Discussion Paper Series

IZA DP No. 18770

July 2026

Proximity to Fast-Food Outlets and Adolescent z-BMI: Accounting for Persistent Health Dynamics

Yu Aoki-Beattie
University of Aberdeen
and IZA@LISER

Wiji Arulampalam
University of Warwick,
CAGE and IZA@LISER

Neil Lloyd
University of St Andrews

IZA@LISER NETWORK
ISSN: 2365-9793

IZA DP No. 18770 July 2026

Proximity to Fast-Food Outlets and Adolescent z-BMI: Accounting for Persistent Health Dynamics*

Abstract

We estimate the causal effect of exposure to fast-food outlets on adolescent z-BMI using the UK Millennium Cohort Study linked to comprehensive geospatial data. We develop a novel approach to modelling z-BMI persistence by clustering early childhood z-BMI trajectories, allowing us to distinguish baseline obesity susceptibility from contemporaneous environmental effects. Identification exploits the near-universal transition from primary to secondary school, generating plausibly exogenous variation in exposure to fast-food outlets around schools. We find that adolescents with at least one major-brand outlet near school have, on average, a 0.158 standard-deviation higher z-BMI. Effects decline with distance, are limited around the home, and do not extend to other food outlets.

JEL classification

I12, I18, R14

Keywords

adolescent obesity, body mass index, fast food, school food environment

Corresponding author

Yu Aoki-Beattie
y.aoki-beattie@abdn.ac.uk

* We thank the UCL Centre for Longitudinal Studies for collecting and managing the data, and the UK Data Service for providing access. We thank Sushil Mathew for excellent research assistance in data preparation and the construction of GIS-based measures of fast-food proximity. Responsibility for the analysis and interpretation remains solely with the authors. We thank participants at the American-European Health Economics Study Group (Barcelona), the SUFE-SDU-Warwick Workshop (Shanghai), the West China School of Public Health, the Scotland and Northern England Applied Microeconomics Conference (Crieff), and seminars at Aberdeen and Warwick for valuable comments. Financial support from CAGE, Department of Economics, University of Warwick, and the British Academy (Grant No. SG2122/211337) is gratefully acknowledged.

1 Introduction

The global increase in childhood obesity and its association with fast food have received considerable attention in both the health and economics literature (Jiang et al., 2023; UNICEF, 2025). Given the well documented persistence of obesity throughout life (Simmonds et al., 2016), the global increase in pre-adolescent obesity represent not only an immediate public health concern, but also a warning signal for future adolescent and adult health outcomes.

The persistence of obesity also poses a fundamental empirical challenge for identifying environmental determinants of adolescent and adult obesity. While a large literature has examined the relationship between fast-food availability and obesity, commonly measured by body mass index (BMI), much of the existing evidence does not explicitly account for dynamic persistence in BMI. This raises concerns about identifying the causal effect of fast-food availability on adolescent obesity, when the availability is correlated with pre-existing weight trajectories. This paper addresses these concerns by developing a new approach to modelling persistence in BMI, and combines with a research design that exploits plausibly exogenous variation in adolescent exposure to fast-food outlets around schools.

Our first contribution is methodological. Using data from the UK Millennium Cohort Study (MCS), we capture persistence in BMI z-score (z-BMI) by constructing obesity risk profiles based on observed z-BMI trajectories in early childhood, applying unsupervised statistical learning methods. Specifically, we use k-means clustering to identify distinct groups of pre-adolescent z-BMI trajectories, allowing us to classify individuals according to their history of overweight or obesity. These clusters summarise biologically and behaviourally persistent components of weight dynamics—such as adipocyte accumulation and habit formation (MacLean et al., 2015; Cleo et al., 2017; Kluser and Pons, 2026). Incorporating these obesity risk profiles into the empirical framework allows us to distinguish baseline susceptibility from contemporaneous environmental effects. Moreover, by constructing the clusters using outcomes observed prior to the estimation period, we mitigate concerns related to endogenous initial conditions that commonly arise in studies of persistent outcomes. A further advantage of our approach is that it accounts for persistence without requiring a panel with equally spaced observations, as is necessary when accounting for persistence using a dynamic panel data model. Consequently, the methodology is applicable to unequally spaced longitudinal datasets in which continuous outcomes, such as cognitive test scores, are observed at irregular intervals.

Our second contribution concerns identification. Existing studies typically rely on within-individual or within-area variation in fast-food exposure, often using two-way fixed-effects models that do not explicitly model dynamic persistence in BMI. In contrast, we model the persistence, and exploit plausibly exogenous variation in school-based exposure induced by the transition from primary to secondary school. This transition creates variation in school-level fast-food exposure that is likely unrelated to individual or household preference for fast food, enabling us to isolate the causal effect of fast-food proximity on adolescent z-BMI.[1]

Our empirical analysis combines MCS's confidential school and residence location information with the geo-coded food-outlet locations from Ordnance Survey. We classify fast-food outlets into three categories that differ in branding and business models: major-brand outlets (e.g., McDonald's), non-major-brand outlets (e.g., kebab shops), and fish and chip shops. Exposure is measured using road-network buffers.

Our main results indicate that adolescents exposed to at least one major-brand fast-food outlet within 400 metres of their secondary school have z-BMI levels that are approximately 0.158 standard deviations higher than adolescents without such exposure. In contrast, proximity to fish and chip shops or non-major-brand fast-food outlets has no statistically significant effect. The estimated effects are highly localised: namely, we find significant effects only within a 400-metre buffer zone, corresponding roughly to a five-minute walk, and no statistically significant effects at larger distances. These findings are consistent with evidence from the United States reported by Currie et al. (2010), suggesting that immediate proximity to major-brand fast food plays an important role in shaping adolescent weight outcomes. The results are robust to a wide range of alternative specifications, including models incorporating individual fixed effects, alternative measures of fast-food exposure, and alternative sample specifications.

The findings are also consistent with evidence that increased access to fast food increases consumption among children and adolescents (Jia et al., 2021), potentially through mechanisms related to temptation, visual cues, and self-control problems (Laibson, 2001; Boswell and Kober, 2016). Moreover, fast-food consumption among adolescents often has a social dimension (van

[1] In food environment research, proximity refers to the spatial closeness between an individual's residence or school and the food outlet, typically operationalised using distance or travel time measures. In contrast, access is commonly used to encompass both geographical accessibility and economic affordability.

Rongen et al., 2020), potentially reinforcing habitual consumption patterns and amplifying the effects of local access.

The existing causal literature on fast-food availability and child obesity generally follows one of three approaches.[2] First, several studies use observational panel data to exploit within-individual or within-area variation in exposure to fast food near homes or schools (Powell, 2009; Currie et al., 2010; Abrahamsson et al., 2023; Libuy et al., 2023). These studies broadly report positive effects of fast-food proximity on body-weight outcomes. Second, a smaller literature uses natural experiments or policy reforms: Dunn et al. (2021) exploit school transitions in Arkansas and find insignificant effects of school-related fast-food exposure, whereas Han et al. (2020) exploit random allocation into public housing in New York City and find positive effects of fast-food exposure on childhood obesity. Evaluating planning restrictions on new fast-food outlets in England, Xiang et al. (2024) find that the policy reduced obesity prevalence only in deprived areas. Third, several studies employ instrumental-variable strategies (Alviola et al., 2014; Qian et al., 2017; Asirvatham et al., 2019), generally finding positive effects of fast-food exposure on children's BMI, except for Dolton and Tafesse (2022), who report no statistically significant effect.

Despite this growing literature, relatively little attention has been paid to the dynamic persistence of BMI. Most existing studies estimate two-way fixed-effects models without explicitly modelling lagged BMI or heterogeneous BMI trajectories. If exposure to fast food is correlated with underlying weight trajectories, failing to account for these trajectories may result in a biased estimator. The scarcity of dynamic approaches are partly due to data limitations. Reliable estimation of standard dynamic models typically requires long and regularly spaced panels, whereas child cohort studies tend to cover a relatively short period and contain observations collected at irregular intervals.[3] Our clustering approach provides an alternative way to capture persistent heterogeneity in BMI using short and unequally spaced longitudinal data.

Our study is also related to previous work using the MCS but differs in several important respects. Green et al. (2021) examine correlations between BMI and fast-food availability around the residence using aggregated geographic units, whereas our focus is on causal effects

[2]We do not discuss purely correlational studies here; see Jiang et al. (2023) for a systematic review.

[3]In many studies, such as Currie et al. (2010), the height and weight data used for estimation are obtained from schools, restricting the scope for incorporating anthropometric measurements taken before school entry into the analysis.

arising from school-based exposure. Libuy et al. (2023) exploit within-individual variation in exposure to fast food, but do not explicitly account for dynamic persistence in BMI, which is evident in our data and well-documented in the literature (e.g., Adair, 2008). In contrast, we model persistent BMI trajectories directly and focus on adolescents experiencing plausibly exogenous changes in school-level exposure following the transition to secondary school.[4]

The remainder of the paper is organised as follows. Section 2 describes the data sources and the construction of our key variables, including BMI z-scores and measures of proximity to fast-food outlets. Section 3 outlines the empirical strategy, whereas Section 4 discusses sample construction. Section 5 presents the main results, while Section 6 reports robustness checks and results from alternative specifications. Section 7 discusses the policy implications and concludes.

2 Data and main variables

2.1 Data sets

The data for this study are obtained from two sources: (i) UK Millennium Cohort Study (MCS), a nationally representative longitudinal survey of children born in the UK, and (ii) Points of Interest (PoI) data and the Integrated Transport Network (ITN) data, both from Ordnance Survey.

Millenium Cohort Study

The initial MCS sample contained 18,818 children born in the UK over a 17-month period (September 2000 to January 2002) and were living in the UK at age nine months when the data collection started. So far, the data have been collected across further seven survey sweeps when the cohort members were approximately at 9 months, and 3, 5, 7, 11, 14, 17 and 23 years old. This ongoing multidisciplinary survey gathers extensive information on a wide range of topics concerning both the cohort members and their parents; including demographic, health, developmental, and behavioural characteristics. A list of the MCS data files used in this

[4] Specifically, Libuy et al. (2023) pool observations from ages 7, 11, and 14, combining primary- and secondary-school years in the same analysis. Another notable difference includes the use of age- and sex-standardised BMI z-scores in our analyses rather than raw BMI, which is particularly important during adolescence because of rapid physical development.

study is provided in Appendix C. Further details of the MCS, including its objectives, sample design, recruitment, response and attrition rates, and survey contents can be found in Joshi and Fitzsimons (2016).

Ordnance Survey Points of Interest

The PoI dataset is a comprehensive geospatial database containing over 4 million locations across Great Britain, covering private and public businesses, educational institutions, and leisure cites. Each PoI is georeferenced and classified into a hierarchical strucobture comprising nine major categories and over 600 subcategories, facilitating small-area level spatial analysis. The data draw on sources including public records, commercial directories and business submissions, and is currently updated quarterly. Each record typically includes name, address, postcode, coordinates and a classification code.

The PoI dataset provides a robust basis for the food environment research, and is widely used in applications including research on obesogenic environments (e.g., Hobbs et al., 2019; Libuy et al., 2023; Boskovic et al., 2026). Detailed information on its scope, sources and coverage can be found in the PoI product information (Ordnance Survey, 2025).[5]

Ordnance Survey Integrated Transport Network

ITN data is a detailed digital representation of the road network in Great Britain, containing all roads, paths and associated infrastructure, such as roundabouts. Each road segment is georeferenced and coded with attributes including road name, classification (e.g. motorway) and routing restrictions (e.g. one-way systems), making it valuable in mapping connectivity between places and estimating travel times.[6] The ITN data is regularly updated by Ordnance Survey and is used operationally by public bodies and navigation systems, reflecting its practical reliability (Ordnance Survey, 2019). Overall, ITN provides a granular foundation for modelling the spatial structure of transport networks in Great Britain.

---

[5] The PoI data were validated through street audits, confirming both the accuracy of the data and the reliability of outlet classifications (Wilkins et al., 2017).

[6] For comprehensive detail on ITN's structure, refer to Ordnance Survey (2010).

2.2 Main variables

BMI z-score

Our outcome of interest is standardised BMI, where BMI is defined as weight in kilograms divided by height in metres squared. At each survey wave, trained enumerators measured the height and weight of every cohort member (Hardy et al., 2019). For adults, whose height is stable, BMI provides an appropriate measure of relative weight. However, for children and adolescents, the distribution of BMI varies nonlinearly with development (see Figures in the online document). To account for this, we standardise BMI using the age- and sex-specific UK–World Health Organisation (WHO) growth reference charts (Cole et al., 2012).[7] While standardisation will account for the non-linear trajectories observed in child and adolescent BMI, it does not resolve the issue of persistence which concerns the serial correlation of z-BMI at the individual-level. Clinically, children are classified as overweight if their BMI z-score is at least 1.34 or 91st centile, and obese if it is at least 2.05 or 98th centile according to the UK clinical thresholds (OHID, 2026).

Proximity to fast-food outlets

Using the PoI data, we construct our main explanatory variable of interest, proximity to a fast-food outlet. To identify outlets serving fast food, we rely on the PoI establishment classification scheme, consisting of over 600 classes. In particular, we extract the location information of outlets under the classes 'Fast food and takeaway outlets'; 'Fast food delivery services'; and 'Fish and chip shops'. We then categorise the outlets under those classes into three broad groups: (i) Major-brand fast food, (ii) Non-major-brand fast food, and (iii) Fish and chips. The major-brand fast food outlets comprise major fast-food chains, such as McDonalds, KFC and Burger King, following the Wilkins et al. (2019) classifications, whereas the non-major-brand outlets comprise fast-food outlets, such as kebab and fried chicken shops.

We distinguish between the three groups to capture potential heterogeneous effects of fast food on z-BMI, in contrast to existing studies that typically aggregate all fast-food types into

[7] We standardise BMI using the `zanthro` Stata package (c.f., Vidmar et al., 2004). This command calculates z-scores for anthropometric measures in children and adolescents according to US, UK, WHO and composite UK-WHO reference growth charts.

a single measure of exposure (e.g., Currie et al., 2010, Dunn et al., 2021, Libuy et al., 2023). This distinction is important because fast-food outlets differ in their branding and business models. Major-brand outlets, such as McDonald's, operate as national chains with strong brand recognition, promotional offers, and longer opening hours, all of which might appeal to adolescents. In contrast, outlets for non-major-brand fast food and fish and chips are more likely to be independently owned, smaller in scale, less intensively marketed, and typically open for shorter hours. We treat fish and chips as a separate category because they occupy a distinctive place in the British culture, where consuming fish and chips on Fridays remains a popular tradition.[8] In addition, we also extract the location information for other take-away or delivery-oriented food outlets, as alternative take-away or delivery options may also affect adolescent z-BMI. These alternative food outlets correspond to supermarkets, convenience stores, bakeries and cafes (see Appendix A for details).

Using information on the locations of cohort members' schools, along with the road network and its associated features (e.g., routing restrictions), we construct the measure of proximity to specific types of food outlets based on road-network buffers.[9] Specifically, we measure proximity based on a road-network buffer of 400 metres around each participant's school, although we will also explore different distance measures in Section 6. This is a buffer size commonly used in retail food environment research (Wilkins et al., 2019), capturing proximity of an approximately five-minute walk. We then construct binary proximity variables indicating the presence of at least one outlet of a given type within the 400-metre buffer zone.[10] We repeat the same exercise to construct the proximity measure around the residence of each cohort member. Full methodological details on the construction of proximity measures are provided in Appendix A.

[8] The tradition of eating fish on Fridays originates in a Christian custom of no meat on Fridays. The custom has then evolved during the Industrial Revolution, was reinforced during the World War II as fish and chips were exempt from rationing, and remains popular as the dish is affordable, filling, and widely available (Panayi, 2014).

[9] Compared to Euclidean buffers, road-network buffers better reflect actual travel paths and accessibility, avoiding the overestimation of reachable areas that can lead to misclassification of environmental exposure (Bivoltsis et al., 2018; Chen et al., 2022).

[10] Previous studies have employed a variety of measures of fast-food exposure, including counts of outlets within predefined neighbourhoods or buffers, binary indicators for the presence of one or more outlets, and density measures normalised by area or population. Our focus on a relatively small buffer size yields limited variation along the intensive margin (i.e., outlet counts), making a binary proximity measure more appropriate. This specification also avoids imposing a linear relationship between outlet counts and z-BMI—a common but potentially restrictive assumption in studies using count- or density-based measures of fast-food access (e.g., Libuy et al., 2023).

3 Accounting for persistence: A new approach

It is well established in the medical and health literature that BMI and obesity exhibit substantial persistence over time (Adair, 2008; Bartle et al., 2013). Childhood BMI, and even *in utero* metrics, strongly predict adolescent and adult BMI and obesity. This persistence can be the result of individual and environmental factors, both observed and unobserved, that determine lifestyle preferences and behaviours (Kluser and Pons, 2026). However, physiological mechanisms could also contribute to the persistence of excess weight by making sustained weight loss difficult. Weight gain is associated with increases in both the size and number of adipocytes, whereas weight loss primarily reduces adipocyte size while leaving adipocyte number largely unchanged (MacLean et al., 2015). As a result, individuals who have lost weight may remain biologically predisposed to regain it. Recent evidence further suggests that adipose tissue retains an ”obesogenic memory” following weight loss, reflecting persistent cellular and epigenetic changes associated with prior obesity (Hinte et al., 2024).

We adopt a novel approach to modelling this persistence in the context of the relationship between adoloscent BMI and fast food availability. A large number of studies employ static panel data models to account for unobserved individual-level heterogeneity (e.g., Currie et al., 2010; Alviola et al., 2014; Qian et al., 2017; Asirvatham et al., 2019; Han et al., 2020; Dolton and Tafesse, 2022; Abrahamsson et al., 2023; Libuy et al., 2023). These models do not account for persistence in the outcome which affects both the level and trajectory of BMI from childhood to adolescents. While dynamic panel data models incorporating lagged dependent variables would ordinarily be preferred for modelling persistence, they require a sufficiently long, equally spaced panel. These requirements are not met by the relatively short and irregularly spaced MCS panel used in this study.

Our modelling approach is informed by the design and scale of the MCS. Measurements were taken at approximately 9 months, 3 years, 5 years, and 7 years for sweeps 1 to 4, and at 11, 14, and 17 years for sweeps 5 to 7. Consequently, sweeps 1 to 4 are spaced at roughly two-year intervals, whereas sweeps 5 to 7 occur at approximately three-year intervals. We focus on the three sweeps collected at three-year intervals that span the adolescent period.

We start with a linear, static baseline specification of z-BMI. As the MCS is a cohort study, let t denote the age in each sweep. For the adolescent years (age-11 to age-17 sweeps), z-BMI is modelled as,

$$Y_{it} = F'_{it}\beta + X'_{it}\gamma + \varepsilon_{it} \qquad\qquad t = 11, 14, 17 \tag{1}$$

where F is a vector of dummy variables denoting the availability of fast food within a specified geographical area. Within this area, for each category of fast food, we include a separate dummy variable for availability at the school as well as the home. We exclude any interactions between fast-food indicators. The vector of covariates, X, contains time-invariant as well as time-varying child, mother and family characteristics, location specific characteristics.[11]

Standard estimators, such as Ordinary Least Squares (OLS), in equation (1) will be inconsistent if the unobserved factors underlying persistence, such as lagged outcome variables and time-invariant unobservables, are correlated with household school-location choices. For example, secondary schools located near city centres and shopping districts - where major fast-food chains tend to cluster - may disproportionately attract pupils from nearby inner cities with higher rates of childhood obesity. This may induce a correlation between changes in fast-food exposure over time, particularly during the transition from primary to secondary school, and prior z-BMI trajectories.

We therefore take an approach to account for persistence by borrowing from the recent literature on group-level heterogeneity (Bonhomme and Manresa, 2015). We treat the unmodelled persistence as a form of unobserved heterogeneity and specify the error term with two components; a group-specific term, $\psi_{c(i)t}$, and an idiosyncratic shock:

$$\varepsilon_{it} = \psi_{c(i)t} + \upsilon_{it}. \tag{2}$$

The unobserved persistence is assumed to be unique up to a group (or cluster) level and is therefore captured in both the level and trend components of $\psi_{c(i)t}$. This approach is thus more flexible than the standard approach of controlling for time-invariant individual-level heterogeneity as it allows for cluster-specific age profiles, or time trends, in z-BMI.

[11] Specifically, we control for: (i) time-invariant characteristics, including the child's gender, maternal characteristics at birth (ethnicity, educational attainment, smoking and drinking before and during pregnancy, and pre-pregnancy BMI), and family social class at birth; (ii) time-varying household characteristics, including whether the cohort member is an only child, whether the mother is a lone parent, maternal employment status, and housing tenure; and (iii) area characteristics, including country of residence, the logarithm of the LSOA population, and quintile indicators for the Index of Multiple Deprivation (IMD). We additionally control for the availability of supermarkets, convenience stores, bakeries, and cafés in the vicinity of both the school and the residence. Further details are provided in Appendix A Table A1.

Next, we assume that there is a finite number of time-invariant clusters which can be identified from the *unconditional* distribution of childhood outcomes (age-3, age-5 and age-7 sweeps),

$$c(i) = h\left(\{Y_{it}\}_{t=3,5,7}\right) \tag{3}$$

This assumption is significant because we do not condition on the past values of the regressors when identifying the cluster groups. It also allows the cluster group assignments to be estimated independently of the main equation in a two-step procedure. By including cluster-by-time fixed effects in the estimating equation, we effectively condition on the entire history through cluster-specific trajectories of z-BMI.[12] This approach provides the flexibility to estimate the cluster groupings using MCS sweeps that are observed at different intervals from those used in the main analyses.

We can extend this model to allow for individual-level time-invariant heterogeneity, which is then removed through a within transformation (as in Extension 1 of Bonhomme and Manresa, 2015).

$$\varepsilon_{it} = \underbrace{\alpha_i + \nu_{c(i)t}}_{\psi_{c(i)t}} + \upsilon_{it}. \tag{4}$$

The within-individual transformation changes the source of variation used to identify β in the model. With only cluster-level heterogeneity in the model, identification of β relies on both within-cluster (cross-sectional) and within-individual (over-time) variation in fast-food availability. The within-individual transformation removes the former source of variation, leaving identification to rely solely on within-individual changes in fast-food availability over time.

Under the maintained assumptions above, identification of β hinges on the strict exogeneity of the fast-food variables conditional on the full set of covariates, cluster-fixed effects, and cluster-by-time fixed effects. This assumption may not hold for two reasons. First, unobserved location-specific factors may remain imperfectly controlled for, resulting in correlation between

[12] An alternative approach is to discretise z-BMI into weight-status categories and define persistence using category-transition indicators. However, the number of possible trajectories increases exponentially with the number of survey waves, limiting scalability. Moreover, this approach performs substantially worse empirically, yielding an R^2 around half that of the cluster-based specification. The clustering approach exploits similarities in continuous z-BMI trajectories across individuals and survey waves and preserves within-category variation that is lost when z-BMI is discretised.

the fast-food variables and the error term. Second, the fast-food availability measures may be subject to misclassification error. We discuss both issues in detail below.

Inadequate control of location-specific variables

Although we control for neighbourhood deprivation and population size, we cannot fully account for all forms of location-specific heterogeneity. These controls capture key but not exhaustive aspects of local context. Dietary behaviours may still vary with unobserved neighbourhood characteristics such as local economic conditions influencing both food prices and the availability of different food outlets.

To address these concerns, our identification strategy exploits the transition from primary to secondary school at around age 11, motivated by Dunn et al. (2021) who similarly exploit educational transition.[13] Information collected in the age-11 sweep reflects the final year of primary education, while secondary school information is captured in the age-14 sweep. In our sample, approximately 97% of children move to a new school site for secondary education, generating plausibly exogenous variation in school-based fast-food exposure.

The secondary-school environment differs markedly from the primary-school setting. Adolescents typically have greater autonomy over their food choices, experience less direct adult supervision, and have more freedom to travel independently within their neighbourhoods. As a result, the proximity of fast-food outlets is likely to play a different role in shaping dietary behaviours and, in turn, z-BMI, between secondary- and primary-school pupils, whose mobility and purchasing autonomy are more constrained. Additionally, we do not have exogenous variation in the exposure to fast-food outlets in the age-11 primary-school sweep. Hence, our main analyses focus on the secondary-school period. However, in the supplementary analysis, we also include the age-11 sweep and examine whether it produces any notable differences in the results.

While this school transition generates useful variation in exposure at school, we do not observe analogous variation in exposure at home. We therefore include home-based exposure

[13] Although Dunn et al. (2021) also exploit the educational transition, their empirical design differs from ours in several important respects. Most notably, they retain both children who transition to a new school and those who remain at the same school, whereas our analysis is restricted only to children who transition to a new school site (c.f., Section 4.2). In addition, they examine fast-food availability along children's commuting routes, whereas we focus on fast-food exposure in the vicinity of the school.

to fast-food outlets as controls—recognising that home-based food choice may reflect household-level food preferences—and test robustness of our main findings by restricting the sample to individuals who did not change residence.

Turning to the age-17 sweep, the period between ages 14 and 17 spans a major educational transition, during which students complete their General Certificate of Secondary Education (GCSE) examinations or equivalent and move into either academic Advanced Level (A-level) programmes or technical and vocational pathways.[14] Unfortunately, the location information of adolescents who pursue the technical and vocational routes is incomplete in our data. As selection into A-level study correlates strongly with socio-economic background, we decided not to analyse this transition. Instead, data from the age-17 sweep are used to investigate medium-run effects of earlier fast-food exposure.

In summary, we restrict our attention to the age-14 sweep in our main analyses. Conditional on changing schools, we assume that secondary-school choice is independent of local fast-food availability. Under this assumption, the transition to secondary school provides plausibly exogenous variation in school-level fast-food exposure. This motivates the following estimation equation for individual i's z-BMI at age 14:

$$Y_i = F_i'\beta + X_i'\gamma + \psi_{c(i)} + \upsilon_i \tag{5}$$

Potential misclassification of fast-food 'availability' variables

The second reason for the strict exogeneity of fast-food exposure variables might fail, even after accounting for the persistence in z-BMI, is the potential misclassification of the binary indicators used to measure exposure to fast-food outlets within a given geographical area. In contrast to the classical measurement error framework, the measurement error is mechanically negatively correlated with the latent "true" variable when the variable of interest is binary. Moreover, the conditional covariance between the measurement error and the true variable is generally non-zero (Bollinger, 1996, 2010). Under standard assumptions, the OLS estimator therefore continues to suffer from attenuation bias, unless the misclassification probabilities

[14] In the UK, children attend primary school from roughly ages five to 11 (Key Stages 1–2) and secondary school from roughly ages 11 to 16 (Key Stages 3–4), culminating in GCSE examinations or equivalent at age 16 (Department for Education, 2026).

are symmetric: that is, unless the probability of misclassifying 0 as 1 equals the probability of misclassifying 1 as 0. Although our fast-food availability measures are not self-reported and are carefully constructed from an independent data source, this symmetry condition need not hold, particularly for small businesses. Specifically, unlike major-brand fast-food outlets, which typically operate as highly organised businesses with relatively stable locations, non-major-brand outlets and fish and chip shops are often independent, single- or family-owned businesses. These outlets tend to experience higher rates of entry, exit, and ownership turnover. As a result, true availability may vary within the measurement interval, leading to temporal misalignment in the measurement of outlet availability. We revisit this in the discussion of our results in Section 5.

A cautionary note on interpretation

Our model examines the relationship between z-BMI and fast-food availability in the vicinity of the school, conditional on group-specific unobserved heterogeneity and adolescent, maternal, and family characteristics. The estimated equation can be interpreted as a reduced-form representation of the underlying structural relationship between z-BMI and fast-food consumption.[15] Consequently, a null estimate does not necessarily imply that fast-food consumption has no effect on z-BMI. Rather, it may reflect the absence of a first-stage relationship between fast-food availability and fast-food consumption, although the existing literature generally supports the presence of such a relationship (e.g., Svastisalee et al., 2016).

4 Sample selection and descriptive statistics

Our sample excludes observations with missing values for key variables, including BMI, sex, age at interview and country of residence. We also drop cohort members who are not singletons or who belong to multiple-birth families.[16] Following these restrictions, further sample selection

[15]The MCS does not provide the information required to credibly estimate the implicit first-stage relationship between proximity to fast-food outlets and fast-food consumption within our empirical framework. The relevant survey item—"How often, if at all, do you eat fast food such as McDonald's, Burger King, KFC, or other fast food like that?"—offers no definition of "fast food," leaving scope for heterogeneous interpretation. In addition, the measure is self-reported, does not differentiate between home and school environments, and is recorded only as an ordered categorical frequency. Given these limitations, we do not incorporate this variable into the analysis.

[16]Multiple birth families have more than one eligible child with different birthdates, born during the survey selection period.

proceeds in two steps.

4.1 Sample for identifying cluster groups

In the first step, we construct a balanced panel of 10,011 cohort members with valid BMI (or z-BMI) measurements observed at ages 3, 5 and 7.[17] The Kernel density plot for the z-BMI distribution by sweep is provided in Figure 1, indicating a leftward shift in the distribution during the childhood period as they grow older. However, when cohort members transition from primary school in the age-11 sweep to secondary school in the age-14 sweep, the average z-BMI begins to increase to 0.539, as shown in Figure 2.

While Figures 1 and 2 are informative about changes in the overall distributions, they do not show how individual cohort members move within these distributions over time. Insight into individual-level dynamics is instead provided by the correlation matrix of z-BMI across sweeps, reported in Appendix B Table B1. The reported correlations indicate a high degree of persistence in z-BMI, with correlations between adjacent sweeps typically in the range of 0.7–0.8. We aim to capture this persistence by clustering cohort members based on their observed unconditional childhood z-BMI trajectories, using a k-means clustering approach, which proceeds as follows.

From the z-BMI measurements at ages 3, 5, and 7, we identify seven clusters, ensuring that each contains at least 5% of the total observations.[18] This requirement meant that each cluster needed a minimum of around 500 cohort members. It is important to note that in some cases, the clustering algorithm may not classify every individual into a group. In our case, 9,727 of the 10,011 cohort members were successfully assigned to clusters, with the smallest cluster containing 676 cohort members (7.0%).[19]

Figure 3 shows the average z-BMI values for the seven clusters identified across the age-3

[17]We use the age-3 to age-7 sweeps to construct z-BMI, since the UK–WHO growth reference charts for BMI-for-age z-scores are available only for age higher than two.

[18]We use Stata's `cluster` package to partition individuals into seven clusters, imposing a minimum cluster size of 5% of the sample. The K-means algorithm assigns each observation to the nearest cluster centroid based on Euclidean distance, iteratively updating the centroids until convergence. The 5% threshold ensures that each cluster contains a sufficient number of observations while preserving meaningful heterogeneity in childhood z-BMI, thereby supporting reliable estimation.

[19]See La Cruz et al. (2020) for an application of k-means clustering method. Their aim was to evaluate the k-means clustering algorithm using anthropometric measurements for the classification of subjects with overweight/obesity and abnormal body fat percentage.

to age-7 sweeps. Although the clusters were derived solely from z-BMI measurements at ages 3, 5, and 7, the average z-BMI values at later sweeps (ages 11 to 14) largely follow the same developmental patterns established in early childhood. Children in Cluster 1 maintain average z-BMI values within the *obese* range throughout the study period, while those in Clusters 2 and 3 hover around the *overweight* threshold. In contrast, the average z-BMI values for Clusters 4–7 shift over time but consistently remain within the *normal-weight* range. As discussed later, the persistence of these trajectories helps explain why the clusters account for approximately 46% of the variation in z-BMI at ages 11 to 14, or 41% at age 14 specifically.

4.2 Sample for estimation

The second step involves the construction of the estimation sample. Specifically, we select a balanced panel of cohort members observed at both the age-11 and age-14 sweeps. This enables us to identify individuals who changed schools between the two sweeps. The timing of the MCS interviews is particularly well suited to this purpose: the age-11 sweep captures information while the vast majority of cohort members are still enrolled in primary school, whereas the age-14 sweep records outcomes after the transition to secondary school. This transition is institutional and largely exogenous, with over 97% of cohort members changing schools as confirmed by differences in school postcodes across the two sweeps. Our estimation sample further excludes cohort members residing in Northern Ireland, as fast-food location data from PoI are not available for this country. Applying these restrictions yields a final unweighted estimation sample of 6,626 cohort members. The summary statistics for all variables included in our model are reported in Appendix A Table A1.

Table 1 reports the proportion of cohort members exposed to fast-food outlets within a 400-metre network buffer of their secondary school, by fast-food category and childhood BMI cluster. As described in Section 2.2, outlets are classified into three categories: major-brand, non-major-brand, and fish and chip outlets. This distinction reflects differences in branding and business models. Major-brand outlets, such as McDonald's, are national chains that typically operate from carefully selected high-street, shopping-centre, or drive-through locations. In contrast, non-major-brand outlets (e.g., kebab shops) and fish and chip shops are typically independently owned, smaller in scale, and more widely distributed across both residential neighbourhoods and commercial areas.

Table 1 is consistent with this pattern. Non-major-brand outlets are the most prevalent category across all cluster groups, with shares ranging from 21% to 29%. Fish and chip shops are the second most common outlet type, with availability varying between 11% and 16% across clusters. In contrast, major-brand fast-food outlets are the least prevalent, with relatively low availability across clusters (3%–8%). Cluster 4—comprising cohort members who are, on average, in the normal-weight range—accounts for the largest share of the estimation sample (26%, or 1,735 individuals), whereas Cluster 1—those in the obese range—represents the smallest share (7%, or 404 individuals). Overall, across the full cohort, 23% of members have non-major-brand outlets, 13% have fish and chip shops, and 5% have major-brand fast-food outlets in the vicinity of their schools.

5 Empirical findings

5.1 Main specification

Table 2 reports our main results, where the parameters of interest are the β coefficients in equation (5), capturing the effect of proximity to fast-food outlets on adolescent z-BMI. Our benchmark specification measures fast-food availability using indicator variables for the presence of an outlet within a 400-metre road-network buffer around the school, corresponding to an approximate five-minute walk. Alternative distance measures will be explored in Section 6.

Our preferred specifications restrict the estimation sample to the age-14 sweep (see Section 3), the first sweep following the transition to secondary school, although we will also conduct the analysis incorporating the age-11 sweep. Identification exploits plausibly exogenous variation in proximity to fast-food outlets at age 14, under the assumption that school choice is independent of fast-food availability, conditional on cluster group membership. Identifying variation therefore comes from adolescents with and without exposure to a fast-food outlet, within the same cluster group, *ceteris paribus*. All models are estimated by OLS, and we report robust standard errors.

Columns (1) and (2) present estimates based on the full age-14 estimation sample, whereas column (3) excludes adolescents in cluster groups 1 and 7 (1,052 individuals), corresponding to those with very high or low childhood z-BMI. Column (4) uses the same age-14 estimation sample, but measures the outcome —adolescent z-BMI— at age 17, capturing potential medium-term effects of school-based exposure to fast-food outlets following the transition to secondary school.

Column (1) reports estimates that control for food-outlet availability and additional covariates, but exclude cluster group fixed effects. Note, this would be the model commonly estimated if one only had access to cross-sectional data. The results suggest positive effects of proximity to both major-brand outlets and fish and chip shops on adolescent z-BMI, although these effects are not statistically significant. However, once cluster group dummies are included in column (2), the estimated effect of major-brand fast-food outlets increases to approximately 0.158 standard deviations and becomes statistically significant.[20] A comparison of columns (1) and (2) underscores the importance of accounting for cluster-level heterogeneity: namely, omitting cluster fixed effects, constructed from the pre-adolescence z-BMI, attenuates the estimated coefficients and substantially reduces explanatory power, as reflected in the markedly lower R-squared in column (1).

In column (3), we evaluate whether extreme trajectories disproportionately influence the results. The base model is re-estimated excluding Cluster 1 (persistently obese on average) and Cluster 7 (persistently very low average z-BMI). These two clusters represent atypical growth patterns that differ substantially from the developmental variability observed in the remaining groups. Their exclusion reduces the risk of outlier-driven bias and potential violations of model assumptions related to variance and linearity. The finding in column (3) is similar to our main finding in column (2), indicating that the primary effects are not driven by these extreme z-BMI profiles. The estimated effect of major-brand outlets also remains robust when using age-17 z-BMI as the outcome variable in column (4). Across specifications that include cluster fixed effects, the effects of exposure to major-brand outlets on adolescent z-BMI are therefore stable and robust.

In contrast, the results for exposure to non-major-brand outlets are weaker, less consistent and generally imprecisely estimated. Estimates for fish-and-chip outlet exposure also tend to be insignificant, although a positive and statistically significant effect emerges when age-17 z-BMI

[20] The full estimation results for this baseline specification, including coefficient estimates for all covariates, are reported in Appendix B Table B2.

is used as the dependent variable. Taken together, these findings indicate that proximity to major-brand fast-food outlets near the school is the most robust predictor of adolescent z-BMI, whereas the effects of other outlet types are less stable across specifications.

This distinctive effect of proximity to a major-brand fast-food outlet can potentially be explained by a combination of marketing, affordability, and social factors. Branded chains such as McDonald's, KFC, and Burger King possess substantial advertising power and strong brand recognition, which may reduce the cognitive costs of food choice and encourage habitual consumption. These chains also employ pricing strategies such as value menus and promotional offers, alongside extended opening hours, thereby enhancing both affordability and convenience. Moreover, fast-food consumption has a social dimension (van Rongen et al., 2020), as it is normalised within adolescent peer groups, which may further reinforce patterns of frequent use. Unlike many independent fast-food outlets and fish-and-chip shops, which are primarily takeaway oriented, major branded chains typically provide indoor seating areas where adolescents can spend time. As a result, these outlets may function not only as places to purchase food but also as social venues, possibly increasing adolescents' exposure to fast food.

In contrast to major-brand outlets, which operate as highly organised businesses, non-major-brand outlets and fish and chip shops are typically independent, single- or family-owned establishments. These outlets tend to exhibit high entry and exit rates and frequently change ownership. As a result, true availability may vary within the measurement interval, leading to temporal misalignment in the measurement of outlet availability. The sensitivity of the estimated effects may therefore be due to attenuation bias arising from measurement error, as discussed in Section 3.

5.2 Heterogeneity analysis

Table 3 presents heterogeneity analyses examining whether the effects of school-level fast-food proximity on adolescent z-BMI vary across key socio-demographic groups. Our main result from Table 2 is copied to column (1) for comparison.

In column (2), we interact fast-food proximity indicators with a dummy variable for being female. Panel A reports coefficients on the main (un-interacted) fast-food proximity variables, representing effects for males, while panel B reports the interaction terms capturing any *ad-*

ditional effect for females. Consistent with the baseline estimates in column (1), proximity to major-brand fast-food outlets is the only statistically significant fast-food measure. For males, the estimated coefficient is 0.210, whereas the corresponding effect for females is 0.100 (= 0.210 − 0.110). However, the gender difference captured by the interaction term, is not significantly different from zero, presenting insufficient evidence that the effect truly differs by gender. No gender differences are found for other outlet types either, and the joint test of the interaction terms fails to reject the null hypothesis of no gender heterogeneity (panel C).

In column (3), adolescents are classified by maternal ethnicity (White vs. non-White). Proximity to major-brand outlets has a larger effect on z-BMI among adolescents with White mothers, 0.184, than among adolescents with non-White mothers, 0.048 (= 0.184 − 0.136), but the difference between the two groups, captured by the interaction term, is insignificant. Interestingly, proximity to non-major-brand outlets has a larger effect for adolescents with non-White mothers, 0.184 (= 0.003 + 0.181), than White mothers, 0.003, suggesting that adolescents with non-White mothers may be more responsive to availability of non-branded outlets near schools.

In column (4), adolescents are grouped according to whether their mother has at least a lower secondary school qualification, corresponding to the GCSE or equivalent in the UK. The result for major-brand outlets indicates no meaningful heterogeneity. However, the effect of proximity to non-branded outlets is larger for adolescents with lower-educated mothers, 0.179 (= 0.006 + 0.173), pointing to a socioeconomic gradient in vulnerability to the local food environment.

In column (5), interaction terms are included for adolescents residing in deprived areas (IMD quintiles 1–2) versus less deprived areas (IMD quintiles 3–5).[21] The effect of proximity to major-brand outlets on z-BMI is positive for both groups (0.163 and 0.136, respectively), with no statistical difference in magnitude. This suggests that exposure to major-brand outlets near schools is a shared risk factor rather than one concentrated among adolescents in more deprived areas.

We additionally explore heterogeneity by prior weight history and school urbanicity. Ado-

[21] The MCS provides country-specific Index of Multiple Deprivation (IMD) quintiles, derived from the official deprivation indices for England, Wales, and Scotland. Each index ranks small geographical areas according to multiple domains of socioeconomic deprivation within its respective country, with lower quintiles indicating greater deprivation.

lescents with a prior history of overweight or obesity may face greater future weight gain due to adipocyte persistence and habit formation (c.f., Section 3). As a result, the effects of fast-food availability could vary by prior weight history. Likewise, as rural students generally have less access to food outlets, a fast-food outlet may represent a greater increase in local food accessibility, potentially leading to larger effects in rural areas. The results (not reported) provide no evidence of significant heterogeneity along either dimension.

In summary, the positive effect of proximity to major-brand fast-food outlets on adolescent z-BMI is broadly consistent across gender, ethnicity, maternal education, neighbourhood deprivation, prior weight history, and school urbanicity. Heterogeneity is limited to a few specific patterns: non-major-brand outlets appear more relevant for adolescents with lower-educated mothers and with non-White mothers, while other outlet types show little systematic variation. Overall, these results suggest that major-brand fast-food availability near schools represents a general risk factor for higher adolescent z-BMI rather than one confined to specific subgroups.

6 Further analyses

We next present additional analyses in Tables 4 to 6 to assess the robustness of our findings to alternative model specifications and identification strategies. Table 4 evaluates sensitivity to different sources of variation in fast-food outlet proximity, and to the inclusion of a potentially important confounder, student ability. Table 5 examines alternative measures of fast-food proximity, to identify the spatial scale most relevant for adolescent z-BMI. Finally, Table 6 considers the role of home proximity to fast-food outlets in explaining adolescent z-BMI.

Table 4 extends the benchmark specification by pooling observations from the age-11 and age-14 sweeps. This specification implicitly assumes that the coefficients in equation (5) are invariant across the two survey sweeps. To account for common time shocks, column (2) includes sweep fixed effects. It also relaxes the assumption of common BMI trajectories by adding cluster-by-sweep fixed effects, thereby allowing weight trajectories to evolve differently across clusters. Relative to the benchmark specification in column (1), the pooled model exploits both cross-sectional variation across individuals and longitudinal variation arising from repeated observations on the same individuals. Column (3) further includes individual fixed effects, a commonly estimated specification. The identification in column (3) comes from plausibly exogenous within-individual changes in fast-food exposure as they transition from primary to secondary school.

Starting from column (2), the estimated effect of proximity to a major-brand fast-food outlet is very similar to that obtained from the benchmark model. That said, the null hypothesis that all coefficients are equal across the two sweeps is rejected ($p < 0.01$).[22] This supports the expectation that the equations for the primary- and secondary-school sweeps capture different relationships, possibly reflecting the differences in direct adult supervision, autonomy over their food choices, and freedom to travel independently within their neighbourhoods.

Moving to the results obtained using the within-group (WG) estimator in column (3), the estimated effect of proximity to major-brand outlets remains statistically significant, although the point estimate is smaller (0.095; SE = 0.039). This reduction may reflect the limited within-individual time variation in the binary indicators of fast-food exposure (see Appendix B Table B3), which is the source of identifying variation for the WG estimator. In addition, the WG estimates may be more susceptible to attenuation bias if the binary fast-food indicators contain non-classical measurement error, as discussed in Section 3. In particular, differencing can amplify noise arising from spurious transitions in mismeasured binary regressors, possibly attenuating the estimated coefficients.

Next, we consider the possibility that the obesogenic environment around schools is correlated with unobserved factors affecting adolescent z-BMI. For example, high-performing schools—often attended by high-achieving students—tend to be located in more affluent neighbourhoods, which typically have fewer fast-food outlets. In such cases, even if school choice is unrelated to fast-food availability, proximity to fast-food outlets may still correlate with unobserved determinants of z-BMI, such as aspects of student ability or health literacy. To address this concern, in addition to controlling for neighbourhood characteristics including IMD quintiles, columns (3) and (4) extend the model to include measures of adolescents' cognitive and non-cognitive abilities.[23] The inclusion of these additional controls does not substantively

[22]We also test whether only the coefficients on the availability of the three types of fast-food outlets (major-brand, non-major-brand, and fish and chips outlets) are equal. The null hypothesis of equal coefficients is rejected at the 1% significance level ($p < 0.01$).

[23]Cognitive ability is measured using the Word Reading and Pattern Construction tests from the British Ability Scales and the National Foundation for Educational Research Progress in Maths assessment. Non-cognitive ability is measured using the Child Social Behaviour Questionnaire (CSBQ), capturing independence and self-regulation,

change the estimated effects, indicating that our results are unlikely to be driven by omitted variation in student ability.

Table 5 examines the sensitivity of our results to alternative measures of proximity to fast-food outlets. We consider larger buffer zones of 800 metres and 1,600 metres, corresponding to roughly 10- and 20-minute walk distances, respectively, which are commonly used in the literature (Wilkins et al., 2019). Columns (1) to (3) report estimates based on a 400-metre buffer (benchmark), 800-metre buffer, and 1,600-metre buffer, respectively. The results show that the magnitude of the estimated coefficients declines monotonically as the buffer size increases, becoming statistically insignificant once the buffer exceeds 400 metres. This pattern suggests that only very close proximity—within roughly a five-minute walk—matters for adolescent z-BMI. The finding is consistent with evidence from the United States reported by Currie et al. (2010) and may reflect the relatively constrained mobility of adolescents during the school day, which limits their ability to travel far from school. It is also in line with empirical evidence showing that greater immediate access to fast-food outlets is associated with higher fast-food consumption among children (see Jia et al., 2021 for a review), potentially through increased temptation driven by visual cues (Laibson, 2001; Boswell and Kober, 2016) and the activation of self-control problems (DellaVigna, 2009).

In column (4), we assess exposure to fast-food outlets located within the same Lower Super Output Area (LSOA) as the cohort member's school. LSOAs are population-based administrative units used to report small-area statistics (e.g., Censuses), and typically contain between 1,000 and 3,000 residents (Office for National Statistics, 2021).[24] Estimates based on LSOA-level exposure are also statistically insignificant, reinforcing the conclusion that it is immediate, school-adjacent access, rather than broader neighbourhood exposure, that appears most relevant for adolescent z-BMI.

Table 6 investigates the potential influence of fast-food outlets near adolescents' homes, in addition to the effects of school proximity. In our baseline analyses, we control for home proximity, recognising that the dietary environment around the home is likely to affect z-BMI. Our baseline result from Table 2 is copied in column (1), which includes both school exposure

emotional dysregulation, and cooperation. All assessments were administered at age seven and are available in MCS. Scores are standardised across tests due to differing scales.

[24]Beyond their role in official statistics, LSOAs are important units for local policy planning, with decisions often guided by the demographic and socioeconomic characteristics observed at the LSOA level.

and home exposure to fast-food outlets. Home exposure exhibits a markedly different pattern from school exposure. Among the home-based measures, only proximity to fish and chip shops is statistically significant (coefficient = 0.096, SE = 0.035), with an effect less than two-thirds the magnitude of that associated with proximity to a major-brand outlet near schools. One possible explanation is that food purchasing decisions around the home are more likely to be mediated by other household members, such as parents, whose preferences may differ from those of adolescents. In addition, fish and chip shops are more prevalent in residential neighbourhoods than major-brand outlets, potentially leading to more exposure to these outlets in the residential environment.

Home proximity, like school proximity, may also be subject to endogeneity concerns. Two mechanisms are particularly relevant: (i) openings or closures of fast-food outlets near the home, and (ii) residential mobility that alters exposure to existing outlets. These processes are unlikely to be random and may be correlated with unobserved neighbourhood characteristics that also influence adolescent z-BMI. Although we control for a set of neighbourhood characteristics, including IMD quintiles, residual endogeneity may persist. Moreover, potential measurement error—particularly for non-major-brand outlets in residential areas—may contribute to the attenuation and lack of robustness of the estimated home-based effects.

Unlike school exposure, we do not have plausibly exogenous variation in fast-food proximity around the home. To assess whether this limitation induces bias in the school-based OLS estimator, column (2) excludes all home proximity measures. The estimated coefficients on school-level fast-food exposure remain virtually unchanged, suggesting that school and home proximity are unlikely to be strongly correlated.

Finally, the baseline analysis implicitly assumes that residential mobility between the age-11 and age-14 sweeps is unrelated to fast-food availability. To relax this assumption, column (3) restricts the sample to adolescents who did not move residences between the two waves, thereby excluding approximately 15% of the sample. The results remain qualitatively unchanged, indicating that our findings are robust to potential endogeneity arising from residential mobility.

Overall, the results indicate that school-based exposure to fast-food outlets—particularly major-brand outlets—has a robust and independent effect on adolescent z-BMI that is not driven by residential fast-food availability. In contrast, home proximity to fast-food outlets plays a more limited and less robust role, with only weaker evidence for an effect of fish and chip shops. These findings reinforce the central importance of the school food environment in shaping adolescent weight outcomes, and support our identification strategy that exploits variation in fast-food exposure around schools.

7 Summary and conclusion

This paper examines the role of exposure to fast-food outlets in shaping adolescent BMI. While a substantial literature studies the relationship between fast-food availability and body weight, much of this evidence fails to account for the persistence of body weight, posing challenges for identification when fast-food exposure is correlated with pre-existing weight trajectories. This study provides new evidence on the causal effect of the food environment on adolescent BMI, by explicitly modelling BMI persistence and exploiting plausibly exogenous variation in fast-food exposure around schools.

We develop a novel approach to modelling persistence in adolescent z-BMI, based on a clustering of *pre-adolescent* z-BMI trajectories using k-means clustering. We then incorporate this persistence component into our estimation of the effects of fast-food availability around schools, allowing us to distinguish baseline obesity susceptibility from contemporaneous environmental effects. Our two-step approach enables persistence in continuous outcomes to be modelled using unequally spaced panel data, and is readily applicable to other persistent outcomes, such as test scores. For identification, we exploit the transition from primary to secondary school, which generates near-universal and largely exogenous changes in school location. This transition induces variation in fast-food exposure around schools that is plausibly unrelated to individual or household fast-food preferences, enabling us to isolate the causal effect of fast-food proximity on adolescent z-BMI.

Our empirical findings suggest that close proximity within 400 metres, approximately a 5-minute walk distance, of major-brand fast-food outlets around schools leads to a statistically and economically meaningful increase in adolescent z-BMI. Specifically, the baseline model estimates indicate that adolescents with at least one major-brand fast-food outlet near the school have, on average, a 0.158 standard-deviation higher z-BMI than those without such exposure, *ceteris paribus*. Given that the average BMI in our sample of adolescents is 21.34 with a standard deviation of 4.04 (c.f., Table 1), the estimate of 0.158 translates to an increase in BMI from 21.34 to 21.98 ($= 21.34 + 0.158 \times 4.04$). Although this is modest at the individual level, it is meaningful in terms of shifting the population BMI distribution, especially near clinical cut-offs.[25]

Proximity to non-major-brand outlets, in contrast, is not statistically significant at any road-network buffer distance. Proximity to fish and chip shops is occasionally significant but not robust, highlighting the distinctive influence of major-brand outlets. Importantly, our findings are robust to a wide range of alternative specifications, including individual fixed effects, alternative spatial measures of fast-food proximity, and samples excluding residential movers. Omitting early-life weight trajectories, however, attenuates the estimated effect of exposure to major-brand fast-food outlets and substantially reduces the model's explanatory power, underscoring the importance of accounting for persistence in weight trajectories when examining environmental determinants of obesity.

We find little evidence that the estimated effects differ across population subgroups. Specifically, we find little evidence of heterogeneity by gender, ethnicity, maternal education, neighbourhood deprivation, early-life obesity risk, or school urbanicity. The sole exception is the estimated effect of proximity to non-major-brand fast-food outlets near schools on z-BMI, which is larger among adolescents with non-White mothers and with mothers who have lower levels of education.

We also examine the effects of fast-food exposure at larger distances and around the home, which appear limited and less robust. These findings suggest that very local, school-based exposure is the spatial scale most relevant for adolescents, consistent with restricted mobility during the school day, and with evidence that fast-food outlet exposure can increase consumption (Svastisalee et al., 2016) through, for example, heightened temptation and self-control challenges (Laibson, 2001; DellaVigna, 2009; Boswell and Kober, 2016). Taken together, our findings support the rationale for policies adopted by some UK local authorities that restrict the opening of new fast-food outlets near schools.

[25] In terms of obesity prevalence, a 0.158 standard-deviation increase in the z-BMI distribution shifts the proportion of adolescents above the obesity threshold of the 95th centile (1.645), used for population monitoring purposes (OHID, 2025), from 5% to approximately 6.9%. Applied to a population of one million adolescents, this corresponds to roughly 20,000 additional adolescents with obesity.

That said, these findings should be viewed in the context of the much stronger influence of earlier-life weight trajectories. Although proximity to fast-food outlets is a statistically significant contributor to adolescent z-BMI, it explains only 0.4% of the variation, whereas persistence from earlier childhood—captured by cluster-group fixed effects—accounts for 41%. These findings suggest that pathways to overweight and obesity in adolescence are, at least partly, established earlier in life, and that interventions aimed at preventing excess weight gain in early childhood may be effective in reducing the prevalence of overweight and obesity during adolescence.

References

Abrahamsson, Sara Sofie, Aline Bütikofer, and Krzysztof Karbownik. 2023. "Swallow this: Childhood and adolescent exposure to fast food restaurants, BMI, and cognitive ability." Working Paper 31226, National Bureau of Economic Research.

Adair, Linda S. 2008. "Child and adolescent obesity: epidemiology and developmental perspectives." *Physiology & behavior* 94 (1): 8–16.

Alviola, Pedro A., Rodolfo M. Nayga, Michael R. Thomsen, Diana Danforth, and James Smartt. 2014. "The effect of fast-food restaurants on childhood obesity: A school level analysis." *Economics & Human Biology* 12 110–119.

Asirvatham, Jebaraj, Michael R. Thomsen, Rodolfo M. Nayga, and Anthony Goudie. 2019. "Do fast food restaurants surrounding schools affect childhood obesity?" *Economics & Human Biology* 33 124–133.

Bartle, Naomi C, Claire Hill, Laura Webber, Cornelia HM van Jaarsveld, and Jane Wardle. 2013. "Emergence and persistence of overweight and obesity in 7-to 11-year-old children." *Obesity Facts* 6 (5): 415–423.

Bivoltsis, Androniki, Emanuela Cervigni, Georgina Trapp, Matthew Knuiman, Paul Hooper, and Gina L. Ambrosini. 2018. "Food environments and dietary intakes among adults: Does the type of spatial exposure measurement matter? A systematic review." *International Journal of Health Geographics* 17 (1): 19.

Bollinger, Christopher R. 1996. "Bounding mean regressions when a binary regressor is mismeasured." *Journal of Econometrics* 73 (2:): 387–399, https://doi.org/10.1016/S0304-4076(95)01730-5.

Bollinger, Christopher R. 2010. "misclassification in binary variables." In *The New Palgrave Dictionary of Economics*, 1–5, Springer.

Bonhomme, Stéphane, and Elena Manresa. 2015. "Grouped patterns of heterogeneity in panel data." *Econometrica* 83 (3): 1147–1184.

Boskovic, Alexandra, Thomas Burgoine, and Jody Chantal Hoenink. 2026. "The Socioeconomic Patterning of Great Britain's Fast-Food Outlets and Supermarkets: A Repeated Cross-Sectional Study of Area-Level Deprivation and Food Outlet Density from 2011 to 2024." *Health & Place* 97 103603.

Boswell, Rebecca G., and Hedy Kober. 2016. "Food cue reactivity and craving predict eating and weight gain: A meta-analytic review." *Obesity Reviews* 17 (2): 159–177.

Chen, Xiao, Tao Pei, Ci Song et al. 2022. "Accessing public transportation service coverage by walking accessibility to public transportation under flow buffering." *Cities* 125 103646.

Cleo, G., E. Isenring, R. Thomas, and P. Glasziou. 2017. "Could habits hold the key to weight loss maintenance? A narrative review." *Journal of Human Nutrition and Dietetics* 30 (5): 655–664.

Cole, Tim J., Charlotte M. Wright, and Alan F. Williams. 2012. "Designing the New UK–WHO Growth Charts to Enhance Assessment of Growth Around Birth." *Archives of Disease in Childhood* 97 (2): 117–121.

Currie, Janet, Stefano DellaVigna, Enrico Moretti, and Vikram Pathania. 2010. "The effect of fast food restaurants on obesity and weight gain." *American Economic Journal: Economic Policy* 2 (3): 32–63.

DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature* 47 (2): 315–372.

Department for Education. 2026. "The National Curriculum: Overview." https://www.gov.uk/national-curriculum, Accessed 3 February 2026.

Dolton, Peter J., and Wiktoria Tafesse. 2022. "Childhood obesity, is fast food exposure a factor?" *Economics & Human Biology* 46 101153.

Dunn, Richard A, Jr Nayga, Rodolfo M, Michael R Thomsen, and Heather L Rouse. 2021. "A longitudinal analysis of fast-food exposure on child weight outcomes: Identifying causality through school transitions." *Q Open* 1 (1): qoaa007.

Green, Mark A., Matthew Hobbs, Ding Ding, Michael Widener, John Murray, Lindsey Reece, and Alex Singleton. 2021. "The Association between Fast Food Outlets and Overweight in Adolescents Is Confounded by Neighbourhood Deprivation: A Longitudinal Analysis of the Millennium Cohort Study." *Int J Environ Res Public Health* 18 (24): 1–15.

Han, Jeehee, Amy Ellen Schwartz, and Brian Elbel. 2020. "Does proximity to fast food cause childhood obesity? Evidence from public housing." *Regional Science and Urban Economics* 84 103565.

Hardy, R, J Johnson, A Park, and D O'Neill. 2019. "CLOSER work package 1: Harmonised height, weight and BMI user guide (revised)." *London: CLOSER*.

Hinte, Laura C, Daniel Castellano-Castillo, Adhideb Ghosh et al. 2024. "Adipose tissue retains an epigenetic memory of obesity after weight loss." *Nature* 636 (8042): 457–465.

Hobbs, M., A. Green M., E. Wilkins, K.E. Lamb, J. McKenna, and C. Griffiths. 2019. "Associations between food environment typologies and body mass index: Evidence from Yorkshire, England." *Social Science & Medicine* 239 112528.

Jia, Peng, Miyang Luo, Yamei Li, Ju-Sheng Zheng, Qian Xiao, and Jiayou Luo. 2021. "Fast-food restaurant, unhealthy eating, and childhood obesity: A systematic review and meta-analysis." *Obesity Reviews* 22 (S1): e12944.

Jiang, Jun, Patrick W C Lau, Yanhui Li et al. 2023. "Association of fast-food restaurants with overweight and obesity in school-aged children and adolescents: A systematic review and meta-analysis." *Obesity Review* 24 (3): 1–17.

Joshi, Heather, and Emla Fitzsimons. 2016. "The UK Millennium Cohort Study: the making of a multi-purpose cohort." *Longitudinal & Life Course Studies* 7 (4): 409–430.

Kluser, Frédéric, and Martina Pons. 2026. "The Apple Does Not Fall Far From the Tree: Intergenerational Persistence of Dietary Habits." *Review of Economics and Statistics*.

La Cruz, Alexandra, Erika Severeyn, Jesús Velásquez, Héctor Herrera, and Sara Wong. 2020. "Assessment of Anthropometric Measurements for Obesity and Abnormal Body Fat Percentage Diagnosis Using k-means as Clustering Technique." In *Conference on Information and Communication Technologies of Ecuador*, 177–191, Springer.

Laibson, David. 2001. "A Cue-Theory of Consumption." *The Quarterly Journal of Economics* 116 (1): 81–119.

Libuy, Nicolás, David Church, George Ploubidis, and Emla Fitzsimons. 2023. "Fast food proximity and weight gain in childhood and adolescence: Evidence from Great Britain." *Health Economics* 33 (3): 449–465.

Libuy, Nicolás, David Church, and Emla Fitzsimons. 2021. "Millennium Cohort Study: Linkage with the Point of Interest Data. User Guide (Version 1)."Technical report, UCL Centre for Longitudinal Studies, London, Published by UCL Centre for Longitudinal Studies, Economic and Social Research Council Resource Centre.

MacLean, Peter S, JA Higgins, ED Giles, VD Sherk, and MR Jackman. 2015. "The role for adipose tissue in weight regain after weight loss." *Obesity reviews* 16 45–54.

Office for Health Improvement and Disparities (OHID). 2025. "Definition of BMI Categories in Children and Adults." https://fingertips.phe.org.uk/profile/obesity-physical-activity-nutrition/supporting-information/definitions, Accessed June 18, 2025.

Office for National Statistics. 2021. "Statistical geographies." https://www.ons.gov.uk/methodology/geography/ukgeographies/statisticalgeographies, Accessed: 2025-06-30.

Ordnance Survey. 2010. "OS MasterMap Integrated Transport Network Layer v2.0." https://digimap.edina.ac.uk/help/gis/os_user_guides/?utm_source=chatgpt.com, Accessed: 2025-06-24.

Ordnance Survey. 2019. "Managing London's transport infrastructure with OS data." https://www.ordnancesurvey.co.uk/customers/case-studies/london-traffic-data, Accessed: 2025-06-24.

Ordnance Survey. 2025. "Points of Interest – Product Overview." https://docs.os.uk/os-downloads/products/addresses-and-names-portfolio/points-of-interest, Accessed: 2026-02-20.

Panayi, Panikos. 2014. *Fish and Chips: A History*. London: Reaktion Books.

Powell, Lisa M. 2009. "Fast food costs and adolescent body mass index: Evidence from panel data." *Journal of Health Economics* 28 (5): 963–970.

Qian, Yiwei, Michael R. Thomsen, Rodolfo M. Nayga, and Heather L. Rouse. 2017. "The effect of neighborhood fast food on children's BMI: Evidence from a sample of movers." *The B.E. Journal of Economic Analysis & Policy* 17 (4): 1–15.

van Rongen, Sofie, Maartje P. Poelman, Lukar Thornton, Gavin Abbott, Meng Lu, Carlijn B. M. Kamphuis, Kirsten Verkooijen, and Emely de Vet. 2020. "Neighbourhood fast food exposure and consumption: the mediating role of neighbourhood social norms." *International Journal of Behavioral Nutrition and Physical Activity* 17 (61): .

Simmonds, M., A. Llewellyn, C. G. Owen, and N. Woolacott. 2016. "Predicting adult obesity from childhood obesity: a systematic review and meta-analysis." *Obesity Reviews* 17 (2): 95–107.

Svastisalee, C., T. Pagh Pedersen, J. Schipperijn, S.E. Jørgensen, B.E. Holstein, and R. Krølner. 2016. "Fast-food intake and perceived and objective measures of the local fast-food environment in adolescents." *Public Health Nutrition* 19 (3): 446–455.

UNICEF. 2025. "Obesity Exceeds Underweight for the First Time Among School-Age Children and Adolescents Globally." https://www.unicef.org/press-releases/obesity-exceeds-underweight-first-time-among-school-age-children-and-adolescents, Accessed on 20 February 2026.

Vidmar, Suzanna, John Carlin, Kylie Hesketh, and Tim Cole. 2004. "Standardizing anthropometric measures in children and adolescents with new functions for egen." *The Stata Journal* 4 (1): 50–55.

Wilkins, Emma L., Duncan Radley, Michelle A. Morris, and Claire Griffiths. 2017. "Examining the validity and utility of two secondary sources of food environment data against street audits in England." *Nutrition Journal* 16 (82): .

Wilkins, Emma, Duncan Radley, Michelle Morris, Matthew Hobbs, Alex Christensen, Windi Lameck Marwa, Adele Morrin, and Claire Griffiths. 2019. "A systematic review employing the GeoFERN framework to examine methods, reporting quality and associations between the retail food environment and obesity." *Health & Place* 57 186–199.

Xiang, Huasheng, Louis Goffe, Viviana Albani, Nasima Akhter, Amelia A. Lake, and Heather Brown. 2024. "Planning policies to restrict fast food and inequalities in child weight in England: a quasi-experimental analysis." *Obesity* 32 (12): 2345–2353.

Figures & Tables

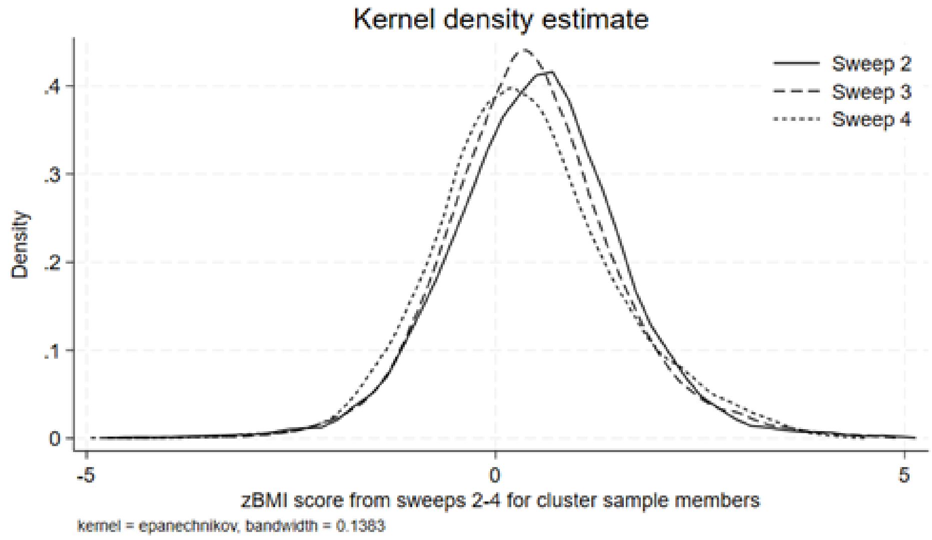

Figure 1: Mean z-BMI in Childhood

Notes: There are 10,011 cohort members, each contributing a balanced panel across the age-3, age-5, and age-7 sweeps (Sweeps 2–4). The mean (standard deviation) of z-BMI is 0.492 (1.06) at age 3, 0.432 (1.05) at age 5, and 0.365 (1.10) at age 7.

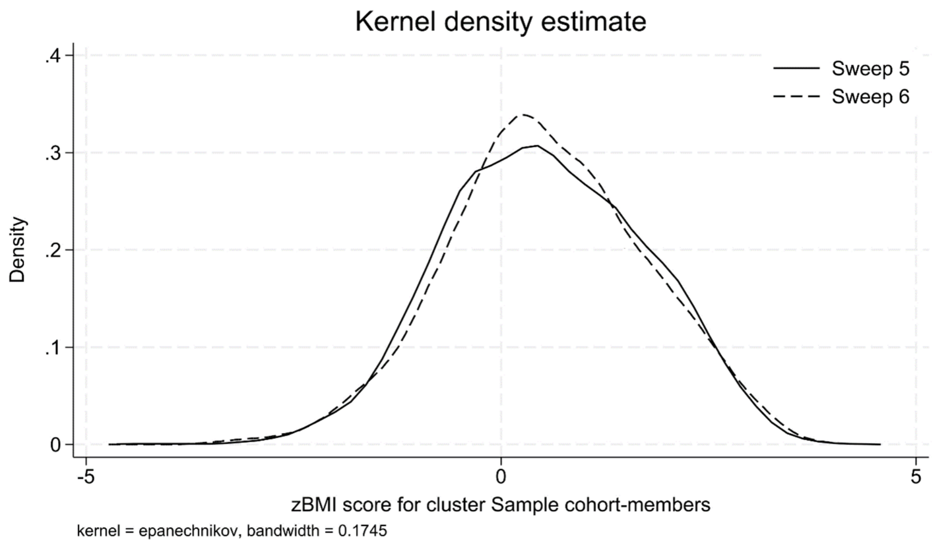

Figure 2: Mean z-BMI in Adolescence

Notes: The estimation sample comprises 6,626 cohort members, who transitioned from primary school to secondary school between the age-11 and age-14 sweeps (Sweeps 5 and 6).

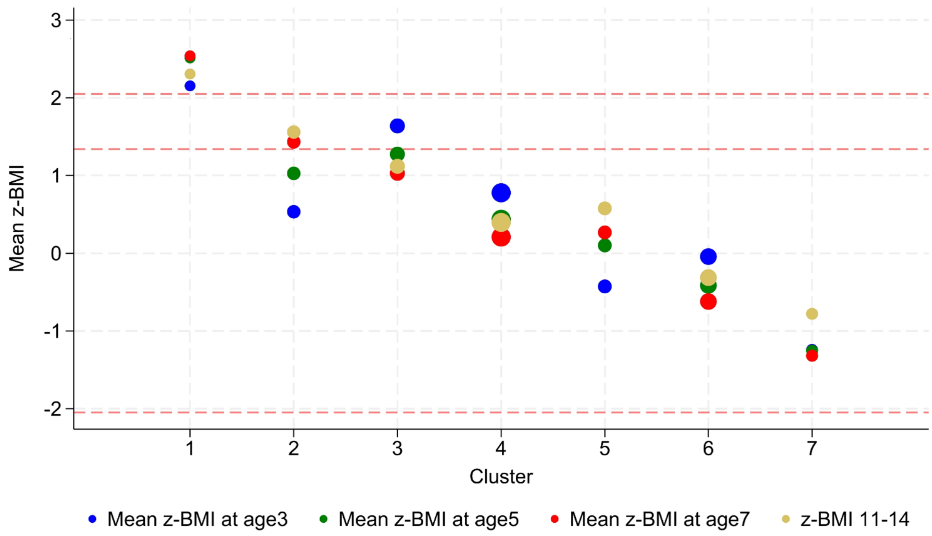

Figure 3: Mean z-BMI by Cluster

Notes: The three dotted horizontal lines indicate the clinical thresholds for obesity, overweight, and underweight, respectively. The sample used for the plot consists of 9,727 cohort members with a valid cluster identifier. The number of children (% sample) in each cluster is: Cluster 1, 676 (7.0%); Cluster 2, 1,138 (11.7%); Cluster 3, 1,447 (14.9%); Cluster 4, 2,500 (25.7%); Cluster 5, 1,233 (12.7%); Cluster 6, 1,860 (19.1%); and Cluster 7, 873 (9.0%).

Table 1: BMI, z-BMI, and Fast-food Exposure around Schools by Cluster

	BMI	z-BMI	Major-brand fast food	Non-major-brand fast food	Fish and chips	N (%)
Cluster group:						
1	28.07	2.23	0.05	0.24	0.15	404
	(4.83)	(0.81)				(7)
2	24.61	1.52	0.03	0.21	0.12	716
	(4.13)	(0.89)				(11.7)
3	22.84	1.10	0.05	0.21	0.11	964
	(3.31)	(0.89)				(14.9)
4	20.61	0.41	0.05	0.21	0.12	1735
	(2.64)	(0.86)				(25.7)
5	21.32	0.60	0.05	0.23	0.14	815
	(3.32)	(0.98)				(12.7)
6	18.98	-0.24	0.04	0.23	0.13	1344
	(2.37)	(0.94)				(19.1)
7	18.13	-0.68	0.08	0.29	0.16	648
	(2.60)	(1.13)				(9)
Total	21.34	0.53	0.05	0.23	0.13	6626
	(4.04)	(1.21)				(100)

Notes: Means, standard deviations (for continuous variables only), and numbers of observations are reported by cluster. The sample consists of 6,626 cohort members in the age-14 sweep. Binary indicators for availability of an outlet are recorded as 1 if at least one outlet is available within 400m of the school, and 0 otherwise.

Table 2: Estimated Effects of Fast-Food Proximity on Adolescent z-BMI

	(1)	(2)	(3)	(4)
	No clusters	With clusters	Excl. extreme childhood BMI	Age-17 zBMI
Availability within 400m from school				
Major-brand fast food	0.096	0.158***	0.159***	0.162**
	(0.068)	(0.054)	(0.057)	(0.069)
Non-major-brand fast food	-0.009	0.039	0.067*	-0.001
	(0.046)	(0.036)	(0.037)	(0.045)
Fish & chips	0.053	0.030	0.017	0.103**
	(0.050)	(0.040)	(0.043)	(0.051)
Cluster fixed effects	No	Yes	Yes	Yes
Home fast-food proximity	Yes	Yes	Yes	Yes
Child, mother, family controls	Yes	Yes	Yes	Yes
Area controls	Yes	Yes	Yes	Yes
Observations	6,626	6,626	5,574	5,393
R-squared	0.113	0.455	0.342	0.377

Notes: (i) Fast-food availability variables are binary indicators. (ii) Control variables include sex, single-child status, single-parent household, housing tenure, mother's employment status, mother's education at birth, mother's smoking and drinking status, ethnicity, household social class, log of LSOA population, Index of Multiple Deprivation (quintiles), number of other food outlets within 400 metres of the cohort member's school and home, and cluster fixed effects. (iii) Column (3) excludes cluster groups 1 and 7, corresponding to children with extremely high and low childhood z-BMI, respectively. (iv) In column (4), the dependent variable is z-BMI at age 17 for cohort members with valid age-17 sweep observations; all other variables are measured at the age-14 sweep. (v) All equations are estimated by ordinary least squares using 6,626 cohort members in the age-14 sweep, and robust standard errors are reported in parentheses. (vi) *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 3: Adolescent z-BMI and Fast-food Proximity to school: Heterogeneity Analysis

	(1)	(2)	(3)	(4)	(5)
A. Base group:	Baseline	Male	White	High maternal education	Low deprivation
Availability within 400m from school					
Major-brand fast food	0.158***	0.210***	0.184***	0.169***	0.136**
	(0.054)	(0.077)	(0.059)	(0.057)	(0.067)
Non-major-brand fast food	0.039	0.044	0.003	0.006	0.018
	(0.036)	(0.049)	(0.038)	(0.038)	(0.042)
Fish & chips	0.030	-0.008	0.042	0.042	0.081
	(0.040)	(0.058)	(0.046)	(0.043)	(0.051)
B. Interactions:		Female	Non-White	Low maternal education	High deprivation
Major-brand fast food		-0.110	-0.136	-0.034	0.027
		(0.106)	(0.131)	(0.153)	(0.108)
Non-major-brand fast food		-0.010	0.181**	0.173**	0.034
		(0.060)	(0.076)	(0.083)	(0.059)
Fish & chips		0.078	-0.091	-0.070	-0.092
		(0.076)	(0.094)	(0.101)	(0.076)
C. Joint significance of interactions					
F-statistic		0.66	1.98	1.48	0.50
p-value		0.58	0.12	0.22	0.69
Cluster fixed effects	Yes	Yes	Yes	Yes	Yes
Home fast-food proximity	Yes	Yes	Yes	Yes	Yes
Child, mother, family controls	Yes	Yes	Yes	Yes	Yes
Area controls	Yes	Yes	Yes	Yes	Yes

Notes: (i) See Table 2 notes for the sample and model specifications. (ii) Column titles indicate the dimensions of heterogeneity examined. (iii) High maternal education is defined as mothers with A-level, O-level, GCSE or equivalent qualifications or above. (iv) Low deprivation refers to the Index of Multiple Deprivation quintiles being 3 or above. (v) The F-statistic tests the joint significance of the interaction terms (df = 3, 6,625). (vi) *** p<0.01 and ** p<0.05.

Table 4: Adolescent z-BMI and Fast-food Proximity: Alternative Specifications

	(1)	(2)	(3)	(4)	(5)
	Baseline	Uses age-11 & age-14 sweeps	WG estimation - controls for individual FE	Baseline + cognitive test scores	Baseline + non-cognitive test scores
Availability within 400m from school					
Major-brand fast food	0.158***	0.157***	0.095**	0.146***	0.148***
	(0.054)	(0.039)	(0.039)	(0.038)	(0.038)
Non-major-brand fast food	0.039	0.019	-0.014	0.016	0.017
	(0.036)	(0.023)	(0.025)	(0.025)	(0.023)
Fish & chips	0.030	0.050*	-0.019	0.056*	0.056**
	(0.040)	(0.027)	(0.032)	(0.029)	(0.028)
Cluster fixed effects	No	Yes	Yes	Yes	Yes
Individual fixed effects	No	No	Yes	No	No
Home fast-food proximity	Yes	Yes	Yes	Yes	Yes
Child, mother, family controls	Yes	Yes	Yes	Yes	Yes
Area controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.455	0.493	0.025	0.497	0.496
Observations	6,626	13,252	13,252	6,453	6,506

Notes: (i) See the notes to Table 2. (ii) Columns (2) and (3) use the sample consisting of age-11 and age-14 sweeps and additionally include time (sweep) fixed effects and cluster-by-time fixed effects; (iii) Column (3) uses the within-group estimator; (iv) Columns (4) and (5) include cognitive and non-cognitive test scores, respectively. (v) *** p<0.01, ** p<0.05, * p<0.10.

Table 5: Adolescent z-BMI and Fast-food Proximity: Alternative Proximity Measures

	(1)	(2)	(3)	(4)
Availability within:	Baseline 400m	800m	1600m	LSOA
Major-brand fast food	0.158***	0.023	0.004	0.036
	(0.054)	(0.033)	(0.026)	(0.041)
Non-major-brand fast food	0.039	-0.002	0.026	0.004
	(0.036)	(0.029)	(0.045)	(0.028)
Fish & chips	0.030	0.034	0.010	0.046
	(0.040)	(0.027)	(0.034)	(0.033)
Cluster fixed effects	Yes	Yes	Yes	Yes
Home fast-food proximity	Yes	Yes	Yes	Yes
Child, mother, family controls	Yes	Yes	Yes	Yes
Area controls	Yes	Yes	Yes	Yes
Observations	6,626	6,626	6,626	6,626
R-squared	0.455	0.454	0.454	0.454

Notes: (i) See the notes to Table 2. (ii) The models use the baseline specification with alternative measures of proximity to fast-food outlets: 400m-buffer, 800m-buffer, and 1,600m-buffer zones, and within the same Lower Super Output Area (LSOA). (iii) *** p<0.01.

Table 6: Adolescent zBMI and Fast-food Proximity: Role of Availability near Home

	(1)	(2)	(3)
	Baseline	(1) without home exposure	(1) excluding house movers
Availability within 400m from school			
Major-brand fast food	0.158***	0.152***	0.174***
	(0.054)	(0.054)	(0.058)
Non-major-brand fast food	0.039	0.041	0.031
	(0.036)	(0.036)	(0.039)
Fish & chips	0.030	0.034	0.039
	(0.040)	(0.040)	(0.043)
Availability within 400m from home			
Major-brand fast food	-0.038		-0.078
	(0.062)		(0.066)
Non-major-brand fast food	-0.015		0.004
	(0.033)		(0.037)
Fish & chips	0.096***		0.074**
	(0.035)		(0.038)
Cluster fixed effects	Yes	Yes	Yes
Child, mother, family controls	Yes	Yes	Yes
Area controls	Yes	Yes	Yes
Observations	6,626	6,626	5,597
R-squared	0.455	0.454	0.455

Notes: (i) See the notes to Table 2. (ii) Column (2) uses the baseline specification but excludes the home proximity variables. (iii) Column (3) is estimated using the subset of 5,597 cohort members in the baseline sample who remained at the same residence between the age-11 and age-14 sweeps. (iv) *** p<0.01 and ** p<0.05.

Appendix A

Constructing Proximity to Fast Food

This section outlines the construction of our key variable of interest—proximity to fast-food outlets. While our approach is informed by Libuy et al. (2021), it differs in important respects in both the definition of the proximity measures and the classification of food outlets, as detailed below.

Step 1: Defining buffers around residences and schools

Using the Ordnance Survey's ITN dataset and postcode centroids for both residence and school locations, we construct road-network buffers around each cohort member's residence and school addresses using the Geographic Information System (GIS) method. Our primary specification uses a 400-metre buffer, corresponding to an approximate five-minute walking distance, although we also try alternative buffer sizes.[26]

Step 2: Identifying food outlets

We extract the location of establishments that allow food to be taken away or delivered from the Ordnance Survey's PoI dataset. These include not only fast-food outlets but also supermarkets, convenience stores, bakeries, and cafes, but exclude establishments serving only seated meals. Food establishments are identified using the PoI classifications, as explained in Section 2.2, and name searches within relevant PoI classes.[27] Since our analyses use the MCS survey age-11 and age-14 sweeps, conducted around 2012 and 2015, we use the September extracts of PoI data from the preceding years, 2011 and 2014. respectively.

Step 3: Categorising food outlets

We aggregate identified outlets into five categories to capture potential heterogeneous effects on z-BMI:

1. Large-brand fast-food outlets: McDonald's, KFC, Burger King, Wimpy, Subway, Pizza Hut, Pizza Express, Domino's Pizza, Dixie Chicken, Chicken Cottage, Papa John's, Southern Fried Chicken, Five Guys, Harry Ramsden's, Little Chef, etc. (Wilkins et al., 2019).

2. Small-brand fast-food outlets: Independent kebab shops, fried chicken outlets, and other chicken shops, identified by keyword searches such as kebab, fried chicken, or chicken.

[26] The finest location information of residence and school available to us is at the postcode level, where each postcode typically contains about 15 residences or, in some cases, one large establishment (e.g. a hospital, company, or school).

[27] Refer to the PoI product information for further details of the classification scheme (Ordnance Survey, 2025).

3. Fish and chip shops.
4. Supermarkets and convenience stores: Tesco, Sainsbury's, Asda, Waitrose, McColl's, Morrisons, Aldi, Lidl, Marks & Spencer, Iceland, etc.
5. Cafes, snack bars, tea rooms, bakeries, confectioneries, and delicatessens: Starbucks, Costa, Caffe Nero, Greggs, Pret a Manger etc.

Step 4: Constructing the proximity variable

We construct our variable of interest which is a binary indicator, recording whether there is at least one food outlet in each of the five categories, within the 400-meter buffer of a school or residence, as defined in Step 1.

Table A1: Summary Statistics by Cluster

Cluster group:	1	2	3	4	5	6	7	Total
Food-outlet availability								
Major-brand fast food – *School*	0.05	0.03	0.05	0.05	0.05	0.04	0.08	0.05
Non-brand fast food – *School*	0.24	0.21	0.21	0.21	0.23	0.23	0.29	0.23
Fish & chips – *School*	0.15	0.12	0.11	0.12	0.14	0.13	0.16	0.13
Café – *School*	0.21	0.18	0.17	0.17	0.18	0.18	0.23	0.18
All supermarkets – *School*	0.45	0.38	0.36	0.40	0.42	0.40	0.47	0.40
Major-brand fast food – *Home*	0.07	0.03	0.04	0.03	0.06	0.05	0.06	0.05
Non-brand fast food – *Home*	0.32	0.28	0.26	0.25	0.29	0.26	0.32	0.27
Fish & chips – *Home*	0.20	0.19	0.18	0.15	0.18	0.16	0.17	0.17
Café – *Home*	0.21	0.19	0.19	0.19	0.19	0.18	0.23	0.19
All supermarkets – *Home*	0.51	0.46	0.44	0.46	0.50	0.46	0.50	0.47
Body Mass Index (BMI)								
Adolescent zBMI	2.23	1.52	1.10	0.41	0.60	-0.24	-0.68	0.53
	(0.81)	(0.89)	(0.89)	(0.86)	(0.98)	(0.94)	(1.13)	(1.21)
Mother's BMI pre-pregnancy	23.06	21.63	21.22	21.06	20.53	19.81	17.99	20.65
	(10.76)	(9.45)	(9.18)	(8.40)	(9.16)	(8.41)	(9.24)	(9.05)
Mother's BMI missing	0.14	0.13	0.13	0.11	0.14	0.13	0.19	0.13
	(0.34)	(0.33)	(0.34)	(0.31)	(0.35)	(0.33)	(0.39)	(0.34)
Adolescent characteristics								
Female	0.43	0.49	0.47	0.51	0.51	0.53	0.51	0.50
Only child living at home	0.16	0.18	0.12	0.12	0.14	0.13	0.12	0.13
Mother's characteristics								
Ethnicity - White	0.76	0.86	0.89	0.91	0.81	0.83	0.64	0.84
No smoking pre-pregnancy	0.63	0.62	0.67	0.69	0.66	0.72	0.70	0.68
No smoking during pregnancy	0.69	0.70	0.73	0.75	0.74	0.79	0.77	0.75
No drink during pregnancy	0.68	0.67	0.60	0.60	0.64	0.63	0.66	0.63
Single parent	0.22	0.23	0.20	0.20	0.24	0.21	0.21	0.21
Employed	0.69	0.73	0.76	0.79	0.72	0.75	0.66	0.74
Mother's education								
Higher Education	0.23	0.27	0.34	0.35	0.28	0.33	0.28	0.31
A-/O-/GCSE Level	0.57	0.56	0.51	0.53	0.53	0.51	0.45	0.52
No qual's/Other	0.20	0.17	0.15	0.12	0.18	0.16	0.27	0.17
Age at the birth of the child								
Aged < 25	0.25	0.24	0.23	0.23	0.27	0.22	0.25	0.23
Aged 26–30	0.33	0.32	0.31	0.32	0.30	0.30	0.30	0.31
Aged 31–35	0.25	0.27	0.31	0.31	0.27	0.32	0.29	0.30
Aged 36+	0.18	0.16	0.15	0.15	0.17	0.17	0.16	0.16

Table continues

Table A1 – continued

Cluster group:	1	2	3	4	5	6	7	Total
Family characteristics								
Family social class								
Unemployed, Not in paid work	0.22	0.24	0.22	0.19	0.25	0.22	0.28	0.22
Self-employed, Lower supervisory, Routine	0.32	0.26	0.26	0.25	0.27	0.24	0.26	0.26
Managerial, Professional, Intermediate	0.46	0.50	0.52	0.56	0.49	0.55	0.46	0.52
Family housing tenure								
Owned outright	0.11	0.10	0.12	0.11	0.12	0.13	0.15	0.12
Partly owned	0.53	0.56	0.60	0.64	0.57	0.61	0.58	0.60
None of the others	0.25	0.24	0.19	0.17	0.22	0.17	0.19	0.19
Private renters	0.11	0.10	0.08	0.09	0.08	0.09	0.08	0.09
Area characteristics								
Country of residence								
England	0.75	0.72	0.71	0.71	0.76	0.75	0.82	0.74
Wales	0.16	0.17	0.16	0.15	0.14	0.13	0.10	0.14
Scotland	0.09	0.12	0.13	0.14	0.11	0.12	0.08	0.12
Index of multiple deprivation								
1st quintile	0.27	0.22	0.18	0.16	0.23	0.18	0.27	0.20
2nd quintile	0.23	0.21	0.19	0.18	0.20	0.16	0.19	0.19
3rd quintile	0.17	0.19	0.20	0.19	0.21	0.19	0.13	0.19
4th quintile	0.15	0.16	0.18	0.22	0.18	0.21	0.19	0.19
5th quintile	0.18	0.22	0.25	0.24	0.19	0.26	0.22	0.23
Total population	1611	1551	1561	1567	1616	1567	1638	1580
	(413)	(417)	(438)	(444)	(449)	(429)	(417)	(434)
N	404	716	964	1735	815	1344	648	6626

Notes: For continuous variables, we report means with standard deviations in parentheses; for binary variables, we report proportions. All summary statistics are computed for the 6,626 cohort members in the base estimation sample who transitioned from primary to secondary school between the age-11 and age-14 sweeps.

Appendix B

Table B1: Persistence in Observed z-BMI across Sweeps: Correlation Coefficients

Sweep	2	3	4	5	6
2	1.00				
3	0.71	1.00			
4	0.63	0.81	1.00		
5	0.47	0.64	0.78	1.00	
6	0.43	0.58	0.70	0.83	1.00

Notes: Sweeps 2-6 are measured at approximately ages 3, 5, 7, 11, and 14, respectively. Correlations for sweeps 2–4 are based on a balanced panel of 10,011 cohort members. Correlations involving sweeps 5–6 use the same cohort but include only non-missing z-BMI, hence the contributing sample may vary across correlations.

Table B2: Full Estimation Results

Dependent variable: Adolescent z-BMI Mean = 0.53 (SD 1.21)			
Food-outlet availability		**Mother/family characteristics** - cont.	
Major-brand fast food (school)	0.158***	Single parent	0.072**
	(0.054)		(0.030)
Non-brand fast food (school)	0.039	Employed	0.005
	(0.036)		(0.030)
Fish & chips (school)	0.030	Education (ref: Degree)	
	(0.040)	A-/O-/GCSE level	-0.026
Café (school)	-0.073*		(0.027)
	(0.038)	No qual / Others	-0.095**
All supermarkets (school)	-0.001		(0.042)
	(0.028)	No smoking pre-pregnancy	0.000
Major-brand fast food (home)	-0.038		(0.040)
	(0.062)	No smoking at birth	-0.076*
Non-brand fast food (home)	-0.015		(0.042)
	(0.033)	No drinking during preg.	0.015
Fish & chips (home)	0.096***		(0.024)
	(0.035)	BMI before pregnancy	0.044***
Café (home)	0.006		(0.003)
	(0.035)	BMI missing	1.115***
All supermarkets (home)	0.014		(0.077)
	(0.027)	*Family social class*	
Cluster group (ref: Group 1)		*(ref: Unemployed/not in paid work)*	
Group 2	-0.633***	Self-employed/routine	-0.007
	(0.050)		(0.036)
Group 3	-1.007***	Managerial / professional	-0.086**
	(0.048)		(0.036)
Group 4	-1.666***	Housing (ref: Owned)	
	(0.044)	Partly owned	0.013
Group 5	-1.513***		(0.035)
	(0.051)	None of the others	0.099**
Group 6	-2.295***		(0.047)
	(0.047)	Private renters	0.151***
Group 7	-2.742***		(0.054)
	(0.059)		
Adolescent characteristics		**Area characteristics**	
Female	0.184***	*IMD (ref: 1st quintile)*	
	(0.022)	2nd quintile	-0.014
Only child at home	0.096***		(0.039)
	(0.034)	3rd quintile	-0.034
Mother/family characteristics			(0.040)
Ethnicity: White	-0.106***	4th quintile	-0.014
	(0.039)		(0.040)
Age (ref: < 25)		5th quintile (least dep.)	-0.063
26–30	-0.045		(0.040)
	(0.033)	Log population (LSOA)	-0.043
31–35	-0.047		(0.059)
	(0.035)	*Country (ref: England)*	
36+	-0.008	Wales	0.082**
	(0.040)		(0.033)
		Scotland	-0.041
			(0.056)

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B3: Changes in Fast-food Outlet Availability within 400m of Schools during the Transition from Primary to Secondary School, by Cluster

Cluster	Fast food availability status	Total	*Major-brand* fast food				*Non-major-brand* fast food				Fish & chips			
			0–0	0–1	1–1	1–0	0–0	0–1	1–1	1–0	0–0	0–1	1–1	1–0
1	Proportion	1.000	0.941	–	0.027	–	0.708	0.072	0.166	0.054	0.812	0.047	0.106	0.035
	N	404	380	–	11	–	286	29	67	22	328	19	43	14
2	Proportion	1.000	0.950	–	–	–	0.728	0.067	0.142	0.063	0.834	0.041	0.080	0.46
	N	716	680	–	–		521	48	102	45	597	29	57	33
3	Proportion	1.000	0.933	0.027	0.025	0.016	0.716	0.075	0.133	0.077	0.830	0.035	0.077	0.56
	N	964	899	26	24	15	690	72	128	74	800	34	74	58
4	Proportion	1.000	0.925	0.029	0.025	0.021	0.724	0.067	0.141	0.068	0.832	0.046	0.070	0.52
	N	1,735	1,605	50	44	36	1,257	116	244	118	1,444	79	122	90
5	Proportion	1.000	0.935	0.028	0.025	0.012	0.693	0.071	0.156	0.080	0.815	0.042	0.102	0.042
	N	815	762	23	20	10	565	58	127	65	664	34	83	34
6	Proportion	1.000	0.946	0.022	0.019	0.013	0.690	0.069	0.161	0.080	0.808	0.041	0.089	0.063
	N	1,344	1,271	29	26	18	927	93	217	107	1,086	55	119	84
7	Proportion	1.000	0.903	0.045	–	–	0.622	0.076	0.219	0.083	0.792	0.065	0.094	0.049
	N	648	585	29	–	–	403	49	142	54	513	42	61	32
Overall	Proportion	1.000	0.933	0.026	0.025	0.016	0.702	0.070	0.155	0.073	0.820	0.044	0.084	0.052
	N	6,626	6,182	172	166	106	4,649	465	1,027	485	5,432	292	559	303

Notes: Share and number (N) of observations are reported by cluster group and fast-food availability status, for each fast-food outlet category. Availability of an outlet is recorded as 1 if available, and 0 otherwise. The sample used for calculations consists of a balanced panel of 6,626 cohort members who transitioned their schools from primary (age-11 sweep) to secondary school (age-14 sweep). For reasons of confidentiality, the entries are suppressed when the count < 10. In addition, complementary suppression also has been applied in some of the relevant cells due to disclosure control rules.

Appendix C

Data List

The table below lists the data files from the Millennium Cohort Study used in our analyses.

Data file
Centre for Longitudinal Studies. 2001. *Millennium Cohort Study: Age 9 Months, Sweep 1, 2001*. DOI: 10.5255/UKDA-SN-4683-1. SN: 4683.
Centre for Longitudinal Studies. 2004. *Millennium Cohort Study: Age 3, Sweep 2, 2004*. DOI: 10.5255/UKDA-SN-5350-1. SN: 5350.
Centre for Longitudinal Studies. 2006. *Millennium Cohort Study: Age 5, Sweep 3, 2006*. DOI: 10.5255/UKDA-SN-5795-1. SN: 5795.
Centre for Longitudinal Studies. 2008. *Millennium Cohort Study: Age 7, Sweep 4, 2008*. DOI: 10.5255/UKDA-SN-6411-1. SN: 6411.
Centre for Longitudinal Studies. 2012. *Millennium Cohort Study: Age 11, Sweep 5, 2012*. DOI: 10.5255/UKDA-SN-7464-1. SN: 7464.
Centre for Longitudinal Studies. 2015. *Millennium Cohort Study: Age 14, Sweep 6, 2015*. DOI: 10.5255/UKDA-SN-8156-1. SN: 8156.
Centre for Longitudinal Studies. 2018. *Millennium Cohort Study: Age 17, Sweep 7, 2018*. DOI: 10.5255/UKDA-SN-8682-1. SN: 8682.
Centre for Longitudinal Studies. 2020. *Millennium Cohort Study: Sweep 7 Geographical Identifiers Using 2001 Census Boundaries: Secure Access*. DOI: 10.5255/UKDA-SN-8758-1. SN: 8758.
Centre for Longitudinal Studies. 2023. *Harmonised Height, Weight and BMI in Five Longitudinal Cohort Studies: Millennium Cohort Study*. DOI: 10.5255/UKDA-SN-8550-1. SN: 8550.
Centre for Longitudinal Studies. 2025. *Millennium Cohort Study: Sweeps 1–8, 2001–2025: Longitudinal Files*. DOI: 10.5255/UKDA-SN-8172-1. SN: 8172.