Neuroticism Mediates the Relationship Between Industrial History and Modern-Day Regional Obesity Levels

Michael Daly[1], Martin Obschonka[2], Michael Stuetzer[3,4], Angelina Sutin[5], Leigh Shaw-Taylor[6], Max Satchell[7], Eric Robinson[8]

[1] Department of Psychology, Maynooth University, Maynooth, Co. Kildare, Ireland

[2] QUT Business School, Queensland University of Technology, Brisbane, QLD, Australia

[3] Baden-Wuerttemberg Cooperative State University, Mannheim, Germany

[4] Faculty of Economic Sciences and Media, Ilmenau University of Technology, Germany

[5] Department of Behavioral Sciences and Social Medicine, Florida State University College of Medicine Tallahassee, FL USA

[6] Cambridge Group for the History of Population and Social Structure, Faculty of History, University of Cambridge, United Kingdom

[7] Department of Geography, University of Cambridge, Cambridge, United Kingdom

[8] Department of Psychological Sciences, Eleanor Rathbone Building, University of Liverpool, Liverpool, UK

Corresponding author details:
Dr. Michael Daly
Room 1.1.7 Education House
Maynooth University, Maynooth
Co. Kildare
Email michael.a.daly@mu.ie
Phone +353 1 474 7742

Abstract

Objective: The historical factors and contemporary mechanisms underlying geographical inequalities in obesity levels remain uncertain. In this study we examine whether modern regional variation in obesity is partly a result of the impact of large-scale industry on the personality traits of those living in regions once at the center of the Industrial Revolution.

Method: Exposure to the effects of the Industrial Revolution was assessed using unique historical data from English/Welsh counties (N=111). Specifically, we examined the relationship between the regional employment share in large-scale coal-based industries in 1813-1820 and contemporary regional obesity levels (2013-2015). The Big Five personality traits and regional unemployment levels were examined as potential mediators of this association.

Results: The historical regional employment share in large-scale industries positively predicted the modern-day regional prevalence of obesity. Mediation analysis showed that areas exposed to the decline of large-scale industries experienced elevated neuroticism and unemployment levels that explained almost half of the association between the historical dominance of large-scale industry and modern-day obesity levels.

Conclusions: Our results provide initial evidence that raised regional neuroticism levels may play a key role in explaining why exposure to the rapid growth and subsequent decline of large-scale industries forecasts modern-day obesity levels.

Keywords: *social deprivation; personality; obesity; industrialization*

Introduction

Obesity has become a health pandemic that is alarming because of the wide range of diseases adiposity contributes to and the estimated 4 million deaths per year caused by adiposity (Abarca-Gómez et al., 2017; GBD 2015 Obesity Collaborators, 2017). The obesity pandemic is also alarming because it has emerged very recently and has spread at a rapid pace. In countries like the US and UK, the prevalence of obesity was relatively low in the 1980s, but by the early 21st century obesity prevalence had more than doubled (Ogden, Yanovski, Carroll, & Flegal, 2007; Rennie & Jebb, 2005). Although country level prevalence statistics of obesity highlight the scale of the obesity pandemic, these statistics hide a striking feature; there are prominent geographical and regional differences in obesity prevalence even within individual countries (Bethell, Simpson, Stumbo, Carle, & Gombojav, 2010; Bixby et al., 2019).

In the present research we propose that vulnerability to the modern day obesity pandemic is in part explained by the spatial clustering of personality traits and adverse economic circumstances set in motion by major societal and technological changes occurring over 150 years ago; the Industrial Revolution of the 18th and early 19th century. The concentration of people into industrial regions during this period led to populated areas characterized by enduring adversity due to poor economic conditions and amplified by the outflux of more resilient individuals (Abdellaoui et al., 2019; Beatty, Fothergill, & Powell, 2007; Kemper et al., 2019; Obschonka et al., 2018). As such, the long-term effects of industrialization on the economic conditions and personality of local populations may have exacerbated the growth of obesity in contemporary society.

Obesity is a multi-faceted health problem characterised by a range of social, psychological and genetic determinants (Hu, 2008). Although genetics explain variation in obesity proneness between individuals, there is little evidence that genetic variation can

explain substantial regional differences in obesity (Amador et al., 2017). In contrast, economic adversity is likely to play a key role in explaining regional variability in health conditions like obesity (McLaren, 2007). People from lower socioeconomic status (SES) are at particularly high risk of developing obesity and this is likely to be due to the lack of physical and psychological resources (e.g. the time, money and knowledge to eat healthily), unhealthy environments (e.g. opportunities to eat healthily) and psychosocial adversity associated with living in poverty (Drewnowski & Specter, 2004; Monteiro, Moura, Conde, & Popkin, 2004).

Personality traits are enduring patterns of thoughts, feelings, and behavior, that may play an important role in explaining regional inequalities in health. Like many traits personality varies between geographic regions (Rentfrow, Jokela, & Lamb, 2015) and aspects of personality can also be protective against, or increase one's risk of developing obesity. Higher conscientiousness is protective against obesity and this is likely to be due to the greater self-discipline and ability to plan and stick to long-term plans that characterizes conscientious individuals (Gerlach, Herpertz, & Loeber, 2015). Conversely, higher neuroticism is associated with an increased risk of developing obesity (Gerlach et al., 2015; Heaven et al., 2001; Sutin, Ferrucci, Zonderman, & Terracciano, 2011). This may be due to the effect of impulsivity (a facet of neuroticism) on the intake of tempting and inexpensive foods and the uptake of sedentary leisure activities in contemporary society (Terracciano et al., 2009). Further, negative emotions are often coped with in maladaptive ways (e.g. comfort eating, disordered eating) by those high in trait neuroticism (Heaven et al., 2001).

Therefore, historical events that have had enduring effects on the socioeconomic conditions, and on personality characteristics of large concentrations of people via selective migration and local socialization effects (see Rentfrow, Gosling, & Potter's, 2008, theory on the emergence, persistence, and manifestations of geographical psychological variation), may

be key to understanding previously unexplained geographic disparities in obesity. While a growing number of studies already indicate that such macro-psychological imprinting effects of major historical events do exist (e.g., Abdellaoui et al., 2019; Fritsch et al., 2020; Obschonka et al., 2018), here we aim to empirically estimate how the rise and decline of large-scale industries in regions once at the centre of the Industrial Revolution created vulnerability to the emergence of obesity in old industrial regions, with a special focus on regional distributions of personality traits and economic circumstances.

The Industrial Revolution of the 18th and 19th centuries began in Great Britain and was a major social and technological turning point in the history of humankind. Central to the rise of large-scale industry and mass production was the rise in importance of coal energy. In particular, the steam age period of the Industrial Revolution (1780s-1914) was characterized by rapid growth, with thousands of factories and industries powered by coal (Mathias, 2013). Due to geographic variability in the availability of coal and costs of transporting this natural resource, regions which were close to coalfields became centres of the Industrial Revolution (Crafts & Mulatu, 2005). As a result, both employment prospects and the division of labour in these regions changed, with a large fraction of the workforce providing manual labour in factories (Mathias, 2013). However, as technology and industry advanced and coal reserves diminished in the post First World War period, these regions that were once the heart of the Industrial Revolution experienced downturn, low economic growth, and mass unemployment (Müller, Finka, & Lintz, 2006).

In addition, the lasting economic problems experienced within old industrial regions until today have been inspiring a wave of selective migration (Beatty et al., 2007) with more resilient people (e.g. low in neuroticism) seeking alternative economic opportunities elsewhere. Research indeed found, for example, that people that grew up in old coal regions and then moved out of these regions score lower in neuroticism than people that stayed in old

coal regions (Obschonka et al., 2018). Abdellaoui et al., (2019) replicated this finding with genetic data (genetic makers of neuroticism). The ongoing economic hardship experienced by many people in the old coal regions until today (local socialization effect, Rentfrow et al., 2008) and such selective migration patterns with an outflux of more resilient people may have been amplifying the clustering of psychosocial adversity in industrial regions in recent decades (Kemper et al., 2019; Obschonka et al., 2018). As such, the adverse economic and psychological consequences of the Industrial Revolution occurring through socialization and historical migration processes may have produced a local vulnerability to obesity.

To test this idea, we draw on present day personality data and historical data from the birth place of the Industrial Revolution, Great Britain. We examine whether the present day obesity prevalence can be traced back to this major technological turning point in the history of humankind. Whilst few present-day participants within our study will have worked in large-scale coal-based industries (e.g. coal mining, iron mill), we predicted that the long-term regional adversity created following de-industrialization would be economically and psychologically scarring and rife for the exacerbation of obesity levels.

Method

Datasets

To examine the association between large-scale industry, personality, and obesity across regions we draw on data from three datasets: (i) $N = 381,916$ adults from England and Wales who took part in the British Broadcasting Corporation (BBC) Lab U.K project from 2009 to 2011, (ii) $N = 446,958$ adults from England who completed the Sport England Active People Survey (2013-2015), and (iii) $N = 25,560$ adults from Wales who took part in either the 2013-2104 or 2014-2015 waves of the Welsh Health Survey.

Personality. Between November 2009 and April 2011 participants in the BBC Lab U.K Project completed the *"BBC Big Personality Test"* an Internet-based survey designed to

measure personality in the UK population. Those who took part were informed that the survey aimed to assess personality and that by clicking on the link to continue to the survey they were giving their consent to participate. Participants reported the country and postcode of the place where they lived at the time of completing the survey. This allowed the county where the participant currently resided to be identified. In total 111 counties of England and Wales were included in our regional comparisons. We restricted the sample to England and Wales because industry structure data from 1813-1820 was not available for Scotland. Prior research has shown that regional demographic profiles derived from the BBC personality data are broadly aligned with those from the 2011 Census, as indicated by strong region-level correlations ($r = .79$ to $.95$) between population size and age and ethnicity levels (Rentfrow et al., 2015).

Personality was assessed at the individual-level using the Big Five Inventory (BFI; (John & Srivastava, 1999). Participants indicated the extent to which they agreed with 44 short statements using a 5-point Likert scale ranging from 1(*disagree strongly*) to 5(*agree strongly*). The BFI trait scales all showed satisfactory internal reliability (*Conscientiousness*, $α = .78$; *Neuroticism*, $α = .78$; *Openness*, $α = .73$, *Extraversion*, $α = .84$; *Agreeableness,* $α = .71$). Individual-level trait scores were matched and aggregated to the county level using postcode information detailing the participant's current residence.

Obesity. The dependent variable in our analyses was obesity prevalence – participants with a Body Mass Index (BMI) of 30.0 kg/m^2 or greater were categorized as having obesity.

To assess obesity levels at the county-level in England we drew on data collected from mid-January 2013 to mid-January 2016 as part of a large national survey of sports participation and active recreation– the Active People Survey (APS) (Reilly, 2012). The APS is a telephone survey that includes a random sample of participants from each local authority in England (mean sample size $= 1,371$; median $= 1368$) contacted using Random Digit

Dialling. The weighted survey results are designed to be representative of the population at local authority level. Participants reported their height and weight and this information was used to produce individual obesity estimates that were aggregated to the local authority and county level. Public Health England used objective BMI data from the Health Survey for England to account for measurement error in BMI levels based on self-reported data. After applying height and weight adjustment formulas and survey weights the APS dataset allows robust estimates of obesity prevalence to be made at the county level.

Obesity in Welsh counties was assessed using the Welsh Health Survey, a population-based survey of individuals in private households sampled from a randomly-selected sample of addresses from the Post Office's Postcode Address File in Wales. The sample was stratified at the county-level. Obesity data was derived from 25,560 individual height and weight reports recorded as part of the 2013-2014 and 2014-2015 waves and aggregated to the county-level.

In additional analyses we examined obesity levels based on responses from a subset of BBC Lab U.K. participants (N = 181,622) who recorded their current weight and height. We calculated body mass index (BMI) at the time of survey and removed a small set of extreme values (< 0.2%) outside the typically observed BMI range (10 – 80 kg/m^2) leaving 181,324 individuals in the analytic sample. The county sample sizes ranged from 92 in Merthyr Tydfil County Borough in Wales to 25,749 in Greater London (mean sample size = 1503; median = 765).

Industry Structure and Coalfields

Employment share in large-scale industry. Our primary independent variable is the regional male employment share in large-scale coal-based industries from 1813 to 1820. Our employment data was derived from occupational records from the Church of England baptism registries that have been collated to produce what has been described as the earliest

comprehensive "Census of adult male employment" in the UK (Kitson et al., 2012). Coal-based industries comprised those employed in coal mining and industries relying heavily on coal to power steam engines for textile, brick and pottery, and metal manufacturing. Details on how occupations were summarized into industries, and how industries were classified as 'large scale' and 'coal-based' using historic records of plant size and industry steam use can be found in Obschonka et al. (2018). On average the employment share in coal-based industries was 8.7% (SD = 12.4) in England and Wales over the period from 1813 to 1820. There was substantial regional variation with 24 regions having less than a 1% share of employment in coal-based industries and 14 regions having an employment share of over 20% in these industries.

Coalfields. We computed the minimum distance of each county to a coalfield by GIS using a digitized map of coalfields before 1700 (Hatcher, 1993).

Historical Confounds

We adjusted our analyses for a set of historical control variables that could explain the localization of large-scale industries and/or contribute to regional differences in obesity levels.

Energy supply. Prior to the advent of the steam age, watermills were the most important large-scale source of power. Watermills contributed energy supply to the initial stages of industrialisation and alongside coalfields likely shaped the establishment of large-scale industries. As such, to isolate the effect of coalfields we adjusted for the number of watermills in British regions around 1800, using data from Kanefsky (1979) who derives this information from detailed historical county maps showing the location of watermills. The number of watermills showed marked variation across counties (M = 65.46, SD = 81.02, Min. = 1, Max. = 454) and we used log-transformation to normalize this variable for subsequent analysis.

Market potential. Lage-scale industries require demand for the products they produce which is facilitated by access to markets and customers. We therefore adjusted for the *market potential* of regions which may contribute to both where large-scale industries are concentrated and the future health of the population of those regions. The market potential of each county was assessed by producing a distance-weighted sum of the employment levels in all other regions. We assume that regions close to many areas of high employment have greater access to a large customer base and strong market potential.

Geology and climate. The pre-industrialization wealth of regions depended substantially on land productivity which we proxy using three variables: *soil quality limiting agricultural use* (e.g. gravelly, lithic, or sordic soil), *soil depth to rocks* (larger values indicating greater suitability for agriculture), and *climate* as gauged by *mean July temperature* (higher levels indicate richer harvests) using 1960-1990 as the reference period. In addition, we include an indicator of *terrain ruggedness* (distance between the maximum and minimum elevation in the county) which may curtail regional wealth by hampering agricultural use and trade with other regions. Soil data was taken from the European Soil Project (Panagos, Van Liedekerke, Jones, & Montanarella, 2012).

Population density. Regional variation in population density is highly stable over time ($r = .68$ between 1811 and 2011) and is associated with a range of structural factors (e.g. availability of workers and transport infrastructure). Further, recent evidence suggests that the global growth in obesity over the past three decades occurred to the largest extent in low density rural areas (Bixby et al., 2019). As such, population density may contribute to both the establishment of heavy industry and obesity levels across UK regions. The population density level variable was non-normally distributed as density levels were several times mean levels in small set of highly populated areas such as Greater London and Portsmouth. We therefore used log-transformation to normalize this variable for subsequent analysis.

Statistical Analysis

To assist in the interpretation of the regression coefficients the distance to coalfields, employment share in large-scale industry, and obesity prevalence outcome variables were log transformed. As such, regression coefficients relating to these variables can be interpreted in terms of percent change (e.g. percent change in employment in large industry when distance to coalfields increases by one percent).

Ordinary Least Squares (OLS) model. First, we specified an OLS regression where regional obesity levels in England and Wales in 2013-2015 were predicted by the presence of large-scale industries during the Industrial Revolution. To rule out alternative explanations we adjusted for a set of regional control variables (described above) assessing historic energy supply, market potential, geology and climate, population density, and a country indicator.

Instrumental Variable (IV) approach. Next, we aimed to account for potential endogeneity using an IV strategy. This approach represents a key methodology that enables researchers to rule out reverse causality and account for the possibility that omitted variables may bias the association of interest, in our case the relationship between large-scale industries and regional obesity levels.

In 18th and 19th century Great Britain the availability of cheaply transported coal near mining areas was a key determinant of growth in large-scale industries that relied on coal for steam power (Stuetzer et al., 2016). We therefore used the distance of a county to the nearest coalfield as a source of exogenous variation that can explain the localization of large-scale coal-based industries. Within the IV approach a credible exogenous instrument (distance to nearest coalfield) must be correlated with the instrumented endogenous variable (employment share in large-scale industries from 1813-1820) but not associated with the dependent variable beyond its correlation with the explanatory variable. As such the spatial proximity of

11

a region to coalfields must predict body mass and obesity levels only by influencing the employment share in large-scale industries.

In support of the validity of our instrument the distance of a county to a coalfield was negatively associated with the regional employment share in large-scale industry ($r(109) = -0.77$, $P < .001$) and predictive of obesity levels ($b = -0.024$, $SE = 0.008$, $P = 0.002$) in a model that adjusted for historical confounding variables. Further, distance to coalfield was no longer related to regional obesity levels after controlling for large-scale industry employment share ($b = 0.002$, $SE = 0.011$, $P = 0.88$). In terms of effect size this relationship was reduced from a meaningful negative relationship ($\beta = -0.34$) to a small non-significant positive link ($\beta = 0.02$). These analyses provided initial support for the idea that the spatial proximity of coalfields may have driven the emergence and persistence of coal-based industries but does not directly influence contemporary obesity levels.

Our formal IV analysis proceeded as follows. In the first-stage we predicted the employment share in large-scale industries explained by the distance to a coalfield variable (*Largescale$_{pred}$*). Variation in *Largescale$_{pred}$* is accounted for by the instrument (distance to coalfield) and can be used to rule out endogeneity when predicting contemporary obesity levels in the second-stage regression. We also adjusted for historic control variables that may shape the localization of coal-based industries. We estimated Huber-White robust standard errors in all IV-regressions to avoid biased standard errors due to unspecified nonconstant error variance.

Personality traits and economic hardship as mediators. An association between the historical concentration of large-scale industries and contemporary regional differences obesity levels may be due to the lasting economic legacy of the decline in these once thriving industries. To account for this possibility, we considered the potential mediating effect of economic hardship in UK regions as proxied by the regional unemployment rate in 2001

taken from the 2001 Census. In addition, we examined the regional growth in employment from 1931 to 2011 to capture the long-term performance of the regional economy. Employment data were collated for this project from the 2011 UK Census and the digitized 1931 British Census produced by the Great Britain Historical GIS Project. Alongside the economic consequences of the Industrial Revolution we anticipated that a legacy of psychological adversity expressed in regional differences in personality trait differences in neuroticism and conscientiousness may mediate between the employment share in large-scale industry and today's obesity levels.

We selected candidate variables for mediation based on two initial criteria: the variable must be robustly predicted by the independent variable, employment share in large-scale industry, and the variable must account for variation in the dependent variable, obesity levels. To estimate the magnitude and statistical significance of the mediating effects we conducted formal mediation analyses using a non-parametric bootstrapping approach with 95% bias-corrected bootstrapped confidence intervals estimated using 10,000 bootstrap samples to test whether indirect effects were statistically different from zero.

To tackle endogeneity we repeated our mediation analyses using $Largescale_{pred}$ as the independent variable for the prediction of regional differences in obesity prevalence. This supplementary analysis aimed to identify if the same pattern of mediation results was observed when the indirect effects estimated relied exclusively on the portion of variation in the historical employment share in large-scale industries predicted by spatial proximity to coalfield ($Largescale_{pred}$).

Finally, given that prior research suggests that economic hardship may be a channel through which large-scale industry generated regional differences in psychological adversity (Obschonka et al., 2018) we tested an additional serial mediation analysis linking heavy industry to obesity via economic conditions and personality traits (see Figure 1).

Recent selective migration. Drawing on migration data from the BBC Lab U.K. Project we first examined differences in the obesity levels between those who remained in old industrial regions (N = 36,859), those who left to live in non-industrial regions (N = 11,639), and those who migrated from non-industrial to industrial regions (N = 8,844). We defined industrial regions as counties where the employment share in large-scale industry in 1813-1820 was above average (i.e. 8.7%). We anticipated that those who left industrial regions may be less vulnerable to obesity than those who remained.

Of the 181,324 participants with complete obesity data in the sample we identified 97,114 individuals who continued to live in the region where they grew up. We excluded migrants and repeated our main analyses to test whether the association between large-scale coal-based industry and obesity levels was observed among those continuing to reside in the region where they grew up.

Results

Large-Scale Industry and Obesity Prevalence

In 2013-2015 the average prevalence of obesity was 24.7% (SD = 3.4) and obesity levels showed marked variation across regions ranging from a minimum of 17.6% (Ceredigion, Wales) to a maximum of 32.4% (Halton, England). The average regional employment share in coal-based industries was 8.7% (SD = 12.4) in England and Wales over the period from 1813 to 1820. Descriptive statistics for the remainder of our sample characteristics are displayed in Table S1 (Supplementary Materials) and predicted levels of obesity and employment in large-scale industry in each region (adjusted for historical confounding variables) are mapped in Figure S1.

In line with our predictions, the historical regional employment share in large-scale coal-based industries (1813-1820) was positively associated with contemporary regional

differences in obesity prevalence ($b = 0.062$, $P < 0.001$) in an OLS model that adjusted for confounding factors, as shown in Table 1. Importantly, the association between large-scale industry and obesity levels was meaningful in magnitude ($\beta = 0.46$) and because of the rich historical data we have assembled we could show that this association was independent of key historical economic conditions and population characteristics that could affect the relationship between large-scale industry and body weight (historical energy supply, market potential, geology, climate, population density). The OLS regression results suggested that a 10% increase in the regional employment share in large-scale industries from 1813-1820 predicted a 0.6% increase in the prevalence of obesity in 2013-2015. In the full regression model, our pre-1820 historical variables could explain 33.7% of the contemporary regional variance in obesity levels in England and Wales. These findings were replicated using the regional obesity measure derived from the BBC Lab U.K. Project data, as shown in Table S2 in the Supplementary Materials.

We were able to address endogeneity concerns by using an instrumental variable (IV) estimation strategy to assess the contribution of large-scale coal-based industries to obesity levels (see Figure S1 for location of coal-fields). Our first-stage regressions (first stage F-statistic of instrument = 118.69) showed that the employment share in large-scale coal-based industries was negatively related to the distance to the nearest coalfield in an unadjusted model ($b = -0.399$, $P < 0.001$) and a model that adjusted for alternative historical explanations for the localization of large-scale industries ($b = -0.403$, $P < 0.001$) (Table 2, first stage). This regression coefficient indicated that a 10% increase in the distance of a region to a coalfield was associated with a 4% reduction in employment in large-scale industries

The results of the second stage IV regressions relied exclusively on the portion of variation in the employment share in large-scale coal-based industries that can be explained

by the distance of the region to coalfields (*Largescale*_{pred}). Consistent with our OLS estimates our IV results suggest that a larger regional share of employment in coal-based industries can be attributed to the proximity of the region to coalfields and this contributes to subsequent regional differences in the prevalence of obesity ($b = 0.061$, $P < 0.001$) (Table 2, second stage). A 10% increase in the employment share in large-scale industries from 1813-1820 from average levels predicted a 0.6% increase in the prevalence of obesity. We replicated these findings closely using regional obesity estimates from the BBC Lab UK Project Data as can be seen in Table S3 (Supplementary Materials).

Economic Hardship and Personality Traits as Mediators

First, we sought to identify initial evidence for mediation by evaluating whether our independent variable (employment share in large-scale industries) predicted the potential mediating variables and whether these in turn were related to obesity levels. Regions where there was large share of employment in coal-based industries (1813-1820) experienced subsequent economic adversity as indexed by low levels of regional growth in employment from 1931-2011 and a high unemployment rate in 2001 (see Table S4). However, only the contemporary regional unemployment rate was predictive of (raised) contemporary obesity levels, as shown in Table S5 (model 2).

Regions where there was a historical concentration of large-scale industry were associated with high levels of neuroticism (Table S4 in the Supplementary Materials) which in turn was related to elevated obesity prevalence (Table S5, model 4). A large regional employment share in coal-based industry during the Industrial Revolution also forecasted low levels of conscientiousness (Tables S4). However, conscientiousness was not significantly associated with regional obesity prevalence. Neither extraversion, agreeableness, nor openness to experience were related to our historic indicator of employment in large-scale

industries. This pattern of results was replicated when examining obesity levels in the BBC Lab UK Project dataset as the outcome variable (Table S6).

As such, our initial analyses pointed to regional differences in the unemployment rate and levels of neuroticism as potential channels through which large-scale industry may influence body weight. We employed an IV strategy to ensure that the localization of coal-based industries robustly predicted the potential mediating variables. Our IV estimates indicated that the historical concentration of heavy industry predicted contemporary regional differences in unemployment ($b = 0.699$, $P < 0.001$) and neuroticism ($b = 0.012$, $P = 0.012$), (see Table S7).

To summarize, regions once central to the Industrial Revolution experienced elevated unemployment and neuroticism levels in recent years which in turn predicted raised obesity levels. To estimate the magnitude and relative importance of these channels we ran parallel mediation analyses examining the impact of employment share in large-scale industries on obesity prevalence via unemployment and neuroticism. We found a statistically significant indirect effect between employment in large-scale industry and obesity levels through our indicator of economic hardship, regional differences in unemployment in 2001 ($b = 0.017$, 95% CI = 0.006–0.029), as shown in Table 3 (upper panel). In addition, we observed significant indirect effects of large-scale industry on obesity prevalence through regional neuroticism levels ($b = 0.014$, 95% CI = 0.003–0.029). The total indirect effect of unemployment and neuroticism explained 48.9% of the association between large-scale industry employment share and contemporary obesity (total indirect effect: $b = 0.030$, 95% CI = 0.016–0.048). A significant direct effect of large-scale industry on obesity (direct effect: $b = 0.032$, 95% CI = 0.009–0.055) remained after adjustment for contemporary unemployment and neuroticism levels. Our mediation analyses therefore provided evidence that regions where coal-based industries dominated during the Industrial Revolution went on

to experience economic adversity and raised neuroticism that translated into today's raised obesity levels. Together our regional variables including the level of employment in large-scale coal-based industries, other historical variables, and mediating factors explained 55% of the contemporary regional variation in obesity levels in England and Wales.

We also tested a serial mediation model to identify whether neuroticism may act as a channel through which the economic hardship that followed the decline of large-scale industry impacts regional differences in obesity prevalence (see Figure 1). This analysis indicated a significant indirect effect of large-scale industry on obesity via unemployment and neuroticism ($b = 0.006$, 95% CI = 0.002–0.012), as shown in Table 3 (lower panel). These additional analyses indicated that the data are consistent with the idea that regions characterized by a high concentration of large-scale industry due to their proximity to coalfields suffered greater unemployment which in turn was associated with developing increased levels of neuroticism, that ultimately predicted raised levels of body weight in today's adults. Our parallel and serial mediation findings were replicated using regional obesity data from the BBC Lab U.K. Project data (see Table S8).

In our final set of path analyses we repeated our mediation analyses with only the variation in the employment share in large-scale industry that can be attributed exclusively to the distance of the region to coalfields ($Largescale_{pred}$) as the independent variable. Our parallel mediation model identified indirect effects of $Largescale_{pred}$ on regional obesity prevalence through unemployment and neuroticism (Table S9). Similarly, our serial mediation model identified robust evidence for an indirect effect of $Largescale_{pred}$ on obesity levels via both unemployment and neuroticism, as highlighted in Figure 1.

Additional Mediation Analysis

We sought to further probe the role of regional unemployment levels in explaining the relationship between the employment share in large-scale industry and regional obesity

levels. Specifically, we tested whether regional economic conditions prior to and after the rapid decline in UK manufacturing employment from the 1970s (ONS, 2019) were of greater importance in explaining the link between industrial heritage and regional obesity levels. This analysis revealed an indirect effect of large-scale industry on obesity prevalence through both regional differences in unemployment in 1971 ($b = 0.013$, 95% CI $= 0.002–0.027$) and the change in unemployment levels between 1971 and 2011 ($b = 0.016$, 95% CI $= 0.006–0.030$), as shown in Table S10. In addition, we observed a significant direct effect of large-scale industry on obesity prevalence after adjustment for both the 1971 regional unemployment rate and the change in unemployment from 1971 to 2011 ($b = 0.034$, 95% CI $= 0.011–0.056$). This analysis suggested that two things. First, the economic circumstances in 1971 before the massive deindustrialization can forecast contemporary obesity prevalence which should be expected as unemployment due to any reason is likely detrimental. Second, the economic impact of the dramatic period of deindustrialization that followed also played a significant role in explaining why large-scale industrial history forecasts regional contemporary obesity prevalence. However, it is important to note that this pathway only partly explains this association, accounting for 47% of the link.

Adjustment for Contemporary Regional Characteristics

We also conducted an additional analysis to test whether the relationship between large-scale industry and regional obesity levels identified in the current study occurs independently of contemporary characteristics of regional populations. Specifically, we included regional information from 2011 on the following variables as controls in our main analysis: (i) population size in each region, (ii) percentage share of regional population in urban areas, (iii) percentage share of the regional population that are male, and (iv) population density in the region. Descriptive statistics for each of these characteristics are displayed in Table S1.

Adjustment for these factors did not markedly attenuate our OLS estimate of the relationship between large-scale industry and regional obesity levels as shown in Table S11.

Recent Selective Migration Patterns

An important caveat of our previous analyses is the omission of migration. During the Industrial Revolution the industrializing regions were centers of in-migration but especially since the 1970s these regions have been centers of out-migration. While we cannot analyse the impact of the former migration pattern, we are able to shed light on the role of the latter. Logistic regression analyses showed that those who left old industrial regions were less likely to be classified as obese than those who stayed ($OR = 0.895$, 95% CI: 0.844-0.949, $P < 0.001$). Adjustment for age polynomials (i.e. age, age-squared, age-cubed) showed that this relationship did not appear to be due to a correlated pattern of change in migration levels and obesity across the lifespan ($OR = 0.726$, 95% CI: 0.684-0.771, $P < 0.001$). We did not find robust evidence that those who migrated to live in industrial regions were at differential risk of obesity. These analyses suggested that a pattern of selective out-migration among those with a reduced risk of obesity may explain part of the association between the historical employment share in large-scale industries and contemporary obesity levels.

Next, drawing on migration data from the BBC Lab U.K. dataset we examined only those who continue to live in the same region where they grew up (i.e. omitting migrants from our models). As in our main analyses, the historical regional employment share in large-scale industries positively predicted obesity prevalence of non-migrants in both OLS ($b = 0.089$, $SE = 0.020$, $P < 0.001$) and IV regression models ($b = 0.118$, $SE = 0.024$, $P < 0.001$). Importantly, we saw no decline in the strength of the relationship between coal-based industry and obesity levels in this analysis. This finding indicating that recent selective

migration (the reduced risk of obesity among those who left old industrial regions) may not be a key driver of the association between industrialization and obesity prevalence.

Discussion

The industrialized world has seen dramatic mass population weight gain for over thirty years and the emergence of stark differences in obesity prevalence between regions within individual countries (Bethell et al., 2010). Our study aimed to trace the potential roots of the geographical distribution of the obesity epidemic from the era of coal-based industrialization in the 18th and 19th centuries through to the present day. We also sought to test whether the clustering of personality traits associated with economic adversity in regions affected by the rise and fall of coal-based industries could explain the striking regional differences in obesity seen today. To do this, we compiled a unique fine-grained spatial dataset of geological and historical conditions in British regions alongside high-quality indicators of modern-day economic circumstances and personality traits and regional variation in body weight.

Based on these data, we found that people living in areas where a large share of employment was previously in coal-based industries experienced higher levels of obesity over 150 years after the surge of industrialization, as well as raised neuroticism and adverse economic conditions. Furthermore, it was possible to disentangle the potential effect of industrial heritage from a set of alternative explanatory factors (e.g. population density, historical energy supply, climate and soil quality), to identify a robust association with industrialization. In addition, we made use of exogenous variation (in the spatial proximity of coalfields) to apply a more causal approach (instrumental variable estimation) to provide further evidence that the historical concentration of large-scale industries associated with local coalfields robustly forecasted modern-day obesity levels.

We also studied how the economic consequences of the decline of large-scale

21

industries may have promoted modern day obesity. The Industrial Revolution resulted in mass migration and a concentration of workers providing a restricted set of highly specialized and often low-skilled tasks. Although beneficial to the growth of industry at the time of the Industrial Revolution, this will have limited employment and entrepreneurship opportunities available following the decline of large industries (Stuetzer et al., 2016). We uncovered evidence that the persistent economic struggles experienced by those living in old industrial regions, assessed using the regional unemployment rate, in part explained why large-scale industries contributed to the development of regional disparities in obesity. Findings that are in line with contemporary research linking lower SES to increased obesity risk (McLaren, 2007).

A striking finding in the present research was the presence of raised trait neuroticism levels also played a role in explaining why obesity is more prevalent in old industrial regions. Consistent with this, US research has linked higher state levels of neuroticism to increased obesity prevalence (McCann, 2011). Here we showed that those living in ex-industrial regions were more likely to report high levels of neuroticism, which in turn predicted higher levels of obesity. Building on these findings we also identified empirical support for a chronologically ordered path model (Figure 1) which implicated the economic repercussions of industrialization in the development of neuroticism ultimately predicting increased obesity among adults today. These findings lend support to recent suggestions that a key reason why living in relative poverty is associated with increased risk of obesity is not simply due to environmental or economic conditions, but the distress associated with poverty (Spinosa et al., 2019). However, we also have to note that selective migration patterns play an important role as well. For example, the increased levels of neuroticism in regions with higher levels of economic hardship (e.g., unemployment) may have been amplified by selective out-migration

from these region (e.g., people disposed towards lower levels of neuroticism may have left the region; Abdellaoui et al., 2019; Obschonka et al., 2018).

There are limitations to our approach. First, there are a complex set of potential channels that we were unable to measure but may stem from the Industrial Revolution and economic downturn and would likely promote obesity, such as job strain and harsh working conditions (Brunner, Chandola, & Marmot, 2007), rates of violent crime (Theall, Chaparro, Denstel, Bilfield, & Drury, 2019), lack of green space (Lachowycz & Jones, 2011) and air pollution (An, Ji, Yan, & Guan, 2018). Indeed, even after accounting for economic circumstances and neuroticism levels, large-scale industry still showed a pronounced direct relationship with obesity that may be explained by such unobserved factors. Thus, there is room for future research investigating the impact of other channels. Such research could also shed light on the robustness of the role of neuroticism as a mediator of the association between regional unemployment levels and the contemporary prevalence of obesity. For example, it is possible that neuroticism may act as a proxy for correlated factors such as regional differences in deprivation, environmental characteristics, or physical health not assessed in the current study (Rentfrow et al., 2015).

The role of physical activity and diet will be important to consider in future research, both as independent mediators and as potential downstream consequences of raised neuroticism levels. For example, work in factories was often physical and demanding (in terms of energy use), so it is plausible that cultural norms in old industrial regions favor energy dense foods – a cultural norm that is no longer adaptive in modern day and instead promotes weight gain (Rouhani, Haghighatdoost, Surkan, & Azadbakht, 2016). Likewise, recent research has linked the psychological experience of economic adversity to increased energy intake, as resource scarcity appears to promote food choice strategies that favor

maximising calories (Cheon & Hong, 2017). Therefore, the food and eating cultures in old industrial regions may play an important role in explaining raised present-day obesity levels.

Alongside socialization processes resulting from persistent economic hardship following the Industrial Revolution it is likely that selective migration magnified the vulnerability of those living in coal-based industrial regions to obesity. Local populations in old industrial regions experienced an outflow of people with a reduced susceptibility to psychological adversity (Obschonka et al., 2018) and we find this is also the case for obesity. Therefore, selective migration among those at reduced risk of adversity and obesity may partly explain why industrialization forecasts obesity risk. Yet, when we restricted our analyses to those remaining in regions where they grew up (i.e. omitting migrants) we found a robust relationship between the presence of large-scale industry and obesity levels, indicating that selective migration may not drive our results. We also do not know whether the present findings would extend to other countries that have experienced the 'boom' and economic downturn of major industries (e.g. United States) and future work would benefit from examining the generalisability of our findings.

While industrialization brought major progress to humanity beginning over 150 years ago, regions at its center continue to suffer relative adversity today. In particular, we identify a large-scale personality-related vulnerability to the emergence of obesity that originates in the first industrial age once concentrated in these regions. Coupled with the modern shift towards obesogenic environments this vulnerability may have exacerbated the obesity epidemic, placing major costs on today's society and populations. It is therefore important that attempts to address the obesity crisis consider the wide-scale effects of industrialization and tackle underlying mechanisms, including socioeconomic deprivation and distress. Policy changes and interventions that address poverty (Chetty, Hendren, & Katz, 2016), target the obesogenic environment (Swinburn et al., 2011), or promote psychological resilience on a

large scale (e.g. reducing the negative consequences of experiencing neuroticism) in old industrial regions may diminish the enduring impact of industrialization on modern-day obesity levels.

References

Abarca-Gómez, L., Abdeen, Z. A., Hamid, Z. A., Abu-Rmeileh, N. M., Acosta-Cazares, B., Acuin, C., . . . Aguilar-Salinas, C. A. (2017). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128· 9 million children, adolescents, and adults. *The Lancet, 390*, 2627-2642.

Abdellaoui, A., D. Hugh-Jones, L. Yengo, K.E. Kemper, M.G. Nivard, L. Veul, L., Y. Holtz, B.P.Zietsch, T.M. Frayling, N.T. Wray J. Yang K.J.H. Verweij and P.M.Visscher (2019): Genetic correlates of social stratification in Great Britain. *Nature Human Behaviour, 3*, 1332-1342.

Amador, C., Xia, C., Nagy, R., Campbell, A., Porteous, D., Smith, B. H., . . . Haley, C. S. (2017). Regional variation in health is predominantly driven by lifestyle rather than genetics. *Nature Communications, 8*, 801. doi:10.1038/s41467-017-00497-5

An, R., Ji, M., Yan, H., & Guan, C. (2018). Impact of ambient air pollution on obesity: a systematic review. *International Journal of Obesity, 42*, 1112-1126.

Beatty, C., Fothergill, S., & Powell, R. (2007). Twenty years on: has the economy of the UK coalfields recovered? *Environment and Planning A, 39*, 1654-1675.

Bethell, C., Simpson, L., Stumbo, S., Carle, A. C., & Gombojav, N. (2010). National, state, and local disparities in childhood obesity. *Health Affairs, 29*, 347-356.

Bixby, H., Bentham, J., Zhou, B., Di Cesare, M., Paciorek, C. J., Bennett, J. E., . . . Carrillo-Larco, R. M. (2019). Rising rural body-mass index is the main driver of the global obesity epidemic in adults. *Nature, 569*, 260-264.

Brunner, E. J., Chandola, T., & Marmot, M. G. (2007). Prospective effect of job strain on general and central obesity in the Whitehall II Study. *American Journal of Epidemiology, 165*, 828-837.

Cheon, B. K., & Hong, Y.-Y. (2017). Mere experience of low subjective socioeconomic status stimulates appetite and food intake. *Proceedings of the National Academy of Sciences, 114*, 72-77.

Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review, 106*, 855-902.

Crafts, N., & Mulatu, A. (2005). What explains the location of industry in Britain, 1871–1931? *Journal of Economic Geography, 5*, 499-518.

Drewnowski, A., & Specter, S. E. (2004). Poverty and obesity: the role of energy density and energy costs. *The American Journal of Clinical Nutrition, 79*, 6-16.

Fritsch, M., Obschonka, M., Wahl, F., & Wyrwich, M. (2020). The deep imprint of Roman sandals: Evidence of long-lasting effects of Roman rule on personality, economic performance, and well-being in Germany. *Hohenheim Discussion Papers in Business, Economics and Social Sciences, No. 05-2020*, Available at http://nbn-resolving.de/urn:nbn:de:bsz:100-opus-17436

GBD 2015 Obesity Collaborators, G. O. (2017). Health effects of overweight and obesity in 195 countries over 25 years. *New England Journal of Medicine, 377*, 13-27.

Gerlach, G., Herpertz, S., & Loeber, S. (2015). Personality traits and obesity: a systematic review. *Obesity Reviews, 16*, 32-63.

Hatcher, J. (1993). The history of the British coal industry: Vol. 1. Before 1700: Towards the age of coal. . In. Oxford, UK: Clarendon Press.

Heaven, P. C., Mulligan, K., Merrilees, R., Woods, T., & Fairooz, Y. (2001). Neuroticism and conscientiousness as predictors of emotional, external, and restrained eating behaviors. *International Journal of Eating Disorders, 30*, 161-166.

Hu, F. (2008). *Obesity epidemiology*: Oxford University Press.

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, & O. P. John (Eds.), Handbook of personality: Theory and research (pp. 102-138). New York: Guilford Press.

Kanefsky, J. W. (1979). *The diffusion of power technology in British industry. 1760-1870.* University of Exeter

Kemper, K.E., Nivard, M.G... Visscher, P.M. (2019) Genetic correlates of social stratification in Great Britain. *Nature Human Behavior.* doi.org/10.1038/s41562-019-0757-5

Kitson, P. M., Shaw-Taylor, L., Wrigley, E. A., Davies, R. S., Newton, G., & Satchell, A. M. (2012). The creation of a 'census' of adult male employment for England and Wales for 1817. *[Online] Retrieved from http://www.geog.cam.ac.uk/research/projects/ occupations/britain19c/ papers/paper4.pdf.*

Lachowycz, K., & Jones, A. P. (2011). Greenspace and obesity: a systematic review of the evidence. *Obesity Reviews, 12*, e183-e189.

Mathias, P. (2013). *The first industrial nation: The economic history of Britain 1700–1914*: Routledge.

McCann, S. J. (2011). Personality and American state differences in obesity prevalence. *The Journal of Psychology, 145*, 419-433.

McLaren, L. (2007). Socioeconomic status and obesity. *Epidemiologic Eeviews, 29*, 29-48.

Monteiro, C. A., Moura, E. C., Conde, W. L., & Popkin, B. M. (2004). Socioeconomic status and obesity in adult populations of developing countries: a review. *Bulletin of the World Health Organization, 82*, 940-946.

Müller, B., Finka, M., & Lintz, G. (2006). *Rise and decline of industry in Central and Eastern Europe: A comparative study of cities and regions in eleven countries*. Berlin, Germany: Springer.

Obschonka, M., Stuetzer, M., Rentfrow, P. J., Shaw-Taylor, L., Satchell, M., Silbereisen, R.

    K., . . . Gosling, S. D. (2018). In the shadow of coal: How large-scale industries

    contributed to present-day regional differences in personality and well-being. *Journal*

    of Personality and Social Psychology, 115, 903-927.

Ogden, C. L., Yanovski, S. Z., Carroll, M. D., & Flegal, K. M. (2007). The epidemiology of

    obesity. *Gastroenterology, 132*, 2087-2102.

ONS (Office of National Statistics). (2019). Changes in the economy since the 1970s.

    Retrived from https://www.ons.gov.uk/economy/economicoutputandproductivity/

    output/articles/changesintheeconomysincethe1970s/2019-09-02

Panagos, P., Van Liedekerke, M., Jones, A., & Montanarella, L. (2012). European Soil Data

    Centre: Response to European policy support and public data requirements. *Land Use*

    Policy, 29, 329-338.

Reilly, D. (2012). Gender, Culture, and Sex-Typed Cognitive Abilities. *PLOS One, 7*,

    e39904. doi:10.1371/journal.pone.0039904

Rennie, K. L., & Jebb, S. A. (2005). Prevalence of obesity in Great Britain. *Obesity Reviews,*

    6, 11-12.

Rentfrow, P. J., Gosling, S. D., & Potter, J. (2008). A theory of the emergence, persistence,

    and expression of geographic variation in psychological characteristics. *Perspectives*

    on Psychological Science, *3*(5), 339-369.

Rentfrow, P. J., Jokela, M., & Lamb, M. E. (2015). Regional personality differences in Great

    Britain. *PLOS One, 10*, e0122245. doi.org/10.1371/journal.pone.0122245

Rouhani, M. H., Haghighatdoost, F., Surkan, P. J., & Azadbakht, L. (2016). Associations

    between dietary energy density and obesity: A systematic review and meta-analysis of

    observational studies. *Nutrition, 32*, 1037-1047.

Spinosa, J., Christiansen, P., Dickson, J. M., Lorenzetti, V., & Hardman, C. A. (2019). From

    scioeconomic disadvantage to obesity: The mediating role of psychological distress

    and emotional eating. *Obesity, 27*, 559-564.

Stuetzer, M., Obschonka, M., Audretsch, D. B., Wyrwich, M., Rentfrow, P. J., Coombes, M.,

    & Satchell, M. (2016). Industry structure, entrepreneurship, and culture: An empirical

    analysis using historical coalfields. *European Economic Review, 86*, 52-72.

Sutin, A. R., Ferrucci, L., Zonderman, A. B., & Terracciano, A. (2011). Personality and

    obesity across the adult life span. *Journal of Personality and Social Psychology, 101*,

    579-592.

Swinburn, B. A., Sacks, G., Hall, K. D., McPherson, K., Finegood, D. T., Moodie, M. L., &

    Gortmaker, S. L. (2011). The global obesity pandemic: shaped by global drivers and

    local environments. *The Lancet, 378*, 804-814.

Terracciano, A., Sutin, A. R., McCrae, R. R., Deiana, B., Ferrucci, L., Schlessinger, D., . . .

    Costa Jr, P. T. (2009). Facets of personality linked to underweight and overweight.

    Psychosomatic Medicine, 71, 682-689.

Theall, K. P., Chaparro, M. P., Denstel, K., Bilfield, A., & Drury, S. S. (2019). Childhood

    obesity and the associated roles of neighborhood and biologic stress. *Preventive*

    Medicine Reports, 14, 100849. doi:https://doi.org/10.1016/j.pmedr.2019.100849

Table 1. OLS Regression Model Examining the Relationship Between the Employment Share in Large-Scale Coal-Based Industries in 1813-1820 and Regional Differences in the Prevalence of Obesity in 2013-2015 (N = 111 Regions in England and Wales).

	b	*SE*	*P*
Employment in large scale industries 1813-1820 (%)	0.062	0.013	< 0.001
Watermills around 1800	-0.022	0.011	0.045
Market potential 1811	0.000	0.000	0.343
Limits to agricultural use	-0.151	0.074	0.045
Depth to rock	0.012	0.014	0.404
Mean July temperature	-0.010	0.019	0.601
Terrain ruggedness	-0.000	0.000	0.247
Population density 1811	-0.053	0.017	0.003
R^2		0.337	

Table 2. Instrumental Variable Regressions Using Distance to Nearest Coalfield as an Instrument for Employment Share in Large-Scale Coal-Based Industries in 1813-1820 (N = 111 Regions in England and Wales).

	First stage			Second stage regression		
	Employment in large industries 1813-1820			Obesity (2013-2015)		
	b	*SE*	*P*	*b*	*SE*	*P*
Distance to coalfield	-0.403	0.035	< 0.001			
Employment share in large industries 1813-1820 (%)				0.061	0.016	< 0.001
Watermills around 1800	-0.065	0.053	0.227	-0.021	0.009	0.017
Market potential 1811	-0.000	0.000	0.711	-0.000	0.000	0.248
Limits to agricultural use	-0.842	0.199	< 0.001	-0.153	0.083	0.065
Depth to rock	-0.061	0.060	0.316	0.012	0.015	0.438
Mean July temperature	-0.157	0.096	0.105	-0.011	0.016	0.497
Terrain ruggedness	-0.000	0.000	0.699	-0.000	0.000	0.140
Population density 1811	0.306	0.093	0.001	-0.052	0.018	0.004
F-statistic of instrument		118.690				
R^2		0.671			0.337	

Table 3. Mediation Models of the Indirect Effect of Employment Share in Large-Scale Coal-Based Industries in 1813-1820 on Obesity Prevalence in 2013-2015 through Unemployment and Neuroticism (N = 111 Regions in England and Wales)

	b	95% CI Lower ; Upper		β^a	Effect ratio
Employment in large industries =>					
Obesity (parallel mediation model)					
Total effect	0.062	0.040	0.085	0.459	
Direct effect	0.032	0.009	0.055	0.236	0.51
Indirect effect (via unemployment[b])	0.017	0.007	0.032	0.122	0.27
Indirect effect (via neuroticism)	0.014	0.004	0.031	0.101	0.22
Total indirect effect	0.030	0.016	0.049	0.223	0.49
Employment in large industries =>					
Obesity (serial mediation model)					
Total effect	0.062			0.459	
Direct effect	0.032	0.009	0.055	0.236	0.51
Indirect effect (via unemployment[b])	0.017	0.006	0.030	0.122	0.27
Indirect effect (via neuroticism)	0.008	-0.002	0.021	0.057	0.13
Indirect effect (via unemployment[b] + neuroticism)	0.006	0.002	0.013	0.044	0.10
Total indirect effect	0.030	0.016	0.048	0.223	0.49

Note: Models were adjusted for historical control variables (energy supply, market potential, geology, climate, population density) and country.

[a] Standardized ordinary least squares regression coefficient.

[b] Regional unemployment rate in 2001.
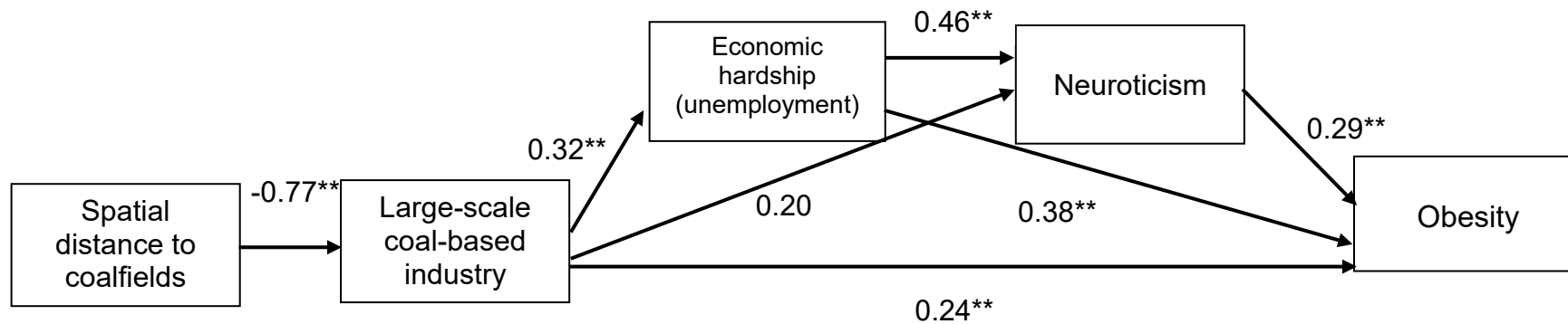
Figure 1. Conceptual Diagram of a Path Model linking the Employment Share in Large-scale Industries (1813-1820) to Contemporary Regional

Differences in Obesity Prevalence (2013-2015) via Unemployment (2001) and Neuroticism (2009-2011). Values are standardized path

coefficients. Model adjusts for historical cofounds. ** $P < 0.01$. * $P < 0.05$.

Supplementary Materials

Table S1. Descriptive Statistics (N = 111 Regions in England and Wales).

	M (SD)	Min. / Max.
Historic and geographic variables		
Distance of region to coalfields (in km) before 1700	35.923 (47.281)	0 / 170.723
Percentage share of male employment in large scale industries 1813-1820	8.709 (12.402)	0 / 65.699
Number of watermills around 1800	65.460 (81.019)	1 / 454
Market potential in 1811[a]	78519 .843 (15913.41)	38347.75 / 124378.9
Limits to agricultural use (%)[b]	2.703 (0.163)	0 /100
Depth to rock[c]	2.378 (1.062)	1 / 4
Mean July temperature (in °C)[d]	15.180 (0.924)	12.32 / 17
Terrain ruggedness[e]	453.650 (259.699)	120 / 1100
Population density in 1811 (per km^2)	118.625 (172.813)	23.179 / 1070.155
Contemporary variables		
Obesity prevalence (%), 2013-2015	24.665 (3.386)	17.578 / 32.438
Obesity prevalence (log %)	3.196 (0.141)	2.867 / 3.479
Unemployment rate 1971 (%)	4.047 (1.306)	2.155 / 7.732
Unemployment rate 2001 (%)	5.249 (1.910)	2.155 / 11.176
Unemployment rate 2011 (%)	6.542 (2.058)	3.542 / 12.898

Conscientiousness, 2009-2011	3.649 (0.05)	3.459 / 3.749
Neuroticism, 2009-2011	2.978 (0.041)	2.852 / 3.105
Openness, 2009-2011	3.654 (0.045)	3.537 / 3.826
Extraversion, 2009-2011	3.234 (0.033)	3.152 / 3.352
Agreeableness, 2009-2011	3.742 (0.023)	3.682 / 3.815
Population size in 2011 (N)	508,465.2 (88,8831.5)	37,369 / 8,222,451
Percentage share of population in urban areas in 2011 (%)	56.430 (45.018)	0 / 100
Percentage share of male population in 2011 (%)	49.215 (0.474)	48.163 / 50.617
Population density in region in 2011 (per km^2)	1132.09 (1397.093)	25.666 / 5230.304

Note: See main text for full description and sources of all variables listed.

[a] Computed as the distance-weighted sum of the employment levels in all other regions.

[b] Gauges the extent of the presence of gravelly, lithic, or sodic soil. Soil data was taken from the European Soil Project (Panagos, Van Liedekerke, Jones, & Montanarella, 2012).

[c] 1= Shallow (<40cm); 2= Moderate (40-80cm), 3= Deep (80-120cm), 4= Very deep (>120cm). Soil data was taken from the European Soil Project (Panagos, Van Liedekerke, Jones, & Montanarella, 2012).

[d] Using 1960-1990 as the reference period.

[e] Difference between the maximum and minimum elevation in the region.

Table S2. OLS Regression Model Examining the Relationship Between the Employment

Share in Large-Scale Coal-Based Industries in 1813-1820 and Regional Differences in

Obesity in the BBC Lab U.K. Project Data (N = 111 Regions in England and Wales).

	b	*SE*
Employment in large scale industries 1813-1820 (%)	0.079***	0.018
Watermills around 1800	-0.046**	0.015
Market potential 1811	-0.000	0.000
Limits to agricultural use	-0.154	0.104
Depth to rock	0.009	0.020
Mean July temperature	-0.013	0.027
Terrain ruggedness	-0.000	0.000
Population density 1811	-0.086***	0.024
R^2	0.382	

Outcomes is the regional prevalence of obesity (%).

*$P < 0.5$. **$P < 0.01$. ***$P < 0.001$.

Table S3. Instrumental Variable Regressions Using Distance to Nearest Coalfield as an Instrument for Employment Share in Large-Scale Coal-Based Industries in 1813-1820 and Predicting Obesity Levels in the BBC Lab U.K. Project Dataset (N = 111 Regions in England and Wales).

| | First stage | | | Second stage regression | | |
| | Employment in large industries (%) | | | Obesity (BBC Lab U.K. data) | | |
	b	SE	P	b	SE	P
Distance to coalfield	-0.403	0.035	< 0.001			
Employment share in large industries 1813-1820 (%)				0.083	0.021	< 0.001
Watermills around 1800	-0.065	0.053	0.227	-0.047	0.013	< 0.001
Market potential 1811	-0.000	0.000	0.711	-0.000	0.000	0.103
Limits to agricultural use	-0.842	0.199	< 0.001	-0.151	0.155	0.330
Depth to rock	-0.061	0.060	0.316	0.009	0.020	0.655
Mean July temperature	-0.157	0.096	0.105	-0.011	0.023	0.639
Terrain ruggedness	-0.000	0.000	0.699	-0.000	0.000	0.324
Population density 1811	0.306	0.093	0.001	-0.089	0.029	0.002
R^2		0.671			0.381	

Table S4. OLS Regression Model Examining the Relationship Between the Employment Share in Large-Scale Coal-Based Industries in 1813-1820 and Regional Differences in Economic Hardship Indicators and Personality Traits (N = 111 Regions in England and Wales).

	Economic growth (1931-2011)	Unemployment (2001)	Conscientious-ness	Neuroticism	Openness	Extraversion	Agreeable-ness
Large-scale industry[a]	-36.03**	0.586***	-0.0117**	0.0136**	-0.00825	-0.00641	0.00110
	(10.61)	(0.163)	(0.00420)	(0.00413)	(0.00419)	(0.00345)	(0.00230)
Watermills around 1800	9.407	-0.572***	0.0161***	-0.00932**	0.00561	0.000239	0.00611**
	(8.656)	(0.133)	(0.00343)	(0.00337)	(0.00342)	(0.00281)	(0.00187)
Market potential 1811	0.00375***	-4.35e-05***	6.28e-07*	-4.50e-07	-6.55e-07*	9.93e-08	2.55e-08
	(0.000667)	(1.03e-05)	(2.64e-07)	(2.60e-07)	(2.63e-07)	(2.17e-07)	(1.44e-07)
Limits to agricultural use	-11.91	-0.346	-0.0204	-0.00536	0.0670**	0.0382	-0.00869
	(60.72)	(0.935)	(0.0240)	(0.0236)	(0.0240)	(0.0197)	(0.0131)
Depth to rock	-15.32	0.115	-0.000117	-0.000208	-0.00219	-0.00190	-0.00287
	(11.49)	(0.177)	(0.00455)	(0.00448)	(0.00454)	(0.00373)	(0.00249)
Mean July temperature	-4.463	-0.520*	0.000983	-0.00483	0.00375	0.00260	-0.00247
	(15.55)	(0.240)	(0.00616)	(0.00606)	(0.00614)	(0.00505)	(0.00337)
Terrain ruggedness	-0.0900	-0.00151	-4.60e-05	2.00e-06	3.00e-05	2.51e-06	-2.83e-05
	(0.0669)	(0.00103)	(2.65e-05)	(2.60e-05)	(2.64e-05)	(2.17e-05)	(1.45e-05)
Population density 1811	-44.45**	0.249	-0.0219***	-0.00140	0.0291***	0.00766	-0.00368
	(14.22)	(0.219)	(0.00563)	(0.00554)	(0.00561)	(0.00462)	(0.00308)
R^2	0.405	0.432	0.459	0.211	0.336	0.168	0.226

Unstandardized regression coefficients presented. Standard errors in parentheses.

[a] Employment share (%) in large-scale coal-based industries 1813-1820.

* $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

Table S5. OLS Regression Model Examining the Relationship Between Economic Hardship Indicators and Personality Traits and Obesity Prevalence (N = 111 Regions in England and Wales).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Large-scale industry[a]	0.0575***	0.0334**	0.0627***	0.0426***	0.0485***	0.0512***	0.0621***
	(0.0137)	(0.0115)	(0.0136)	(0.0122)	(0.0112)	(0.0118)	(0.0131)
Watermills around 1800	-0.0202	0.0107	-0.0220	-0.00793	-0.0121	-0.0211*	-0.0227*
	(0.0107)	(0.00989)	(0.0118)	(0.00981)	(0.00909)	(0.00948)	(0.0112)
Market potential 1811	-2.74e-07	1.81e-07	-7.98e-07	-1.23e-07	-1.88e-06**	-6.07e-07	-7.84e-07
	(9.36e-07)	(6.88e-07)	(8.44e-07)	(7.39e-07)	(7.12e-07)	(7.31e-07)	(8.21e-07)
Limits to agricultural use	-0.153*	-0.110	-0.151*	-0.144*	-0.0391	-0.0854	-0.150*
	(0.0744)	(0.0616)	(0.0751)	(0.0663)	(0.0653)	(0.0677)	(0.0749)
Depth to rock	0.00975	0.00344	0.0118	0.0121	0.00814	0.00852	0.0124
	(0.0142)	(0.0117)	(0.0142)	(0.0126)	(0.0119)	(0.0126)	(0.0142)
Mean July temperature	-0.0106	0.0118	-0.0100	-0.00297	-0.00371	-0.00550	-0.00951
	(0.0191)	(0.0160)	(0.0192)	(0.0170)	(0.0162)	(0.0171)	(0.0192)
Terrain ruggedness	-0.000108	-1.84e-05	-9.41e-05	-9.84e-05	-4.52e-05	-9.11e-05	-8.98e-05
	(8.26e-05)	(6.84e-05)	(8.36e-05)	(7.30e-05)	(6.98e-05)	(7.32e-05)	(8.39e-05)
Population density 1811	-0.0594**	-0.0633***	-0.0528**	-0.0514**	-0.00468	-0.0402*	-0.0527**
	(0.0182)	(0.0144)	(0.0188)	(0.0155)	(0.0166)	(0.0158)	(0.0176)
Employment growth 1931-2011	-0.000135						
	(0.000122)						
Unemployment rate in 2011		0.0431***					
		(0.00618)					
Conscientiousness			0.0302				
			(0.310)				
Neuroticism				1.457***			
				(0.279)			
Openness					-1.678***		

40

					(0.261)		
Extraversion						-1.731***	
						(0.335)	
Agreeableness							0.200
							(0.566)
R^2	0.345	0.554	0.337	0.479	0.530	0.476	0.338

Unstandardized regression coefficients presented. Standard errors in parentheses.

[a] Employment share (%) in large-scale coal-based industries 1813-1820.

* $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

Table S6. OLS Regression Model Examining the Relationship Between Economic Hardship Indicators and Personality Traits and Obesity Levels in the BBC Lab U.K. Project Dataset (N = 111 Regions in England and Wales).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Large-scale industry[a]	0.0777***	0.0482**	0.0831***	0.0573**	0.0602***	0.0637***	0.0782***
	(0.0193)	(0.0176)	(0.0189)	(0.0179)	(0.0159)	(0.0166)	(0.0182)
Watermills around 1800	-0.0453**	-0.0115	-0.0515**	-0.0309*	-0.0329*	-0.0450**	-0.0487**
	(0.0150)	(0.0152)	(0.0164)	(0.0144)	(0.0129)	(0.0133)	(0.0156)
Market potential 1811	-1.50e-06	-5.87e-07	-1.83e-06	-8.92e-07	-3.07e-06**	-1.37e-06	-1.61e-06
	(1.31e-06)	(1.05e-06)	(1.17e-06)	(1.09e-06)	(1.01e-06)	(1.03e-06)	(1.14e-06)
Limits to agricultural use	-0.155	-0.110	-0.147	-0.146	-0.00391	-0.0651	-0.150
	(0.104)	(0.0945)	(0.104)	(0.0975)	(0.0927)	(0.0952)	(0.104)
Depth to rock	0.00819	-0.000240	0.00865	0.00894	0.00369	0.00415	0.0101
	(0.0199)	(0.0179)	(0.0197)	(0.0184)	(0.0169)	(0.0177)	(0.0198)
Mean July temperature	-0.0127	0.0105	-0.0129	-0.00494	-0.00412	-0.00645	-0.0113
	(0.0267)	(0.0246)	(0.0266)	(0.0250)	(0.0229)	(0.0240)	(0.0268)
Terrain ruggedness	-0.000103	-1.94e-05	-8.38e-05	-0.000104	-3.34e-05	-9.49e-05	-8.61e-05
	(0.000116)	(0.000105)	(0.000116)	(0.000107)	(9.89e-05)	(0.000103)	(0.000117)
Population density 1811	-0.0871**	-0.0962***	-0.0778**	-0.0836***	-0.0205	-0.0679**	-0.0839**
	(0.0256)	(0.0221)	(0.0261)	(0.0228)	(0.0235)	(0.0222)	(0.0246)
Employment growth 1931-2011	-2.73e-05						
	(0.000171)						
Unemployment rate in 2011		0.0456***					
		(0.00948)					
Conscientiousness			0.370				
			(0.431)				
Neuroticism				1.576***			
				(0.410)			
Openness					-2.248***		

						(0.371)	
Extraversion						-2.341***	
						(0.472)	
Agreeableness							0.518
							(0.788)
R^2	0.382	0.498	0.386	0.461	0.548	0.504	0.384

Unstandardized regression coefficients presented. Standard errors in parentheses.

[a] Employment share (%) in large-scale coal-based industries 1813-1820.

* $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

Table S7. Instrumental Variable Regressions Using Distance to Nearest Coalfield as an Instrument for Employment Share in Large-Scale Coal-Based Industries to Predict Regional Differences in Potential Mediator Variables (N = 111 Regions in England and Wales).

	First stage		Second stage regressions	
	Large-scale industry[a]	Large-scale industry[a]	Neuroticism	Unemployment rate (2001)
Distance to coalfield	-0.399***	-0.403***		
	(0.032)	(0.0347)		
Watermills around 1800		-0.0650	-0.00905**	-0.596***
		(0.0535)	(0.00313)	(0.135)
Market potential 1811		-1.39e-06	-4.41e-07	-4.43e-05***
		(3.74e-06)	(2.62e-07)	(1.08e-05)
Limits to agricultural use		-0.843***	-0.00643	-0.250
		(0.200)	(0.0175)	(0.456)
Depth to rock		-0.0605	-0.000313	0.124
		(0.0601)	(0.00355)	(0.195)
Mean July temperature		-0.158	-0.00528	-0.480*
		(0.0964)	(0.00560)	(0.219)
Terrain ruggedness		-0.000138	1.09e-06	-0.00143
		(0.000356)	(2.42e-05)	(0.000851)
Population density 1811		0.307**	-0.000679	0.184
		(0.0927)	(0.00450)	(0.277)
Large-scale industry[a]			0.0123*	0.699***
			(0.00489)	(0.191)
R^2				

[a] Employment share (%) in large-scale coal-based industries 1813-1820.

Unstandardized regression coefficients presented. Standard errors in parentheses.

* $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$.

Table S8. Mediation Models of the Indirect Effect of Employment Share in Large-Scale Coal-Based Industries in 1813-1820 on Obesity Levels through Unemployment and Neuroticism in the BBC Lab U.K. Project Dataset (N = 111 Regions in England and Wales).

	b	95% CI Lower ; Upper		Effect ratio
Employment in large industries => Obesity				
(parallel mediation model)				
Total effect	0.079	0.043	0.115	
Direct effect	0.048	0.013	0.084	0.61
Indirect effect (via unemployment[a])	0.014	-0.001	0.032	0.18
Indirect effect (via neuroticism)	0.016	0.003	0.037	0.20
Total indirect effect	0.030	0.013	0.053	0.38
Employment in large industries => Obesity				
(serial mediation model)				
Total effect	0.079	0.043	0.115	
Direct effect	0.048	0.013	0.084	0.61
Indirect effect (via unemployment[a])	0.014	-0.001	0.032	0.18
Indirect effect (via neuroticism)	0.009	-0.006	0.026	0.11
Indirect effect (via unemployment[a] + neuroticism)	0.007	0.001	0.017	0.09
Total indirect effect	0.030	0.013	0.053	0.38

Unstandardized regression coefficients presented. Models adjusted for historical control variables (energy supply, market potential, geology, climate, population density) and country.
[a] Regional unemployment rate in 2001.

Table S9. Mediation Models of the Indirect Effect of Portion of Variation in Employment Share in Large-Scale Coal-Based Industries Predicted by Proximity to Coalfields (*Largescale*$_{pred}$) on Obesity Levels in the BBC Lab U.K. Dataset (N = 111 Regions in England and Wales).

	b	95% CI Lower ; Upper		Effect ratio
Employment in large industries => Obesity **(parallel mediation model)**				
Total effect	0.061	0.024	0.099	
Direct effect	0.026	-0.006	0.058	0.43
Indirect effect (via unemployment[a])	0.021	0.007	0.040	0.34
Indirect effect (via neuroticism)	0.014	0.002	0.035	0.23
Total indirect effect	0.035	0.016	0.061	0.57
Employment in large industries => Obesity **(serial mediation model)**				
Total effect	0.061	0.024	0.099	
Direct effect	0.026	-0.006	0.058	0.43
Indirect effect (via unemployment[a])	0.021	0.007	0.040	0.34
Indirect effect (via neuroticism)	0.005	-0.005	0.023	0.08
Indirect effect (via unemployment[a] + neuroticism)	0.009	0.003	0.19	0.15
Total indirect effect	0.035	0.016	0.061	0.57

Unstandardized regression coefficients presented. Models adjusted for historical control variables (energy supply, market potential, geology, climate, population density) and country.

[a] Regional unemployment rate in 2001.

Table S10. Mediation Models of the Indirect Effect of Employment Share in Large-Scale

Coal-Based Industries in 1813-1820 on Obesity Prevalence in 2013-2015 through Regional

Unemployment in 1971 and the Change in Unemployment Levels from 1971 to 2011

(N = 111 Regions in England and Wales)

	b	95% CI		Effect ratio
		Lower ; Upper		
Employment in large industries =>				
Obesity (parallel mediation model)				
Total effect	0.062	0.040	0.085	
Direct effect	0.034	0.011	0.056	0.55
Indirect effect (via the unemployment rate in 1971)	0.013	0.002	0.027	0.21
Indirect effect (via the change in the unemployment rate from 1971 - 2011)	0.016	0.006	0.030	0.26
Total indirect effect	0.029	0.015	0.046	0.47

Note: Models were adjusted for historical control variables (energy supply, market potential,

geology, climate, population density) and country.

Table S11. OLS Regression Model Examining the Relationship Between the Employment Share in Large-Scale Coal-Based Industries in 1813-1820 and Regional Differences in the Prevalence of Obesity in 2013-2015 after Adjustment for 2011 Regional Characteristics (N = 111 Regions in England and Wales).

	b	*SE*	*P*
Employment in large scale industries 1813-1820	0.050	0.014	= 0.001
Population size[a]	0.000	0.000	.286
Percentage share of population in urban areas[a]	0.001	0.0004	.020
Percentage share of male population[a]	-0.011	0.031	.719
Population density[a]	0.000	0.000	.491
R^2		0.337	

Note: Model adjusts for watermills around 1800, market potential 1811, limits to agricultural use, depth to rock, mean July temperature, terrain ruggedness, population density 1811, and a country dummy variable.
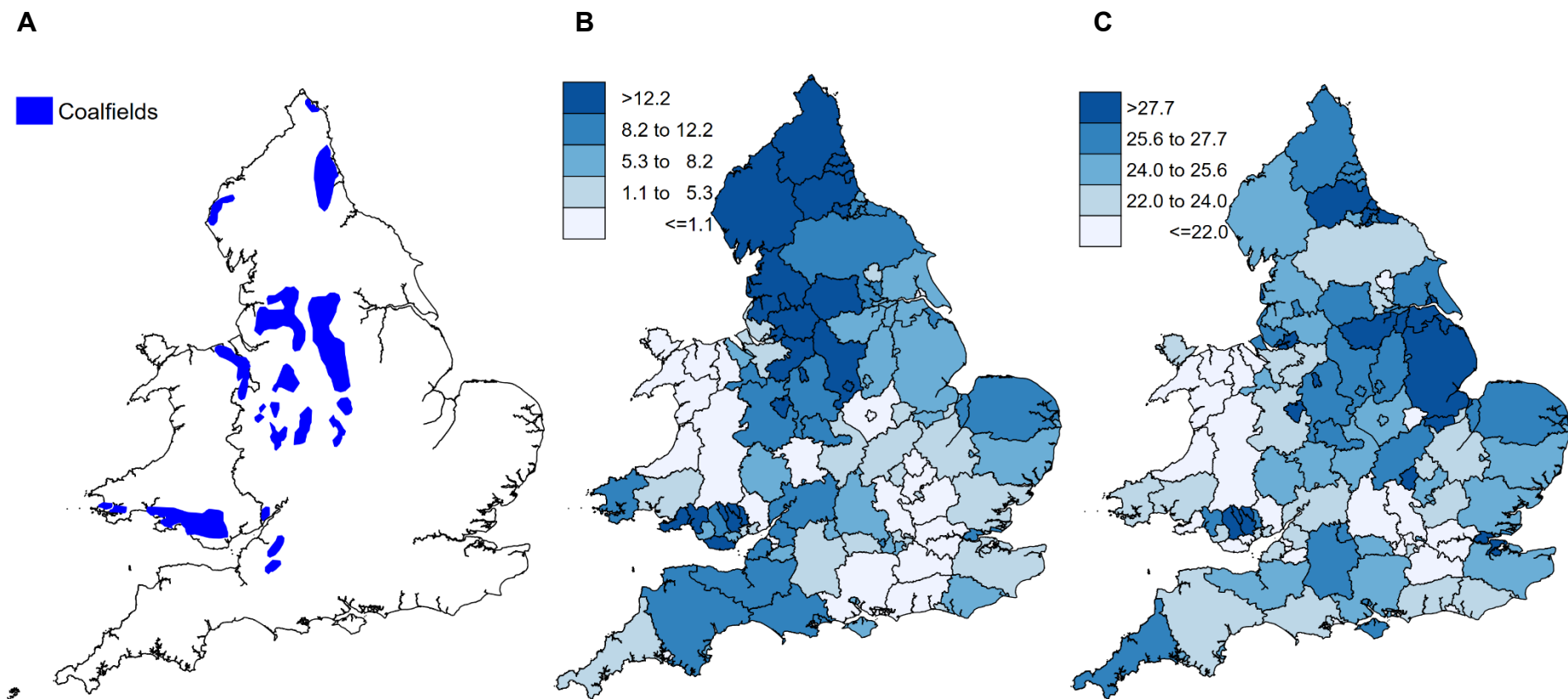
[a] Assessed for each region in 2011.

Figure S1. Coalfields before 1700 (A), predicted employment in large-scale coal-based industries (%) in 1813-1820 (B), and predicted obesity prevalence (%) in 2013-2015 (C) in England and Wales (N = 111 regions).