Online Appendix

Poverty, Religious Differences and Child Mortality in the Early Twentieth Century: The Case of Dublin

Dylan Shane Connor

1. Differences Between Sample and Population by "Relationship to Head"

Street	Immigrant to Dublin				Born in Dublin			
	Total	Wife	Lodger	Other	Total	Wife	Lodger	Other
All	4717	120 (91%)	100 (1%)	100 (8%)	9469	150 (90%)	110 (1%)	130 (9%)
> 50% Immigrants	683	70 (87%)	70 (2%)	60 (11%)	311	80 (87%)	0 (1%)	40 (12%)
<= 50% Immigrants	4034	130 (92%)	100 (1%)	100 (7%)	9158	150 (90%)	120 (1%)	130 (9%)

Appendix Table 1. Child Deaths (per thousand) and Mothers' "Relationship to Head"

The couple-level observations used in this study were constructed by pairing husbands and wives based on the "relationship to head" variable in the 1911 census. This technique could produce a sample in which mortality is either overestimated or underestimated compared to the population of parents as a whole. This would be the case if mortality was correlated with the return on "Relationship to Head" return in the census. In light of the findings of this study, one notable instance of this could be where immigrants to the city rent a room in the house of a person who is not their husband or wife. This practice was relatively common in the early 20th century. These people may declare themselves as "Lodgers" and would be excluded from my sample.

In the main analysis, I show that mortality is lower for immigrants to the city. Thus, in Appendix Table 1 I examine differences in the mortality rate of women who claims to be a "Wife," "Lodger" or "Other." Further, I compare this between streets with majority and minority shares of immigrants. Appendix Table 1 shows that mortality was much lower on majority immigrant streets. However, mortality was only slightly higher among natives to the city.

Appendix Table 1 suggests that women classified as "Wife" experienced higher mortality. Thus, mortality is likely higher in my sample, which is limited to clearly identifiable husbands and wives, than in the population as a whole. However, this tradeoff is necessary to ensure reliable construction of couples. Further, Appendix Table 1 also indicates that the classification of women as wives or lodgers did not vary much between immigrants and natives. The fact that immigrants to the city have a similar likelihood to natives to be classified as a "Lodger" or a "Wife" suggests that classification is not a major concern for my study.

2. Missing Data Imputation

I investigated several sources of bias in constructing the sample. Many husbands erroneously answered or were assigned the fertility information of their wives. Where possible wives with blank fertility information were reassigned the values of their husbands. I analyzed the distribution of missing values for the thirty five couples that did not state their number of years married, the sixty six that did not state their number of children born, and the 456 not stating their number still alive. For those missing responses on children born and years of marriage questions, there was overrepresentation among illiterates and husbands with lower class occupations, those likely to have higher rates of child mortality.

	Child Alive	
Children Born	Missing	Not Missing
One	0.57	0.25
Two	0.20	0.18
Greater than two	0.22	0.57

Appendix Table 2. Parity and Missing Mortality Information

The 'children alive' question is more problematic. Appendix Table 2 shows that younger couples with only two or less children ever born were more likely to not report to number of children still alive. It appears that many mothers whose children had all died left this response blank. This is consistent with a greater share of missing values among mothers with fewer children born. I assigned a zero value to mothers with complete information on parity but missing values for mortality. Following this imputation, the distribution of remaining non-response was similar to other martial fertility variables and appears to have been randomly distributed.

This imputation procedure did not influence any of the major conclusions of this study. I provide a sensitivity analysis for this decision in Appendix Table 3. Although there are some changes in effect size, the direction and significance of key coefficients remain unchanged. The decision not to impute would lower the gap in Catholic mortality but this is expected due to their higher rate of infant mortality. Further, failure to undertake this imputation would lead to the underestimation of Catholic infant mortality.

	Omit Missing	**Imputed Missing**
(Intercept)	1.162***	1.419***
	(0.051)	(0.065)
Age of Wife	0.113***	0.072***
	(0.003)	(0.004)
Age at Marriage	-0.117***	-0.066***
	(0.004)	(0.005)
Religion [Reference = Catholic]		
Protestant	-0.119**	-0.135**
	(0.038)	(0.049)
Jewish	-0.546***	-0.672***
	(0.110)	(0.142)
Birthplace [Reference = Dublin]		
Outside Dublin	-0.089***	-0.065
	(0.027)	(0.034)
Outside Ireland	-0.084	-0.065
	(0.052)	(0.066)
Husband's Occupation (HISCAM)	-0.042***	-0.048**
	(0.012)	(0.016)
Husband Can Read and Write	-0.173***	-0.207***
	(0.041)	(0.052)
Wife Can Read and Write	-0.150**	-0.229***
	(0.046)	(0.058)
SES of Street [Reference = High]		
Low Status Street	0.084**	0.134**
	(0.032)	(0.042)
Mid Status Street	0.047	0.040
	(0.029)	(0.039)
Wife Works	0.135**	0.170**
	(0.042)	(0.053)
Marital Fertility	0.348***	0.229***
	(0.013)	(0.016)
Num. obs.	12859	13247
Num. groups: street	1448	1457
Var: street (Intercept)	0.003	0.016
Var: Residual	1.557	2.605

***p < 0.001, **p < 0.01, *p < 0.05

Statistical models

Appendix Table 3. Sensitivity Analysis for Imputation of Missing Mortality Information

3. Geolocation of Streets

In this study, I rely on a sample of couples extracted from the 1911 census and also subset of these observations ("Geolocated Sample"), who could be located on a street in Dublin using GIS. These streets were located by hand using historical maps. Thus, this geolocated sample is biased toward couples living on more easily identified streets. If mortality was higher on difficult to find streets (e.g. in lanes, courts, side-streets), the geolocated sample may not give an entirely representative picture of the distribution of mortality.

The Geolocated Sample contains relative few streets with only one occupant. Appendix Table 4 shows that while 17 percent of the streets in Dublin have one occupant, only around 52 percent of these could be found. In contrast, 80 percent of streets with one hundred or less could be found.

Number of Couples on Street	**All Streets**		**Geolocated Streets**	
	Count	Share	Count	Share
One	249	17.1%	131	11.1%
Five or less	789	54.2%	546	46.3%
Twenty or less	1300	89.2%	1024	86.9%
Fifty or less	1428	98.0%	1150	97.5%
One hundred or less	1454	99.8%	1176	99.7%
More than one hundred	3	0.2%	3	0.3%
All	1457	100.0%	1179	100.0%

Appendix Table 4. Comparison of the Population Distributions of Geolocated Streets to All Streets

The difficulty of locating a street is correlated with mortality and population size, and to a lesser extent, its Catholic share. Appendix Table 5 presents the results of a logistic regression where the outcome variable is whether a street was found (1) on a historic map or not (0). Column 1 suggests that mortality was higher on streets that could be geolocated. Column 2 adds controls for the total population and the Catholic share of streets. Streets with larger populations were considerably more likely to be found. This suggests that mortality is lower in the Geolocated Sample and the population of individual streets is larger. This finding may suggest that the influence of hazards may be slightly underestimated in my analysis.

	Model 1	**Model 2**
(Intercept)	0.852^{***}	0.851^{***}
	(0.015)	(0.014)
Average Mortality Index	-0.042^{***}	-0.041^{***}
	(0.010)	(0.010)
SES of Street	0.014	0.006
	(0.010)	(0.011)
Catholic Share of Street		-0.026^{*}
		(0.011)
Population Street		0.106^{***}
		(0.010)
Deviance	221.861	204.938
Num. obs.	1457	1457

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Statistical models

Appendix Table 5. Logit of Finding Street on Historic Map

4. Mixed Marriages

	1	2	3	4	5	6
(Intercept)	1.440***	1.429***	1.630***			
	(0.047)	(0.048)	(0.028)			
Wife Catholic	0.192***	0.111				
	(0.049)	(0.092)				
Husband Catholic		0.092				
		(0.088)				
Marriage Composition (ref = both Catholic)						
Wife Catholic: Husband non-Catholic			0.015			
			(0.106)			
Wife non-Catholic: Husband Catholic			0.103			
			(0.150)			
Both non-Catholic			-0.220***			
			(0.051)			
Wife Dubliner				0.074*	0.044	
				(0.034)	(0.035)	
Husband Dubliner					0.074*	
					(0.034)	
Marriage Composition (ref = both Dubliners)						
Wife Dubliner: Husband in-migrant						-0.042
						(0.044)
Wife in-migrant: Husband Dubliner						-0.003
						(0.050)
Both in-migrants						-0.127*
						(0.039)
Num. obs.	13247	13247	13247	13247	13247	13247
Num. groups: street	1457	1457	1457	1457	1457	1457
Num. groups: ded	15	15	15	15	15	15

***p < 0.001, **p < 0.01, *p < 0.05

Statistical models

Appendix Table 6. Regression Showing Mixed-Marriage Effects on Infant and Child Mortality

The religion of each couple is one of the main variables of interest in this study. This was defined using the stated religion of the wife/mother and mixed marriages (which were rare) were ignored. To explore whether this could have influenced my results, I regress the mortality index on measures of wife's religion, husband's religion and on cases of mixed marriage. These results are presented in Appendix Table 6. These models included the full battery of independent variables and these effects are robust to the mixed-marriage specification.

Whether or not I focus on the religion of the husband or the wife, my models produce consistent results. Column 1 shows higher mortality among Catholic wives, while Column 2 shows that including husband's religion partitions this effect almost equally between the husband and wife. Column 3 shows that Catholic wives with non-Catholic husbands experience similar levels of mortality to Catholic wives with Catholic husbands. Column 6 provides similar reassurance that this not of serious consequence to define migrant status using only the characteristics of the wife. Moreover, in the analysis, I show very high rates of endogamy (95% and up), so any deviations based on mixed marriages would have only a minor influence on mortality outcomes.

5. Robust for Distribution of the Response Variable

	Outcome (Y) = log(mortality index)		Outcome (Y) = Standard mortality index	
	1	**2**	**3**	**4**
(Intercept)	-2.384***	-1.967***	1.090***	1.419***
	(0.028)	(0.098)	(0.017)	(0.065)
Religion [Reference = Catholic] Protestant	-0.636***	-0.316***	-0.288***	-0.135**
	(0.074)	(0.074)	(0.046)	(0.049)
Jewish	-1.343***	-1.353***	-0.726***	-0.672***
	(0.213)	(0.214)	(0.132)	(0.142)
Age of Wife		0.289***		0.072***
		(0.006)		(0.004)
Age at Marriage		-0.299***		-0.066***
		(0.007)		(0.005)
Birthplace [Reference = Dublin]				
Outside Dublin		-0.124*		-0.065
		(0.052)		(0.034)
Outside Ireland		-0.110		-0.065
		(0.100)		(0.066)
Husband's Occupation (HISCAM)		-0.079**		-0.048**
		(0.024)		(0.016)
Husband Can Read and Write		-0.310***		-0.207***
		(0.079)		(0.052)
Wife Can Read and Write		-0.256**		-0.229***
		(0.088)		(0.058)
Wife Works		0.135		0.170**
		(0.080)		(0.053)
Marital Fertility		0.717***		0.229***
		(0.024)		(0.016)
Num. obs.	13247	13247	13247	13247
Num. groups: street	1457	1457	1457	1457
Var: street (Intercept)	0.088	0.020	0.035	0.016
Var: Residual	7.020	5.991	2.685	2.605

***p < 0.001, **p < 0.01, *p < 0.05

Statistical models

Appendix Table 7. Comparison of effects in models with and without logged dependent variable

The mortality index is bounded between zero and positive infinity and is right-skewed. A large body of research has shown that this skewness is not a cause for concern in regression analyses (see text for sources). One way to adjust for this skewness is to (log) transform the response variable. For robustness, I include this alternate approach in Appendix Table 7.

The main conclusions with respect to religion - the primary variable of interest in this study - do not vary between the logged and standard response models. In Appendix Table 7, Columns 1 and 3 show that mortality is lower among Protestant and Jewish couples relative to Catholic couples. The inclusion of the full battery of control variables in Columns 2 and 4 lead to reductions in the Protestant-Catholic mortality gap of around 50 percent and negligible reductions in the Jewish difference. These consistent results across both sets of models provides quite strong evidence that alternate specifications of the response variable would not change these results and that the distribution of the response (as specified in the main analysis) is not cause for concern.

6. Wald Tests of Estimates by Model

In the article, my significance testing of parameters relies on equivalents of the likelihood ratio tests. However, tests using a chi-squared distribution, referred to as Wald Tests, are often recommended as a more formal means of significance testing for parameters or groups of parameters (as in categorical variables with n-1 dummy contrasts) (Goldstein 2011, p. 41). In this section, I provide the results of these Wald tests for each set of parameters in the four main models used in this analysis.

These estimates provide very similar results to those estimated from the likelihood ratio tests:

Table 5. Model 3.	Chisq	DF	Pr(>Chisq)	
Age of Wife	220.6903	1	< 2.2e-16	***
Age at Marriage	109.0189	1	< 2.2e-16	***
Religion	29.0159	2	5.00E-07	***
Birthplace	4.8498	2	0.088487	.
Husband's Occupation	7.3587	1	0.006674	**
Wife Can Read and Write	16.3302	1	5.32E-05	***
Husband Can Read and Write	15.4975	1	8.26E-05	***
Status of Street	11.6425	2	0.002964	**
Wife Works	10.9719	1	0.000925	***
Marital Fertility	99.9883	1	< 2.2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix Table 10a.

Table 6. Model 3	Chisq	DF	Pr(>Chisq)	
Age of Wife	146.1486	1	< 2.2e-16	***
Age at Marriage	69.1432	1	< 2.2e-16	***
Husband's Occupation	8.838	1	0.00295	**
Wife Can Read and Write	12.7773	1	0.000351	***
Husband Can Read and Write	12.0782	1	0.00051	***
Status of Street	6.3301	2	0.042212	*
Religion/Birthplace	38.7992	5	2.61E-07	***
Diversity	6.4076	1	0.011363	*
Religion/Birthplace X Diversity	14.4713	5	0.012877	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix Table 10b.

Table 7. Model 5.	Chisq	DF	Pr(>Chisq)	
Age of Wife	69.6838	1	< 2.2e-16	***
Birthplace	11.942	2	0.002552	**
Catholic	13.2456	1	0.000273	***
Status of Street	17.0092	2	0.000203	***
Husband's Occupation	14.3342	1	0.000153	***
Diversity	11.411	1	0.00073	***
Catholic x Status of Street	3.398	2	0.182868	
Catholic x Husband's Occupation	0.5197	1	0.47099	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix Table 10c.

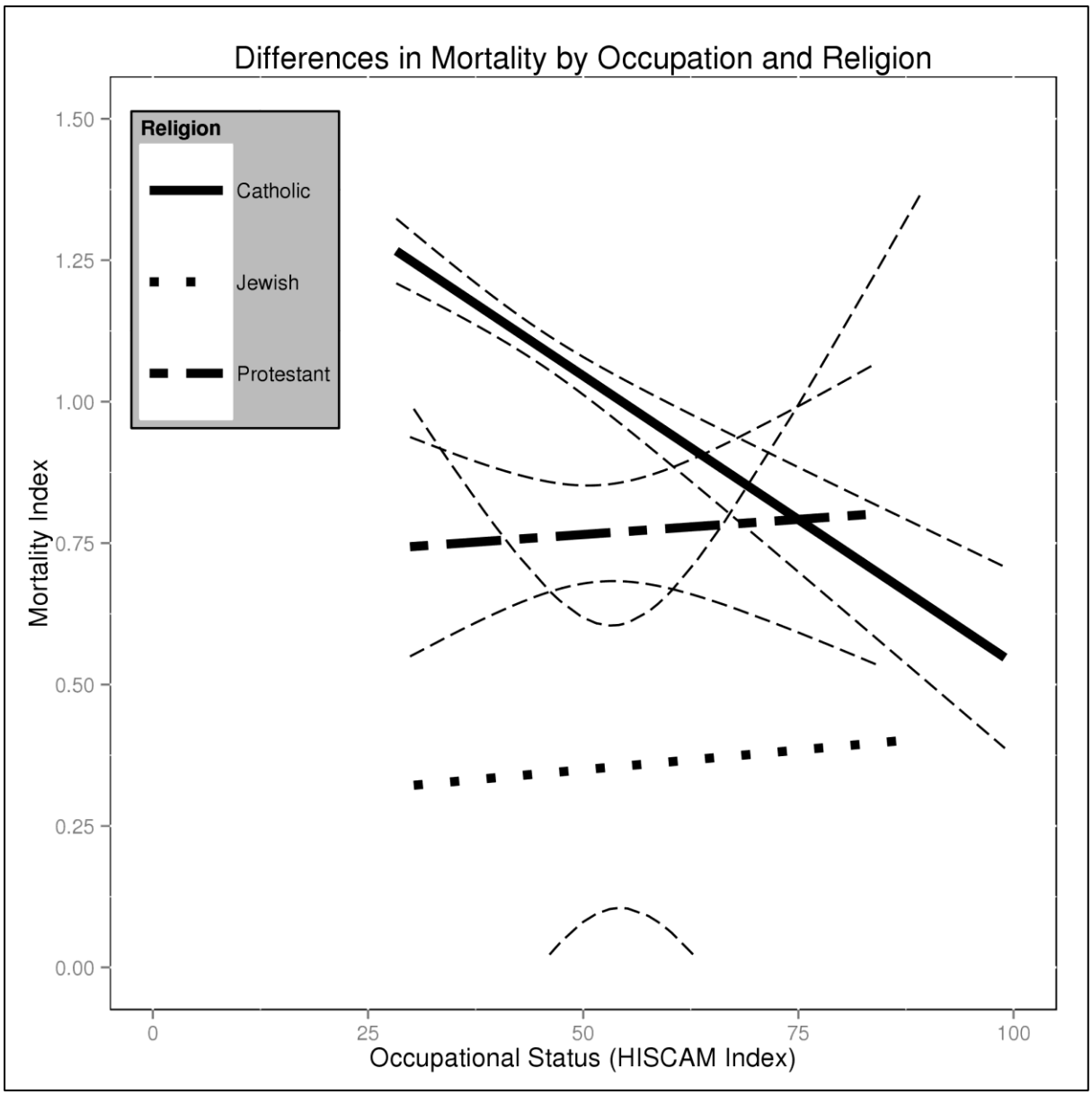
Table 8. Model 4.	Chisq	DF	Pr(>Chisq)	
Age of Wife	58.0201	1	2.60E-14	***
Catholic	19.1652	1	1.20E-05	***
Husband's Occupation	13.1524	1	0.000287	***
Status of Street	6.4745	2	0.039271	*
Prox. to River Liffey	1.7416	1	0.186935	
Prox. to Typhoid	5.6438	1	0.017517	*
Prox. to Clean Milk	3.3734	1	0.066258	.
Quality Housing	3.6609	1	0.055703	.
Diversity	5.2282	1	0.022224	*
Catholic X Prox. to River Liffey	0.0575	1	0.81057	
Catholic X Prox. to Typhoid	0.2458	1	0.620032	
Catholic X Prox. to Clean Milk	3.1147	1	0.07759	.
Catholic X Quality Housing	0.0786	1	0.779147	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Appendix Table 10d.

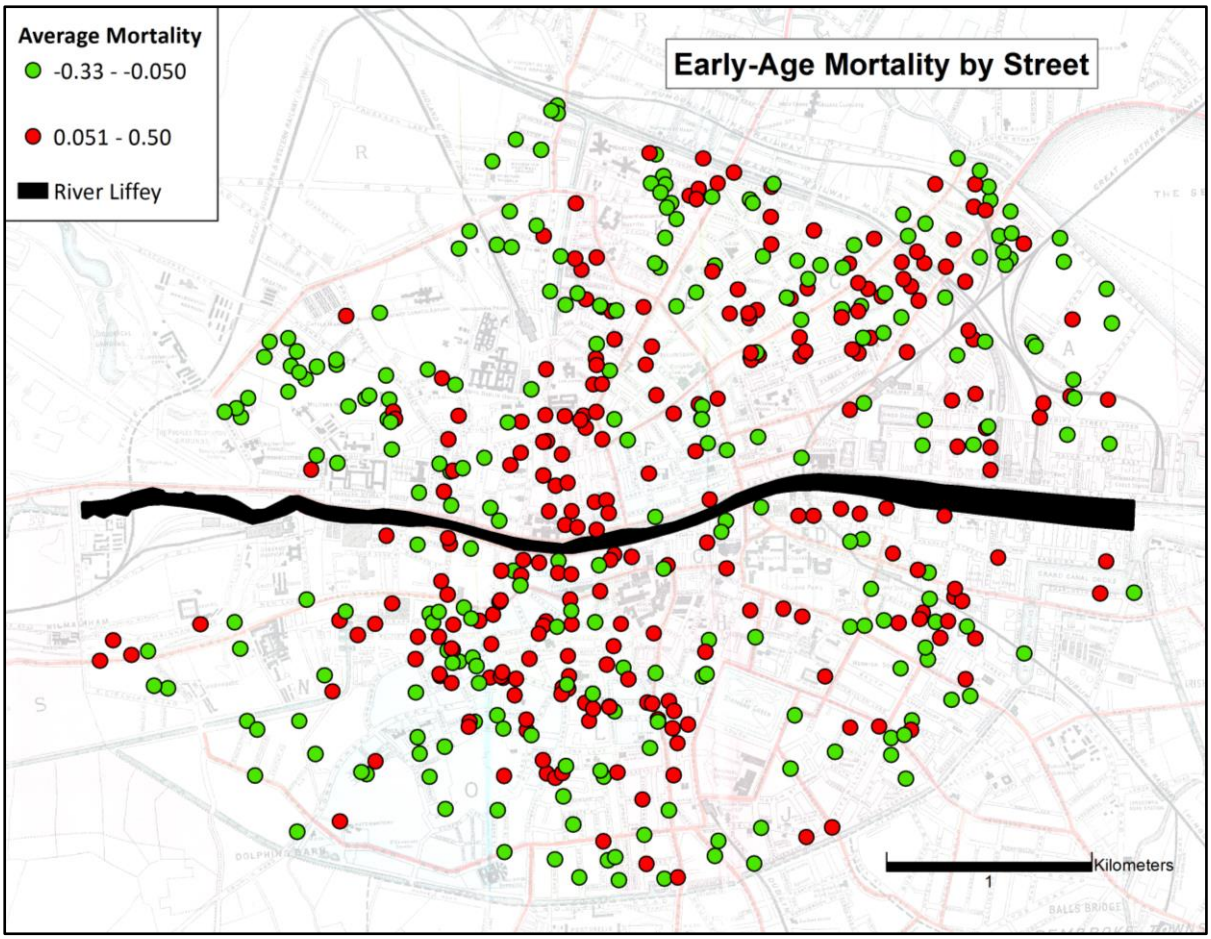
7. Appendix Figures

Appendix Figure 1. Age at Marriage



Appendix Figure 2. Differences in Mortality by Occupation and Religion

Appendix Figure 2 suggests that there is a strong occupational gradient in mortality, but only for Catholics. The relationship between occupational status and mortality is weak for Jewish and Protestant couples. These confidence interval bands are calculated at the 95 percent level and these predictions are made from a linear model with basic demographic characteristics.



Appendix Figure 3. Early-Age Mortality by Street (Point Estimates)

Appendix Figure 3 depicts the street-level estimates which were used to generate the interpolated surface estimate of mortality, which are presented in the main paper. Red depicts streets with higher than average mortality and green represents streets that are below the average. These estimates show the same pattern, albeit in a slightly more cluttered fashion, and lead to the same conclusions as the surface estimates.