
Article:
Davillas, Apostolos and Jones, Andrew Michael (2020) Regional inequalities in adiposity in England:distributional analysis of the of the contribution of individual-level characteristics and the small area obesogenic environment. Economics and Human Biology. 100887. ISSN: 1570-677X

https://doi.org/10.1016/j.ehb.2020.100887

Regional inequalities in adiposity in England: distributional analysis of the contribution of individual-level characteristics and the small area obesogenic environment

Apostolos Davillas[¶] and Andrew M Jones[†]

[¶] Health Economics Group, University of East Anglia, Norwich, UK
[†] Department of Economics and Related Studies, University of York, York, UK and Centre for Health Economics, Monash University, Australia

Abstract

We use nationally representative English data to examine regional variations in body mass index (BMI) and waist circumference (WC), and to explore their underlying sources. Beyond our "at the mean" analysis, Shapley decomposition combined with unconditional quantile regression analysis allow us to explore the relative contribution of small-area level proxies of the obesogenic environment as opposed to our set of individual-level characteristics, across the whole adiposity distribution . We find that the regional BMI differences, that are more evident towards the right tails of its distribution, are fully accounted for by the neighbourhood obesogenic environment. The latter exerts an independent contribution to excess adiposity over and above the potential mediating role of individual-level lifestyle and socio-economic position (SEP). Overall, the relative contribution of demographics (age and gender) becomes less evident moving to higher quantiles of the BMI distribution, while that of obesogenic environment, individual-level lifestyle and SEP measures becoming more relevant. The neighbourhood obesogenic environment is also much more relevant in the tails of the WC distribution. The role of the obesogenic environment on excess adiposity is more pronounced for women than men. Overall, our results highlight that policies that aim to tackle excess adiposity should address both people and places.

Keywords: adiposity; obesogenic environment; regional inequality; decomposition analysis
JEL codes: I14; I12

Acknowledgments

Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. Andrew Jones acknowledges funding from the Leverhulme Trust Major Research Fellowship (MRF-2016-004). The funders, data creators and UK Data Service have no responsibility for the contents of this paper.

1. Introduction

Although progressive improvements in life expectancy and the health of the population have been evident over recent decades, preventable inequalities in health (including obesity) persist between regions within England (Baker, 2018; Newton et al., 2015; NHS Digital, 2018a; Shelton et al., 2009). Evidence for the period between 2008 and 2010 shows that England has some of the largest regional health inequalities in Europe. For example, the life expectancy difference for women between the poorest English regions (the North East and North West) and the more affluent (London and the South East) was similar to the life expectancy gap between the former West and post-communist East Germany in the mid-1990s (Bambra et al., 2014). Understanding the geography of health is complex, given the regional variations and interactions of the different underlying health determinants at the individual, contextual and environmental level. The importance of mitigating regional differences in health is emphatically stated in the recent study by Newton et al. (2015): "if levels of health in the worst performing regions in England matched the best performing ones, England would have one of the lowest burdens of disease of any advanced industrialised country".

Excess adiposity is one of the leading risk factors for preventable ill-health and disability (Davillas and Pudney, 2019; Newton et al., 2015), being associated with several chronic conditions (GBD 2015 Obesity Collaborators, 2017; Must et al., 1999). Existing evidence shows significant regional differences in adiposity across the English regions. For example, obesity prevalence, defined as Body Mass Index (BMI) ≥ 30 kg/m^2, is about 23% in London as opposed to 33% in the West Midlands (NHS Digital, 2018b). Regional inequalities in excess adiposity are of interest themselves, given the important socio-economic ramifications of obesity and the implied clustering of the health risks and disadvantage at the regional level. They are also important because they may influence

resource allocation and to infer the success of area-based policies to tackle obesity. From the perspective of health policy, better understanding of the underlying sources of the regional inequalities in excess adiposity are useful for local authorities, given their enhanced role as leaders for local population health.

We use a nationally representative dataset for the UK, Understanding Society: the UK Household Longitudinal Study (UKHLS) to examine the relative contribution of the neighbourhood-level obesogenic environment and individual-level characteristics to adiposity variations in England. We use two measures of adiposity: BMI and waist circumference (WC). We contribute to the literature in a number of ways.

First, we explore the presence of regional differences in adiposity at the nine Government Office Regions (GORs) of England. The GOR-level inequalities in obesity are frequently monitored in public health policy reports (e.g., NHS Digital, 2018a), with recent research aiming to explain regional-level health (including obesity) inequalities (Newton et al., 2015; Vallejo-Torres and Morris, 2010). Measurement of inequalities at GOR level has been of interest and policy relevance because these regions mostly coincide with administrative structures aimed to address English regional imbalances (i.e., the Strategic Health Authorities), up to their abolishment in favour of smaller scale and more local units of governance (Bambra et al., 2014; DHSC, 2013)[1]. These inequalities between regions are attributed to within region differences in the obesogenic environment at the small-area neighbourhood level is a timely issue, given the recently enhanced role of local authorities as leaders for local population health (DHSC, 2013). The legal duty for local authorities to commission care and support services is stated by the Department of Health and Social Care (DHSC, 2013): "they use

[1] The Strategic Health Authorities (SHA) are coterminous with Government Office Regions, except that the large South East England region which is divided into the South Central and South East Coast SHA.

their knowledge of their communities to tackle challenges such as smoking, alcohol and drug misuse and obesity."

In this study, we use a detailed set of individual-level characteristics along with small-area level proxies of the obesogenic environment to explore their relative contributions to the observed variations in our adiposity measures. We use multilevel analysis to explore the relative contribution of the neighbourhood-level effects, capturing the role of unobserved obesogenic characteristics. Beyond this analysis, we also employ a Shapley decomposition to directly decompose the contribution of our rich set of observed neighbourhood-level obesogenic characteristics and understand their relative role, over and above individual-level characteristics. This complements existing research exploiting the US institutional context on regional variations in food prices and taxes (Powell and Chaloupka, 2009; Rahkovsky and Gregory, 2013) or availability of detailed data on neighbourhood-level "food deserts" to better understand inequalities in health and nutritional status (Allcott et al., 2019). Other studies aiming to explore the role of the neighbourhood social conditions and "built" environmental factors on adiposity are often limited to specific geographical regions and selected population groups and, thus, do not provide nationally representative results (for example, Booth et al., 2005; Drewnowski et al., 2016; Lovasi et al, 2009). We focus on local (small-area level), modifiable obesogenic characteristics here, such as proxies for geographical barriers, air quality, criminality, local levels of anxiety, income and education deprivation levels. These latter may affect adiposity via their direct and indirect effects on the energy imbalance that causes excess adiposity (Papas et al, 2007; Stafford et al., 2007; Swinburn, 2011).

Although certain aspects of the built environment may affect obesity via direct biological mechanisms (such as, for example, the association between air pollution and metabolic disorders (An et al., 2018)), the role of obesogenic environment on adiposity is mainly behavioural via affecting consumption and energy expenditure (An et al., 2018; Carroll-Scott, 2013; Santana et al., 2009). Our analysis allows us to explore the extent to which its role is mediated by the individual-level lifestyle and other characteristics, given the interplay between individual-level and environmental characteristics in affecting adiposity (Costa-Font and Gil, 2008; Raftopoulou, 2017; Santana et al., 2009), as well as the potential direct role of the obesogenic environment on adiposity. Identifying that the neighbourhood environment plays a systematic role for excess adiposity is of particular importance, indicating that a hypothetical movement of an individual from a less to a more obesogenic neighbourhood may put them at higher risk of excess adiposity. This information is also relevant for the design and planning of interventions to target aspects of the obesogenic environment at the neighbourhood level to tackle the obesity epidemic.

Second, beyond our analysis "at the mean", we also focus at quantiles of the distribution of adiposity. "Beyond the mean" estimation techniques allow us to explore the potentially heterogeneous patterns in the contribution of the explanatory variables across quantiles of the distribution with a focus on the right tails, where higher health care risks and costs are concentrated. To the best of our knowledge, this is the first study that combines Shapley decomposition techniques with unconditional quantile regression techniques (UQR) to explore regional inequalities in adiposity and the underlying individual and neighbourhood-level factors that may contribute to variations in adiposity. Analysis "at the mean" may mask important information in other parts of the adiposity distribution (for example, Green et al., 2016; Stifel and Averett,

2009).Dichotomising (or categorising) our adiposity measures using conventional clinical thresholds may also result in loss of information (Jolliffe, 2011).[2] For example, assuming one of the most extensively used dichotomous indicators, i.e. obesity defined as BMI≥30mk/m^2; does not explicitly explore the distribution of BMI above or below the threshold but treats those who fall into (or out of) obesity homogenously. By categorising BMI using thresholds, masks the fact that people with values well above any chosen threshold may experience significantly higher health risks compared to those close to the threshold. It has been shown that the latter matters when it comes to exploring the association between socioeconomic measures and adiposity (Jolliffe, 2011). Third, in contrast to many previous studies, we use nurse-collected adiposity measures. It has been shown that reporting errors in body weight (or BMI) are non-classical, meaning that they depend on individual characteristics and may create misclassification biases in obesity prevalence (Cawley et al., 2015).

The rest of the paper is organised as follows. Section 2 describes our regression and decomposition analysis and section 3 introduces the data. Our results are presented in section 4, and section 5 concludes and summarises our findings.

2. Methods

2.1 Regression and Decomposition Methods

Our adiposity measures are initially modelled by linear regression, estimated using ordinary least squares (OLS). Regression models are first estimated using age, gender and regional dummies (Specification 1). We then enhance this specification by adding obesogenic small-area level characteristics (Specification 2). Our full specification is

[2] It has been shown also that the prevalence of excess adiposity may increase over time, despite the mean BMI levels being unchanged, as a result of changes in the BMI distribution, with a greater numbers of people above certain excess adiposity thresholds (Madden, 2012).

further augmented by individual-level socio-economic position (SEP) and lifestyle covariates (Specification 3).

We also estimate unconditional quantile regression models (UQR) that allow us to consider the entire distribution of the adiposity measures and to investigate the potentially differential associations at different points of their distribution (Firpo et al., 2009). Unlike conventional quantile regression models, which explore the effect of covariates on the conditional quantiles of the outcome, the UQR technique estimates *unconditional* quantile partial effects.

The estimation of the UQR is based on the Recentered Influence Function (RIF) (Firpo et al., 2009). The RIF can be estimated directly from the data by computing sample quantiles (q_τ) of the adiposity measure (H_i) and then estimating the density of the distribution of adiposity measure at that quantiles using kernel density methods. Specifically, for an observed quantile (q_τ), a RIF is generated as:

$$RIF(H_i; q_\tau) = q_\tau + \frac{\tau - 1[H_i \leq q_\tau]}{f_H(q_\tau)} \quad (1)$$

where, q_τ is the observed quantile, $1[H_i \leq q_\tau]$ is an indicator that equals to one if the observed adiposity value is less than or equal to the observed quantile q_τ and zero otherwise. $f_H(q_\tau)$ is the estimated kernel density of the adiposity measure at the τ^{th} quantile. The RIF is then regressed on our different sets of covariates as defined in Specifications 1, 2 and 3 above. Our analysis is weighted using UKHLS sample weights to account for survey non-response and attrition, making the sample representative of the English population.

We then build on recent inequalities in health research that combines quantile regression techniques and Shapley decomposition techniques (Davillas and Jones, 2020). The Shapley decomposition (Shorrocks, 2013) is used to explore the contribution of each of the explanatory variables to the variance in adiposity measures explained (expressed in relative terms as a ratio to the overall variance) by our model specifications[3]. This allows us to explore the potentially heterogeneous patterns in the contribution of the explanatory variables across quantiles of the adiposity distribution with a focus on the right tail, implying higher health risks.

In order to estimate the contribution of each of the explanatory variables to the variance in adiposity measures, the Shapley decomposition calculates the marginal effect to the variance in adiposity explained by eliminating each of the explanatory variables ($\boldsymbol{X}$) in sequence and, for each explanatory variable, assigns the average of its marginal contributions in all possible elimination sequences.[4]

Unlike other decomposition techniques, the Shapley decomposition is both exactly additive and path (order) independent, indicating that the sum of factor (Xs)

[3] This is equivalent to the contribution of each covariate to the R-squared of our models focusing at the mean (OLS) or at different quantiles of the distribution (UQR).

[4] Specifically, as an illustration example, there is no loss of generality to assume that there are only two covariates in our model specifications (X_1, X_2) and that (I) is the indicator of interest to be decomposed (here, the relative explained variance in adiposity). The marginal effect of each variable on the indicator of interest can be computed in two ways, by first eliminating one variable and then the other, or vice versa. To compute the marginal contribution of X_1 for example, one way is to subtract from the overall (I) indicator, the one obtained when this variable is omitted from the regression. The second way is to estimate (I) by a regression with X_1 only and, then, subtract the derived (I) from a regression with both variables are omitted; the latter is zero here as we are interested in the contribution of the explanatory variables to the explained variance in adiposity. Then, the Shapley contribution for X_1 can be given by averaging these possibilities: $C_1 = \frac{1}{2}[I(X_1, X_2) - I(X_2) + I(X_1) - I(.)]$; and analogously for X_2. Allowing for more covariates makes things more complex and computationally intensive as all possible permutations of the explanatory variables need to be estimated, i.e., 2^k, where k is the number of explanatory variables. A more formal description of the Shapley decomposition of the relative variance explained by the model specification (equivalent to R-squared) can be found elsewhere (Israeli, 2007).

contributions adding up to the total variance explained. Typically, in most decomposition methods, the value of the contribution assigned to any given covariate depends on the order in which each covariate is eliminated from the model to calculate its contribution; as such, the different factors are not treated symmetrically. The path independence property of the Shapley decomposition ensures that we are able to explore the contribution of each factor robustly, irrespective of this ordering (Shorrocks, 2013).[5] However, as also noted by Ferreira and Gignoux (2014), Shapley decomposition techniques should only be interpreted as evidence on the relative importance of the different set of covariates. For example, the correlation between different aspects of the obesogenic environment at the small-area level may indicate that the coefficients might suffer from multicollinearity, which may bias our results and decomposition analysis.

2.2 Multilevel random intercept models

A multilevel random intercepts model may be a useful alternative to our analysis "at the mean". A random intercepts model that allows for grouping of adiposity levels within small neighbourhood areas (LSOA-level) is estimated here. In practice, this is equivalent to a random effects model, as it allows for separate error components for each LSOA. This analysis provides an alternative way to explore the contribution of the neighbourhood level obesogenic environment to the explained variance in adiposity.

Specifically, we estimate random intercept models without explanatory covariates, as:

$$H_{i,j} = \beta + \theta_j + u_{i,j} \quad (2)$$

where, $H_{i,j}$ is the adiposity measure for an individual i in neighbourhood j, β is an unknown fixed intercept, θ_j stands for the neighbourhood-level random effects and $u_{i,j}$ is

[5] Shapley decomposition techniques have been used recently to explore to what extent (and what kind of) distributional changes in the BMI distribution may result in changes in obesity prevalence over time (Madden, 2012; Pak et al., 2016).

the error term. Given that this model allows for neighbourhood-level random intercepts, it captures the overall contribution of the neighbourhood obesogenic environment on adiposity (including unobserved factors), without the need to include specific neighbourhood-level obesogenic characteristics. The intra-class correlation coefficient (ICC) estimates the correlation between adiposity levels within neighbourhoods and, thus, the proportion of the total variance in adiposity that is attributed to neighbourhood characteristics. We then augment the models with GOR dummies and, subsequently, with our set of demographic, individual-level SEP and lifestyle covariates to explore whether the proportion of the total variance in adiposity that is attributed to neighbourhood characteristics is reduced.

Although this analysis measures the share of the total adiposity variance that is attributed to the neighbourhood level, it is not helpful for understanding the neighbourhood-level obesogenic characteristics that exert the largest contribution. As such, these results may be considered as useful comparisons to our Shapley decomposition analysis "at the mean", providing an alternative estimate for the total neighborhood-level contributions to adiposity.

3. Data

The UKHLS is a large, nationally representative UK study. For this paper, we employ the General Population Sample (GPS), a random sample of the general population. Adiposity measures were collected for the GPS as part of the UKHLS wave 2 data collection (2010-2011). We focus on the English sub-sample, given the absence of comparable cross-national neighbourhood-level data for the rest of the UK that can be linked at the small-area level with UKHLS. Specifically, to obtain our neighbourhood-level obesogenic data we have linked the UKHLS Wave 2 data (2010-2012) at the Lower

Layer Super Output Area (LSOA) level with selected sub-domains of the 2010 English Indices of Deprivation (EID2010). The LSOAs are lower layer geographies, taking into account population size, mutual proximity and social homogeneity; they have on average 1,500 residents and 650 households. Following common practise (Flouri et al., 2013) and given data availability on linkage, respondents' LSOAs at the UKHLS wave 2 collection are used to define neighbourhood areas.

Given that we focus on adults we have excluded those individuals below the age of 20 years old to overcome puberty-related body weight growth concerns (Davillas and Benzeval, 2016; Power et al., 1997). After excluding missing data on all variables used in our analysis, our working sample reduced from 13,162 adults aged 20+ (potential sample) to 12,271 (working sample)[6]. We use two adiposity measures: BMI and WC. BMI is calculated as the weight (kilograms) over the square of height (metres). Body weight, height and WC are measured by trained nurses using standard protocols. WC (in cm) was measured twice, or three times if the two original measurements differed by more than 3 cm. The mean of the valid measurements (the two closest, if there were three) is used (McFall, 2014).

3.1 Aggregated regional-level (GOR) adiposity differences

To explore the regional differences in adiposity levels, nine dummies for the GORs for England are included (South East, London, North West, East of England, West Midlands, South West, Yorkshire and the Humber, East Midlands and North East). GORs are the highest regional layer level for England, and are used here to explore regional adiposity differences at the aggregated level. One of the aims of our paper is to

[6] Comparisons between the raw means of the full sample and our working sample show similar results, suggesting that the impact of item missingness may be limited.

explore the extent that regional differences in adiposity are attributable to the obesogenic environment and individual-level characteristics and behaviours.

3.2 Neighbourhood-level obesogenic characteristics

We use a set of modifiable small-area level characteristics to proxy the neighbourhood obesogenic environment. Our analysis allows us to explore the extent to which their role is mediated by individual-level characteristics and lifestyle, which are also accounted for in our model specifications (sub-section 3.3), as well as the potential direct role of the obesogenic environment on adiposity. Exploring the direct effect of these small-area level characteristics on adiposity levels (over and above individual characteristics) and to what extent these explain the adiposity inequalities at the aggregate regional level (GOR) is important for local-area policymaking.

A proxy of neighbourhood air quality levels is included here. Several mechanisms may link air pollution to excess body weight (An et al., 2018): a) directly, as air pollution may result in metabolic dysfunction (via a number of biological mechanisms, such as increased oxidative stress and adipose tissue inflammation) that may result in excess adiposity; b) indirectly, as air pollutants have been linked to decreased lung function, elevated blood pressure, and other cardiovascular and respiratory symptoms, resulting in impaired exercise capacity and performance. In addition, air pollution may prevent people from engaging in regular physical activity, outdoor activities and promote sedentary lifestyles (An et al., 2018). A number of studies found systematic associations between air quality and excess adiposity (for example, An et al., 2018, Barrea et al., 2017, Chaparro et al., 2018). In this study, we use sulphur dioxide concentrations at the LSOA level (the EID2010 sulphur dioxide indicator); a frequently used measure of air quality that is mainly attributable to metal processing, smelting facilities and motor

vehicles. This is defined as the average concentration of the pollutant by LSOA, divided by the WHO safe guideline levels for sulphur dioxide that are related with severe health risks (McLennan et al., 2011). It is, therefore, a ratio-scale indicator, and the higher the value, the closer the levels are to the WHO threshold, indicating worse air quality levels at the neighbourhood level.

Neighbourhood crime levels may affect excess adiposity via several mechanisms. Neighbourhoods with excess crime levels may experience disinvestment and declining community resources and, thus, an environment which is unappealing and unsafe for fostering physical activity (Yu and Lippert, 2016). Beyond this community level mechanism, it has been shown that high neighbourhood crime rates are linked to a greater stress levels for the local population, which may lead to stress eating behaviours (Stafford et al., 2007; Torres and Nowson, 2007); more directly chronic stress initiates key physiological processes that increase the risk of obesity, and abdominal obesity in particular (Kahn et al. 1998). The composite IMD crime domain index is used to proxy criminality levels at the neighbourhood level. This is a derived index that measures the rate of recorded crime levels (expressed per 1000 at-risk population or properties, accordingly) for four major crime types – violence, burglary, theft and criminal damage (McLennan et al., 2011)[7]. We have created a dummy variable taking the value of one for the most crime-deprived neighbourhoods (ranked at the higher tertile of the crime deprivation index) and zero otherwise.

To explore the role of neighbourhood-level stressors in adiposity, a proxy of the community stress levels is also included in our analysis. Exposure to environmental, life

[7] Given the differences in units of measurement, these indicators are standardised by ranking, and combined by the Office for National Statistics using weights to define a composite crime domain that can be used to order neighbourhoods with respect to their level of criminality (McLennan et al., 2011).

or work-related stressors is associated with a greater preference for high-sugar and fat foods, with existing evidence of a causal effect of stress on weight gain (Torres and Nowson, 2007). The EID2010 mood and anxiety disorder indicator derived index combines data from different administrative sources and can be used to rank neighbourhoods with respect to anxiety levels (McLennan et al., 2011). We have created a dummy variable to capture those neighbourhoods that have the highest anxiety levels (ranked at the highest tertile of the EID2010 anxiety index).

We also account for the road distance to a GP to capture another aspect of the built environment. We used the road distance to a GP indicator of the EID2010, which captures the mean LSOA distance to the closest GP (in kilometres). Beyond being a proxy of the geographical barriers to access healthcare, this may be linked to excess adiposity via a direct link. In the UK, NHS Health Checks are largely conducted by GPs, with eligible individuals being offered a free screening test every five years. Obesity measurements (along with other cardiovascular related biomarkers) are fundamental parts of these checks; those individuals who are at risk are warned about their excess adiposity levels. Proximity to GPs may affect uptake (Burgess et al., 2015) and, thus, the possibility of tailored feedback on their adiposity levels, which may lead to behavioural changes and weight reduction.

To capture SEP-related deprivation at the neighbourhood level, we have used the EID2010 income deprivation and the adults' qualifications sub-domain. More prosperous neighbourhoods and those in which people with higher qualifications are clustered are more likely to have better built infrastructure, access to parks, and increased food options, with these resulting in more physical activity and/or reduced calories intake (Lopez, 2007; Santana et al., 2009).

The income deprivation domain measures the proportion of the population in an area that live in income-deprived families, defined as families claiming income support, income-based jobseeker's allowance or pension credit. The skills deprivation sub-domain captures the proportion of adults at LSOA-level with no qualifications or with qualifications below National Vocational Qualification (NVQ) level 2 (McLennan et al., 2011). Both indexes are ratio-scale variables with a natural interpretation. The higher values mean that a larger part of the LSOA-level population are income deprived or have low qualifications.

Finally, we also account for fast-food density (per 1,000 population) at the local authority level. We match the most proximal administrative local authority-level data on fast-food restaurants (mid-2014) to our UKHLS wave 2 data[8]. The availability and proximity to fast-food restaurants is hypothesised to influence obesity rates (Dunn et al., 2012; Lhila, 2011).

3.3 Individual-level characteristics

Following the literature (Baum and Ruhm, 2009; Davillas and Benzeval, 2016), we use a set of individual-level covariates that are typically associated with adiposity. These are factors that affect the process by which energy balances (calories in versus calories out) are translated into changes in adiposity (Chou et al., 2004; Cutler et al., 2003).

We include five age-group dummies (20-34, 35-49, 50-64, 65-79, 80+) for each gender. Three individual-level measures of SEP are included: levels of educational attainment

[8] It should be noted that data availability prevents us from including fast-food density data at the LSOA level and, thus, it is the only obesogenic characteristic in our analysis that is merged at the local authority rather than the LSOA level (Public Health England, 2014).

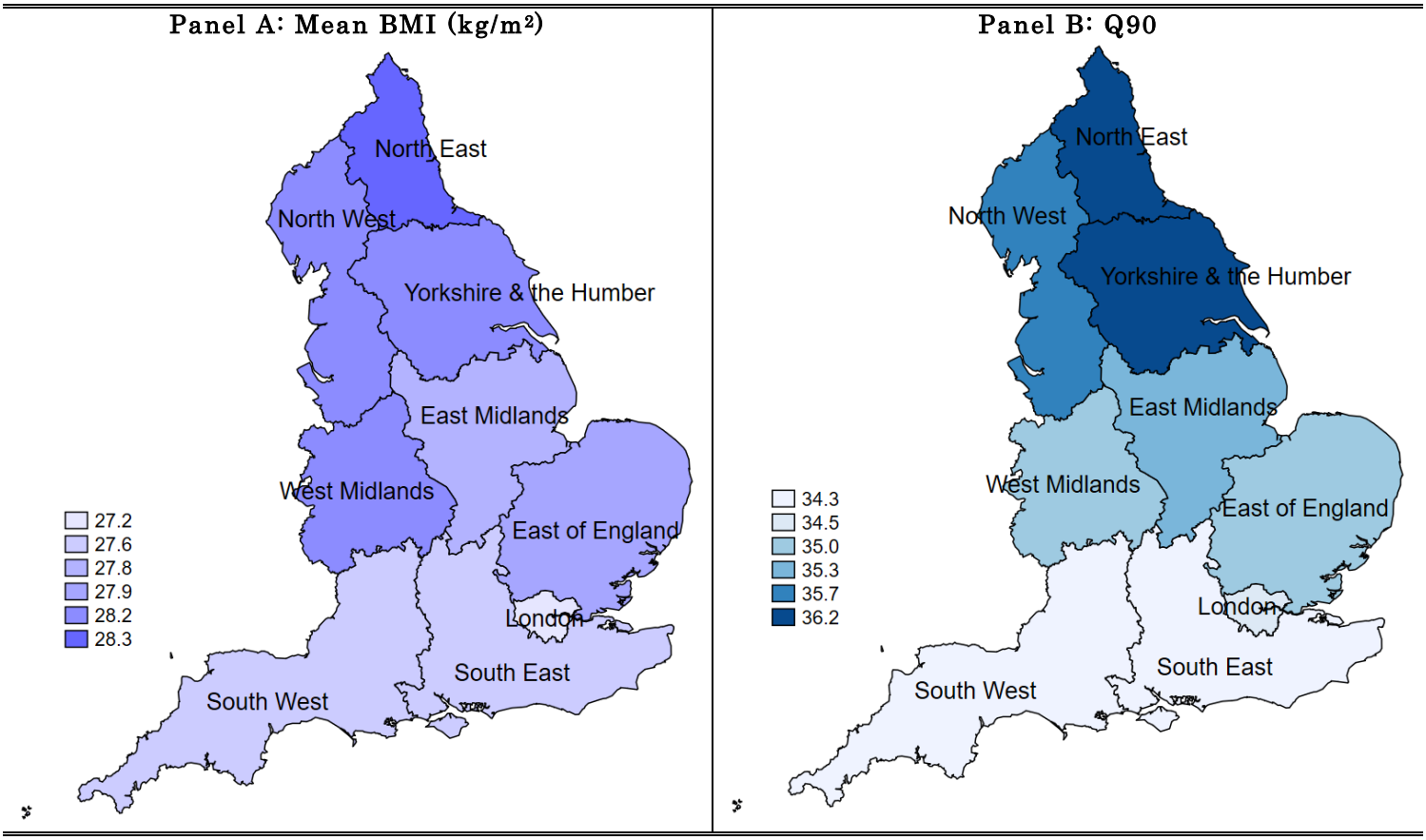
(degree (reference category), post-compulsory but not tertiary qualification, secondary school-leaving qualification, basic/no-qualifications), a home ownership dummy and household income (equivalised using the OECD scale and log transformed). We also include individual-level indicators of health-related lifestyles: a) physical activity, proxied by sports activities (three of more times per week, at least monthly, less frequently/not at all (reference category)); b) a dummy for complying with public health recommendations to consume five or more fruits/vegetables per day to proxy healthy dietary habits (Davillas and Benzeval., 2016); and c) commuting to work, captured by a 4-category variable ordered by the physical activity levels that each commuting method requires; no commuting (reference category), private transport, public transport and active transport). It has been shown that commuting patterns are associated with adiposity levels (Flint et al., 2014).

4. Results

Figure 1 presents unconditional (without accounting for covariates) mean BMI levels (Panel A) and the 90th quantile of the BMI distribution (Panel B) across the nine regions (GOR) in England. Overall, there is a gradually increasing pattern in the mean BMI levels from the South to the North of England (p-value for the joint equality of the mean BMI across regions: 01e-5); however, there are no clear breaks to group regions into different classes by mean BMI (Figure 1, Panel A). Specifically, there are moderate differences in mean BMI between those regions with the higher and the lower mean BMI levels, although specific pairwise comparisons in the mean BMI levels across regions may not be sizable. The highest between-regions difference in the mean BMI levels are observed between London (the GOR with the lowest mean BMI) and North East (the one with the higher BMI), which is 1.10 kg/m^2, equivalent to a difference of 3 kg in body weight for someone on average height (1.67m).

More pronounced are the regional differences at the higher quantiles of the BMI distribution, where the health risks are higher (Figure 1, Panel B). For example, there are sizable differences in the BMI distribution for Yorkshire or North East compared to the two GORs at the South of England (South East and South West). The highest 10% of the BMI distribution (Q90) for Yorkshire and North East corresponds to higher BMI values (36.2 kg/m2 and above) compared to the relevant threshold for the South East and South West (34.3 kg/m^2); this is a difference of about 2 BMI points (kg/m^2), which is 5.6kgs of body weight for an individual of average body height (1.67m tall). These results indicate that the BMI distribution for Yorkshire and North East seems to have longer, heavier tails compared to those for South East and South West.

Figure 1. Mean BMI and the 90th quantile of its distribution by region (GOR).

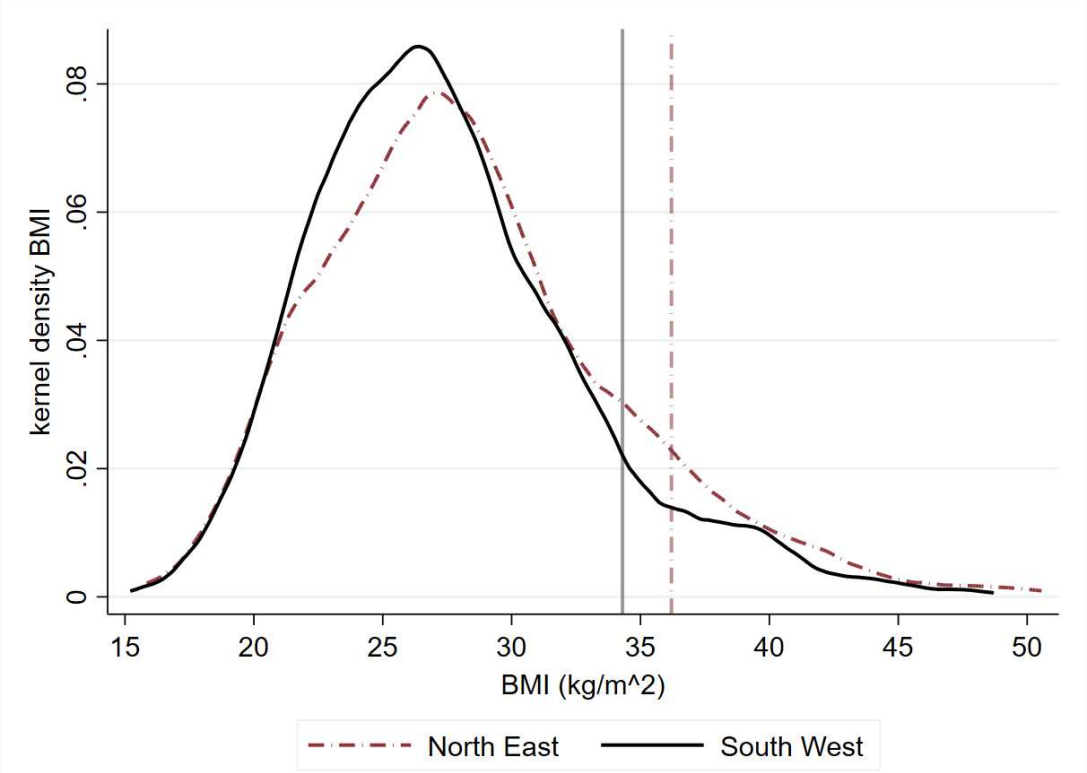


Note: The colour palette is selected to allow for wider colour differences in the case of the Q90 graphs (Panel B, where regional differences are more evident) as opposed to those based on mean levels (Panel A).

As an illustrative example to compare the whole BMI distribution by GOR, Figure 2 presents the BMI distribution for two of the regions with the higher and lower Q90 in Figure 1 (Panel B). In accordance with the findings above, the BMI distribution of North East deviates from its for South West: the former has less mass around moderate BMI levels (20-28kg/m²) and more mass with high and extreme BMI values (compared to South West's BMI distribution), also reflected by the higher BMI value that corresponds to the 90th quantile of the distribution (vertical lines).[9]

[9] Two-sample Kolmogorov-Smirnov tests for equality of distribution functions reject the null hypothesis of equality in distributions across the two regions (p-value: 0.000).

Figure 2. Distribution of BMI for selected regions.



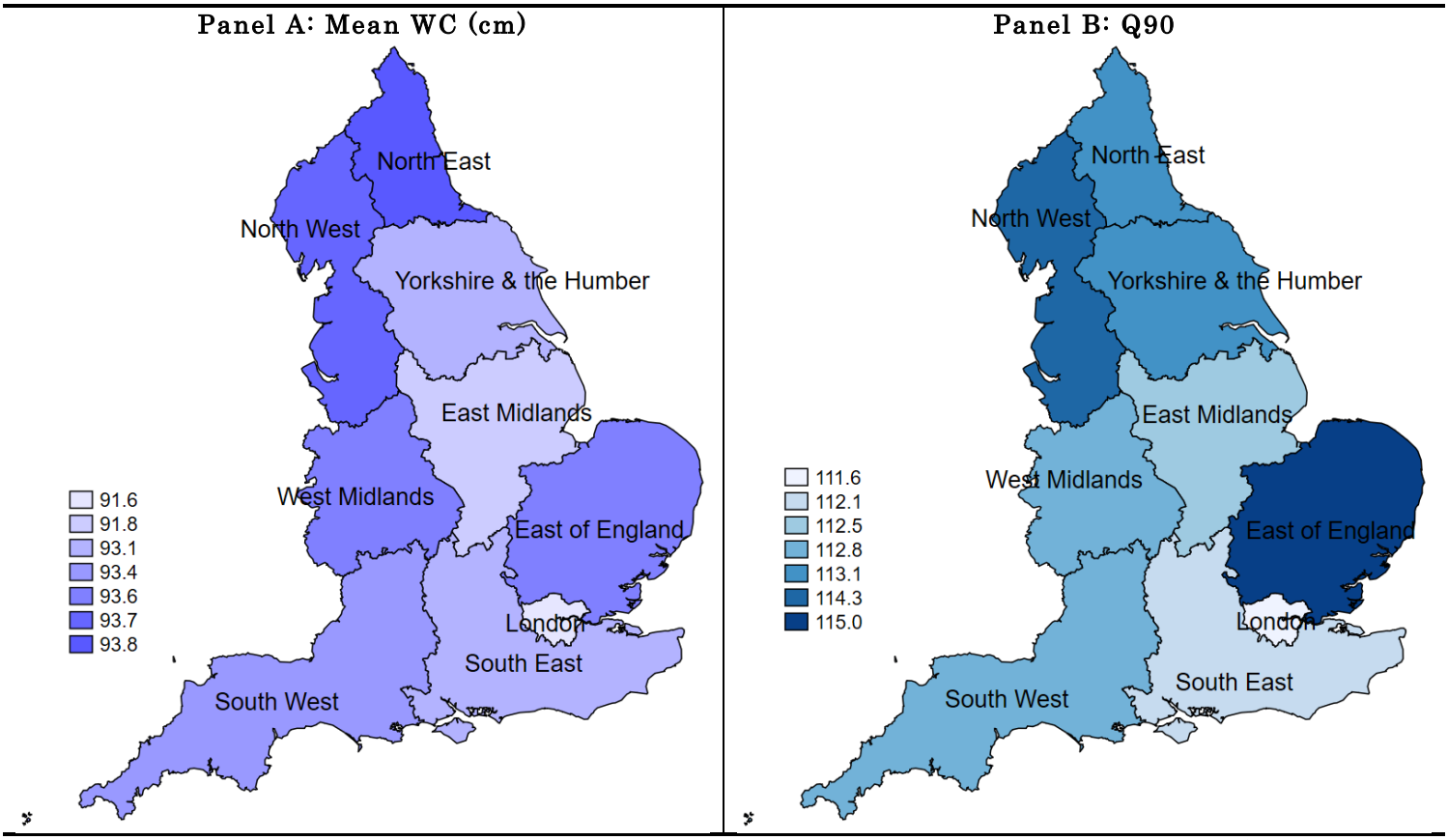
Note: Vertical lines show the 90th percentiles of the two distributions.

Figure 3 shows regional variations in the mean and the 90th quantiles of the WC distribution. Overall, we found only moderate between-region mean WC differences, which are gradually increasing when comparisons are made between regions with the lowest mean WC values and those with increasingly higher mean WC level. The larger differences in mean WC are observed between London and North East (Figure 3, Panel A); this is about 2.2 cm, which is equivalent to the one fifth of the standard deviation of the WC. Turning to the distributional statistics (Q90; Figure 3, Panel B), it is evident that it is not always the regions with the higher mean WC levels (Figure 3, Panel A) that have a WC distribution with the heavier tails and, thus, higher WC values that correspondent to Q90 as well as vice versa. The largest difference is observed between London and East of England, with the WC values at the 90th quantile of distribution

differs by about 3.4 cm between these regions. The WC distribution for East of England is more to the right, has more mass at higher WC values and heavier tails compared to this for London (Figure 4).[10]

Overall, these findings suggest the importance of looking at the whole distribution of BMI and WC measures, rather than solely at mean. Regional variations are more pronounced at the right tails of the adiposity distributions, where health risks are more likely. We explore what are the underlying factors of the observed regional differences in adiposity at the mean and across the whole distribution as well as what drives within-regional variations in adiposity.

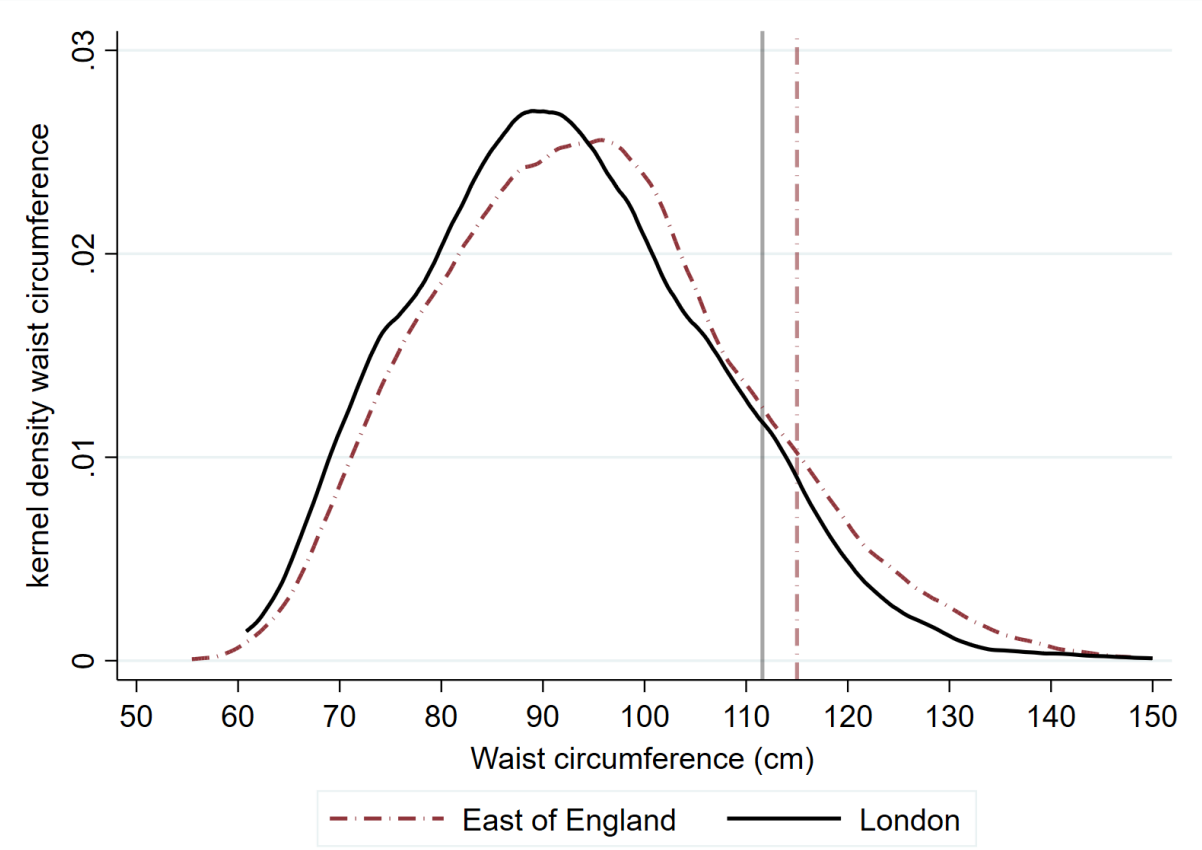
Figure 3. Mean waist circumference and the 90th quantile of its distribution by region.



Note: The colour palette is selected to allow for wider colour differences in the case of the Q90 graphs (Panel B, where regional differences are more evident) as opposed to those based on mean levels (Panel A).

[10] The two-sample Kolmogorov-Smirnov test rejects the null hypothesis of equality of the WC distributions between East of England and London at the 1% level.

Figure 4. Distribution of waist circumference for selected regions.



Note: Vertical lines show the 90th percentiles of the two distributions.

4.1 Analysis at the mean

Table 1 shows the Shapley decomposition results for the contribution of each of the explanatory variables to the variance in BMI explained (expressed in relative terms as a ratio to the overall variance) by our model specifications. After adjusting solely for demographics, GOR-level inequalities in BMI account for about 11.5% of the total explained variance (Specification 1). In other words, about one tenth of the explained variation in mean BMI is attributed to regional differences, after accounting for the effect of pure demographic variations[11].

[11] It should be noted here that the contribution of demographics remains practically unchanged when polynomials in age (interacted with gender, when statistically significant) are used instead of age-gender dummies. This is also the case for our "beyond the mean" analysis presented below

Table 1. Contribution to the share of the variance in BMI explained: Shapley decomposition.

	Specification 1		Specification 2		Specification 3	
	Shapley value	%	Shapley value	%	Shapley value	%
Demographics (age-gender)	0.032	88.57	0.034	71.89	0.030	48.66
Regional dummies (GOR)	0.004	11.43	0.002	4.47	0.002	3.00
Neighbourhood-level characteristics						
Criminality index			0.001	0.80	0.001	0.49
Anxiety index			0.001	2.00	0.001	1.48
Air quality			0.002	3.86	0.002	2.38
Distance to GP			0.001	1.24	0.001	0.69
Income deprivation			0.002	4.53	0.002	2.49
Skills deprivation			0.005	10.50	0.004	5.72
Fast food density			0.000	0.71	0.000	0.65
Individual-level characteristics						
SEP						
Household income					0.001	0.71
Education					0.008	12.74
House ownership					0.001	2.21
Lifestyle						
Fruits/vegetables:5 day					0.000	0.20
Sports activity					0.005	8.04
Commuting behaviour					0.006	10.42
Share of the total variance explained	0.036	100	0.047	100	0.061	100

Augmenting Specification 1 with measured area-level obesogenic characteristics fully accounts for the contribution of the regional differences. The percentage contribution of the regional-level inequalities in BMI is reduced in magnitude from around 11.5% to 4.5%, and the relevant regional dummies are not any more statistically significant (see Table A1, Appendix; p-value=0.168); this indicates that the observed regional variations in BMI are fully accounted for by the role of neighbourhood-level characteristics. Specifically, the Specification 2 shows that the neighbourhood-level SES deprivation, air-quality and proximity of the small-area in terms of distance to GP are the main contributors.

and when WC is used as our adiposity outcome measure. Statistical tests for equality of the age dummies rejected the null hypothesis, indicating the presence of systematic differentiation of the age effect across age groups.

Further augmenting our model with individual-level SEP and lifestyle indicators (Specification 3), reduces the aggregate percentage contribution of all the small-area level obesogenic characterises from 24% (specification 2) to 14% (specification 3). This indicates that individual-level characterises only partially account for the total contribution of the obesogenic environment, with the latter still having a systematic effect (joint significance test, p-value=0.000). Specification 3 shows that the individual-level SEP and lifestyle indicators respectively contribute to about 16% and 19% of the explained variance in BMI, along with the increase in the absolute magnitude of the total variance explained. There is also a reduction in the contribution of demographics from 72% to 49% between specifications 2 and 3. Specifically, individual education is the dominant SEP contributor and physical activity proxies (frequency of sports participation and active commuting) exert the largest contribution from our set of lifestyle factors. Overall, these results indicate that much of the between-region BMI variation (GOR-level BMI differences; Specification1) can be attributed to the differences in the small-area level obesogenic environment, while individual-level SEP and lifestyle also exert an independent role in explaining overall variation in adiposity (after accounting for demographics).

Regarding the mean level of central adiposity (Table 2), demographic factors dominate and regional inequalities in WC account for a much smaller proportion (around 1%) of the explained variation in WC across all model specifications (although regional dummies are statistically significant at the 1% level; Table A2, Appendix). This indicates that despite being statistically significant and unlike our results for BMI, the corresponding regional differences in central adiposity are not large enough to explain a considerable part of the total variation in central adiposity. However, as in the case of BMI, individual-level educational attainment and physical activity are the SEP and

lifestyle factors that make the largest contribution. Unlike the BMI results, the area-level obesogenic environment explains a trivial part of the variation in central adiposity as our analysis "at the mean" shows.

Table 2. Contribution to the share of the variance in WC explained: Shapley decomposition.

	Specification 1		Specification 2		Specification 3	
	Shapley value	%	Shapley value	%	Shapley value	%
Demographics (age-gender)	0.185	98.66	0.188	95.63	0.179	83.21
Regional dummies (GOR)	0.003	1.34	0.002	1.02	0.002	0.88
Neighbourhood-level characteristics						
Criminality index			0.000	0.08	0.001	0.05
Anxiety index			0.001	0.49	0.001	0.34
Air quality			0.001	0.35	0.001	0.27
Distance to GP			0.001	0.39	0.001	0.27
Income deprivation			0.001	0.54	0.001	0.27
Skills deprivation			0.003	1.44	0.002	0.78
Fast food density			0.000	0.05	0.000	0.03
Individual-level characteristics						
SEP						
Household income					0.001	0.21
Education					0.010	4.63
House ownership					0.002	0.74
Lifestyle						
Fruits/vegetables:5 day					0.001	0.26
Sports activity					0.009	4.08
Commuting behaviour					0.009	3.98
Share of the total variance explained	0.188	100	0.196	100	0.215	100

4.2 Random intercept models

Table 3 shows the ICC from our random intercept models. About 14% of the variance in BMI is attributed to unobserved neighbourhood-level effects, which is practically unchanged after accounting for GOR-level differences. The contribution of the neighbourhood-level effects is reduced to about 12% when all individual-level characteristics are accounted for. These results broadly accord with the total contribution of our set of observed neighbourhood-level obesogenic characteristics in the Shapley decomposition (Table 1, Specification 3), confirming the role of the neighbourhood-level characteristics.

Turning to WC, about 12% to 13% of the variance attributed to the total role of the neighbourhood-level characteristics, captured by the LSOA-level random effects (Table 3). These results further confirm the role of the neighbourhood environment, over and above the role of all other characterises accounted in our analysis.

Table 3. Properties of the random intercept models

	No covariates	Accounting for GOR	Accounting for all individual-level characteristics
BMI			
ICC	0.141	0.138	0.118
	[0.117; 0.171]	[0.114; 0.167]	[0.094; 0.148]
Waist circumference			
ICC	0.121	0.118	0.130
	[0.099; 0.146]	[0.097; 0.144]	[0.112; 0.148]

Notes: Intra-class correlation coefficient (ICC). 95% confidence intervals in brackets.

4.3 "Beyond the mean" analysis

Table 4 shows Shapley decomposition results regarding the explained variance, expressed in relative terms as a ratio to the overall variance (equivalent to the R-squared of linear regressions on the RIF), across selected BMI quantiles. There are three main findings. First, adjusting for gender and age only, GOR-level BMI inequalities exert an increasing and quantitatively important contribution of the total BMI variation explained, when moving to higher quantiles of BMI distribution. Specially, the relative contribution of the regional dummies (as a percentage of the total variance explained) increases from 5% to 19% across the BMI distribution. This indicates that, in the case of excess BMI (Q90), about one quarter of the total explained variance is attributed to regional BMI inequalities. Second, augmenting our models with small-area level obesogenic characteristics (Specification 2), fully accounts for the GOR-level differences in BMI (joint test of regional dummies; p-value>0.10). The latter is

more evident towards the right tails of the BMI distribution; for example, in the case of Q90, the contribution of the obesogenic environment is around 50% and the direct contribution of the GOR dummies fall from 19% (Specification 1) to 5% (Specification 2). A more detailed look on the role of the particular neighbourhood-level characteristics at Q90 reveal that the small-area level SEP is the dominant contributor, while smaller but still important contributions are attributable to air quality and anxiety levels (detailed results available upon request). Third, accounting for individual level SEP and lifestyle characteristics, we find that part of the role of the neighbourhood obesogenic environment may be partially mediated by the individual-level characteristics. Specifically, Specification 3 shows that the contribution of SEP and lifestyle increases towards the right tails of the BMI distribution, with the role of demographics being much less evident. For example, the percentage contribution of SEP (lifestyles) increases from 11% (9%) at the 25th quantile to 18% (19%) at the 90th BMI quantile. This indicates that the underlying sources of excess BMI is not solely attributed to age and gender inequalities in BMI, but the neighbourhood environment has an independent role over and above the role of SEP and lifestyle factors.

Table 4. Contribution of covariates across quantiles of the BMI distribution.

	Q25	Q50	Q75	Q90
Specification 1	%	%	%	%
Demographics	94.7	92.1	85.0	81.3
Regional differences (GOR)	5.3	7.9	15.0	18.7
Total	**100.0**	**100.0**	**100.0**	**100.0**
Specification 2				
Demographics	89.5	80.8	61.6	43.0
Regional differences (GOR)	3.7	3.9	5.0	3.7
Neighbourhood-level obesogenic characteristics	6.8	15.3	33.4	53.3
Total	**100.0**	**100.0**	**100.0**	**100.0**
Specification 3				
Demographics	72.0	56.3	36.3	30.1
Regional differences (GOR)	3.1	3.0	3.1	2.6
Neighbourhood-level obesogenic characteristics	5.0	9.4	16.5	30.7
SEP	11.3	16.0	21.2	17.7
Lifestyle	8.7	15.4	22.9	18.9
Total	**100.0**	**100.0**	**100.0**	**100.0**

As in the case of analysis at the mean, results for central adiposity differ from those for BMI in terms of the contribution of regional differences (Table 5). After accounting for our set of covariates, the regional inequalities in waist circumference account for a small percentage of variance in central adiposity, as in the case of analysis at the mean. Age and gender seem to explain a larger proportion of the total variance in WC, compared to BMI, even in the case of our full model specification. However, the percentage contribution of age and gender reduced at higher WC quantiles, with the contribution of the neighbourhood-level obesogenic characteristics becoming more evident. These have a direct association over and above the role of individual-level SEP and lifestyle. Of particular interest, unlike our analysis "at the mean" (Table 2), the neighbourhood environment seems to be much more relevant at the tails of the WC distribution (with a percentage contribution of about 12%), independent from all other factors accounted for (Specification 3). The neighbourhood-level SEP, air quality and anxiety levels are the first (individual contribution of 6%), second (around 2%) and third (around 2%) more important contributors, respectively.

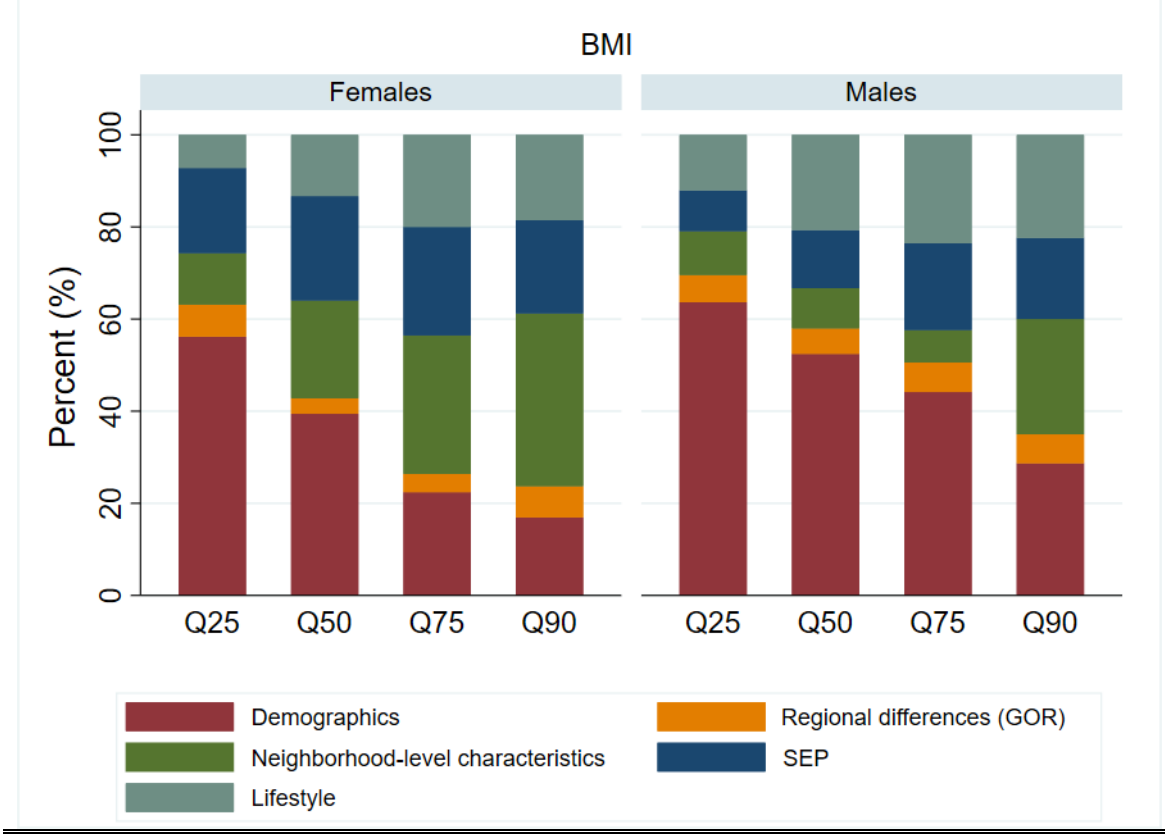
Table 5. Contribution of covariates across quantiles of the WC distribution.

	Q25	Q50	Q75	Q90
Specification 1	%	%	%	%
Demographics	98.7	98.9	99.0	95.1
Regional differences (GOR)	1.3	1.1	1.0	4.9
Total	**100.0**	**100.0**	**100.0**	**100.0**
Specification 2				
Demographics	97.7	96.5	92.9	75.8
Regional differences (GOR)	1.4	0.9	0.5	2.9
Neighbourhood-level obesogenic characteristics	1.0	2.7	6.5	21.3
Total	**100.0**	**100.0**	**100.0**	**100.0**
Specification 3				
Demographics	89.4	83.6	73.3	56.6
Regional differences (GOR)	1.5	0.8	0.4	1.9
Neighbourhood-level obesogenic characteristics	0.6	1.5	3.5	12.1
SEP	4.4	6.0	7.0	10.2
Lifestyle	4.1	8.2	15.8	19.2
Total	**100.0**	**100.0**	**100.0**	**100.0**

Analysis by gender

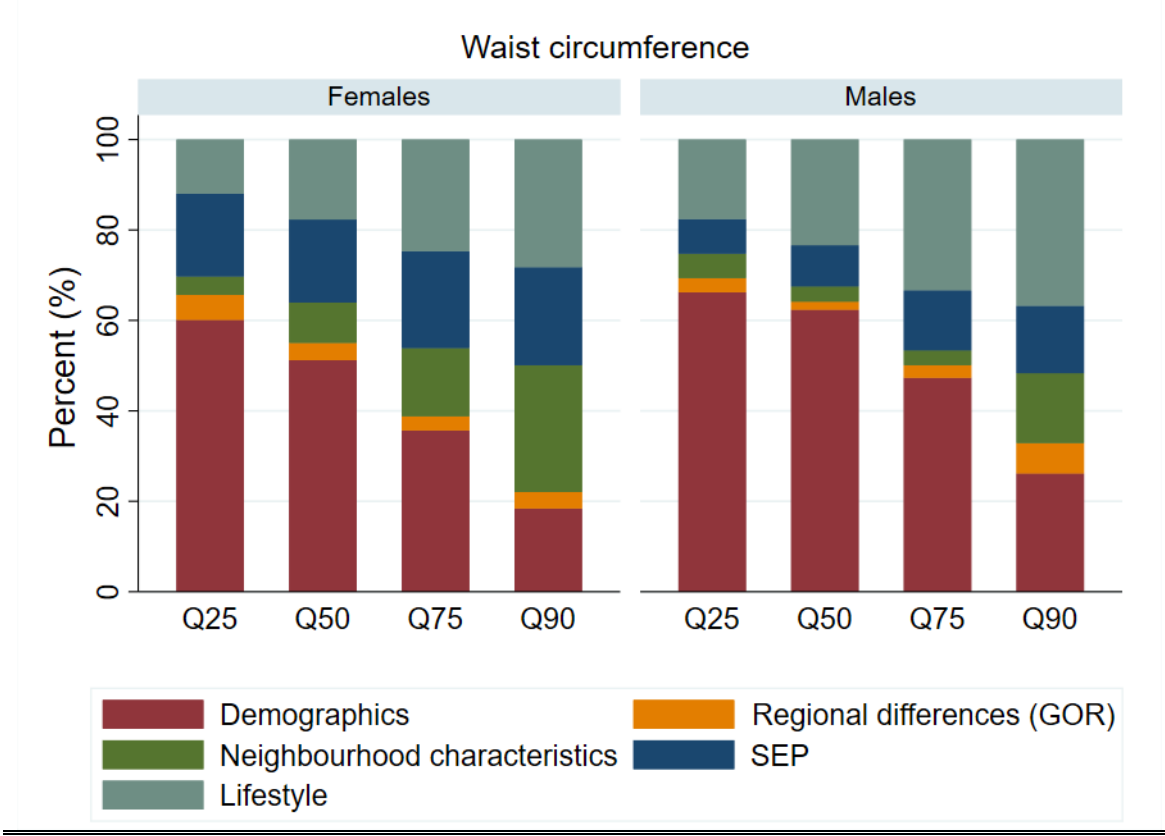
Figures 5 and 6 present the Shapley decomposition results across quantiles of the BMI and WC distributions for our full model specification (Specification 3), separately by gender. Turning to the results for BMI (Figure 5) , the contribution of demographics become less relevant towards the right tails of the BMI distribution for both males and females. On the other hand, the neighbourhood-level obesogenic factors are more relevant at the higher tails of the distribution; their contribution is larger for females, accounting for about 40% of the total variance explained at Q90, as opposed to 25% for the case of males (Figure 5). Similarly, individual's SEP and lifestyle exert higher contributions towards the right tails of the BMI distribution, which are similar in magnitude across gender.

Figure 5. Contribution of covariates across quantiles of the BMI distribution: Analysis by gender.



The relevant results for WC (Figure 6) show that, as in the case of our pooled sample (Table 5), the role of demographics decreases moving towards to higher quantiles of the distribution, while the role of the obesogenic environment, individual's SEP and lifestyle become more evident for both men and women. The contribution of the obesogenic environment is more relevant for females at higher WC quantiles as opposed to males, while the opposite is the case for lifestyle. Overall, despite some gender differences in the magnitude of the contribution of the different factors, the gender-stratified analysis confirms our pooled sample results.

Figure 6. Contribution of covariates across quantiles of the WC distribution: Analysis by gender.



5. Conclusion

Using representative English data, we explore the contribution of individual-level characteristics and small-area level obesogenic environment on adiposity. The observed GOR-level differences in BMI, that are more evident towards the right tails of its distribution (explaining up to 19% of the total BMI variance), are fully accounted for by the neighbourhood obesogenic environment. Although part of the role of the neighbourhood environment is accounted for by the individual-level differences in lifestyle and SEP, the former still exerts an independent role of about 32% of the explained variance at the right tails of the BMI distribution. Regarding central adiposity, the neighbourhood obesogenic environment is much more relevant at the tails of the WC distribution, where accounts for about 12% of the explained variance. Our evidence that the role of obesogenic environment on excess adiposity is more pronounced for women than men complements existing literature that has argued that women are more vulnerable to obesogenic environmental risks than men (Lovasi et al., 2009; Shapira, 2013).

Our results show that the neighborhood obesogenic environment exerts a sizable contribution to variation in adiposity, even though its role partially explained by the observed individual-level characteristics and lifestyles; the latter suggests that these individual-level characteristics may be key mediators. Exploring the association of the obesogenic environment with adiposity levels (over and above other individual characteristics) is important for local-area policymaking, as they are modifiable environmental characteristics. We find that neighborhoods with disadvantaged environments, such as those of higher SEP deprivation, geographic barriers and lower air quality levels may influence individuals' adiposity levels, especially at the higher tails of its distribution. Our evidence accords with existing research that has found

significant associations between adiposity measures and environmental risk factors (e.g., An et al., 2018, Chaparro et al., 2018, Lopez et al., 2007, Raftopoulou et al., 2017, Santana et al., 2009) and extends them by quantifying the relative contribution of the obesogenic environment, as opposed to individual-level characteristics, across the whole distribution of the adiposity measures.

Moving towards the right tails of BMI and WC distribution, the contribution of the neighbourhood-level obesogenic environment become much more pronounced, while the relative contribution of demographics is reduced in magnitude. This highlights the importance of considering analysis across the whole distribution of adiposity as these results would have been masked if focused solely at the mean. The set of neighborhood-level characteristics that play the most important role at the right tails of the adiposity distribution are environmental constraints that may reduce individuals' physical activity, promote unhealthy or stress eating patterns. For example, our results suggesting that being a resident of a more SEP-deprived neighborhood is positively associated with higher adiposity, accord with arguments that the better built infrastructure and increased food options (available to more affluent neighborhoods) may result in more physical activity and/or reduced calorie intake (Lopez, 2007; Santana et al., 2009). Given the modifiable nature of these characteristics, consideration may need to be given to the built and social environment to further encourage active lifestyles and healthy eating and, thus, reduce the risk of obesity.

Our findings are relevant to the English local authorities, in their relatively new role (as from 2014) as leaders for local population health. This highlights that, at least for BMI, inequalities within regions are more evident compared to inequalities between them; the latter is consistent with recent results on disability and mortality for England (Newton

et al., 2015). Using data collected between 2010 and 2012, just before the abolishment of the strategic health authorities and the enhanced role of the local authorities, these results are supportive to the need for focusing at the smaller-area level to reduce regional-level inequalities in adiposity in particular. Given our findings, a recent call for action at the local authority level to develop policies and initiatives for healthier food environments as well as planning restrictions for more activity-friendly and walkable neighbourhoods need to be further stimulated (Public Health England, 2017).

We also find that individual-level SEP and lifestyle exert independent contributions, over and above the role of the obesogenic environment, with their relative contribution becoming more pronounced towards the right tails of adiposity distribution. These results accord and extend existing research showing a steeper SEP-health gradient at the tails of the distribution, where health risks are more evident (e.g., Davillas et al., 2019; Carrieri and Jones, 2017). Overall, our results suggest that policies that aim to tackle excess adiposity should address both people and places in order to be more fruitful (Blüher, 2019; Santana et al., 2009). Excess adiposity has multi-factorial causes beyond any biological causes and mechanisms. Although our analysis confirms existing literature on the prominent role of individual level characteristics as risk factors for excess adiposity (e.g., Brunello et al., 2013; Davillas and Benzeval, 2016; Madden, 2012), the neighbourhood-level environment still plays an independent role. Efforts to reduce the obesity burden, therefore, require approaches that combine individual-based interventions with changes in small-area and neighbourhood environmental characteristics.

There are some limitations to our study, however. First, our analysis does not aim to address the potential endogeneity biases or establish any causal links between our

explanatory variables and adiposity measures. Second, although our results are supportive to the need for focusing at the small area level to reduce regional-level inequalities in adiposity. They are based on data collected just before the abolishment of the strategic health authorities and the enhanced role of the local, smaller area-level authorities. Future research, when more up to date adiposity data linked to detailed neighbourhood-level characteristics are available is needed to assess the effectiveness of the small-area local authorities in their new duties as leaders for the local population health. Third, unlike the other neighbourhood-level characteristics, fast food density is available at an aggregate level rather than at the small area-level. We are thus unable to disentangle whether its limited role observed in our study is an artefact of an aggregation effect or accords with existing studies found weak or non-systematic associations between excess adiposity and local availability of fast-food outlets (Dunn et al., 2012; Feng et al., 2010; Salois, 2012)[12].

References

Allcott, H., Diamond, R., Dubé, J. P., Handbury, J., Rahkovsky, I., Schnell, M. (2019). Food deserts and the causes of nutritional inequality. *The Quarterly Journal of Economics*, 134(4), 1793-1844.

An, R., Ji, M., Yan, H., Guan, C. (2018). Impact of ambient air pollution on obesity: a systematic review. *International Journal of Obesity*, 42(6), 1112-1126.

Baker, C., 2018. Obesity statistics. Briefing paper 3336. House of Commons, London, UK.

Bambra, C., Barr, B., Milne, E. (2014). North and South: addressing the English health divide. *Journal of Public Health*, 36(2), 183–186.

Barrea, L., Savastano, S., Di Somma, C., Savanelli, M. C., Nappi, F., Albanese, L., ..., Colao, A. (2017). Low serum vitamin D-status, air pollution and obesity: A dangerous liaison. *Reviews in Endocrine and Metabolic Disorders*, 18(2), 207-214.

[12] It should be noted here that we found virtually no change to our results (and the contribution of all other covariates) in the case that fast food outlet density is excluded from our set of covariates.

Baum II, C.L., Ruhm, C.J., 2009. Age, socioeconomic status and obesity growth. *Journal of Health Economics*, 28, 635-648.

Blüher, M. (2019). Obesity: global epidemiology and pathogenesis. Nature Reviews Endocrinology, 15(5), 288.

Booth, K. M., Pinkston, M.M., Poston, W.S.C. (2005). Obesity and the built environment. *Journal of the American Dietetic Association*, *105*(5), 110-117.

Brunello, G., Fabbri, D., Fort, M. (2013). The causal effect of education on body mass: Evidence from Europe. *Journal of Labor Economics*, 31(1), 195-223.

Burgess, C., Wright, A. J., Forster, A. S., Dodhia, H., Miller, J., Fuller, F., ..., Gulliford, M. C. (2015). Influences on individuals' decisions to take up the offer of a health check: a qualitative study. *Health Expectations*, 18(6), 2437-2448.

Carrieri, V., Jones, A.M. (2017). The income–health relationship 'beyond the mean': New evidence from biomarkers. *Health economics*, *26*(7), 937-956.

Carroll-Scott, A., Gilstad-Hayden, K., Rosenthal, L., Peters, S. M., McCaslin, C., Joyce, R., Ickovics, J.R. (2013). Disentangling neighborhood contextual associations with child body mass index, diet, and physical activity: the role of built, socioeconomic, and social environments. *Social Science & Medicine*, 95, 106-114.

Cawley, J., Maclean, J. C., Hammer, M., Wintfeld, N., (2015). Reporting error in weight and its implications for bias in economic models. *Economics & Human Biology*, 19, 27-44.

Chaparro, M.P., Benzeval, M., Richardson, E., & Mitchell, R. (2018). Neighborhood deprivation and biomarkers of health in Britain: the mediating role of the physical environment. BMC Public Health, 18(1), 801.

Chou, S.Y., Grossman, M., Saffer, H., (2004). An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. *Journal of Health Economics*, 23, 565-587.

Costa-Font, J., Gil, J. (2008). What lies behind socio-economic inequalities in obesity in Spain? A decomposition approach. *Food policy*, *33*(1), 61-73.

Cutler, D.M., Glaeser, E.L., Shapiro, J.M., (2003). Why have Americans become more obese? *Journal of Economic Perspectives*, *17*, 93-118.

Davillas, A., Benzeval, M., 2016. Alternative measures to BMI: Exploring income-related inequalities in adiposity in Great Britain. *Social Science & Medicine*, 166, 223-232.

Davillas, A., Jones, A. M., Benzeval, M. (2019). The income-health gradient: evidence from self-reported health and biomarkers in understanding society. In Panel data econometrics (pp. 709-741). Academic Press.

Davillas, A., Jones, A.M., (2020). Ex ante inequality of opportunity in health, decomposition and distributional analysis of biomarkers. *Journal of Health Economics*, 69, 102251.

Davillas, A., Pudney, S. (2019). Biomarkers as precursors of disability. *Economics & Human Biology*, 100814.

DHSC, 2013. Guidance: The health and care system explained. https://www.gov.uk/government/publications/the-health-and-care-system-explained/the-health-and-care-system-explained

Drewnowski, A., Aggarwal, A., Tang, W., Hurvitz, P. M., Scully, J., Stewart, O., & Moudon, A. V. (2016). Obesity, diet quality, physical activity, and the built environment: the need for behavioral pathways. *BMC Public Health*, *16*(1), 1153.

Dunn, R.A., Sharkey, J.R., Horel, S. (2012). The effect of fast-food availability on fast-food consumption and obesity among rural residents: an analysis by race/ethnicity. *Economics & Human Biology*, 10(1), 1-13.

Feng, J., Glass, T. A., Curriero, F. C., Stewart, W. F., Schwartz, B. S. (2010). The built environment and obesity: a systematic review of the epidemiologic evidence. *Health & Place*, 16(2), 175-190.

Ferreira, F.H.G., J. Gignoux. (2014). The measurement of educational inequality: achievement and opportunity. *World Bank Economic Review*, 28, 210–246.

Firpo, S., Fortin, N., Lemieux, T., (2009). Unconditional quantile regressions. *Econometrica*, 77, 953-973.

Flint, E., Cummins, S., Sacker, A.,2014. Associations between active commuting, body fat, and body mass index: population based, cross sectional study in the United Kingdom. *BMJ*, 349, g4887.

Flouri, E., Mavroveli, S., Midouhas, E., (2013). Residential mobility, neighbourhood deprivation and children's behaviour in the UK. *Health & Place*, 20, 25-31.

GBD 2015 Obesity Collaborators, (2017). Health effects of overweight and obesity in 195 countries over 25 years. *New England Journal of Medicine*, *377*(1), 13-27.

Green, M.A., Subramanian, S.V., Razak, F., (2016). Population-level trends in the distribution of body mass index in England, 1992–2013. *Journal of Epidemiology and Community Health*, 70(8), 832-835.

Israeli, O. (2007). A Shapley-based decomposition of the R-square of a linear regression. *The Journal of Economic Inequality*, 5(2), 199-212.

Jolliffe, D., (2011). Overweight and poor? On the relationship between income and the body mass index. *Economics & Human Biology*, 9, 342-355.

Kahn, H. S., Tatham, L. M., Pamuk, E.R., Heath Jr, C.W. (1998). Are geographic regions with high income inequality associated with risk of abdominal weight gain? *Social Science & Medicine*, 47(1), 1-6.

Lhila, A. (2011). Does access to fast food lead to super-sized pregnant women and whopper babies? *Economics & Human Biology*, 9(4), 364-380.

Lopez, R.P. (2007). Neighborhood risk factors for obesity. *Obesity*, 15(8), 2111-2119.

Lovasi, G. S., Hutson, M. A., Guerra, M., & Neckerman, K. M. (2009). Built environments and obesity in disadvantaged populations. Epidemiologic Reviews, 31(1), 7-20.

Madden, D. (2012). A profile of Obesity in Ireland, 2002–2007. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(4), 893-914.

McFall, S. L., Petersen, J., Kaminska, O., Lynn, P. (2014). Understanding society: the UK household longitudinal study. Waves 2 and 3 Nurse Health Assessment, 2010 – 2012. Guide to Nurse Health Assessment. Institute for Social and Economic Research, University of Essex, UK

McLennan, D., Barnes, H., Noble, M., Davies, J., Garratt, E., Dibben, C., (2011). The English indices of deprivation 2010. London: Department for Communities and Local Government.

Must A, Spadano J, Coakley E.H., Field A.E., Colditz G., Dietz W.H., (1999). The disease burden associated with overweight and obesity. *JAMA*, 282, 1523-1529.

Newton, J.N., Briggs, A.D., Murray, C.J., Dicker, D., Foreman, K. J., Wang, H., ... Vos, T., (2015). Changes in health in England, with analysis by English regions and areas of deprivation, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 386, 2257-2274.

NHS Digital, (2018a). Statistics on Obesity, Physical Activity and Diet-England. NHS Digital. https://files.digital.nhs.uk/publication/0/0/obes-phys-acti-diet-eng-2018-rep.pdf

NHS Digital, (2018b). Health Survey for England 2017. NHS Digital. https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2017#related-links

Pak, T. Y., Ferreira, S., Colson, G. (2016). Measuring and tracking obesity inequality in the United States: Evidence from NHANES, 1971-2014. *Population Health Metrics*, 14(1), 12.

Papas, M. A., Alberg, A. J., Ewing, R., Helzlsouer, K. J., Gary, T. L., Klassen, A. C. (2007). The built environment and obesity. *Epidemiologic Reviews*, 29(1), 129-143

Powell, L. M., Chaloupka, F.J. (2009). Food prices and obesity: evidence and policy implications for taxes and subsidies. *The Milbank Quarterly*, 87(1), 229-257.

Power, C., Lake, J.K., Cole, T.J. (1997). Body mass index and height from childhood to adulthood in the 1958 British born cohort. *The American Journal of Clinical Nutrition*, *66*(5), 1094-1101.

Public Health England (2014). Density of fast food outlets in England: data by local authority and ward.

Public Health England (2017). Guidance. Health matters: obesity and the food environment.

Raftopoulou, A., (2017). Geographic determinants of individual obesity risk in Spain: A multilevel approach. *Economics & Human Biology*, 24, 185-193.

Rahkovsky, I., Gregory, C.A. (2013). Food prices and blood cholesterol. *Economics & Human Biology*, *11*(1), 95-107.

Salois, M.J. (2012). Obesity and diabetes, the built environment, and the 'local'food economy in the United States, 2007. *Economics & Human Biology*, 10(1), 35-42.

Santana, P., Santos, R., Nogueira, H. (2009). The link between local environment and obesity: a multilevel analysis in the Lisbon Metropolitan Area, Portugal. *Social Science & Medicine*, 68(4), 601-609.

Shapira, N. (2013). Women's higher health risks in the obesogenic environment: a gender nutrition approach to metabolic dimorphism with predictive, preventive, and personalised medicine. *EPMA Journal*, *4*(1), 1.

Shelton, N.J., (2009). Regional risk factors for health inequalities in Scotland and England and the "Scottish effect". *Social Science & Medicine*, 69, 761-767.

Shorrocks, A.F., (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality*, 11, 1-28.

Stafford, M., Cummins, S., Ellaway, A., Sacker, A., Wiggins, R. D., Macintyre, S. (2007). Pathways to obesity: identifying local, modifiable determinants of physical activity and diet. *Social Science & Medicine*, 65(9), 1882-1897.

Stifel, D.C., Averett, S.L. (2009). Childhood overweight in the United States: A quantile regression approach. *Economics & Human Biology*, 7(3), 387-397.

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(9793), 804-814.

Torres, S.J., Nowson, C.A., (2007). Relationship between stress, eating behavior, and obesity. *Nutrition*, 23, 887-894.

Vallejo-Torres, L., Morris, S., (2010). The contribution of smoking and obesity to income-related inequalities in health in England. *Social Science & Medicine*, 71, 1189-1198.

Yu, E., Lippert, A.M. (2016). Neighborhood crime rate, weight□related behaviors, and obesity: a systematic review of the literature. *Sociology Compass*, 10(3), 187-207.

Appendix

Table A1. BMI regression models: OLS estimates

	Specification 1	Specification 2	Specification 3
North East	0.971***	0.417	0.403
	(0.257)	(0.284)	(0.283)
North West	0.783***	0.337	0.300
	(0.194)	(0.227)	(0.227)
Yorkshire	0.522**	-0.064	-0.117
	(0.206)	(0.237)	(0.236)
East Midlands	0.256	-0.114	-0.166
	(0.200)	(0.211)	(0.210)
West Midlands	0.706***	0.254	0.197
	(0.214)	(0.218)	(0.216)
East of England	0.194	-0.011	-0.044
	(0.193)	(0.196)	(0.195)
London	0.043	0.040	0.176
	(0.223)	(0.249)	(0.250)
South West	-0.181	-0.266	-0.296
	(0.188)	(0.194)	(0.194)
Neighborhood-level characteristics			
Crime deprived		0.065	0.062
		(0.153)	(0.152)
High anxiety levels		0.058	0.037
		(0.166)	(0.165)
Sulphur dioxide level		6.872***	6.562***
		(2.121)	(2.096)
Distance to GP		0.099***	0.077**
		(0.036)	(0.036)
Income deprivation		1.295	0.345
		(1.175)	(1.196)
Skills deprivation		0.022***	0.016***
		(0.006)	(0.006)
Fast food density		0.261	0.394
		(0.307)	(0.302)
Individual-level characteristics			
Ln(income)			0.148
			(0.133)
Post-compulsory/ no tertiary			0.545***
			(0.162)
Secondary qualification			0.739***
			(0.142)
No qualification			0.892***
			(0.204)
House ownership			-0.627***
			(0.146)
Fruits/vegetables:5/day			-0.088
			(0.126)
Sports activity: 3/week			-0.655***
			(0.151)
Sports activity: monthly			-0.372***
			(0.127)
Commuting: private			0.153
			(0.151)
Commuting: public			-0.917***
			(0.262)
Commuting: active			-0.940***
			(0.218)
Joint significance tests			
Regional dummies (p-values)	0.000	0.168	0.122
Age-gender dummies (p-values)	0.000	0.000	0.000
Sample size	12,271	12,271	12,271

***P < 0.01; **P < 0.05; *P < 0.10.

Notes: Estimates are weighted using UKHLS nurse visits sample weights.

Table A2. Waist circumference regression models: OLS estimates

	Specification 1	Specification 2	Specification 3
North East	0.829	-0.440	-0.437
	(0.626)	(0.685)	(0.681)
North West	1.233**	0.246	0.167
	(0.487)	(0.555)	(0.551)
Yorkshire	-0.039	-1.188**	-1.316**
	(0.507)	(0.579)	(0.571)
East Midlands	-1.072**	-1.866***	-1.967***
	(0.506)	(0.531)	(0.525)
West Midlands	1.133**	0.141	-0.008
	(0.517)	(0.528)	(0.524)
East of England	0.183	-0.258	-0.380
	(0.488)	(0.493)	(0.491)
London	-0.248	-0.128	0.079
	(0.538)	(0.604)	(0.605)
South West	0.058	-0.188	-0.234
	(0.474)	(0.488)	(0.485)
Neighborhood-level characteristics			
Crime deprived		0.160	0.110
		(0.363)	(0.360)
High anxiety levels		0.415	0.336
		(0.404)	(0.398)
Sulphur dioxide level		13.393***	12.506**
		(5.167)	(5.094)
Distance to GP		0.209**	0.156
		(0.094)	(0.095)
Income deprivation		3.085	-0.323
		(2.782)	(2.807)
Skills deprivation		0.049***	0.033**
		(0.013)	(0.014)
Fast food density		0.121	0.476
		(0.725)	(0.708)
Individual-level characteristics			
Ln(income)			0.555
			(0.352)
Post-compulsory/ no tertiary			1.018**
			(0.399)
Secondary qualification			1.966***
			(0.351)
No qualification			2.394***
			(0.496)
House ownership			-1.706***
			(0.359)
Fruits/vegetables:5/day			-0.484
			(0.309)
Sports activity: 3/week			-3.073***
			(0.389)
Sports activity: monthly			-1.572***
			(0.314)
Commuting: private			-0.490
			(0.372)
Commuting: public			-2.644***
			(0.669)
Commuting: active			-2.595***
			(0.531)
Joint significance tests			
Regional dummies (p-values)	0.001	0.007	0.003
Age-gender dummies (p-values)	0.000	0.000	0.000
Sample size	12,271	12,271	12,271

***P < 0.01; **P < 0.05; *P < 0.10.

Notes: Estimates are weighted using UKHLS nurse visits sample weights.