## SUPPLEMENTARY MATERIAL

## Supplemental Sample

The annual purchases of fruits and vegetables by a household are high frequency items compared to with low frequency products such as clothes and furniture. However, in our raw study sample, we still observed that, for example, 1,078 and 763 households spent less than $10 for fruits and vegetables, respectively, in 2010. In addition, 14 and 11 households spent more than $ 1,000 for fruits and vegetables, respectively. To avoid undue influence by some outlying points on the regression, we excluded households with extremely large or small expenditures for both fruits and vegetables together. However, households with extremely large or small expenditures for fruits but not for vegetables, or vice versa, were still included. We defined extreme values of expenditure for fruits or vegetables as those below the 2nd percentile or above the 98th percentiles. We did not use “below 1st percentile” and “above the 99th percentile” as cut off points because there would still have been some households with zero purchase of fruits or vegetables with a 1st percentile cut-off point; thus, the narrower cut-offs did not facilitate the log transformation of expenditures.

We used only the data of the magnet households that reported nonstandard UPC products, which included random-weighted (loose) items such as fruits, vegetables, meats, and in-store baked goods; households that did not report nonstandard UPC products were excluded from our models. Thus, we needed to ensure enough overlap (e.g., balance in covariates) between magnet and non-magnet households. If there was enough overlap, the estimated densities of the probability of being magnet household versus non-magnet household would not have too much mass around 0 or around 1([1](#_ENREF_1)). We generated the probabilities of being magnet (i.e., propensity scores) for the two subgroups and plotted the probabilities on the same graph. There were enough overlaps between magnet households and non-magnet households (Supplemental Fig. S1), suggesting that all covariates were largely balanced. Thus, we did not necessarily need to exclude some of the magnet households in the expenditure for fruits or vegetables models because of significantly different household-level sociodemographic characteristics. We used kdens command in STATA 14 to draw the plots.



Supplemental Fig. S1. Kernel density estimate: probability of being magnet households

## Supplemental Measures

### Purchase of fresh fruits and vegetables

The purchases of fresh fruits and vegetables using standard UPC codes were categorized as fresh produce by Nielsen (Supplemental Table S1). By choosing products from the fresh produce category, we explicitly excluded fruits and vegetables that were dried, tinned, bottled, frozen, or refrigerated; these latter products are classified as dry grocery, frozen food, and deli.

The purchases of fresh fruits and vegetables using nonstandard UPC codes were categorized as magnet by Nielsen (Supplemental Table S1). From the products of magnet data, we used brand\_desr= reference card fruits and brand\_desc= reference card vegetables to identify fruits and vegetables, respectively. Other nonstandard UPC (i.e., magnet) products included baked goods, prepared foods, cheese, meat/poultry/fish, coffee, flora, etc., according to the description of products (variable name= brand\_desr) in the product file. Items in the categories “reference card fruits” and “reference card vegetables” were not described in detail, i.e., whether the fruits or vegetables were fresh, dried, tinned, frozen, refrigerated, etc. As magnet data were known for random weighted (loose) products, we assumed that we could ignore the proportion of random-weighted magnet dried, frozen fruits or vegetables. Therefore, we specified all the magnet products with brand\_descr=reference card fruits and brand\_descr=reference card vegetables as fresh fruits and fresh vegetables, respectively.

Supplemental Table S1. Fresh produce type and Nielsen production group module description

|  |  |  |
| --- | --- | --- |
| Food type | Department a | Product module/brand a |
| Fruits using standard UPC codes | Fresh produce | fresh apples, fresh cranberries, fresh grapefruits, fresh kiwi, fresh oranges, fresh strawberries, fresh tomatoes, and fresh fruits-remaining.  |
| Fruits using nonstandard UPC codes | Magnet | Magnet data with brand\_descr=reference card fruits |
| Vegetables using standard UPC codes | Fresh produce | fresh carrots, fresh cauliflower, fresh celery, fresh lettuce, fresh garlic, fresh mushrooms, fresh onions, fresh potatoes, fresh radishes, fresh spinach, fresh sprout, and fresh vegetables-remaining.  |
| Vegetables using nonstandard UPC codes | Magnet | Magnet data with brand\_descr=reference card vegetables |

Note: a Department, product module, and brand are all Nielsen defined product codes.

In calculating the annual expenditures for fruits or vegetables by household, we linked the product file to the purchase file using the UPC numbers (variable name 1=upc, variable name 2=upc\_ver\_uc) as the joint identifying numbers to create a purchase-product file. Upc\_ver\_uc indicated different versions of upc. We then linked the purchase-product file to the trip file using the trip number (variable name=trip\_code\_uc) as the joint identifying number to create a trip-purchase-product file. We linked the trip-purchase-product file to the household sociodemographic file using the household number (variable name=household\_code) as the joint identifying number to create a household-trip-purchase-product file.

To properly classify self-reported, non-magnet expenditures, we used the departmental category (fresh produce) and the product module description (e.g., fresh apple, fresh lettuce) to identify fruits or vegetables; for the self-reported magnet expenditures, we used the departmental category (magnet) and brand description (e.g., reference card fruits, reference card vegetables) to identify fruits or vegetables. We then calculated the self-reported expenditures for fruits and vegetables (separately) as the sum of standard UPC products (non-magnet) and nonstandard UPC products (magnet) by household for 2010; then, we used these estimates to partially address the potential issue of random purchasing behaviors (i.e., impulsive purchases) in a short observational period (e.g., weekly or monthly)([2](#_ENREF_2)).

### SIC codes from ReferenceUSA

See Supplemental Table S2 below for the primary six-digit standard industrial classification (SIC) codes used to classify supermarkets, convenience stores, fast food restaurants, sit-down restaurants, child care services, other food and non-food stores, health-care services, and churches.

Supplemental Table S2. Primary 6-digit a SIC codes from ReferenceUSA used in the analysis for year 2010

|  |  |  |
| --- | --- | --- |
| Food Resource Type | Description | ReferenceUSA primary SIC code |
| Supermarkets | Supermarkets | 541101  |
| Convenience stores | Variety storesSnack products Convenience storesGasoline service stationsGas stations and convenience stores  | 533100541102 541103554100554199 |
| Fast food restaurants | Fast food restaurants and standsPizza restaurants | 581203581206 |
| Sit-down Restaurants | Fine dining restaurantsFamily restaurantsSeafood restaurantsSteak and barbecue restaurants | 581201581205581207581208 |
| Child care services |  | 835101 |
| Churches |  | 866107 |
| Department stores |  | 531102 |
| Retail shops |  | 531104 |
| Wholesale clubs |  | 531110 |
| Miscellaneous general merchandise stores |  | 539900 |
| Offices and clinics of physicians |  | 801101 |
| Offices and clinics of dentists |  | 802101802104 |
| Offices and clinics of doctors of osteopathy |  | 803198 |
| Offices and clinics of other health practitioners |  | 804101804201804301804302804303804901-32, 35-42, 44-47, 50-67, 69, 71-77, 79-81, 83, 84-89, 91-92, 94, 97-99 |
| Health care facilities |  | 805101-02, 805198, 805298, 805901 |
| Hospitals |  | 806904806906806998 |

Note: a ReferenceUSA has created their own 2-digit extension to the original SIC system as a means to update and expend the system so their customers can more precisely define their business classification.

### Regional destination accessibility

The measure of regional destination accessibility is based on a network analysis model that considers the attractiveness (number of employees) of each reachable block group and the travel time between each origin block group and all the destination block groups. The SLD used the employment information in the InfoUSA 2011 and the street network information in the NAVSTREETS to generate measures of job opportunities in each reachable block group and the travel time between each origin block group and all the destination block groups using network analysis models. The SLD then generated the value of regional destination accessibility by decaying the employment at destinations by the distance decay curve and summed for each origin block group([3](#_ENREF_3)). Compared to the traditional measure of the total attractiveness of reachable block groups, such as summing the total number of potential destinations in a certain area, by decaying the attractiveness of destinations using distance decay curve, regional destination accessibility calculated by the SLD is a more accurate measure of total attractiveness.

### Neighborhood destination diversity

We used American Time Use Survey (ATUS) to identify grocery-chained activities. In 2010, the ATUS had 257,193 observations of activities by 13,260 participants. For each of these participants, the ATUS recorded how many minutes a participant spent time on a specific activity during one observational day (00:00-23:59, 24 hours), in what type of activity the participant attended, and where the participant was when he or she attended the activity. To identify which activities were “chained” with grocery shopping, we retained 2,011 participants who reported at least one grocery shopping in grocery stores. Participants who socialized and communicated in grocery stores or who reported grocery shopping but not in grocery stores (for example, via internet or by telephone) were removed from our sample. We used a 0.5-hour time duration window to identify the activities that occurred before (the start time of) grocery shopping in grocery stores. We also used a 0.5-hour time duration window to identify the activities that occurred after (the end time of) grocery shopping in grocery stores. We defined all activities in such a 1-hour time duration window as grocery-chained activity. We then identified and ranked the locations at which such grocery-chained activities occurred. The locations with high occurrence rates (i.e., the number of grocery-chained activities at one specific location divided by the total number of grocery-chained activities was greater than 10%) were: 1) other store/mall (25.4%); 2) other place (19.2%); 3) someone else’s home (14.7%); and 4) restaurants (13.3%), excluding travel activities such as car, truck, motorcycle, walking, bus, etc. According to the activity description linked with such locations, we defined other store/mall as department store, retail shop, and wholesale club. Similarly, we defined locations where people socialized and communicated with others as churches, health care services, and child care services. Similarly, we defined restaurants as fast food and sit-down restaurants. See Supplemental Table S2 above for the primary six-digit standard industrial classification (SIC) codes used to classify fast food restaurants, sit-down restaurants, child care services, churches, department stores, retail shops, wholesale clubs, miscellaneous general merchandise stores, and six types of health care services. We used these three types of locations (other store, other places, and restaurants) as the potential destinations that were chained with grocery shopping. We calculated the total number of outlets under each type and used the entropy function to calculate the neighborhood destination diversity.

### Neighborhood street connectivity

We retrieved the neighborhood street connectivity calculated by the SLD. The SLD used the road network information in the NAVSTREETS to measure neighborhood street connectivity as the total weighted number of street intersections divided by total land area in the block group. The SLD formula to calculate the weighted street connectivity in the block group is as follows:

$$street connectivity=d1\*0.667+d2+d3\*0.667+d4$$

where d1 is the number of multi-modal intersections having three legs per square mile, d2 is the number of multi-modal intersections having four or more legs per square mile, d3 is the number of pedestrian-oriented intersections having three legs per square mile, and d4 is the number of pedestrian-oriented intersections having four or more legs per square mile. To reflect the connectivity for pedestrian and bicycle travel, the SLD assigned a weight of zero to auto-oriented intersections to reflect the fact that they are a barrier to pedestrian and bicycle mobility([3](#_ENREF_3)). Similarly, the SLD assigned lower weights to three-way intersections to reflect the fact that they do not promote street connectivity as effectively as four way intersections([3](#_ENREF_3)).

### Household-level covariates

Households reported the highest degree obtained by the female head of household in one of seven categories. We combined the seven to generate a new education variable with three categories, i.e., high school or below, college or higher, and no female head. We chose the female head of household rather than the male because maternal education has been shown previously to be an important determinant of child diet([4-7](#_ENREF_4)). Households reported their income in one of 20 categories, which we combined into a three-category variable, i.e., less than $20,000, $20,000-$59,999, and $60,000 or more. Households reported their race identity in one of four categories, including white, black, Asian and other. Households reported their sizes in one of nine categories; we combined these nine categories into a four-category variable, i.e., one member, two members, three members, and four members or above. Households reported the marital status of the head as four categories, married, widowed, divorced/separated, and single.

## Supplemental Data Analyses and Results

In the sensitivity analysis, we excluded households located in 1,831 largest ZCTAs (with a land area equal to or greater than 153.5 km2). We used the centroid of residential ZCTA as a household’s residence and ran fruit and vegetable models. We used availability of neighborhood supermarkets and convenience stores, neighborhood destination diversity, availability of neighborhood destinations, and neighborhood street connectivity in the 5-km buffer. The model results are shown in Supplemental Tables S3 and S4.

In the sensitivity analysis, we used the centroid of residential ZCTA as a household’s residence and ran fruit and vegetable models. We used availability of neighborhood supermarkets and convenience stores, neighborhood destination diversity, availability of neighborhood destinations, and neighborhood street connectivity in the 3-km buffer. The model results are shown in Supplemental Tables S5 and S6.

In the sensitivity analysis, we defined an extreme value for expenditure on fruits or vegetables as those below the 3rd percentile or above the 97th percentile. We used the centroid of residential ZCTA as household’s residence and ran fruit and vegetable models (n=21,269 for fruit and n=21,249 for vegetable). We used availability of neighborhood supermarkets and convenience stores, neighborhood destination diversity, availability of neighborhood destinations, and neighborhood street connectivity in the 5-km buffer. The model results are shown in Supplemental Tables S7 and S8.

Supplemental Table S3. Coefficients of cross-sectional associations between annual expenditures (logarithmic-transformed) for fruits purchased by Nielsen households (excluding households located in 1,831 ZCTAs with a land area equal to or greater than 153.5 km2, n=17,420), availability of neighborhood supermarkets and convenience stores (5-km buffer, centroid of residential ZCTA as a household’s residence), broader built environment context characteristics, and household-, neighborhood- and area-level covariates

|  |  |  |
| --- | --- | --- |
| Characteristics | Availability of neighborhood supermarkets and convenience stores only a | Availability of neighborhood supermarkets and convenience stores and broader built environment context a |
| Coefficient (SE) | p-value | Coefficient (SE) | p-value |
| Availability of supermarkets, count, 5-km buffer |  |  |  |  |
|  0 (Ref.) | **---** | **---** | **---** | **---** |
|  1 | 0.011 (0.045) | 0.482 | -0.000 (0.017) | 0.989 |
|  2+ | 0.030 (0.016) | 0.104 | 0.003 (0.020) | 0.866 |
| Availability of convenience stores, 10 counts, 5-km buffer | 0.025 (0.018) | 0.693 | -0.023 (0.007) | **0.002** |
| Broader built environment context |  |  |  |  |
|  Regional destination accessibility: Jobs within 45-min  automobile travel time, 10,000 jobs  |  |  | 0.003 (0.001) | **0.001** |
|  Neighborhood destination diversity: Entropy, 10 percent, 5-km  buffer  |  |  | 0.002 (0.004) | 0.570 |
|  Availability of neighborhood destinations: Total other stores/malls, locations where people socialized and communicated with others, and restaurants, 10 counts, 5-km buffer |  |  | 0.000 (0.000) | 0.106 |
|  Neighborhood street connectivity: 10 intersections per square  mile, 5-km buffer |  |  | 0.014 (0.004) | **<0.001** |

n, number of observations; SE, standard error; Ref., reference category; ZCTA, zip code tabulation area.

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

a Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one child in the family, and number of employed household members (household head excluded), expenditure (logarithmic-transformed) on vegetables, and urbanicity (R x64 3.5.1 and Rstudio 1.1456).

Entries in bold mark statistically significant associations (p<0.05).

Supplemental Table S4. Coefficients of cross-sectional associations between annual expenditures (logarithmic-transformed) for vegetables purchased by Nielsen households (excluding households located in 1,831 ZCTAs with a land area equal to or greater than 153.5 km2, n=17,394), availability of neighborhood supermarkets and convenience stores (5-km buffer, centroid of residential ZCTA as household’s residence), broader built environment context characteristics, and household-, neighborhood- and area-level covariates

|  |  |  |
| --- | --- | --- |
| Characteristics | Availability of neighborhood supermarkets and convenience stores only a | Availability of neighborhood supermarkets and convenience stores and broader built environment context a |
| Coefficient (SE) | p-value | Coefficient (SE) | p-value |
| Availability of supermarkets, count, 5-km buffer |  |  |  |  |
|  0 (Ref.) | **---** | **---** | **---** | **---** |
|  1 | 0.013 (0.039) | 0.373 | 0.017 (0.015) | 0.265 |
|  2+ | 0.015 (0.015) | 0.376 | 0.021 (0.018) | 0.265 |
| Availability of convenience stores, 10 counts, 5-km buffer | -0.005 (0.017) | 0.410 | -0.002 (0.007) | 0.826 |
| Broader built environment context |  |  |  |  |
|  Regional destination accessibility: Jobs within 45-min  automobile travel time , 10,000 jobs  |  |  | -0.001 (0.001) | 0.589 |
|  Neighborhood destination diversity: Entropy, 10 percent, 5-km  buffer  |  |  | -0.010 (0.004) | **0.010** |
|  Availability of neighborhood destinations: Total other stores/malls, locations where people socialized and communicated with others, and restaurants, 10 counts, 5-km buffer |  |  | 0.000 (0.000) | 0.610 |
|  Neighborhood street connectivity: 10 intersections per square  mile, 5-km buffer |  |  | -0.002 (0.003) | 0.608 |

n, number of observations; SE, standard error; Ref., reference category; ZCTA, zip code tabulation area.

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

a Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one child in the family, and number of employed household members (household head excluded), expenditure (logarithmic-transformed) on fruits, and urbanicity (R x64 3.5.1 and Rstudio 1.1456).

Entries in bold mark statistically significant associations (p<0.05).

Supplemental Table S5. Coefficients of cross-sectional associations between annual expenditures (logarithmic-transformed) for fruits purchased by Nielsen households (between 2nd and 98th percentile of the fruit expenditures, n=21,710 a), availability of neighborhood supermarkets and convenience stores (3-km buffer, centroid of residential ZCTA as a household’s residence), broader built environment context characteristics, and household-, neighborhood- and area-level covariates

|  |  |  |
| --- | --- | --- |
| Characteristics | Availability of neighborhood supermarkets and convenience stores only b | Availability of neighborhood supermarkets and convenience stores and broader built environment context b |
| Coefficient (SE) | p-value | Coefficient (SE) | p-value |
| Availability of supermarkets, count, 3-km buffer |  |  |  |  |
|  0 (Ref.) | **---** | **---** | **---** | **---** |
|  1 | 0.036 (0.017) | **0.033** | 0.018 (0.017) | 0.295 |
|  2+ | 0.004 (0.023) | 0.845 | -0.024 (0.023) | 0.309 |
| Availability of convenience stores, 10 counts, 3-km buffer | 0.001 (0.012) | 0.943 | -0.044 (0.014) | **0.002** |
| Broader built environment context |  |  |  |  |
|  Regional destination accessibility: Jobs within 45-min  automobile travel time, 10,000 jobs  |  |  | 0.004 (0.001) | **<0.001** |
|  Neighborhood destination diversity: Entropy, 10 percent, 3-km  buffer  |  |  | 0.001 (0.003) | 0.237 |
|  Availability of neighborhood destinations: Total other stores/malls, locations where people socialized and communicated with others, and restaurants, 10 counts, 3-km buffer |  |  | 0.000 (0.000) | 0.261 |
|  Neighborhood street connectivity: 10 intersections per square  mile, 3-km buffer |  |  | 0.009 (0.003) | **0.001** |

n, number of observations; SE, standard error; Ref., reference category; ZCTA, zip code tabulation area.

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

a We excluded who reported extremely low or high values for purchases of fruits, defined here as less than the 2nd percentile or greater than the 98th percentile.

b Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one child in the family, and number of employed household members (household head excluded), expenditure (logarithmic-transformed) on vegetables, and urbanicity (R x64 3.5.1 and Rstudio 1.1456).

Entries in bold mark statistically significant associations (p<0.05).

Supplemental Table S6. Coefficients of cross-sectional associations between annual expenditures (logarithmic-transformed) for vegetables purchased by Nielsen households (between 2nd and 98th percentile of the vegetable expenditures, n=21,686 a), availability of neighborhood supermarkets and convenience stores (3-km buffer, centroid of residential ZCTA as household’s residence), broader built environment context characteristics, and household-, neighborhood- and area-level covariates

|  |  |  |
| --- | --- | --- |
| Characteristics | Availability of neighborhood supermarkets and convenience stores only b | Availability of neighborhood supermarkets and convenience stores and broader built environment context b |
| Coefficient (SE) | p-value | Coefficient (SE) | p-value |
| Availability of supermarkets, count, 3-km buffer |  |  |  |  |
|  0 (Ref.) | **---** | **---** | **---** | **---** |
|  1 | 0.007 (0.015) | 0.645 | 0.008 (0.016) | 0.610 |
|  2+ | 0.013 (0.021) | 0.526 | 0.013 (0.022) | 0.544 |
| Availability of convenience stores, 10 counts, 3-km buffer | 0.003 (0.011) | 0.804 | 0.002 (0.013) | 0.907 |
| Broader built environment context |  |  |  |  |
|  Regional destination accessibility: Jobs within 45-min  automobile travel time , 10,000 jobs  |  |  | -0.001 (0.001) | 0.274 |
|  Neighborhood destination diversity: Entropy, 10 percent, 3-km  buffer  |  |  | -0.005 (0.002) | **0.035** |
|  Availability of neighborhood destinations: Total other stores/malls, locations where people socialized and communicated with others, and restaurants, 10 counts, 3-km buffer |  |  | 0.000 (0.000) | 0.325 |
|  Neighborhood street connectivity: 10 intersections per square  mile, 3-km buffer |  |  | 0.003 (0.002) | 0.263 |

n, number of observations; SE, standard error; Ref., reference category; ZCTA, zip code tabulation area.

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

a We excluded who reported extremely low or high values for purchases of vegetables, defined here as less than the 2nd percentile or greater than the 98th percentile.

b Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one child in the family, and number of employed household members (household head excluded), expenditure (logarithmic-transformed) on fruits, and urbanicity (R x64 3.5.1 and Rstudio 1.1456).

Entries in bold mark statistically significant associations (p<0.05).

Supplemental Table S7. Coefficients of cross-sectional associations between annual expenditures (logarithmic-transformed) for fruits purchased by Nielsen households (between 3rd and 97th percentile of the fruit expenditures, n=21,269 a), availability of neighborhood supermarkets and convenience stores (5-km buffer, centroid of residential ZCTA as household’s residence), broader built environment context characteristics, and household-, neighborhood- and area-level covariates

|  |  |  |
| --- | --- | --- |
| Characteristics | Availability of neighborhood supermarkets and convenience stores only b | Availability of neighborhood supermarkets and convenience stores and broader built environment context b |
| Coefficient (SE) | p-value | Coefficient (SE) | p-value |
| Availability of supermarkets, count, 5-km buffer |  |  |  |  |
|  0 (Ref.) | **---** | **---** | **---** | **---** |
|  1 | 0.024 (0.014) | 0.103 | 0.012 (0.015) | 0.412 |
|  2+ | 0.040 (0.017) | **0.021** | 0.013 (0.019) | 0.470 |
| Availability of convenience stores, 10 counts, 5-km buffer | 0.004 (0.006) | 0.466 | -0.019 (0.007) | **0.007** |
| Broader built environment context |  |  |  |  |
|  Regional destination accessibility: Jobs within 45-min  automobile travel time , 10,000 jobs  |  |  | 0.003 (0.001) | **<0.001** |
|  Neighborhood destination diversity: Entropy, 10 percent, 5-km  buffer  |  |  | -0.000 (0.003) | 0.960 |
|  Availability of neighborhood destinations: Total other stores/malls, locations where people socialized and communicated with others, and restaurants, 10 counts, 5-km buffer |  |  | 0.000(0.000) | 0.240 |
|  Neighborhood street connectivity: 10 intersections per square  mile, 5-km buffer |  |  | 0.011 (0.003) | **0.001** |

n, number of observations; SE, standard error; Ref., reference category; ZCTA, zip code tabulation area.

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

a We excluded who reported extremely low or high values for purchases of fruits, defined here as less than the 3rd percentile or greater than the 97th percentile.

b Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one child in the family, and number of employed household members (household head excluded), expenditure (logarithmic-transformed) on vegetables, and urbanicity (R x64 3.5.1 and Rstudio 1.1456).

Entries in bold mark statistically significant associations (p<0.05).

Supplemental Table S8. Coefficients of cross-sectional associations between annual expenditures (logarithmic-transformed) for vegetables purchased by Nielsen households (between 3rd and 97th percentile of the vegetable expenditures, n=21,249 a), availability of neighborhood supermarkets and convenience stores (5-km buffer, centroid of residential ZCTA as household’s residence), broader built environment context characteristics, and household-, neighborhood- and area-level covariates

|  |  |  |
| --- | --- | --- |
| Characteristics | Availability of neighborhood supermarkets and convenience stores only b | Availability of neighborhood supermarkets and convenience stores and broader built environment context b |
| Coefficient (SE) | p-value | Coefficient (SE) | p-value |
| Availability of supermarkets, count, 5-km buffer |  |  |  |  |
|  0 (Ref.) | **---** | **---** | **---** | **---** |
|  1 | 0.014 (0.013) | 0.298 | 0.017 (0.014) | 0.206 |
|  2+ | 0.014 (0.016) | 0.363 | 0.017 (0.017) | 0.326 |
| Availability of convenience stores, 10 counts, 5-km buffer | 0.021 (0.005) | 0.968 | -0.001 (0.007) | 0.936 |
| Broader built environment context |  |  |  |  |
|  Regional destination accessibility: Jobs within 45-min  automobile travel time, 10,000 jobs  |  |  | -0.000 (0.001) | 0.820 |
|  Neighborhood destination diversity: Entropy, 10 percent, 5-km  buffer  |  |  | -0.008 (0.003) | **0.002** |
|  Availability of neighborhood destinations: Total other stores/malls, locations where people socialized and communicated with others, and restaurants, 10 counts, 5-km buffer |  |  | 0.000 (0.000) | 0.371 |
|  Neighborhood street connectivity: 10 intersections per square  mile, 5-km buffer |  |  | 0.001 (0.003) | 0.781 |

n, number of observations; SE, standard error; Ref., reference category; ZCTA, zip code tabulation area.

Food purchase data and household-level sociodemographic data were derived from the Nielsen Homescan Consumer Dataset in 2010. Copyright © 2018, The Nielsen Company.

a We excluded who reported extremely low or high values for purchases of vegetables, defined here as less than the 3rd percentile or greater than the 97th percentile.

b Regressions controlled for percent of zero-car households in the residential census block group, percent of population below poverty level in the residential census tract, household income, race identity of household, household size, marital status, if there is at least one child in the family, and number of employed household members (household head excluded), expenditure (logarithmic-transformed) on fruits, and urbanicity (R x64 3.5.1 and Rstudio 1.1456).

Entries in bold mark statistically significant associations (p<0.05).

## References

1. Busso M, DiNardo J, McCrary J. New evidence on the finite sample properties of propensity score reweighting and matching estimators. Rev Econ Stat. 2014;96(5):885-97.

2. DiSantis KI, Hillier A, Holaday R, Kumanyika S. Why do you shop there? A mixed methods study mapping household food shopping patterns onto weekly routines of black women. International Journal of Behavioral Nutrition and Physical Activity. 2016;13(1):11.

3. Ramsey K, Bell A. Smart location database Version 2.0 User Guide [Internet]. Washington, DC: United States Environmental Protection Agency; 2014 [updated 2014 Mar 14. Available from: <https://www.epa.gov/smartgrowth/smart-location-database-technical-documentation-and-user-guide>.

4. Zhen C, Taylor JL, Muth MK, Leibtag E. Understanding differences in self-reported expenditures between household scanner data and diary survey data: a comparison of Homescan and consumer expenditure survey. Appl Econ Perspect Policy. 2009;31(3):470-92.

5. Crawford PB, Obarzanek E, Schreiber GB, Barrier P, Goldman S, Frederick MM, et al. The effects of race, household income, and parental education on nutrient intakes of 9-and 10-year-old girls NHLBI growth and health study. Ann Epidemiol. 1995;5(5):360-8.

6. Northstone K, Emmett P. Multivariate analysis of diet in children at four and seven years of age and associations with socio-demographic characteristics. Eur J Clin Nutr. 2005;59(6):751.

7. Hendricks K, Briefel R, Novak T, Ziegler P. Maternal and child characteristics associated with infant and toddler feeding practices. J Am Diet Assoc. 2006;106(1):135-48.