CAGE

Working Paper

791/2026
February 2026

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

Yu Aoki-Beattie,
Wiji Arulampalam,
Neil Lloyd,
Sushil Mathew

ISSN: 2978-0276
Grant number: ES/7504701/1

UNIVERSITY OF WARWICK

UKRI Economic and Social Research Council

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

Yu Aoki-Beattie
University of Aberdeen
& IZA

Wiji Arulampalam
University of Warwick
CAGE & IZA

Neil Lloyd
University of St Andrews

Sushil Mathew
Imperial College London

February 2026

Abstract

We examine the causal effect of exposure to fast-food outlets on adolescent z-BMI using data from the UK Millennium Cohort Study. We develop a novel approach to modelling persistence in adolescent BMI by clustering early childhood BMI trajectories, capturing biologically and behaviourally persistent obesity risk profiles. Including these profiles in the model allows us to separate baseline susceptibility from contemporaneous environmental effects. For identification, we exploit the near-universal transition from primary to secondary school in Great Britain, which creates plausibly exogenous variation in exposure to fast-food outlets around schools. Using this variation, we find that adolescents with at least one major-brand outlet within 400 metres of their school have, on average, a 0.158 standard-deviation higher z-BMI. Effects decline at larger distances, are limited around the home, and do not extend to other food outlets.

JEL Codes: I12, I18, L83.

Keywords: Adolescent obesity, Body mass index, Fast food, School food environment.

*We are grateful to the UK Data Services (UKDS) for facilitating access to their Secure Laboratory, and to UCL-CLS for enabling the use of postcode-level data required for our analyses. We also thank the participants at the American-European Health Economics Study Group meeting in Barcelona (2022), SUFE-SDU-Warwick Economics Workshop, China (2024), West China School of Public Health (2024), and Warwick seminar (2023) for their valuable comments. Financial support from CAGE (Department of Economics at 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.[1]

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.[2] 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 issues 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 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 accu-

[1] Overweight in children is defined as a BMI (weight in kilograms divided by height in meter-squared) at or above the 85th percentile but below the 95th percentile, while obesity is defined as a BMI at or above the 95th percentile for their age and sex (Cole et al., 1995).

[2] In food environment research, proximity refers to the spatial closeness between an individual's residence or school and the nearest food outlet, typically operationalised using distance or travel time measures. Access, by contrast, captures the ability of individuals to obtain food from these outlets, incorporating factors such as economic resources and other constraints that may facilitate or hinder food acquisition.

mulation and habit formation (MacLean et al., 2015; Cleo et al., 2017). 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 and omitted dynamic z-BMI trajectories that commonly arise in studies of adolescent and adult BMI.

Our second contribution concerns identification. We exploit the institutional transition from primary to secondary school in Great Britain, which generates near-universal and largely exogenous changes in school location. This transition creates variation in school-level fast-food exposure that is plausibly unrelated to individual or household preference for fast food, enabling us to isolate the causal effect of fast-food proximity on adolescent z-BMI.

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 broad categories that differ in branding, business models, and spatial distribution - major-brand outlets (e.g. McDonald's), non-major-brand outlets (e.g. kebab shops), and fish and chip shops - and measure proximity to each outlet type using road-network buffers.

Our main results show that adolescents with at least one major-brand fast-food outlet near their school have, on average, a 0.158 standard-deviation higher z-BMI than those without such exposure. In contrast, proximity to a fish and chip shop or a non-major-brand outlet has no statistically significant effect. These findings point to a distinctive role for major-brand fast food in shaping adolescent z-BMI. Importantly, we find that our results are robust to a wide range of alternative specifications, controlling for individual fixed-effects, alternative spatial measures of fast-food proximity, and samples that exclude residential movers.

Concerning the relevant area size, significant effects arise only within a 400-meter buffer zone (equivalent of roughly 5-minute walk distance), and no statistically significant effects are detected for 800- or 1600-meter buffer zones. This implies that close proximity to fast-food outlets is what matters for adolescent z-BMI, consistent with the US evidence from Currie et al. (2010). Our findings are also consistent with empirical evidence suggesting that greater local access to fast-food outlets increases consumption among children (see Jia et al. (2021) for a review), possibly by amplifying temptation through visual cues (Laibson, 2001; Boswell and Kober, 2016; Kanoski and Boutelle, 2022) and exacerbating self-control problems (DellaVigna, 2009). In addition, fast-food consumption carries a social dimension (van Rongen et al., 2020), being normalised among groups such as adolescents, potentially reinforcing frequent use.

The existing causal evidence on fast-food availability and child or adolescent BMI is typically based on one of three research designs.[3] First, a strand of literature uses observational panel data to exploit within-individual (Powell, 2009; Dunn et al., 2021; Abrahamsson et al., 2023; Libuy et al., 2023) or within-location (Currie et al., 2010) variation in availability of fast food near school and/or home. Notably, Dunn et al. (2021) also exploit within-student changes in school-related fast-food exposure arising from the transition between primary and intermediate school in Arkansasin the US. They find no significant relationship between fast-food availability along the route between school and home and changes to student z-BMI, conditional on the previous z-BMI. In contrast, other panel studies, including Currie et al. (2010), report strong evidence of a positive relationship between geographical exposure to fast food and BMI or obesity.

Second, there is limited but some evidence from natural experiments and direct policy evaluations. Han et al. (2020) exploit plausibly exogenous variation in fast-food exposure generated by the random allocation of households to public housing in New York City, and find a positive effect on childhood obesity, particularly among children from lower socioeconomic backgrounds. Xiang et al. (2024) evaluate local planning reforms in England that restrict the opening of new fast-food outlets and find a significant reduction in obesity and overweight prevalence in more deprived areas. Third, some studies employ instrumental-variables 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,

[3] Correlation studies are not included in our discussions, but refer to Jiang et al. (2023) for a systematic review of fast-food availability and health-related outcomes in children and adolescents.

although Dolton and Tafesse (2022) find no statistically significant effect.

With the exception of some instrumental-variable studies (e.g., Alviola et al., 2014), the existing literature primarily relies on time variation in fast-food exposure at home or school to identify causal effects. The econometric models typically adopt a two-way fixed effects structure, controlling for time-invariant unobserved heterogeneity and time effects. However, the lag of the outcome variable, BMI or obesity status, is not included as an explanatory variable, and dynamic persistence in BMI is generally ignored. This will introduce an omitted variable bias if the persistence in BMI is time-varying, and the variation in availability of fast-food outlets is correlated with these underlying dynamics in BMI. The scarcity of dynamic panel data models in this literature largely reflects data limitations. Specifically, reliable estimation of dynamic models requires long and regularly spaced panels, whereas studies of child and adolescent BMI typically rely on short panels covering school-age years.[4] Cohort studies such as MCS, used here and in Libuy et al. (2023), include measurements from early childhood but are collected at irregular intervals, limiting the applicability of a standard dynamic model.

Our study is related to prior 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, measured using aggregate units such as Local Authorities, whereas our aim is to identify the causal effects of school-based exposure. Libuy et al. (2023) exploit within-individual variation to estimate the association between exposure to fast-food outlets and BMI, but do not account for dynamic BMI persistence, which is evident in our data and well-documented in the literature (e.g., Adair, 2008). In contrast, we explicitly model BMI persistence and exploit within-cluster variation in school-level fast-food exposure, focusing on adolescents who have transitioned to secondary school and are therefore exposed to a plausibly exogenous change in fast-food availability near their schools, whereas their analyses pool children from both primary and secondary school

[4] In many studies, schools are the source of weight and height measurements used, e.g., Currie et al. (2010).

age groups.[5] [6]

The remainder of the paper is structured 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 estimation strategy, and Section 4 details the sample construction. Section 5 presents the main empirical results, while Section 6 reports robustness checks and additional analyses based on alternative estimators, model specifications, and sample definitions. 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. Further details of the MCS, including its objectives, sample design, recruitment, response and attrition

[5] Specifically, Libuy et al. (2023) pool the age-7, age-11, and age-14 sweeps, the first two sweeps and the last sweep corresponding to primary and secondary schools, respectively.

[6] Another important difference is that we analyse age- and sex-standardised z-BMI, which are more appropriate than the raw BMI used by Libuy et al. (2023) in their main analyses, given the rapid physical growth experienced during adolescence.

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 (Burgoine et al., 2011; Hobbs et al., 2019; Libuy et al., 2023; Boskovic et al., 2025).[7] Detailed information on its scope, sources and coverage can be found in the PoI product information (Ordnance Survey, 2025).

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.[8] 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, 2019a,b). Overall, ITN provides a granular foundation for modelling the spatial structure of transport networks in Great Britain.

[7] To assess the validity of the PoI data, Wilkins et al. (2017) conduct street audits, and confirms its validity and accuracy of outlet classifications.

[8] 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 meters 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 age- and sex-specific UK reference distributions (Cole and Lobstein, 2012).[9] 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.

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 a single measure of exposure (e.g., Currie et al., 2010, Dunn et al.,

[9] 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 UK or US reference growth charts.

2021, Libuy et al., 2023). This distinction is important because fast food businesses vary in their modes of operation and spatial distribution. Major-brand outlets such as McDonald's, which operate as national chains with strict site requirements and formal planning processes, are more commonly located in high streets, shopping centres, or drive-through sites. In contrast, fish and chip shops and non-major-brand fast-food outlets typically have a strong spatial presence in both residential and commercial areas across the UK. These businesses are often independently owned, smaller in scale, and more easily integrated into local neighbourhoods. We treat fish and chip shops as a distinct category due to their unique cultural significance in the UK, where consuming fish and chips on Fridays remains a popular tradition.[10]

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.[11] Specifically, we measure proximity based on a road-network buffer of 400 meters around each participant's school, which captures proximity of an approximately five-minute walk.[12] This is a buffer size commonly used in retail food environment research (Wilkins et al., 2019). We then construct binary proximity variables indicating the presence of at least

[10]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).

[11]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).

[12]We also test larger buffer sizes of 800 meters and 1,600 meters, representing roughly 10- and 20-minute walking distances, respectively. Additionally, we assess proximity within the same Lower Super Output Area (LSOA) as the cohort member's school or residence. LSOAs (or Data Zones in Scotland) 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). 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 (Overman, 2023).

one outlet of a given type within the 400-meter buffer zone.[13] 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.

3 Accounting for persistence: A new approach

It is well established in the medical and health literature that BMI and the incidence of obesity are highly persistent processes (Adair, 2008; Bartle et al., 2013; Lustig et al., 2022). Childhood BMI, and even *in utero* metrics, are strong predictors of adolescent and adult BMI and the incidence of obesity. This persistence can be the result of individual and environmental factors, both observed and unobserved, that determine lifestyle preferences and behaviours. However, there are also physiological factors that make excessive weight *gain* persistent by making it difficult to maintain weight loss (Ochner et al., 2013; MacLean et al., 2015; Lustig et al., 2022; Hinte et al., 2024). When individuals gain excess weight, both the number and the size of their fat cells increase, but when they lose weight it is only the size of the fat cells that shrinks (MacLean et al., 2015). This makes it easier to gain back weight; a process referred to as adipocyte persistence.[14]

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 (Powell, 2009; 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, which include lags of the dependent variable, would be more appropriate, they require a sufficiently long, equally-spaced, longitudinal data structure. This does

[13] By focusing on the small buffer size, we observe little variation along the intensive margin (i.e., outlet count), which makes a binary proximity measure more appropriate. This specification also avoids imposing a linear relationship between outlet count and z-BMI, a common but potentially restrictive assumption used in the previous literature (e.g., Libuy et al., 2023).

[14] See also Hinte et al. (2024) on obesogenic memory, of which adipocyte persistence is one component.

not align with the irregular and relatively short panel structure of the MCS used here.

Our modeling 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 1–4, and at 11, 14, and 17 years for sweeps 5–7. Consequently, sweeps 1–4 are spaced at roughly two-year intervals, whereas sweeps 5–7 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 and mother/family characteristics, location specific characteristics and time fixed effects.[15]

Persistence in the outcome is ignored in this static model. Standard estimators like Ordinary Least Squares (OLS) will be inconsistent if the persistence term(s) in the error is correlated with household location choices. For example, a secondary school near a city centre and shopping malls—where big fast-food chains cluster—may attract students from neighbourhoods with higher rates of childhood obesity. This may induce a correlation between changes in exposure to fast food over time - in particular, during the transition from primary to secondary school - and early childhood trajectories of z-BMI.

[15]More specifically, we control for the following: (i) time-invariant characteristics: gender of the child, birth mother characteristics- (ethnicity, education, drinking and smoking habits pre-pregnancy and during pregnancy, pre-pregnancy BMI), family social class at birth: (ii) time-varying: cohort-member is an only child, mother is a single parent, mother is employed, housing tenure; (iii) Area characteristics: country of residence, logged number of LSOA population, quintile dummies for Index of Multiple Deprivation (IMD). In addition to these variables, we also include a set of indicators for availability of supermarkets, convenience stores, bakeries and cafes near the school and residence, and time fixed effects. Further details are provided in Appendix Table A1.

Our approach borrows from the recent literature on group-level heterogeneity (Bonhomme and Manresa, 2015). We treat the un-modelled 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:

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

We assume that the unobserved persistence is 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 therefore more flexible than the standard approach of controlling for time-invariant individual-level heterogeneity as it allows for group-specific age profiles (or time trends) in z-BMI.

Next, we assume that there is a finite number of time-invariant groups 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 as we do not condition on past values of the regressors when identifying cluster groupings. It also means that we can estimate the cluster-groupings independently of the main equation in a two-step procedure. By including cluster times time-fixed-effects in the estimating equation, we essentially condition on the full history and group-specific trajectory of z-BMI. This approach also permits us the flexibility we need to estimate the clusters using sweeps of the MCS that have a different time interval.

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 group-level heterogeneity in the model, both within-group (i.e.

cross-sectional) and within-individual (i.e. time) variation in fast-food availability are used to identify β. The within-individual transformation limits the identifying variation to within-individual only.

Under the above maintained assumptions, identification of β hinges on the strict exogeneity of the fast-food variables conditional on the full set of covariates and the cluster times time-fixed-effects. This assumption may be violated for at least two reasons. First, residual location-specific unobservables may remain imperfectly controlled, inducing correlation between the fast-food variables and the error term. Second, the fast-food availability measures may be subject to misclassification error. We discuss these 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 builds on Dunn et al. (2021) and exploits the transition from primary to secondary school at around age 11. 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 to fast-food outlets as controls—recognising that home-based food choice may reflect household-level food preferences—and test robustness by restricting the sample to individuals who did not change residential address.

Turning to the age-17 sweep, the transition between the age-14 and age-17 sweeps coincides with students completing their GCSEs, and entering either academic A-level programmes or a range of technical and vocational pathways.[16] 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 longer-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:

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

Mis-classification 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

[16] In the UK, children attend primary school from ages 5–11 (Key Stages 1–2). They then enter secondary school from ages 11–16 (Key Stages 3–4), culminating in GCSE examinations at age 16 (AboutSchools, 2024; Education, 2025).

the binary indicators used to measure the exposure to fast-food outlet 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 nonzero (Bollinger, 1996, 2010). Under standard assumptions, the OLS estimator, therefore continues to suffer from attenuation bias, unless the misclassification probabilities 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 operate as highly organised businesses, fish and chip shops are typically independent, single- or family-owned businesses. 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. We revisit this in the discussion of our results in Section 5.

A cautionary note on interpretation

Our model tests the relationship between z-BMI and the availability of fast food within a neighbourhood of the school, conditional on group-specific unobservable heterogeneity, persistence in z-BMI, and adolescent, mother and family characteristics. This can be thought of as the reduced-form of the implicit structural equation between z-BMI and fast food consumption.[17]

As such, a null result does not necessarily imply no effect of fast-food consumption on z-BMI. A null result may reflect the absence of a first-stage relationship: namely,

[17] 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.

the absence of the effect of fast-food proximity on fast-food consumption, although the literature supports this association (e.g., Svastisalee et al., 2016).

4 Sample selection and descriptive statistics

Our sample excludes observations with missing values for key variables, including sex, age at interview, country of residence, and BMI. We also drop cohort members who are not singletons or who belong to multiple-birth families.[18] Following these restrictions, further sample selection 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.[19] 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 Table A2. 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

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

[19] 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.

that each contains at least 5% of the total observations.[20] [21] 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%).[22]

Figure 3 shows the average z-BMI values for the seven clusters identified across the age-3 to age-7 sweeps. Because the clusters were derived using z-BMI at ages 3, 5 and 7, it is notable that the averages of later measurements at ages 11–14 largely follow the same developmental patterns established earlier. Children in Cluster 1 maintain average z-BMI values well within the *obese* range across all sweeps, while those in Clusters 2 and 3 consistently fall within the *overweight* range. For Clusters 4–7, the average z-BMI values shift only modestly over time and remain within the *normal-weight* range. As discussed later, these stable patterns help explain why the clusters account for approximately 46% of the variation in z-BMI in the age-11 and age-14 sweeps, and 41% of variation in z-BMI in the age-14 sweep.

4.2 Sample for estimation

In the second step, we define the estimation sample. We select a balanced panel of cohort members observed in both age-11 and age-14 sweeps in order to identify individuals who changed schools between these two waves. The timing of the MCS interviews is particularly well suited for 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 and characteristics after the transition to secondary school. This transition is institutional and largely exogenous, with over 97% of cohort members

[20] We use the Stata `cluster` package to identify the seven clusters based on having at least 5 per cent of observations in each cluster. K-means algorithm allocates each observation to a cluster with the nearest centroid using a distance function (in our case Euclidean distance). In the next step, the centroids in each cluster are calculated again and the process is repeated until the centroids are the same between each iteration.

[21] The minimum threshold of 5% was chosen to ensure an adequate number of individuals per cluster. Our aim is to retain enough variability in observed z-BMI to support reliable estimation of the equation.

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

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 Table A1.

Table 1 in particular presents the proportion of cohort members exposed to fast-food outlets within the 400-metre network buffer from their schools by fast-food category and by cluster in the age-14 sweep, the first sweep conducted after cohort members transitioned to secondary school. As described in Section 2.2, fast-food outlets are categorised into three groups —major-brand, non-major-brand, and fish and chip shops. This classification reflects differences in business models and spatial distribution across outlet types. Major-brand outlets, such as McDonald's, operate as national chains with strict site requirements and formal planning processes, and are more likely to be located on high streets, in shopping centres, or at drive-through sites. In contrast, non-major-brand outlets (e.g., kebab shops) and fish and chip shops tend to be independently owned and smaller in scale, with a broader geographic presence across the UK.

Table 1 is consistent with this pattern. Non-major-brand outlets are the most prevalent category across all clusters, with shares ranging from 20.7% to 29.5%. Fish and chip shops are the second most common outlet type, with availability varying between 11.2% and 15.9% across clusters. In contrast, major-brand fast-food outlets are the least prevalent, with relatively low availability across clusters (3.2%–8.3%). Cluster 4—comprising cohort members who are, on average, in the normal-weight range at ages 3 and 7—accounts for the largest share of the estimation sample (25.7%, or 1,735 individuals), whereas Cluster 1—those in the obese range—represents the smallest share (7.0%, or 404 individuals). Overall, across the full cohort, 22.5% of members have non-major-brand outlets, 12.8% have fish and chip shops, and 5.1% 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.

As discussed in Section 3, our preferred specifications restrict the estimation sample to the age-14 sweep, 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 at the age-14 sweep, 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. Column (3) excludes adolescents in cluster groups 1 and 7 (1,052 individuals), corresponding to those with extremely 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 a 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. Once cluster group dummies are included (column (2)), the estimated effect of major-brand fast-food outlets increases to approximately 0.16 standard deviations and becomes statistically significant at the 1% level. A comparison of columns (1) and (2) underscores the importance of accounting for cluster-level heterogeneity: omitting cluster fixed effects 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 dispropotionately influence the results found for the effects for major-brand outlets. 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 stable and 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 similarity of findings between the full (column (2)) and restricted model (column (3)) indicates that the primary effects are robust and 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 (column (4)), with estimated coefficients of comparable magnitude. 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 fast-food 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 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 peer groups such as adolescents, which may further reinforce patterns of frequent use.

In contrast to major-brand outlets, which operate as highly organised businesses, 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 varies across key socio-demographic groups. Column (1) reproduces the baseline specification from Table 2 column (2), for ease of comparison.

In column (2), we interact fast-food proximity indicators with a dummy variable for being female. The upper panel of Panel A reports coefficients on the main (un-interacted) fast-food proximity variables, representing effects for males, while the lower panel reports the interaction terms capturing any *additional* effect for females. Panel C provides the implied combined effects for each subgroup. 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 and significant, whereas the corresponding effect for females is smaller, 0.101, and not significantly different from zero (Panel C). However, the difference between males and females captured by the interaction term, is itself not statistically significant, presenting insufficient evidence that the effect truly differs by gender. No evidence of gender differences is found for other outlet types, and the joint test of the interaction terms fails to reject the null hypothesis of no gender heterogeneity (Panel B, column (2)).

In column (3), adolescents are classified by maternal ethnicity (White vs. non-White). Panel C shows that proximity to major-brand fast-food outlets has a larger effect on z-BMI among White adolescents 0.184 (SE=0.059) than among non-White adolescents 0.048 (SE=0.118), but the difference between the two groups itself, captured by the interaction term, is not statistically significant. Interestingly, proximity to non-major-brand outlets is significant only for non-White adolescents, with an estimated coefficient of 0.185 (SE=0.74). The difference between the two groups is also significant, suggesting that non-White adolescents may be more responsive to availability of non-major-brand fast-food outlets near schools.

In column (4), adolescents are grouped according to whether their mother has at least a lower secondary school qualification (corresponding to O-level/GCSE in the UK). The result for major-brand outlets indicates no meaningful heterogeneity. However, the effect of proximity to non-major-brand outlets is significant only among adolescents with lower-educated mothers (0.173, SD=0.083), pointing to a socioeconomic gradient in vulnerability to the local food environment. One possible explanation is that adolescents with less-educated mothers may have lower nutritional awareness or greater autonomy in food purchasing, leading to increased exposure to nearby unhealthy food options.

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).[23] The effect of proximity to major-brand fast-food outlets on z-BMI is positive for both groups (0.163 and 0.136, respectively), with no statistical difference in magnitude. Overall, these findings suggest that the effect of major-brand fast-food availability does not differ meaningfully by neighbourhood deprivation, indicating that exposure to such outlets near schools may represent a broadly shared risk factor rather than one concentrated solely among adoles-

[23]Index of Multiple Deprivation (IMD) is based on 37 separate indicators measured across seven distinct domains (income, employment, health & disability, education, skills & training, crime, barriers to housing and services, and living environment). This is a relative measure based at the LSOA level in England and Wales, or Data Zones in Scotland, and aggregated using population weights for each domain. Refer to footnote 12 for the descriptions of LSOAs.

cents in more deprived areas.

Interaction terms in column (6) capture whether adolescents with a prior history of overweight or obesity (cluster groups 1–3, based on z-BMI in early childhood) experience differential effects. Adolescents in this high-risk group may face elevated future weight gain due to adipocyte persistence, as discussed in Section 3. The results provide no evidence of statistically significant heterogeneity by risk status.

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, and prior overweight risk. Heterogeneity is limited to a few specific patterns: non-major-brand outlets appear more relevant for adolescents with lower-educated mothers and non-White adolescents, 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 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 a different source 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 whether home proximity to fast-food outlets independently affects adolescent z-BMI.

In Table 4, we first extend our benchmark specification by estimating the model by pooling the age-11 and age-14 sweeps. Pooling the two sweeps implicitly introduces the assumption that β coefficients in equation (5) are the same across the two sweeps. To allow for time-specific shocks and cluster-specific weight trajectories, we add sweep dummies and their interactions with cluster dummies. The results presented in columns (2) and (3) are from commonly estimated specifications, where the data are pooled and estimated including individual fixed effects.[24] This approach exploits both within-individual time variation and within-cluster (cross-sectional) variation, in contrast to the benchmark specification which relies solely on within-cluster variation.

We first turn to column (2) results. The estimated effect of proximity to a major-brand fast-food outlet is very similar to that obtained from the benchmark model, reported in column (1). However, the null hypothesis that all coefficients are equal across the two sweeps (except the effects of cluster dummies) is rejected $p < 0.01$. Also, the null hypothesis that only the coefficients on the availability of the three fast-food outlet within 400m from school (large-brand, small-brand and fish and chips outlets), is also rejected at $p < 0.01$. 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, between primary- and secondary-school periods.

Moving to column (3) results obtained using the within-group (WG) estimator, the estimated effect of proximity to major-brand fast-food outlets remains statistically significant, although the point estimate is smaller (0.095; SE = 0.039). There are several potential explanations for this reduction in magnitude. Fist, estimates based on within-individual variation may be subject to attenuation bias if the binary fast-food indicators contain non-classical measurement error, as discussed in Section 3. In particular, the WG estimator can amplify noise arising from spurious transitions in mismeasured regressors. Second, there is little temporal variation in the binary variables defined for exposure to fast food, possibly diminishing the estimates. (see Appendix Table A4).

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

[24] For example, Libuy et al. (2023) pool the MCS age-7, age-11 and age-14 sweeps and include individual fixed effects. In our model where the individual fixed effects are included (column (3)), the identification comes from plausibly exogenous within-individual changes in fast-food exposure as they transition from primary to secondary school.

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, additionally 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.[25] The inclusion of these additional controls does not substantively 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. Columns (1)–(3) report estimates based on 400-metre, 800-metre, and 1,600-metre buffers, with the 400-metre measure serving as our benchmark. 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; Kanoski and Boutelle, 2022) and the activation of self-control problems (DellaVigna, 2009).

In column (4), we assess exposure to fast-food outlets located within the same LSOA

[25]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, emotional dysregulation, and cooperation. All assessments were administered at age seven and are available in MCS wave 4. Scores are standardised across tests due to differing scales.

as the cohort member's school.[26] 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. Column (1) reports the baseline specification (Table 2 column (2)), which includes both school exposure and home exposure to fast-food outlets. Home exposure exhibits a markedly different pattern from school exposure. Among home-based measures, only proximity to fish and chip shops is statistically significant, with an estimated coefficient of 0.096 (SE=0.035) — just over half the magnitude of the effect associated with proximity to major-brand outlets 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 fish-and-chip outlets in residential areas—may contribute to the attenuation and lack of robustness of the estimated home-based effects.

[26] LSOAs cover approximately four square kilometres on average in England, though their land area varies substantially with population density, as they are defined based on population size rather than geography (Gregory, 2019). See footnote 12 for more details of LSOAs.

Unlike school exposure, we do not have plausibly exogenous variation in fast-food proximity around the home. To assess whether this limitation induces biases in the 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 not strongly correlated and that the school-based estimates are not driven by omitted residential exposure.

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, 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 BMI, 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 susceptibility from contemporaneous environmental effects. 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 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 compared to those without such exposure (SE=0.054), *ceteris paribus*. As per Table 1, the average BMI in our sample of adolescents is 21.34 with a SD of 4.04. Thus, the estimate of 0.158 translates to an increase from 21.34 to 21.98 BMI units.[27] Although this is modest at the individual level, it is meaningful in terms of shifting the population BMI distribution, especially near clinical cutoffs.[28]

Exposure at larger distances and around the home, in turn, 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 school hours, and with evidence that proximity to fast food increases consumption through, for example, heightened temptation and self-control challenges (Laibson, 2001; DellaVigna, 2009; Kanoski and Boutelle, 2022). 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

[27] This is calculated as 21.34 + 0.158*4.04 = 21.98.

[28] In terms of obesity rates, an increase of 0.158 SD in the z-BMI distribution will increase the area under the curve at the obesity cutoff line of 1.645, to about 6.9% from 5%. If there are about 1 million adolescents, one would expect this to approximately translate to 20,000 additional obese adolescents.

major-brand outlets. Importantly, we find that our results are robust to a wide range of alternative specifications, controlling for individual fixed-effects, alternative spatial measures of fast-food access, and samples that exclude residential movers.

We find little evidence of heterogeneity by gender, ethnicity, maternal education, neighbourhood deprivation, or early-life obesity risk. Notably, the absence of heterogeneity by prior z-BMI trajectory suggests that school-level fast-food exposure influences adolescents broadly, rather than disproportionately reinforcing pre-existing weight disadvantage. 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 non-White adolescents and those whose mothers have lower levels of educational attainment.

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. We further find that omitting early-life weight trajectories attenuates the estimated effect of exposure to major-brand fast-food outlets and substantially reduces the model's explanatory power. This underscores the importance of accounting for persistence in weight trajectories when examining environmental determinants of obesity. That said, although proximity to fast-food outlets is a statistically significant contributor to adolescent z-BMI, it is not a dominant one. Fast-food proximity explains only 0.4% of the variation in adolescent z-BMI, whereas persistence from earlier childhood—captured by cluster-group fixed effects—accounts for 46%. 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

AboutSchools. 2024. "Overview of the stages of the English school system." https://www.aboutschools.uk/overview, Accessed: 3 February 2026.

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. 2025. "The socioeconomic patterning of Great-Britain's retail food environment: A repeated cross-sectional study of area-level deprivation and food outlet density from 2011-2024." *medRxiv*.

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.

Burgoine, Thomas, Seraphim Alvanides, and Amelia A. Lake. 2011. "Assessing the obesogenic environment of North East England." *Health & Place* 17 (3): 738–747.

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, and Tim Lobstein. 2012. "Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity." *Pediatric obesity* 7 (4): 284–294.

Cole, Timothy J, Jennifer V Freeman, and Michael A Preece. 1995. "Body mass index reference curves for the UK, 1990.." *Archives of disease in childhood* 73 (1): 25–29.

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.

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.

Education, Ivy. 2025. "Comprehensive Guide to the UK Education System in 2025." https://www.ivyeducation.co.uk/insights/uk-education-system-explained, Accessed: 3 February 2026.

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.

Gregory, Kimberley. 2019. "LSOAs, LEPs and lookups: A beginner's guide to statistical geographies." https://ocsi.uk/2019/03/18/lsoas-leps-and-lookups-a-beginners-guide-to-statistical-geographies/, Accessed: 2025-09-09.

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.

Kanoski, Scott E, and Kerri N Boutelle. 2022. "Food cue reactivity: Neurobiological and behavioral underpinnings." *Reviews in Endocrine and Metabolic Disorders* 23 (4): 683–696.

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.

Lustig, Robert H, David Collier, Christopher Kassotis et al. 2022. "Obesity I: Overview and molecular and biochemical mechanisms." *Biochemical pharmacology* 199 115012.

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.

Ochner, Christopher N, Dulce M Barrios, Clement D Lee, and F Xavier Pi-Sunyer. 2013. "Biological mechanisms that promote weight regain following weight loss in obese humans." *Physiology & behavior* 120 106–113.

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. 2019a. "Harrow Council Optimises Transport for Vulnerable Passengers Using OS MasterMap." https://www.ordnancesurvey.co.uk/customers/case-studies/harrow-transport, Accessed: 2025-06-24.

Ordnance Survey. 2019b. "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.

Overman, Henry. 2023. "Using small scale output data to inform local economic policy." https://whatworksgrowth.org/insights/using-small-scale-output-data-to-inform-local-economic-policy/, Accessed: 2025-08-18.

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, United Nations Children's Fund. 2025. "Obesity exceeds underweight for the first time among school-age children and adolescents globally." https://www.unicef.org.uk/press-releases/obesity-exceeds-underweight-for-the-first-time-among-school-age-children-and-adolescents-globally-unicef/, 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

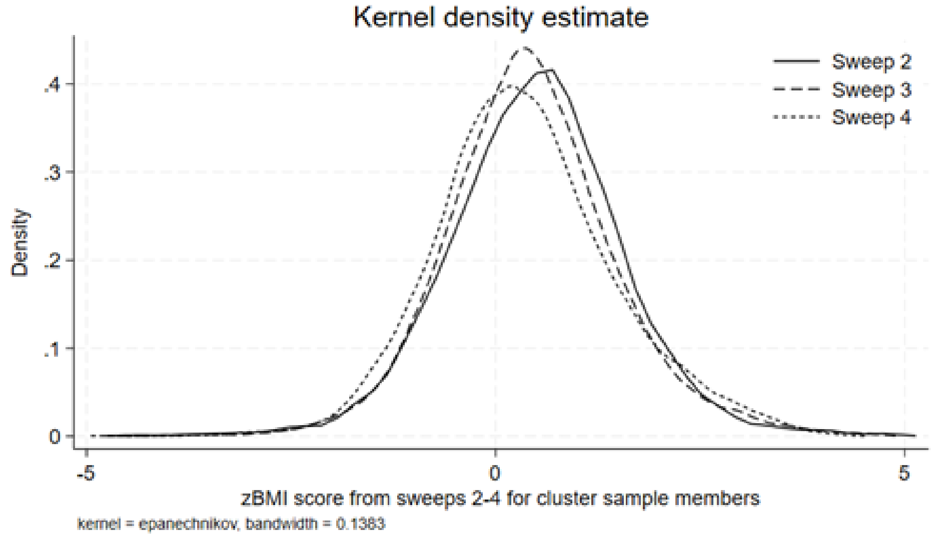

Figure 1: Mean z-BMI in childhood

Note: There are 10,011 Cohort Members each providing a balanced panel over sweeps 2-4 (ages 3, 5 and 7). The mean and standard deviation of z-BMI values for sweeps 2-4 are 0.492 (SD=1.06); 0.432 (SD=1.05); and 0.365 (SD=1.10), respectively.

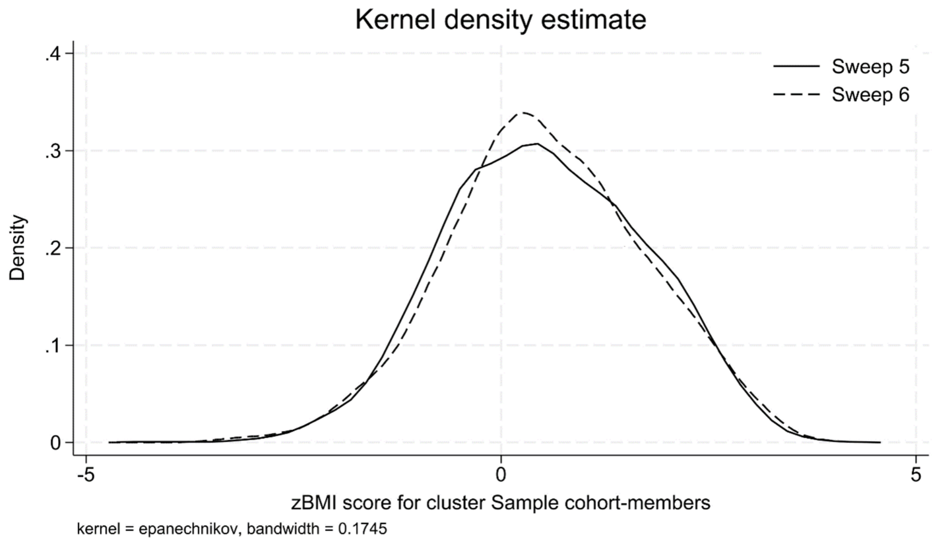

Figure 2: Mean z-BMI in adolescence

Note: The sample used here consists of 6,626 Cohort members (CM) present in the estimation sample. The CMs transitioned to secondary school in between sweeps 5 and 6 (ages 11 and 14).

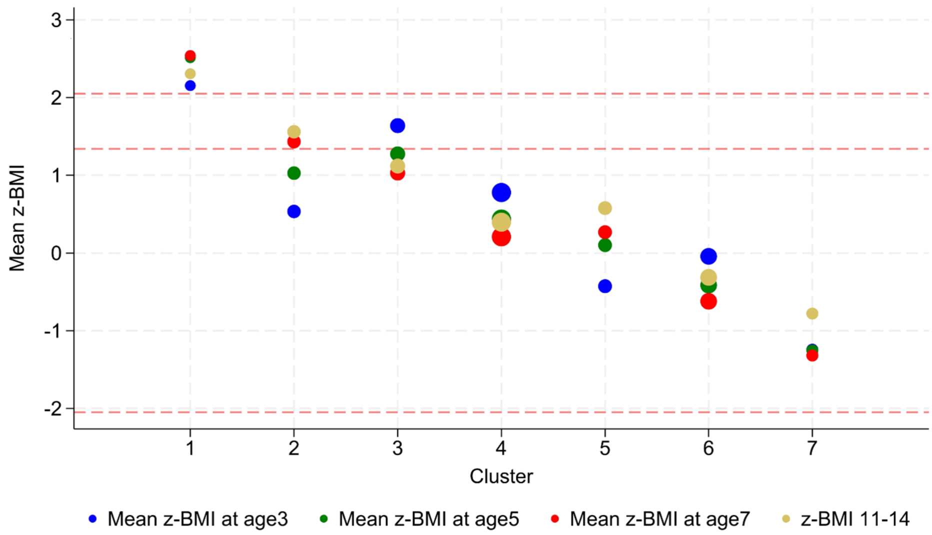

Figure 3: Mean z-BMI by cluster groups

Note: (a) Sample consists of 9,727 cohort members who had a valid cluster group identifier. (b) Number of cohort members (%) in each cluster: 1. 676 (7.0); 2. 1138 (11.7); 3. 1447 (14.9); 4. 2500 (25.7); 5. 1233 (12.7); 6. 1860 (19.1); 7. 873 (9.0).

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

	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)

Note: Binary indicators for availability of an outlet is recorded as 1 if at least one outlet is available within 400m from school, and 0 otherwise. Mean, standard deviation (for continuous variables), and number of observations are reported by cluster group. The sample used for calculations consists of a balanced panel of 6,626 cohort-members in the age-14 sweep.

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	No	Yes	Yes	Yes
Fast food near home	Yes	Yes	Yes	Yes
Covariates	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) All equations are estimated using ordinary least squares, using the age-14 sweep. (iii) Control variables include sex, single-child, single-parent, housing tenure, mother's employment status, mother's education at birth, mother's smoking and drinking status, ethnicity, household social class, logged population of the CM's LSOA, Index of Multiple Deprivation (quintiles), other food outlets (e.g. cafes) within 400m of school and home, and cluster fixed effects. (iv) Column (3) excludes cluster groups 1 and 7, corresponding to those with extremely high or low childhood z-BMI. (v) Column (4) uses cohort members with valid age-17 sweep zBMI as the dependent variable; all other controls are from the age-14 sweep. (vi) Robust standard errors are in parentheses. (vii) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

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

	(1)	(2)	(3)	(4)	(5)	(6)
A. Base group	Baseline	Male	White	Mother high education	Low deprivation	Low obesity risk
Availability within 400 meters from school						
Major-brand fast food	0.158***	0.210***	0.184***	0.169***	0.136**	0.166**
	(0.054)	(0.077)	(0.059)	(0.057)	(0.067)	(0.066)
Non-major-brand fast food	0.039	0.044	0.003	0.006	0.018	0.033
	(0.036)	(0.049)	(0.038)	(0.038)	(0.042)	(0.042)
Fish & Chips	0.030	-0.008	0.042	0.042	0.081	0.054
	(0.040)	(0.058)	(0.046)	(0.043)	(0.051)	(0.048)
Interactions		Female	Non-White	Low education	High deprivation	High risk
Major-brand fast food		-0.110	-0.136	-0.034	0.027	-0.029
		(0.106)	(0.131)	(0.153)	(0.108)	(0.112)
Non-major-brand fast food		-0.010	0.181**	0.173**	0.034	0.017
		(0.060)	(0.076)	(0.083)	(0.059)	(0.062)
Fish & Chips		0.078	-0.091	-0.070	-0.092	-0.079
		(0.076)	(0.094)	(0.101)	(0.076)	(0.079)
Cluster fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Fast food near home	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
B. Joint significance of interactions						
F(3,6625)		0.66	1.98	1.48	0.50	0.41
[p-value]		[0.579]	[0.115]	[0.218]	[0.685]	[0.747]
C. Aggregated coeff's for non-base groups		Females	Non-White	Low education	High deprivation	High risk
Major-brand fast food		0.101	0.048	0.134	0.163*	0.137
		(0.075)	(0.118)	(0.143)	(0.086)	(0.092)
Non-major-brand fast food		0.034	0.185***	0.180***	0.053	0.051
		(0.044)	(0.074)	(0.079)	(0.051)	(0.053)
Fish & Chips		0.070	-0.049	-0.029	-0.012	-0.025
		(0.052)	(0.083)	(0.092)	(0.059)	(0.064)

Notes: (i) See Table 2 notes. (ii) The variables included in the interactions are the ones reported on the table. (iii) Low maternal education is defined as mothers without any qualifications or 'other' qualifications than a degree, A-level and O-level/GCSE. (iv) Most deprived category refers to the Index of Multiple Deprivation quintiles being 1 or 2. (v) 'Hirisk' is defined as those CMs in the top 3 cluster-groups who are on or above the *over-weight* (1+SD) cut-off; (vi) Total number of observations is 6,626.

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 400 meters of 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)
Fast food near home	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Cluster fixed effects	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	No	No	Yes	No	No
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 Table 2 notes. (ii) Columns (2)and (3) use the sample consisting of age-11 and age-14 sweeps and hence include cluster dummies interacted with sweep dummies; (iii) Column (3) results are from the within-group estimator; (iv) Columns (4) and (5) include cognitive and non-cognitive test scores, respectively. See footnote 25 for details.

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)
Fast food near home	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Cluster fixed effects	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 Table 2 notes. (ii) The models estimated are our baseline model but using different measures of proximity to fast food outlets – 400m-buffer, 800m-buffer and 1600m-buffer zones, and within the same Lower Super Output Area (LSOA). See footnote 12 for details.

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 400 meters 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)
Covariates	Yes	Yes	Yes
Cluster fixed effects	Yes	Yes	Yes
Observations	6,626	6,626	5,597
R-squared	0.455	0.454	0.455

Note: (i) See Table 2 notes. (ii) Column (3) draws on the subset of individuals (5,597) from our base sample who remained at the same residence between the age-11 and age-14 sweeps.

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.[29]

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.[30] 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 zBMI:

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).

[29] 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).

[30] Refer to the PoI product information for further details of the classification scheme (?).

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.

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 Group

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

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 (413)	1551 (417)	1561 (438)	1567 (444)	1616 (449)	1567 (429)	1638 (417)	1580 (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 cohort members in the base estimation sample—specifically, those who changed schools between the age-11 and age-14 sweeps, transitioning from primary school to secondary school.

Table A2: Persistence in observed zBMI 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 A3: 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 A4: Changes in fast-food outlet availability within 400m of schools between primary (age-11 sweep) and secondary school (age-14 sweep), by cluster group

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. 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.