eAppendix 1

Food resource densities

We calculated restaurant and food store densities as counts per 10 km secondary and local roadway using StreetMap 2000 (v.9.0) for years 7 (1993) and 10 (1996), StreetMap Pro 2005 (v.5.2) for year 15 (2001) and StreetMap Pro 2010 (v.7.2) for year 20. We used roadway lengths aggregated within either 8km or 3km Euclidean buffers to be consistent with the restaurant (3km) and food store (8km) counts. The resulting food resource densities reflect concentrations of fast food restaurants, sit-down restaurants supermarkets, and convenience stores along streets representing overall commercial activity (Hawkins et al., 2009; Author et al., 2011).

SEM Methods

Latent Variables used in structural equation modeling

Food environment. We created latent factors for each neighborhood food store and restaurant factors type (fast food restaurant, sit-down restaurant, supermarket and convenience stores) at each year using observed indicators: count per 10 km local and secondary roadway, within 8 km (food stores) or 3 km (restaurants) Euclidean buffer and the Z-score of population density.

Diet behaviors. We created four latent diet factors for each year (fast food restaurant-type diet; sit-down restaurant-type diet, supermarket-type diet, and convenience store-type diet) using intake categories of foods we considered, *a priori*, to be acquired at each type of establishment (e.g., fries from fast food restaurants, fruits from supermarkets). We hypothesized that food and beverage groups reflected the types of foods and beverages commonly offered at each specific type of store or restaurant (Cannuscio et al., 2013; Gustafson et al., 2012; Hutchinson et al., 2012; Rundle et al., 2009) and that the restaurants and food stores would be associated with the consumption of these food, as shown in Figure 1. Our approach differs from standard approaches focusing on classifying establishments on the basis of selling "healthy" (Morland et al., 2002a) or "unhealthy" (Smith et al., 2013) foods, given that identical foods can be acquired from a range of stores and restaurants.

eAppendix 2

Confounding

Individual-level confounders

We characterized individual-level sociodemographic and behavioral confounders using data from questionnaires collected at each exam year. Time-invariant sociodemographic variables were sex, race (white/black), baseline age, exam attendance, and center. Time-varying characteristics were maximum reported number of years of schooling completed by the exam year (continuous), and mean household income inflated to U.S. dollars at year 20 (2005-06) using the Consumer Price Index. Income was not collected in year 0, so we used the closest measurement (year 5) for year 0. At each exam, participants completed an interviewer-administered physical activity history designed for CARDIA (Jacobs et al., 1989). Participants reported on 13 different categories of moderate and vigorous recreational sports, exercise, leisure, and occupational activities in the past 12 months and scores were calculated based on frequency and intensity of each activity

Methods to adjust for confounding

As shown in Figure 3, we addressed confounding of associations between: (1) food environment and diet; (2) diet and BMI; and (3) food environment and BMI after excluding diet-BMI confounders that were likely affected by the food environment (Cameron et al., 2013; Levinson, 2012; Valeri and Vanderweele, 2013; Zenk et al., 2006). We controlled for time-varying education and income (food environment-diet); age at baseline, race, sex and time-varying education, physical activity, and income (diet and BMI); time-varying education, and income, center, and the longitudinal neighborhood SES class (food environment-BMI). Since access to diverse destinations may promote physical activity and physical activity may contribute to better weight regulation (Cerin et al., 2011), we controlled for physical activity along the food environment (exposure)- BMI (outcome) pathway. We also modeled associations between covariates to account for dependencies between covariates. For example, individual sociodemographics play a role in physical activity (e.g., young adults are on average more physically active than older adults). Thus, we modeled time-varying physical activity as a function of baseline age, race, sex and current education and income. In addition, we modeled longitudinal neighborhood SES as a function of race, sex, baseline age, education, and income.