*Online Appendix*

*Atmospheric Pollution, Health, and Height in Late Nineteenth Century Britain*

Roy E. Bailey, Timothy J. Hatton, and Kris Inwood

SUPPLEMENTARY TABLES

Table A1

PROPORTIONS OF SERVICEMEN AND POPULATION AT DIFFERENT LEVELS OF COAL INTENSITY

|  |  |  |  |
| --- | --- | --- | --- |
| Coal Intensity in District (Tons Per Annum) | +/– Median | +/– Upper Quartile | +/– Lower Quartile |
| >3.769 | <3.769 | >8.686 | <8.686 | >2.396 | <2.396 |
| Servicemen Sample |
| Average height | 65.69 | 66.42 | 65.58 | 66.21 | 65.88 | 66.59 |
| Number of servicemen | 1121 | 1114 | 561 | 1674 | 1676 | 559 |
| Share of servicemen | 0.502 | 0.498 | 0.251 | 0.749 | 0.750 | 0.250 |
| Population Aged 0 to 9 in the 1901 Census |
| Population 0 to 9 (000s) | 3530.4 | 3541.0 | 1866.5 | 5205.3 | 5261.2 | 18110.3 |
| Share of population | 0.499 | 0.501 | 0.264 | 0.736 | 0.744 | 0.256 |

Those that enlisted in the army could be drawn disproportionately from districts with high or low coal intensity. Table A1 compares the share of servicemen from districts with above and below different levels of coal intensity with the shares of population aged 0–9 in the same districts in the 1901 census. The first panel shows that the share of servicemen and the share of population in districts with coal intensity above 3.37 tons per annum (approximately the median) is almost identical. The shares of servicemen and population are also similar when divided at the upper and lower quartiles of coal intensity. A regression of the ratio of the number of servicemen to population aged 0–9 across registration districts on log coal intensity yields a coefficient of –0.001 with a t-value of 0.07. Thus there is no evidence that the selection of our sample of servicemen is correlated with coal intensity.

*Source*: Constructed from data described in the text.

Table A2

HETEROSKEDASTICITY AND CLUSTERING LEVEL OF STANDARD ERRORS (DEP VAR HEIGHT)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| Log district coal intensity | –0.343\*\*\* | –0.343\*\*\* | –0.343\*\*\* | –0.343\*\*\* |
| (0.128) | (0.092) | (0.116) | (0.121) |
| Log coal intensity in districts to the southwest | –0.116\*\*\* | –0.116\*\*\* | –0.116\*\*\* | –0.116\*\*\* |
| (0.041) | (0.031) | (0.031) | (0.033) |
| Log coal intensity in districts to the northeast | –0.026 | –0.026 | –0.026 | –0.026 |
| (0.039) | (0.022) | (0.028) | (0.034) |
| Clustering level (number of clusters) | None | Region (11) | County (45) | Registration district (472) |
| Cook-Weisberg (χ2(1)) | 0.57 |  |  |  |
| R-squared | 0.082 | 0.082 | 0.082 | 0.082 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |

*\** = Significant at the 10 percent level.

*\*\** = Significant at the 5 percent level.

*\*\*\** = Significant at the 1 percent level.

*Note*: Standard errors in parentheses. Individual and household variables, district variables (population density, share of white collar workers, population growth, coast, windspeed, rainfall, temperature, ruggedness), and regional dummies are included but not reported.

*Source*: Constructed from data described in the text.

If there are spatial correlations across registration districts then this could bias the standard errors. As there is typically more than one observation in each registration district it is not possible to use standard spatial regression to test for error correlations between nearby districts. Table A2 reports the specification in column (3) of Table 4 with different standard errors clustered at different levels. In column (1) there is no clustering and the Cook–Weisberg test indicates no heteroscedasticity. In column (2) the standard errors are clustered at the regional level and in column (3) they are clustered at the county level (as in Table 4). The significance levels are similar in column (4) where standard errors are clustered at the district level.

Table A3

TRUNCATED REGRESSIONS

|  |  |  |  |
| --- | --- | --- | --- |
|  | (1) | (2) | (3) |
|  | Not Truncated | Truncated at 5 Feet 3 Ins. | Truncated at 5 Feet 0 Ins. |
| Log district coal intensity | –0.343\*\*\* | –0.340\*\*\* | –0.353\*\*\* |
| (0.116) | (0.113) | (0.113) |
| Log coal intensity in districts to the southwest | –0.116\*\*\* | –0.148\*\*\* | –0.117\*\*\* |
| (0.031) | (0.041) | (0.032) |
| Log coal intensity in districts to the northeast | –0.026 | –0.014 | –0.031 |
| (0.028) | (0.034) | (0.030) |
| F-statistic | 12.67 | 12.85 | 11.96 |
| R-squared | 0.082 | (0.079) | (0.081) |
| Observations | 2,235 | 2,021 | 2,221 |

*\** = Significant at the 10 percent level.

*\*\** = Significant at the 5 percent level.

*\*\*\** = Significant at the 1 percent level.

*Note*: Standard errors, in parentheses, are clustered at the county level. Individual and household variables, district variables (population density, share of white collar workers, population growth, coast, windspeed, rainfall, temperature, ruggedness), and regional dummies are included but not reported. R-squared in parentheses is calculated from the correlation of actual and predicted values.

*Source*: Constructed from data described in the text.

Minimum height requirements are a potential problem for army samples. Using a sample from WW1 somewhat mitigates those concerns. As noted in the text, the British army became increasingly desperate for recruits as the war wore on; medical examinations were perfunctory and the minimum height standard was largely ignored. Table A3 reports truncated regressions to see how potential left side attenuation affects the key coefficients. As a basis for comparison, column (1) is the same regression as column (3) in Table 4 of the paper. Column (2) reports the coefficients from a truncated regression where the lower bound for height is 5 feet 3 inches, which was the original standard at the beginning of the war. Both coal intensity variables remain negative and significant with some increase in the size of the coefficient on coal intensity to the southwest. Given how much the initial standards were relaxed, perhaps a more plausible minimum height would be 5 feet 0 inches (although there are 14 servicemen shorter than this in our data). Column (3) shows that imposing this cut-off makes little difference to the coefficients.