearance) or death (skull). Red crosses indicate hospitals. Appearance indicates the establishment of a new temporary typhoid hospital; disappearance indicates shutting down a temporary typhoid hospital.

Note: Animation based on early version of the data with no corrections to time of onset and only patients in temporary hospitals included. See section “data and variables” of the paper for details.

Appendix Figures A1 – A3

Figure A1



Figures A2

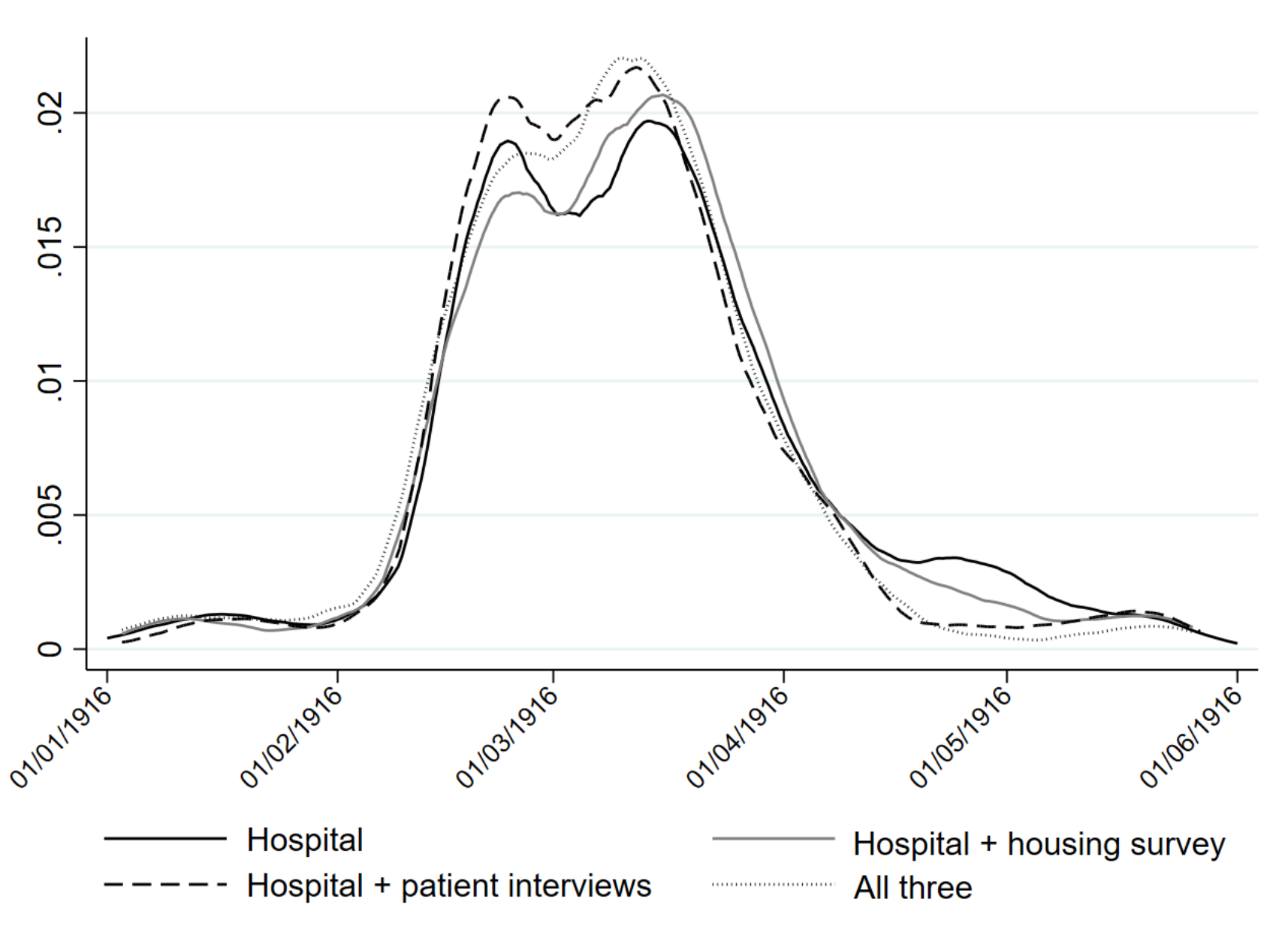
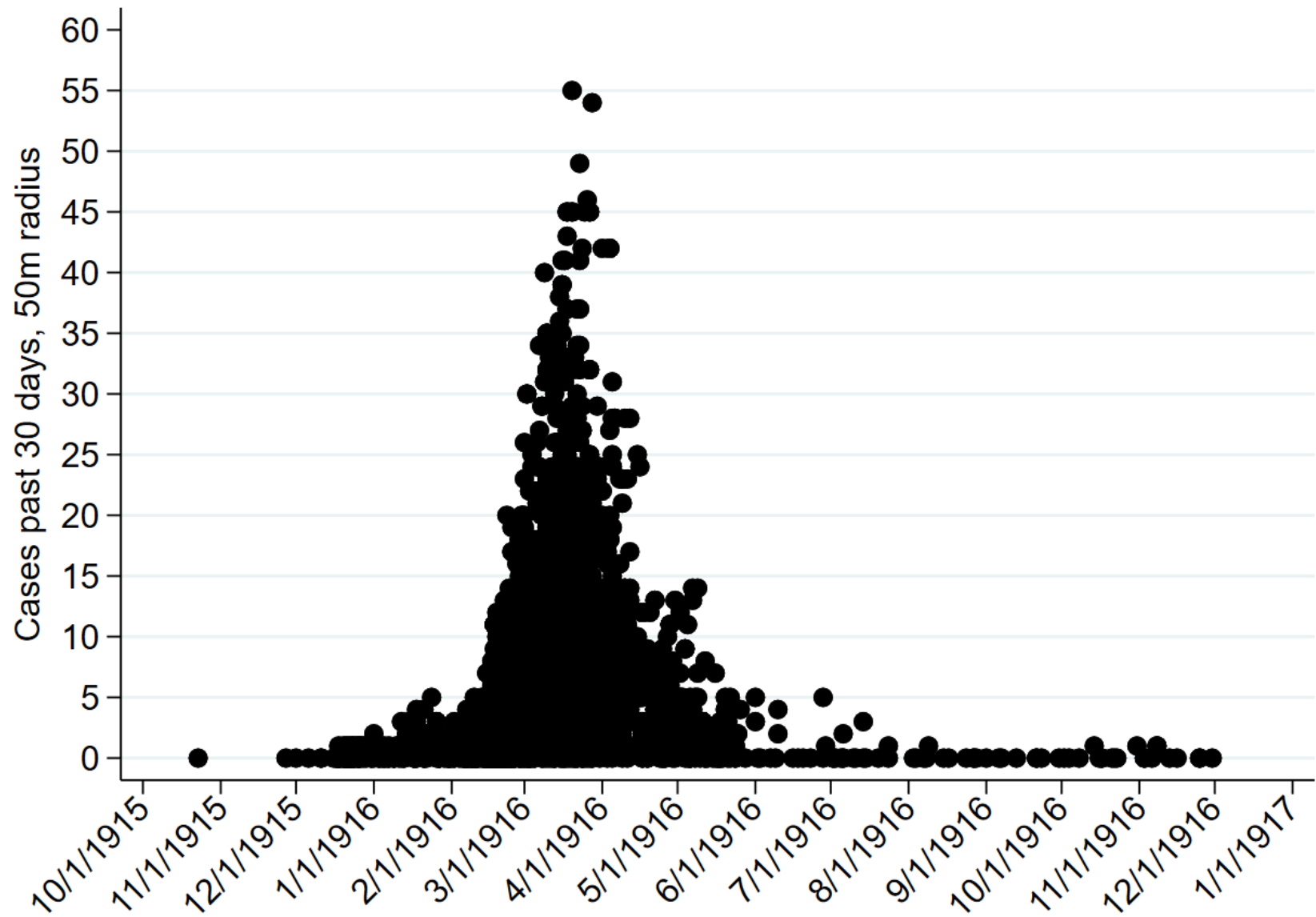


Figure A3



Appendix Tables A1 – A9

Table A1. Characteristics of city districts in Tampere

District	Tax rate per capita	Tap inside: %	WC inside: %	Swedish speaking (high SES proxy). %	Typhoid morbidity / 1910 census population: %
I	1.6	35.3	5.0	5.4	7.2
II	4.2	54.2	18.0	12.4	6.9
III	3.3	69.3	28.4	14.9	5.5
IV	0.8	13.8	2.5	0.9	8.7
V	0.9	26.6	4.7	3.7	8.4
VI	1.9	71.4	22.7	10.1	5.2
VIII	2.4	3.7	0.0	6.2	0.0
IX	1.4	9.0	0.5	2.9	6.7
X	1.6	28.6	6.6	2.3	5.2
XI	2.2	65.8	15.1	7.5	5.9
XII	2.0	84.7	16.2	7.8	5.5
XIII	0.7	1.1	0.0	2.6	9.8
XIV	0.7	59.4	1.3	0.7	10.9
XV	0.8	38.2	0.6	2.0	8.6
XVI	0.9	42.3	3.4	2.4	8.5
XVII	1.5	93.1	59.8	1.7	7.8
Zoned area	1.6	42.9	9.5	5.1	7.5

Sources: Tampereen terveydenhoitolautakunnan kertomus vuodelta 1916; Tampereen rahatoimikamarin kertomus vuodelta 1916; Vuoden 1910 väestönlaskenta Tampereella. SVT VI: 44; 4 (Väenlasku Tampereella joulukuun, 7 p., 1910); Vuoden 1910 väestönlaskenta Tampereella. SVT VI: 50; 4 (Väenlasku Tampereella joulukuun, 7 p., 1910).

Table A2. Cox proportional hazards regressions on daily infection risk among the eventually infected, baseline, occupational, and city district controls for housing survey-linked data only.

Age	0.999 (0.14)
Age^2	1.000 (0.54)
Male	1.019 (0.22)
Opening peaks	3.872 (12.85)***
Distance	1.000 (4.46)***
30d. 50m	1.008 (1.27)
Artisan	0.897 (0.61)
Employee	0.728 (1.87)*
Entrepreneur	0.237 (1.99)**
Farmer	0.919 (0.16)
Industrial worker	0.977 (0.20)
Landlord	2.248 (1.12)
None	0.916 (0.43)
Outworker	0.734 (1.84)*
Primary school	1.076 (0.25)
Professions	0.396 (0.92)
Railways	0.548 (0.78)
Secondary school	1.341 (0.75)
Servant	1.005 (0.03)
Skilled worker	1.117 (0.76)
Small business	1.265 (0.94)
Student	0.770 (0.84)
Univ. student	5.378

	(2.32)**
I	0.859 (0.40)
II	0.819 (0.56)
III	0.884 (0.36)
IV	0.934 (0.21)
V	0.707 (1.02)
VI	1.363 (0.85)
VIII	1.921 (0.61)
<i>N</i>	688

Notes: Proportional hazards, z scores shown in parentheses. * p<0.1; ** p<0.05; *** p<0.01

Table A3. Cox proportional hazards regressions for daily infection risk among the eventually infected, for housing and amenities, for cases without missing values on any housing characteristics only.

Shared kitchen	1.249 (2.27)**				1.188 (1.54)
Tap outside		1.107 (0.81)			1.030 (0.22)
Shared latrines			1.049 (0.38)		0.941 (0.44)
Laundry house				0.802 (2.30)**	0.843 (1.67)*
<i>N</i>	565	565	565	565	565

Notes: All models include baseline, occupational, and city district controls. Modeling is based on time to infection from Nov 15th, 1915. *N* represents patients. Proportional hazards, z scores shown in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A4. Cox proportional hazards regressions on daily infection risk among the eventually infected, housing, and amenities with time-varying coefficients.

		Univariate	Univariate	Multivariate
Main	Shared kitchen	1.387 (2.31)**		1.253 (1.23)
Tvc	Shared kitchen	0.887 (0.72)		0.931 (0.32)
main	Laundry house		0.822 (1.73)*	0.713 (2.02)**
tvc	Laundry house		1.127 (0.88)	1.297 (1.29)
main	Tap outside			1.013 (0.06)
tvc	Tap outside			1.015 (0.06)
main	Shared latrines			0.867 (0.63)
tvc	Shared latrines			1.145 (0.52)
<i>N</i>		688	987	565

Notes: estimates denoted “tvc” are interactions with an indicator for the time being later than March 1st, 1916. “Main” denotes the main effects from the same regressions. Proportional hazards in columns marked “univariate” are from models including only the housing characteristic with estimates. All models include baseline, occupational, and city district controls. Variation in N is based on response rates to individual questions; the results are robust to limiting to smallest N. Proportional hazards, z scores shown in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A5. Cox proportional hazards regressions on daily infection risk among the eventually infected, baseline, occupational, and city district controls for patient survey-linked data only.

Age	1.007 (0.90)
Age^2	1.000 (1.44)
Male	1.043 (0.73)
Opening peaks	3.135 (16.79)***
Distance	1.000 (0.22)
30d. 50m	0.987 (3.36)***
Ag. laborer	0.359 (1.74)*
Artisan	1.054 (0.45)
Employee	0.908 (0.74)
Entrepreneur	0.325 (1.90)*
Farmer	0.901 (0.23)
Industrial worker	1.010 (0.12)
Landlord	1.797 (2.05)**
None	0.933 (0.42)
Outworker	0.861 (1.18)
Primary school	1.204 (0.94)
Professions	1.341 (0.58)
Railways	1.191 (0.83)
Secondary school	1.614 (1.66)*
Servant	1.167 (1.03)
Skilled worker	0.955 (0.47)
Small business	0.995 (0.03)
Student	0.795

	(0.82)
Univ. student	2.843
	(1.04)
I	0.829
	(0.95)
II	1.100
	(0.54)
III	0.924
	(0.47)
IV	1.046
	(0.32)
V	0.879
	(0.83)
VI	1.263
	(1.32)
XI	0.674
	(1.99)**
XII	0.682
	(2.08)**
XIII	0.642
	(1.29)
XIV	0.869
	(0.96)
XV	0.782
	(1.61)
XVI	0.815
	(1.33)
XVII	0.954
	(0.16)
XX	0.419
	(0.85)
XXI	1.086
	(0.30)
<i>N</i>	1,366

Notes: Proportional hazards, z scores shown in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A6. Cox proportional hazards regressions for daily infection risk among the eventually infected, using unboiled water, for cases without missing values on any water use only.

Unboiled (drink)	1.385 (3.75)***			1.246 (2.34)**
Unboiled (dishes)		1.233 (3.33)***		0.968 (0.41)
Unboiled (bath)			1.407 (5.55)***	1.378 (4.34)***
<i>N</i>	1,279	1,279	1,279	1,279

Notes: All models include baseline, occupational, and city district controls. Modeling is based on time to infection from Nov 15th, 1915. *N* represents patients. Proportional hazards, z scores shown in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A7. Cox proportional hazards regressions on daily infection risk among the eventually infected, using unboiled water, with time-varying coefficients.

		Univariate	Univariate	Multivariate
main	Unboiled (drink)	2.244 (7.50)***		1.602 (3.87)***
tvc	Unboiled (drink)	0.378 (5.90)***		0.567 (3.09)***
main	Unboiled (bath)		2.827 (10.00)***	2.467 (7.41)***
tvc	Unboiled (bath)		0.332 (8.62)***	0.399 (5.91)***
main	Unboiled (dishes)			0.987 (0.11)
tvc	Unboiled (dishes)			0.977 (0.15)
<i>N</i>		1,335	1,331	1,279

Notes: estimates denoted “tvc” are interactions with an indicator for the time being later than March 1st, 1916. “Main” denotes main effects from the same regressions. Proportional hazards in columns marked “univariate” are from models including only the housing characteristic with the estimates. All models include baseline, Occupational, and city district controls. Variation in N is based on response rates to individual questions; the results are robust to limiting to smallest N. Proportional hazards, z scores shown in parentheses. * p<0.1; ** p<0.05; *** p<0.01

Table A8. Cox proportional hazards regressions for daily infection risk among the eventually infected, false explanations, for cases without missing values on any water use only.

	Over the whole epidemic		Before May 12th, 1916	
False explanations	1.196 (2.52)**	1.236 (2.98)***	1.029 (0.40)	1.073 (1.00)
Unboiled (drink)		1.284 (2.65)***		1.444 (3.80)***
Unboiled (dishes)		0.986 (0.18)		1.152 (1.78)*
Unboiled (bath)		1.380 (4.35)***		1.488 (5.44)***
<i>N</i>	1,278	1,278	1,243	1,243

Notes: All models include baseline, occupational, and city district controls. Modeling is based on time to infection from Nov 15th, 1915. *N* represents patients. May 12th, 1916 is used to limit the data to the cold period, when patients gave false explanations, to control for a spurious correlation between warmer weather and fewer false explanations. Proportional hazards, *z* scores shown in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A9. Linear probability model (OLS) for reporting the use of unboiled water for washing (personal hygiene) based on housing characteristics.

	Unboiled (bath)	Unboiled (bath)	Unboiled (bath)	Unboiled (bath)
Time (days)	-0.016 (0.004)***	-0.016 (0.004)***	-0.017 (0.004)***	-0.017 (0.004)***
Time (days)^2	0.000 (0.000)***	0.000 (0.000)***	0.000 (0.000)***	0.000 (0.000)***
Shared kitchen	0.061 (0.053)			0.073 (0.058)
Tap outside		0.068 (0.067)		0.135 (0.078)*
Laundry house			0.049 (0.046)	0.094 (0.062)
R^2	0.19	0.21	0.20	0.20
N	388	472	531	360

Notes: All models include baseline, occupational, and city district controls. Variation in N is based on response rates to individual questions; the results are robust to limiting to smallest N. Robust standard errors shown in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$