Food Access and Food Expenditures: A Multidimensional Examination

Jackie Yenerall and Wen You[1]

Introduction

Simultaneous changes in the local food environment, consumption patterns, and diet-related diseases have prompted researchers to investigate possible relationships between these phenomena (Morland, 2015; Papas et al, 2007). Research on the food environment started in the 1990's with an interest in identifying areas that had poor access to grocery stores and eventually these areas became known as food deserts (Walker et al, 2010; Larson et al, 2009; Charreire et al, 2010). In this context the distance from the household to the nearest grocery stores defined access, and longer distances indicated poorer access. This literature consistently revealed disparities in access amongst communities in the United States that have high proportions of low-income and African American residents (Beaulac et al, 2009; Walker et al, 2010; Larons et al, 2009; Cummins and Macintyre, 2005). These findings partially reflect significant changes in the retail food environment since the 1960's, which has resulted in the rapid growth of chain supermarkets, and the movement of store locations to the edges of town (White, 2007; Walker et al, 2010).

Meanwhile researchers identified disparities in healthy food consumption and obesity by race and socio-economic status (White, 2007, Wang and Beydoun, 2007; Flegal et. al, 2010; Wen and Kowaleski-Jones, 2012). Given the disparities in healthy food access and the relationship between healthy food consumption and diet related diseases (such as obesity), the local food environment literature turned towards the relationship between access and individual consumption and health outcomes in the early 2000's (White, 2007; Capsi et al, 2012).

[1] Senior authorship is shared between the two authors. Jackie Yenerall: Ph.D., population health economist, Tennessee Department of Health; Wen You: Ph.D., associate professor, Department of Agricultural and Applied Economics, Virginia Tech. Email: wenyou@vt.edu.

One frequently investigated channel is the access to healthy food. However, the literature examining the effect of healthy food access on expenditure, consumption and health outcomes has found mixed evidence, and when significant relationships are found, the magnitude is often small (Morland, 2015; Ver Ploeg et. al, 2014; Capsi et. al, 2012). Two intertwined challenges are currently recognized in the literature: one is the lack of theoretical guidance in defining access (Papas et al., 2007) and the other is the operationalization of the access definition (i.e., measurement issues) (Thornton et al., 2012; Blitstein et al., 2012).

Access has been conceptualized to include five dimensions: availability, accessibility, affordability, acceptability, and accommodation (Charreire et al, 2010; Capsi et al, 2012). Availability and accessibility capture the geographical component of access, and measure the density and proximity of food retailers respectively. Accessibility is specifically designed to capture the geographical ease with which food can be obtained. Affordability captures costs and perceived budget friendliness, and the remaining two dimensions, acceptability and accommodation, capture the perceptions and quality of the local food environment (Capsi et al, 2012; Powell and Baio, 2009). Acceptability describes attitudes about the local food environment, and accommodation describes how well local food sources meet local needs (Capsi et al, 2012; Charreire et al, 2010). Very few studies attempt to measure acceptability or accommodation since they relate to attributes of the food retailer, such as hour of operation, quality of produce selection, and require more intensive data collection methods such as store audits and surveys (Capsi et. al, 2012).

Although all five dimensions may contribute to the effect of the local food environment on individual behavior, most research only focuses on the effect of a single dimension, with availability and accessibility being the most popular dimensions (Morland, 2015). However the gestalt concept nature of the food environment has been recently recognized (i.e., the food environment effect is greater than the mere summation of its components) and therefore calls for systematic definition, measurement and estimation have been put out (Morland, 2015). We illustrated in this paper that multi-dimensional investigation is important

in this literature due to the fact that single-dimensional examination, even with error-free measurement, cannot capture the core of the problem.

Besides the conceptualization, the most often examined dimensions also suffer measurement errors. Availability is generally measured either at the individual or a more aggregated geographical level. When measured at the individual level a buffer around the individual's home is used to designate the geographic area of interest (the relevant local food environment) and the count or density of food retailers is calculated within the buffer (Capsi et al, 2012). Alternatively, availability can be calculated for a geographic unit, which is generally defined by the census (i.e. block group, or census tract), and the density or count of food retailers is calculated for the geographic unit (Morland, 2015). In either case, a higher density of grocery stores would indicate better availability of healthy food therefore a higher quality local food environment. On the other hand, a higher density of fast-food restaurants would indicate better availability of unhealthy food and a lower quality local food environment.

Accessibility is most often measured at the household level, using the distance from home to the nearest food retailer (Aggarwal et al, 2014; Dunn et al, 2014). In the case of lacking individual home location information, some studies have used the centroid of a census unit to serve as the proxy (Capsi et al, 2012; Inagami et al, 2006). In either case, a longer distance to the retailer indicates poorer accessibility. There are several ways to measure distance: a Euclidean distance will measure the straight line distance from the home to the retailer, while a network distance uses distances traveled on available roads. Availability is generally defined using a one-mile buffer for urban areas and a ten-mile one for rural areas. These choices for buffers are based on the assumptions that one mile is a reasonable distance of an individual to travel on foot in an urban area while and ten-mile is a reasonable distance for a person to travel in a rural area (Ver Ploeg et al, 2014).

One common weakness in the above common measurements is the implicit assumption that the household usually shops at the store that is nearest to their home which is why the nearest store (such as grocery store) distance should captures the majority of healthy food access feature of the local food environment

and the homogeneous buffer choice is reasonable for capturing the degree of exposure (Aggarwal et al, 2014). If the nearest store is not the primary store where the household usually shops, this assumption is easily one of the reasons for the mix findings of the current literature. The core of this issue is the equivalence between local food environment (defined based on the home distance to nearest store and a constant buffer choice for everyone in the same region) and actual consumer food environment (defined based on the home distance to the observed primary store of the consumer's actual choice and heterogenous buffers for different households with different primary store choices). If the consumers' actual primary store choice is not their nearest store, the primary food environment they are exposed to will be larger and differ across households than the local food environment. Therefore the mixed evidence shown in the literature may merely because of not targeting the correct areas.

Furthermore, the household's decision of which store to shop primarily can reflect the effect from unobserved factors that also simultaneously influencing health-related behaviors and outcomes such as often uncontrolled affordability, acceptability and accommodation and those unmeasured part of availability and accessibility.

Our paper will test this implicit assumption through endogenous switching regression models to recognize the potential unobserved factors that simultaneously influence primary shopping store choices and the healthy food expenditures. Further we will compare the results with naïve models that follow the similar specifications as the previous literature where one or two dimensions are studied without recognizing the potential differences of consumer realized vs. local food environment.

Lastly, there are issues to be addressed when operationalizing the access definition and measurement. First is the choice of food retailer of interest to be the focus of capturing food environment. The North American Industry Classification System (NAICS) is often to be used in distinguishing between retailer types, such as grocery store versus supermarket, or fast food versus sit down restaurant (Morland, 2015; Morland et. al, 2002). The current literature often uses grocery store to be the

flagship of healthy food environment and recent studies have expanded to consider other outlets such as fast food restaurants and corner stores to capture poor food environment (Walker et al, 2010; Larson et al, 2009). We will consider both aspects in our study to allow potential off-setting effects from both sides of the coin.

Secondly, the current literature evidence is also sensitive to the chosen outcomes. Fruit and vegetable consumption is often chosen as the healthy diet outcomes. The findings are either conflicting (i.e., longer the distance from a grocery store is correlated with more fruit and vegetable consumption) or no or minimal effect of accessibility and availability of healthy food found (Ball et al, 2009; Larson et al, 2009; Capsi et al, 2012; Dunn et al, 2012; Thornton et al, 2012, Morland, 2015). Our study will focus on household level behaviors such as household total food expenditures and total fruit and vegetable expenditure instead of consumption to avoid the consumption recall errors. Furthermore, household exposure can use home location as the foci while individual exposure will involve even more layers with different work and study locations etc.

Model Framework

The households are decision-making units that choose where to shop primarily for their groceries and what to buy when choose to cook at home and make similar decisions on eating out. No matter it is food at home or food away from home the households will need resource inputs such as money and labor (i.e., time) to produce outcomes such as health. The most suitable economic framework for this type human behavior is Becker's household production model. We modified the standard household production model to suit the context of food environment and household behaviors and illustrated the need for multi-dimensional look at the problem in addition to providing empirical specification guidance.

For simplicity, we group household produced goods to three categories: meals made from home (Z_{FAH}), meals made away from home (Z_{FAFH}) and health (H) and allow a composite market goods (X_o) to directly contribute to utility along with those three categories. Therefore, the household will have the utility function as the following:

(1) $U(Z_{FAH}, Z_{FAFH}, X_o, H; \boldsymbol{K})$,

where $\boldsymbol{K}$ is a vector of standard demographics. The household will have three sets of production functions:

(2) $Z_{FAH} = f(X_g^M, X_g^O, T_g^S, T_g^P, T_g^C; \mu, \boldsymbol{K}, \boldsymbol{E}_g^P)$,

(3) $Z_{FAFH} = f(X_a^F, X_a^R, T_a^S, T_a^C; \mu, \boldsymbol{K}, \boldsymbol{E}_a)$,

(4) $H = f(Z_{FAH}, Z_{FAFH}, X_o, T_{PA}; H^o, \boldsymbol{K})$,

where X_g^M denotes groceries purchased from primary store, X_g^O denotes groceries purchased from other stores, (T_g^S, T_g^P, T_g^C) are time spent in grocery shopping, production and consumption respectively, $\boldsymbol{E}_g^P$ is the food environment vectors related to the primary store which should capture all five dimensions. The assumption here is that the primary store is the one defining the at-home-food groceries environment exposure. For food away from home production function, there are inputs of away from home foods from fast food outlets (X_a^F) and sit-down restaurants (X_a^R) with associated time allocations: shopping (i.e., travel time) (T_a^S) and consumption time (T_a^C). The associated away-from-home food environment vector is $\boldsymbol{E}_a$. There is also nutritional knowledge factor, μ, in both production functions. The health outcome is produced using those meals produced at home, away from home and other composite goods and have energy expenditure sources (denote by time spent in physical activities, T_{PA}) and is conditioned on the initial health stock, H^o.

Meanwhile, two standard resource constraints are:

(5) $p_g^M \cdot X_g^M + p_g^O \cdot X_g^O + p_o \cdot X_o + p_a^F \cdot X_a^F + p_a^R \cdot X_a^R = w \cdot T_w + I$,

(6) $T_g^S + T_g^P + T_g^C + T_a^S + T_a^C + T_{PA} + T_w + T_o = T$,

where the wage, w, and assumed pre-determined time spent at work, T_w, constitute the earned income part and I denotes the unearned income sources.

The household maximization problem then leads to the following indirect input demands:

(7) $\left(X_g^M, X_g^O, X_a^F, X_a^R, T_g^S, T_g^P, T_g^C, T_a^S, T_a^C, T_o\right) = g(p_g^M, p_g^O, p_o, p_a^F, p_a^R, w, I, T, \boldsymbol{E}_g^P, \boldsymbol{E}_a)$.

Several insights can be gained from equation (7). First, examining grocery purchases should also control for away-from-home food environment in addition to the

grocery food environment. Second, within each food environment all five dimensions should be considered to capture the gestalt nature of the problem. Third, usually we do not have specific price information which leads to the substitution of quantity purchased with expenditures (i.e., substitute X_g^M with $p_g^M \cdot X_g^M$). Forth, the affordability dimension of $\boldsymbol{E}_g^P$ after controlling for prices should now have left the perceived affordability component. Lastly, the grocery food environment exposure radius is mostly relevant to primary store choice since people usually mainly go to one store for grocery shopping while it is not a reasonable assumption for eating out. Therefore, the assumption of nearest grocery store is the primary store is an important one to be empirically tested.

Empirical Method

The above theoretical framework assumes that the primary store choices are pre-determined. If this assumption is true, the empirical estimation of equation (7) should be free from endogeneity problem since the primary store associated grocery environment should not be affected by the unobserved factors that also influencing food purchasing. However, similar to the classic returns-to-education by occupation problem, there are suspiciously unobserved factors (e.g., cultural preference which can also be a mirror reflection of acceptability dimension, unobserved taste etc.) that is correlated with primary store choices therefore the primary store associated environment vectors. This causes the concerns for inconsistency. The usual empirical method to explicitly model the endogenous selection process is the so call endogenous switching regression (or Roy Model) (Cameron and Trivedi, 2005).

In general the behavior of a household will be allowed to differ depending on which regime it chooses. There is a criterion function, ξ, to determine the regime choice.

(8) $\xi_i = 1 \quad if\ \gamma Z_i + u_i > 0;\ \xi_i = 0 \quad if\ \gamma Z_i + u_i \leq 0,$

(9) $Regime\ 1: y_i^1 = \beta^1 X_i^1 + \epsilon_i^1,\ if \quad \xi_i = 1\ ,$

(10) $Regime\ 2: y_i^2 = \beta^2 X_i^2 + \epsilon_i^2,\ if \quad \xi_i = 0\ .$

The error terms, $(u_i, \epsilon_i^1, \epsilon_i^2)$ have a trivariate normal distribution with mean vector zero and covariance matrix:

$$\Omega = \begin{bmatrix} \sigma_u^2 & \sigma_{1u} & \sigma_{2u} \\ \sigma_{1u} & \sigma_1^2 & \cdot \\ \sigma_{2u} & \cdot & \sigma_2^2 \end{bmatrix}$$

where $(\sigma_u^2, \sigma_1^2, \sigma_2^2)$ are the variances of the error terms $(u_i, \epsilon_i^1, \epsilon_i^2)$, the covariance between (u_i, ϵ_i^1) is denoted as σ_{1u} and the covariance between (u_i, ϵ_i^2) is denoted as σ_{2u}. The covariance between $(\epsilon_i^1, \epsilon_i^2)$ is not defined since the outcomes from both regimes are never observed simultaneously for one individual. The model is usually estimated using maximum likelihood estimators.

Put into our specific context, the selection equation is whether or not the household chose the nearest store to its home as the primary store. The equation (8) shows that when the latent utility of choosing the nearest store is greater than a certain threshold (normalized to zero for simplicity) the primary store of the household equals to the nearest store and otherwise the primary store choice is not the same as the nearest store. The household total food expenditure and total fruit and vegetable expenditure will then follow equation (9) or (10) depending on the choice. The selection equation will be a function of different components of access. The identification instruments on the selection equation are relative differences in distance from home to primary store and home to the nearest store of the same type, reasons for the primary store choice (i.e., perceived affordability, perceived quality, perceived distance from home), availability of different types of stores and restaurants that use a larger radius (10-mile for both urban and rural regions) to cover a sizable area for store choices.

The continuous equation will have log form of the expenditures (either total food expenditure or total fruit and vegetable expenditure) as dependent variable. The common controls in both selection and continuous equations are households' and primary respondents' characteristics (i.e., household size, income, child presence or not, auto ownership, SNAP participation status, gender, race, education). The continuous expenditure equation will have an additional dummy to indicate if a holiday occurs during the observation week. It will also control distance to the

primary store and control observed availability instead of the constant 10-mile availability as in the selection equation. The rationale is that when choosing stores a large homogenous radius is needed to cover all potential store options; but when the choice is made the frequent trips to that store marks out an observed radius where within this radius the households have larger exposure to the food environment within. Therefore the observed availability is defined as: using the distance to the primary store as the radius to calculate the availability of different stores and restaurants within that window which will be different across households.

The tests of the statistical significance of the correlation coefficients between error of continuous equation and selection equation will provide indication of potential existence of selection effect. The statistical significant correlation coefficient will show that previous literature's assumption of nearest store defines the core of the food environment exposure area is questionable. The data set we will use although rich enough to provide both local and observed food environment in accessibility and availability dimensions and limited perceived affordability and actual price information and limited perceived acceptability, it does not have information on accommodation. The household expenditure results conditional on primary store choices is one way to take into account the missing dimensions without specifically measuring them.

Furthermore, results from the model will be compared to a more traditional form of access model that considers only pre-defined measures of availability (i.e., 1-mile buffer for urban and 10-mile buffer for rural) and uses the distance to the nearest store to measure accessibility. All models will be estimated separately for rural and urban subsamples to allow food environment effects and primary store choice underlying mechanisms to differ. The comparison between naïve model and our model that considers primary store choice and therefore the difference between consumer food environment and local food environment influences can shed light on how the assumption of nearest store is the most relevant store affect results in this context.

Data

This paper utilizes a nationally representative data set that contains information on detail household food expenditures and both the standard local food environment and observed food environment. The data comes from the Food Acquisition and Purchase Survey (FoodAPS), which was administered through the United States Department of Agriculture Economic Research Service (USDA-ERS). This survey includes 4,826 households and is sampled in a way to be not only nationally representative but also be representative of the Supplemental Nutrition Program (SNAP) population.

The survey collected households' detailed expenditure information for seven consecutive days on both food-at-home (FAH) and food-away-from-home (FAFH). Expenditure data was collected though a combination of survey books coupled with interviews, scanning-data, and receipts. For each expenditure event the store location was recorded, which allows for the calculation of the distance from home. The primary respondent was identified in a screening interview as the household's primary food shopper or meal planner. The primary respondents also reported the household's primary store (defined as the store were they make most of their food purchases) and were asked to select (can have multiple selection) three provided reasons that can explain their primary store choice: low prices or good value ('good prices'), good quality food ('good quality') and close to home.

Our study examined two expenditure outcomes: total household FAH weekly expenditures and total household fruit and vegetable weekly expenditures. The FAH events were identified by the following definition: "food and drinks that are brought home and used to prepare meals for consumption at home or elsewhere (for example, food used to make a sandwich that you bring to work)" (USDA ERS FoodAPS documentation 2015). All events (both FAH and FAFH) that were identified by the participants as "free" were excluded along with those households with zero total FAH expenditures for the week (n=158). Fruit and vegetable expenditures were identify from item level data, which recorded information on the items purchased during each FAH event. Dollar amounts spent on items purchased

from the produce or frozen produce sections of the food retailer store (as identified from their IRI aisle classification) were counted as fruit and vegetable expenditures.

In additional to standard household characteristics and respondent's socio demographics, the FoodAPS also collected information on local food environment of each household (FoodAPS Retail Environment Study, 2014). At-home-food (grocery) retailer information was collected through the USDA Store Tracking and Redemption System (STARS), which is a national dataset that contains the location of all SNAP authorized retailers. The types of grocery retailer include supermarket, superstore, grocery store (including medium, large, and combination of grocery store and other store), and convenience store and other. Away-from-home-food (restaurant) retailer information was collected through a private data set called InfoUSA and includes fast-food restaurant and non fast-food restaurant.

The FoodAPS restricted data provides the calculated straight-line distances from the household home location to the nearest retailers for each type of stores and restaurants and the primary store. For availability, the counts of each type of stores and restaurants were recorded for different buffers (i.e., 0.25-mile, 0.5-mile, 1-mile, 2-mile, 5-mile, 10-mile, 20-mile and 30-mile). The observed availability used in our empirical estimation is the FoodAPS provided availability with the nearest buffer size as compared to the distance to the household's primary store (i.e., if the household's chosen primary store is about 2.1 miles in straight-line distance from the home, the observed availability measurements will be the count of different types of stores and restaurants within a 2-mile buffer).

We do not have nutrient intake data available yet therefore could not examining the household production outcomes. However, the input demands for producing health are also important to be examined. One of the policy and programs interventions' common goals is to make healthy food more accessible and promote more choices towards healthy food. Therefore, this paper chose two input demands to study: one is the total household weekly at-home-food expenditures to examine whether there are food environmental influences on increasing relatively healthier (as compared to away-from-home-food) at-home-food spending; second is the total

household weekly fruit and vegetable expenditures to zoom in on the most commonly promoted food category for healthy eating.

Results and Discussion

As discussed in the data section, after removing those households with zero reported at-home-food expenditure of the week, the total sample size is 4,668. The rural subsample contains 1,264 households which is about 27% of the full sample and the urban subsample contains 3,404 households. The sample weighted summary statistics of selected household and individual level variables used in our model are reported in Table 1. The summary statistics were presented for full sample and for rural and urban subsamples. Within the subsample, it is further broken down to the subgroup of households whose primary store of their choice is the nearest store of the same type to their home (Primary = Nearest) (PN group) and the subgroup of households who did not choose the nearest store as their primary store (Primary ≠ Nearest) (PNN group).

The overall average household size is about 2.4 and in rural subsample those PNN households on average have statistically significantly larger household size as compared to those PN groups (2.4 vs. 2.2 with $p<0.05$). About 33% households have children presence and it is similar across urban and rural. The urban subsample exhibits relatively higher average household monthly gross income as compared to rural with national average of about $5,100.00. The urban subsample also has larger percentage of households who participated in the Supplemental Nutrition Assistance Program (SNAP) during the time of the interview than the rural sample with national average of about 13%. Comparing car and home ownership, the rural subsample exhibits relatively higher percentage of ownerships with national average of 90% car owners or leasers and 62% home owners.

The primary respondents were primarily females (67% nationally) and the rural subsample PNN group have significantly larger percentage of females as the primary respondents as compared to PN group (76% vs. 60% with $p<0.01$). There are about 13% primary respondents are African Americans (urban contains higher African American population share) and 80% are Caucasians (rural contains higher

Caucasian population share). There are about 32% primary respondent with college degree or higher and 8% with education of less than high school level and 45% are married.

The dependent variables of interest here are household weekly at-home food expenditure and fruit and vegetable expenditures. The average at-home-food household weekly expenditure is $108.42 and away-from-home-food expenditure is about $60.07 which renders an average of 63% FAH budget share among total food expenditure. The average fruit and vegetables expenditure is only about $11.66. The statistical significant differences between PN and PNN groups for those expenditure items also only are observed in rural subsample: The PNN groups on average spent more on FAH ($114.54 vs. $95.69 with $p<0.001$) and spent more on fruit and vegetables as well ($11.28 vs. $8.5 with $p<0.01$).

Consistent with primary store definition, the primary store at-home-food expenditure budget share among the total FAH expenditure is on average 44% and about 50% total fruit and vegetable spending happened in primary store. Both expenditure shares exhibit statistical significant differences between PN and PNN groups: in both rural and urban subsamples, the PN groups show average higher budget shares in primary store spending as compared to PNN groups (FAH from primary store/total FAH budget share: 57% (49%) rural (urban) PN group vs. 42% (40%) in rural (urban) PNN group. Fruit and vegetable spending in primary store/total fruit and vegetable spending budget share: 68% (55%) in rural (urban) PN group vs. 50% (45%) in rural (urban) PNN group).

Turning to food environment related variables, we reported accessibility of primary store and nearest store of the same type and availability with 1-mile and 10-mile radius summary statistics in Table 2. Those summary statistics were also weighted by the sample weights provided in FoodAPS. Majority of the primary stores are supermarket (47% overall) and more households in PN groups choose supermarket as compared to PNN group in both rural and urban subsamples (56% (61%) in rural (urban) PN group vs. 32% (45%) in rural (urban) PNN group). This trend is reversed when we examine the superstore: on average PNN groups contain more superstore primary store choice than PN groups. In this primary store type

break down summary statistics, we cannot report further break down of the store types besides supermarket and superstore due to data security concerns expressed by USDA ERS therefore we can only group grocery store, convenience store and other stores into "other" category. This "other" store category contains very minimal of the primary store choices: 8.6%. In urban subsample, those households in PNN group select this type more than those in PN group (9.9% vs. 0.8% with $p<0.001$). This may in part due to mainly the convenience store availability abundance in urban group as shown in the later availability numbers.

There are about 30% households choose their nearest store as their primary store for at-home-food purchases. This percentage is slightly higher in rural subsample as compared to urban (34% vs. 29%). The selected reasons for primary store choices show interesting evidence. The reasons of "good prices" (perceived affordability) (52.6% chose this) and "close to home" (perceived accessibility) (53.3% chose this) are the two major ones. Those households in urban subsamples who did not choose the nearest store as their primary store (PNN group) placed relatively more weight on perceived affordability: 58% on average choose "good prices" as the reason in urban PNN group vs. 37% on average choose "good prices" in urban PN group. Consistent with our expectation that those who choose their nearest store as primary store contain much more households who think "close to home" is the reason which is similar in both rural and urban (64% (75%) in rural (urban) PN group vs. 41% (46%) in rural (urban) PNN group). The reason of "good quality" (perceived acceptability) is chosen by about 21% of the sample and is similar across subsamples.

The average straight-line distance from the household's home to the household's reported primary store is about 3.8 miles. This is longer than the traditionally assumed food environment relevant radius for urban (which is 1-mile). Consistent with the expectation, those PNN group's primary store choice is further away from home as compared to PN group (8.6 miles (2.2 miles) in rural (urban) PNN group vs. 4.4 miles (0.9 miles) in rural (urban) PN group). The average distance to primary store in rural is over 6 miles which is shorter than the commonly assumed food environment influence radius of 10 miles for rural and the average distance in

urban is over 1.5 miles which is longer than the commonly assumed relevant radius of 1 mile for urban. The average differences in distance to primary store as compared to the nearest store of the same type is about 1.2 miles with larger difference in rural (2.9 in rural vs. 1.1 in urban).

The availability measures presented here were the counts of different food outlets within 1-mile and 10-mile radius. As expected, the counts increase dramatically from 1-mile to 10-mile. Urban subsamples exhibit more counts for all types of food outlets than rural subsamples as expected. This does not signal better availability since both healthy food and unhealthy food availability are more abundant in urban. Those PNN group in urban subsample on average face more convenience store choices as compared to the PN group (6.3 vs. 4.2 with $p<0.001$).

Household Weekly At-Home-Food Total Expenditure Demand Models

We first start with comparing the continuous expenditure equation results from the endogenous switching regression with the naïve model that ignores primary store choice. The results are reported in Table 3. The standard errors are robust adjustment accounting for the probability weightings of the sample frame of the data. Due to the skewness feature of the expenditure data, we use the log form of the expenditure data as the dependent variable. The Model 1 is the naïve model that treats the nearest store as the relevant base for both accessibility and availability measurement. The Model 2 is the endogenous switching regression model that allows different behavior mechanisms for those households who chose the nearest store as their primary store (PN group) and those households who did not choose the nearest store as the primary store (PNN group). Consistent with previous findings, no statistical significant accessibility effects were found in all models. However it will show later for model 2 those null effects were conditional on the primary store choice.

Among all the availability measures, the counts of convenience stores showed significant negative effect on household at-home-food expenditure for both rural and urban regions (for one more convenience store increase there will be about 0.8% (0.9%) decreases to rural (urban) household at-home-food weekly

expenditure amounts). These significant results disappear when accounting for primary store choice in rural region. For urban region, the switching regression (model 2) finds that the significant negative influence of convenience store availability stays for the PN group but it has positive influence for PNN group. For those households did not choose nearest store as primary store (PNN group), one convenience store increases will induce an average 0.2% increase in the household at-home-food weekly expenditure. The only significant away-from-home food sources is the fast food restaurant availability for PN group where one more increase within the primary-store-distance-based radius will increase expenditure by 2.3%. Those all signals the latent differences between the PN and PNN groups.

The household level variables, household size and monthly income, consistently show positive and significant influences with similar magnitudes throughout models and just lost significance for PN group in rural subsample. The car possession status is statistically significant just for urban subsamples where having an access to a car (either own or lease) resulting in about 22% more at-home-food expenditure . The PNN group behaves similar to the evidence shown in the naïve model while PN group did not show significance. Most of the significant demographic variable results disappear when we look at the PN subgroups in Model 2 for both urban and rural subsample. More specifically, results show that after considering potential primary store choice endogeneity, those commonly examined expenditure influencing factors loose statistical relevance for those rural households who chose the nearest store as their primary store. The food expenditures from households who did not choose the nearest store as their primary store no matter in rural or urban regions seem to still be influenced by availability and can be explained by some household level characteristics and demographics.

Household Weekly Fruit and Vegetable Expenditure Demand Models

Similar models were also estimated using the households' weekly fruit and vegetable expenditure (in log form) as dependent variables and results were presented in Table 4. Only PN group in urban subsample show positive expenditure response to accessibility to primary store. For one mile increases in primary store

straight-line distance, there will be about 22% increase in the fruit and vegetable expenditure. This seems to suggest that once the primary store is chosen the further travel distance will result in larger spending in fruit and vegetables category. However we cannot distinguish between fresh, frozen and canned further yet. The future research should focus on those breakdowns to understand what those increase mainly come from.

The naïve model suggests significant and negative influence of convenience store availability and positive and significant influence of supermarket/superstore on expenditure for rural subsample only. But these results disappear when we consider primary store choice. On the opposite, no significant availability in urban subsample are indicated in the naïve model while switching regression model suggests that for those PN group households convenience store availability exhibits negative influence on fruit and vegetable spending while non-fast food restaurants availability shows positive influence: one more convenience store increase will result in an average 4.2% decrease in fruit and vegetable expenditure and one more non-fast food restaurant increase will result in about 0.6% increase in the spending.

The household size and income show significant positive influence for urban subsamples. The access to car does not show as important role in fruit and vegetable expenditure as in total at-home-food expenditure with no significance found in switching regression. Having college or higher education level consistently exhibit positive influence on fruit and vegetable spending: comparing to those have lower than college education level, those households spend about 29% to 62% more on fruit and vegetable expenditures with everything else held constant. The largest effect happens in those PNN group in rural subsample.

Primary Store Selection

Now let us turn attentions to the primary store selection stage results in the endogenous switching regression model reported in Table 5. The relative accessibility (i.e., the relative distance differences between the distance from home to primary store and distance from home to nearest store) consistently shows significant negative influence on the probability of choosing the nearest store as the

primary store. This is consistent with the expectation in that the further away the two stores are the less likely the households would choose the nearest store. In other words, the primary store choice has more features than merely "nearest to home". This effect is larger in urban subsamples. Holding the relative distance constant, in urban subsample, the higher perceived affordability will decrease the probability of choosing the nearest store of the same type while the higher perceived accessibility will increase the probability.

The only significant local food environment factor in primary store choice is the supermarket/superstore availability in rural when considering at-home-food expenditure: the more availability of supermarket/superstore within 10-mile radius the less likely the households in the rural area will choose their nearest store of the same type as the primary store.

Combined with the primary store choice reasons, it shows that the traditional local food environmental availability did not seem to affect primary store choice much especially in urban subsample. This may explain those field experiments that plan new grocery stores in urban food deserts and did not desirable find behavior changes.

Let us examine the statistical evidence of nonrandom selection indicated by the correlation coefficient between selection equation and continuous expenditure equations. The only significant correlation coefficient is between selection equation error and PN group expenditure equation error for urban subsample when examining at-home-food expenditure. This indicates that those urban households who chose the nearest store as their primary store tend to spend less on at-home-food expenditure than a random household from the sample who has positive at-home-food expenditure, and those households who did not choose the nearest store as primary store do not spend more or less than a random household. Coupled with the relative abundance of store availability, it seems to indicate that those households chose nearest store as their primary store in the food environment that has relatively more choices are more likely to have unobserved preferences that influence at-home-food spending.

Most of the primary store vs. nearest store selection effects do not show statistical significance seem to suggest that most of choice of nearest store as primary store are exogenous to at-home-food and fruit and vegetable expenditure decisions in our sample. These may explain why merely putting stores near to residence area did not work as desired.

Conclusion

The food access literature has long struggled with puzzling null or conflicting findings on accessibility and availability. Although different study design, sampling frame and measurements prevent systematic comparison across studies, there are no theoretical guidance in empirical investigation of food access and there has been no justification of using a constant radius across households for measuring relative food environment coverage.

This study presented a modified household production model to motivate the need for multi-dimensional look at the food access issue. We utilized a nationally representative rich data set, FoodAPS, to examine food environmental effect on household at-home-food spending and fruit and vegetable spending. Although not examining health outcomes directly, we are able to examine household level expenditure decisions that are the direct inputs into individual health production. Further, although this data set has provided more information for the investigation than any other known data set, it does not have direct measurement of other dimensions of the food environment such as acceptability, accommodation and perceived affordability. To circumvent this limitation and guided by the theoretical model framework, we examined the common assumption of nearest store is the primary store of choice in the current literature by endogenous switching regression that allow the underlying mechanism differences and directly model the primary store choice process.

We found that accessibility as measured in straight-line distance only has influences in primary store choice and did not influence actual expenditure amount decisions. This may explain findings of null effect in the current literature. On the contrary, availability of different types of food outlet seems to show some effect on

actual spending amount (i.e., convenience store) but not store choice. Household size and household income affect actual spending amount but not store choices as well. Education level is positively correlated with fruit and vegetable spending but not the total at-home-food expenditure. Urban region exhibits more household level and demographic profile correlations than rural region.

The testing of primary store choice endogeneity in the food expenditure context shows that the store choice seems to be exogeneous. This may explain several null findings of those field experiments that plant stores into food desert areas. Our study can be further improved by examining network distance instead of straight-line distance which is expected to make a bigger difference especially in urban regions. Further household level aggregated health and nutrient intake outcomes should be the next step of investigation as well.

Reference

Aggarwal, Anu, Andrea Cook, Junfeng Jiao, Rebecca Seguin, Anne Vernez Moudon, and Philip Hurvitz." Access to Supermarkets and Fruit and Vegetable Consumption." American Journal of Public Health 104.5 (2014): 917-23.

Baker, Elizabeth, Mario Schootman, Ellen Barnidge, and Cheryl Kelly. "The Role of Race and Poverty in Access to Foods That Enable Individuals to Adhere to Dietary Guidelines." Preventing Chronic Diseases 3.3 (2006): 1-11.

Ball, Kylie, Anna Timperio, and David Crawford. "Neighbourhood Socioeconomic Inequalities in Food Access and Affordability." Health and Place 15 (2009): 578-85.

Beaulac, Julie, Elizabeth Kristjansson, and Steven Cummins. "A Systematic Review of Food Deserts, 1966-2007." Preventing Chronic Disease 6.3 (2009): 1-10.

Becker, G.S. "A Theory of the Allocation of Time." Economic Journal. Vol. 75. (1965): 493-517.

Blitstein, Jonathan, Jeremy Snider, and W Douglas Evans. "Perceptions of the Food Shopping Environment Are Associated with Greater Consumption of Fruits and Vegetables." Public Health Nutrition 15.6 (2012): 1124-129.

Capsi, Caitlin E., Glorian Sorensen, S.V. Subramanian, and Ichiro Kawachi. "The local food environment and diet: A systematic review." Health & Place 18 (2012): 1172-1187.

Charreire, Helene, Romain Casey, Paul Salze, Chantal Simon, Basile Chaix, Arnaud Banos, Dominique Badariotti, Christiane Weber, and Jean-Michel Oppert. "Measuring the Food Environment Using Geographical Information Systems: A Methodological Review." Public Health Nutrition 13.11 (2010): 1773-785.

Chau, Clarice, Jamie Zoellner, and Jennie Hill. "Availability of Healthy Food: Does Block Group Race and Income Matter?" Journal of Hunger and Environmental Nutrition 8 (2013): 22-38.

"Creating Access to Healthy, Affordable Food." Agricultural Marketing Service. Web. 22 Sept. 2015. <http://apps.ams.usda.gov/fooddeserts/fooddeserts.aspx>.

Cummins, Steven, and Sally Macintyre. "Food Deserts"-evidence and Assumption in Health Policy Making." BMJ 325 (2002): 436-38.

Cummins, Steven, and Sally Macintyre. "Food Environments and obesity - neighborhood or nation?" International Journal of Epidemiology 35 (2005): 100-104.

Day, Peter, and Jamie Pearce. "Obesity-Promoting Food Environments and the Spatial Clustering of Food Outlets Around Schools." American Journal of Preventative Medicine 40.2 (2011): 113-21.

Dunn, Richard, Wesley Dean, Cassandra Johnson, Andrew Lediner, and Joseph Sharkey. "The Effect of Distance and Cost on Fruit and Vegetable Consumption in Rural Texas." Journal of Agricultural and Applied Economics 44.4 (2012): 491-500.

Economic Research Service (ERS), U.S. Department of Agriculture (USDA). National Household Food Acquisition and Purchase Survey (FoodAPS). http://www.ers.usda.gov/data-products/foodaps-national-household-food-acquisition-and-purchase-survey.aspx

FoodAPS Retail Environment Study Data. Economic Research Service (ERS), U.S. Department of Agriculture (USDA). National Household Food Acquisition and Purchase Survey (FoodAPS). http://www.ers.usda.gov/data-products/foodaps-national-household-food-acquisition-and-purchase-survey.aspx (July 22, 2014).

Glanz, Karen, James Sallis, Brian Saelens, and Lawrence Frank. "Nutrition and Environment Measures Study in Stores (NEMS-S) Development and Evaluation." American Journal of Preventative Medicine 32.4 (2007): 282-89.

Hamermesh, D.S. “Direct estimates of household production.” Economic Letters 98 (2008) 31-34.

Hill, Jennie, Wen You, and Jamie Zoellner. "Disparities in Obesity among Rural and Urban Residents in a Health Disparate Region." BMC Public Health 14 (2014): 1051-059.

Inagami, Sanae, Deborah Cohen, Brian Finch, and Steven Asch. "You Are Where You Shop. Grocery Store Locations, Weight, and Neighborhoods." American Journal of Preventative Medicine 3.1 (2006).

Kulldorff, Martin. "A Spatial Scan Statistic." Communications in Statistics - Theory and
Methods 26.6 (1997): 1481-496.

Larson, Nicole, Mary Story, and Melissa Nelson. "Neighborhood Environments Disparities in Access to Healthy Foods in the US." American Journal of Preventative Medicine 36.1 (2009): 74-81.

Liao, Tim. "Comparing Social Groups: Wald Statistics for Testing Equality Among

Multiple Logit Models." International Journal of Comparative Sociology 45 (2004): 3-16.

Maddala, G.S. "Disequilibrium, Self-Selection, and Switching Models." Handbook of Econometrics. Vol. 3. Elsevier Science, 1986. 1634-1682.

Michimi, Akihiko, and Michael Wimberly. "Associations of Supermarket Accessibility with Obesity and Fruit and Vegetable Consumption in the Conterminous United States." International Journal of Health Geographic's 9.1 (2010): 49-63.

Morland, Kimberly, ed. Local Food Environment Food Access In America. Boca Raton:
CRC, 2015.

Morland, Kimberly , and Kelly Evenson. "Obesity prevalence and the local food environment." Health & Place 15 (2009): 491-495.

Papas, Mia A., Anthony J. Alberg, Reid Ewing, Kathy J. Helzisouer, Tiffany L. Gary, and Ann C. Klassen. "The Built Environment and Obesity." Epidemiologic Reviews 29 (2007): 129-143.

Pearson, Tim, Jean Russell, Michael Campbell, and Margo Baker. "Do 'food Deserts' Influence Fruit and Vegetable Consumption? A Cross Sectional Study." Appetite 45 (2005): 195-97.

Pollak, Robert, and Michael Wachter. "The Relevance of the Household Production Function and Its Implications for the Allocation of Time." Journal of Political Economy 83.2 (1975): 255-78.

Powell, Lisa, and Yanjun Bao. "Food Prices, Access to Food Outlets and Child Weight."
Economics and Human Biology 7 (2009): 64-72.

Roja, Samina, Changxing Ma, and Pavan Yadav. "Beyond Food Deserts Measuring and
Mapping Racial Disparities in Neighborhood Food Environments." Journal of Planning Education and Research 27 (2008): 469-82.

Sadler, Richard, Jason Gilliland, and Godwin Arku. "An Application of the Edge Effect in Measuring Accessibility to Multiple Food Retailer Types in Southwestern Ontario, Canada." International Journal of Health Geographics 10.1 (2011): 34-49.

Sadler, Richard, Jason Gilliland, and Godwin Arku. "A Food Retail-Based Intervention on Food Security and Consumption." International Journal of Environmental Research and Public Health 10 (2013): 3325-346.

Saelens, Brian, Karen Glanz, James Sallis, and Lawrence Frank. "Nutrition and Environment Measures Study in Restaurants (NEMS-R) Development and Evaluation." American Journal of Preventative Medicine 32.4 (2007): 273-81.

Santana, Paula, Rita Santos, and Helena Nogueira. "The link between local environment
and obesity: A multilevel analysis in the Lisbon Metropolitan Area, Portugal." Social Science & Medicine 68 (2009): 601-609.

Smith, Chery, and Lois Morton. "Rural Food Deserts: Low-income Perspectives on Food
Access in Minnesota and Iowa." Journal of Nutrition Education and Behavior 41.3 (2009): 176-87.

Thornton, Lukar, Jamie Pearce, Laura Maconald, Karen Lamb, and Anne Ellaway. "Does
the Choice of Neighbourhood Supermarket Access Measure Influence Associations with Individual-level Fruit and Vegetable Consumption? A Case Study from Glasgow." International Journal of Health Geographics 11.1 (2012): 29-41.

USDA ERS. "Documentation." FoodAPS National Household Food Acquisition and Purchase Survey. 23 Mar. 2015. Web. 30 Nov. 2015. <http://www.ers.usda.gov/data-products/foodaps-national-household-food-acquisition-and-purchase-survey/documentation.aspx>.

"Urban and Rural Classification." Geography. The United States Census Bureau, 1 July 2015. Web. 25 Sept. 2015.

Ver Ploeg, Michele, Paula Dutko, and Vince Breneman. "Measure Food Access and Food Deserts for Policy Purposes." Applied Economic Perspectives and Policy (2014).

Walker, Renee, Christopher Keane, and Jessica Burke. "Disparities and Access to Healthy Food in the United States: A Review of Food Deserts Literature." Health and Place 16 (2010): 876-84.

Wang, May, Kara MacLeod, Cathy Steadman, Laura Williams, Shene Bowie, Denise Herd, Mia Luluquisen, and Mark Woo. "Is the Opening of a Neighborhood Full-service Grocery Store Followed by a Change in the Food Behavior of Residents?" Journal of Hunger and Environmental Nutrition 21.1 (2007): 3-18.

White, M. "Food Access and Obesity." Obesity Reviews 8.Suppl. 1 (2007): 99-107.

Zachary, Drew, Anne Palmer, Sarah Beckham, and Pamela Surkan. "A Framework for

Understanding Grocery Purchasing in a Low-Income Urban Environment." Qualitative Health Research 23.5 (2013): 665-78.

Table 1. Summary Statistics of Selected Household and Individual Level Variables: Sample Weighted Mean (Sample Weighted Standard Deviations)

	Full Sample	Rural (N=1,264)		Urban (N=3,404)	
Variable Name	**(N=4,668)**	**Primary=Nearest**	**Primary≠Nearest**	**Primary=Nearest**	**Primary≠Nearest**
Household Level Characteristics					
Household size	2.4(0.05)	2.2(0.09)	2.4(0.07)*	2.4(0.10)	2.5(0.07)
Children presence (%)	33.1(1.42)	28.3(4.41)	33.4(2.73)	31.0(2.42)	35.3(1.81)
Household gross income ($1,000/month)	5.1(0.27)	5.21(0.54)	4.52(0.23)	5.11(0.51)	5.28(0.46)
SNAP participation (%)	13.1(1.05)	10.7(1.77)	9.9(1.84)	15.1(2.4)	14.3(1.49)
Own/Lease car (%)	89.6(1.22)	91.7(1.89)	94.4(0.77)	87.6(2.2)	87.2(1.82)
Own home (%)	62.0(2.84)	77.3(4.35)	78.2(4.65)	54.2(3.79)	52.8(3.37)
Primary Respondent Characteristics					
Female (%)	67.4(1.27)	59.8(2.96)	75.9(3.84)**	64.7(2.89)	66.7(1.83)
African American (%)	12.5(2.06)	3.6(1.42)	5.7(1.96)	16.4(3.25)	16.9(2.59)
Caucasian (%)	80.1(2.35)	94.1(1.82)	92.0(2.2)	75.0(3.26)	72.2(2.87)
College and More (%)	31.7(2.16)	32.1(5.28)	25.9(3.68)	36.5(3.74)	31.2(2.46)
Less than high school (%)	8.0(0.89)	10.3(2.37)	6.0(1.86)	7.2(1.16)	9.5(1.22)
Married (%)	44.8(1.61)	49.7(3.4)	50.9(2.84)	40.6(3.40)	41.4(2.07)
Household Level Food Expenditures					
At-home-food ($/week)	108.42(3.38)	95.69(7.05)	114.54(5.56)***	101.63(7.45)	109.68(4.46)
Away-from-home-food ($/week)	60.07(2.88)	53.49(4.43)	53.15(3.96)	66.27(6.46)	61.57(3.61)
Fruit and Vegetables (FV) ($/week)	11.66(0.57)	8.5(0.87)	11.28(0.86)**	11.67(1.10)	12.54(0.71)
FAH/total food expenditure (%)	63.1(0.98)	63.2(2.77)	66.5(1.69)	58.2(2.34)	63.1(1.31)
FAH from primary store/total FAH (%)	43.7(1.51)	57.4(3.35)	42.3(2.67)**	48.6(3.09)	40.0(1.96)**
FV from primary store/total FV (%)	50.2(1.77)	68.4(3.47)	49.6(4.57)**	55.4(3.04)	45.2(2.56)**

Note: * denotes 5% significance level, ** denotes 1% significance level, *** denotes 0.1% significance level. The tests were t-test comparing between primary store is nearest store subgroup with primary store is not nearest store subgroup within rural and urban subsamples.

Table 2. Summary Statistics of Selected Food Environment Variables: Sample Weighted Mean (Sample Weighted Standard Deviations)

	Full Sample	Rural		Urban	
Variable Name		**Primary= Nearest**	**Primary≠ Nearest**	**Primary= Nearest**	**Primary≠ Nearest**
Primary store type					
Supermarket (%)	47.1(4.15)	56.0(6.57)	31.5(7.04)***	60.5(6.19)	45.4(4.38)**
Superstore (%)	44.3(3.78)	39.5(7.32)	54.4(5.23)	38.7(6.09)	44.7(4.71)
Other (%)	8.6(1.42)	4.5(1.71)	14.2(4.75)	0.8(0.47)	9.9(1.55)***
Primary=Nearest (%)	30.4(1.86)	33.9(3.31)		28.8(1.96)	
Reasons for choosing primary store					
Good prices (%)	52.6(2.05)	52.1(8.1)	55.3(5.49)	37.4(2.82)	58.0(2.63)***
Good quality (%)	20.5(1.59)	12.8(4.12)	20.1(3.39)	18.3(2.73)	21.9(2.09)
Close to home (%)	53.3(1.95)	63.8(4.7)	40.8(3.21)***	74.7(2.76)	45.6(2.67)***
Distance to primary store (mile)	3.8(0.65)	4.4(0.84)	8.6(1.39)***	0.9(0.09)	2.2(0.13)***
Distance to nearest store (same type) (mile)	2.5(0.44)	4.4(0.80)	5.6(0.98)***	0.9(0.12)	1.0(0.05)
Δ in Distance (primary vs. nearest) (mile)	1.2(0.23)	-	2.9(0.57)	-	1.1(0.11)
Counts within 1-mile radius					
Supermarket and Superstore	1.5(0.20)	0.3(0.08)	0.2(0.08)	1.8(0.19)	2.2(0.25)
Grocery store	3.0(0.52)	0.6(0.14)	0.5(0.16)	3.7(0.54)	4.4(0.67)
Convenience store	3.9(0.66)	0.3(0.09)	0.4(0.07)	4.2(0.67)	6.3(1.01)***
Fast food restaurant	5.3(0.61)	0.8(0.23)	0.7(0.26)	7.3(0.83)	7.8(0.72)
Non-fast-food restaurant	25.4(4.45)	3.0(0.83)	2.0(0.55)	33.7(6.4)	38.4(6.32)
Counts within 10-mile radius					
Supermarket and Superstore	63.0(12.66)	7.2(1.85)	6.9(1.4)	78.1(13.6)	89.6(14.95)
Grocery store	130.4(36.6)	15.4(4.3)	12.9(2.70)	159.3(40.08)	179.9(46.45)
Convenience store	141.8(27.54)	15.2(3.94)	14.5(3.02)	172.9(36.4)	209.1(33.33)
Fast food restaurant	131.5(20.45)	23.0(7.55)	17.9(4.04)	165.6(23.99)	194.3(23.86)
Non-fast-food restaurant	575.1(95.78)	79.2(24.01)	63.8(15.29)	737.2(121.60)	832.2(104.97)

Note: * denotes 5% significance level, ** denotes 1% significance level, *** denotes 0.1% significance level. The tests were t-test comparing between primary store is nearest store subgroup with primary store is not nearest store subgroup within rural and urban subsamples.

Table 3. Household Weekly At-Home-Food Expenditure (in log form) Demand Models: Coefficient (Probability Weight Robust Standard Errors)

Variable Name	Rural			Urban		
	Model 1	Model 2		Model 1	Model 2	
		Primary= Nearest	Primary≠ Nearest		Primary= Nearest	Primary≠ Nearest
Accessibility (mile)						
Distance to nearest store	0.008(0.01)			0.027(0.06)		
Distance to primary store		0.007(0.03)	-0.004(0.01)		0.080(0.081)	0.016(0.01)
Availability (count)#						
Supermarket/Superstore	0.006(0.01)	-0.047(0.06)	0.006(0.01)	-0.008(0.02)	-0.048(0.04)	-0.004(0.003)
Grocery store	-0.002(0.003)	0.032(0.02)	-0.006(0.004)	0.009(0.01)	0.02(0.02)	0.00004(0.001)
Convenience store	-0.008(0.004)*	0.016(0.01)	0.002(0.004)	-0.009(0.004)*	-0.031(0.01)**	0.002(0.001)*
Fast-food restaurant	0.004(0.003)	-0.025(0.02)	0.002(0.004)	0.004(0.006)	0.023(0.01)*	-0.001(0.001)
Non-fast food restaurant	0.0002(0.001)	0.006(0.01)	-0.0002(0.001)	0.001(0.001)	0.002(0.002)	-0.0001(0.0002)
Household Characteristics						
Household size	0.151(0.04)***	0.097(0.07)	0.134(0.07)*	0.195(0.02)***	0.152(0.05)**	0.192(0.03)***
Income ($1,000/month)	0.026(0.01)*	0.027(0.02)	0.046(0.02)*	0.024(0.01)***	0.037(0.01)***	0.023(0.01)***
At lease one child	-0.004(0.10)	0.125(0.20)	-0.037(0.19)	-0.035(0.07)	0.228(0.15)	-0.076(0.10)
Own/lease car	-0.138(0.13)	-0.170(0.26)	-0.067(0.15)	0.221(0.07)**	0.164(0.15)	0.229(0.11)*
Participate in SNAP	0.008(0.09)	0.053(0.18)	-0.097(0.16)	-0.067(0.07)	-0.045(0.14)	-0.048(0.10)
Primary Respondent Characteristics						
Female	0.378(0.11)**	0.193(0.16)	0.485(0.13)***	0.167(0.12)	0.131(0.13)	0.101(0.09)
Caucasian	-0.299(0.21)	-0.058(0.38)	-0.680(0.18)***	-0.524(0.24)*	0.095(0.13)	0.124(0.08)
African American	-0.511(0.25)*	-0.772(0.44)	-0.743(0.25)**	-0.826(0.36)*	-0.406(0.22)	-0.475(0.12)***
College Education or more	0.066(0.11)	0.182(0.19)	0.149(0.14)	0.392(0.11)**	-0.062(0.13)	0.209(0.08)*
Holiday	0.221(0.16)	0.307(0.23)	0.142(0.26)	0.121(0.18)	-0.016(0.14)	0.197(0.14)

Note: * denotes 5% significance level, ** denotes 1% significance level, *** denotes 0.1% significance level. #: for Model 1, the radius is 1-mile for urban and 10-mile for rural; for Model 2, the radius is the mile that is closest to the primary store distance to the home. Model 1: the naïve model uses nearest store distance as accessibility and constant 1-mile for urban and 10-mile for rural for accessibility radius. Model 2: the continuous expenditure equation in the endogenous switching regression model after considering primary store choices.

Table 4. Household Weekly Fruit and Vegetable Expenditure (in log form) Demand Models: Coefficient (Probability Weight Robust Standard Errors)

Variable Name	Rural			Urban		
	Model 1	Model 2		Model 1	Model 2	
		Primary= Nearest	Primary≠ Nearest		Primary= Nearest	Primary≠ Nearest
Accessibility (mile)						
Distance to nearest store	0.009(0.01)			-0.029(0.08)		
Distance to primary store		0.048(0.03)	0.031(0.02)		0.219(0.11)*	0.002(0.02)
Availability (count)#						
Supermarket/Superstore	0.025(0.01)*	0.053(0.06)	0.028(0.02)	0.004(0.03)	-0.042(0.04)	0.002(0.004)
Grocery store	-0.003(0.01)	0.-010(0.02)	0.0005(0.01)	0.009(0.01)	0.010(0.02)	-0.001(0.001)
Convenience store	-0.018(0.01)*	-0.022(0.02)	-0.013(0.01)	-0.006(0.01)	-0.042(0.02)*	0.001(0.001)
Fast-food restaurant	0.003(0.01)	-0.005(0.02)	0.003(0.01)	-0.015(0.01)	-0.007(0.01)	-0.001(0.001)
Non-fast food restaurant	0.0003(0.001)	0.005(0.01)	-0.001(0.002)	0.001(0.001)	0.006(0.002)**	0.0001(0.0004)
Household Characteristics						
Household size	0.171(0.09)	0.174(0.16)	0.131(0.10)	0.118(0.03)***	0.151(0.05)**	0.114(0.03)***
Income ($1,000/month)	-0.004(0.02)	-0.026(0.03)	0.001(0.03)	0.025(0.004)***	0.031(0.01)**	0.026(0.006)***
At lease one child	-0.419(0.27)	-0.387(0.47)	-0.418(0.32)	-0.060(0.10)	-0.100(0.15)	-0.105(0.12)
Own/lease car	-0.005(0.21)	-0.227(0.45)	0.018(0.30)	0.214(0.10)*	0.148(0.15)	0.250(0.13)
Participate in SNAP	-0.145(0.14)	-0.249(0.27)	-0.289(0.18)	-0.121(0.08)	-0.253(0.16)	-0.016(0.11)
Primary Respondent Characteristics						
Female	0.167(0.12)	0.155(0.21)	0.221(0.20)	0.084(0.09)	0.423(0.14)**	0.052(0.11)
Caucasian	-0.524(0.24)*	-0.207(0.78)	-0.698(0.28)*	-0.123(0.09)	-0.124(0.16)	-0.043(0.13)
African American	-0.826(0.36)*	-0.855(0.88)	-0.857(0.36)*	-0.366(0.14)*	0.155(0.20)	-0.490(0.18)**
College Education or more	0.392(0.11)**	0.426(0.22)*	0.624(0.18)***	0.376(0.09)***	0.290(0.13)*	0.351(0.11)**
Holiday	0.121(0.18)	0.151(0.36)	-0.105(0.31)	0.241(0.14)	0.424(0.19)*	0.146(0.17)

Note: * denotes 5% significance level, ** denotes 1% significance level, *** denotes 0.1% significance level. #: for Model 1, the radius is 1-mile for urban and 10-mile for rural; for Model 2, the radius is the mile that is closest to the primary store distance to the home. Model 1: the naïve model uses nearest store distance as accessibility and constant 1-mile for urban and 10-mile for rural for accessibility radius. Model 2: the continuous expenditure equation in the endogenous switching regression model after considering primary store choices.

Table 5. Primary Store Selection Equation Results: Coefficient (Robust Standard Errors)

	At-Home-Food Expenditure		Fruit and Vegetable Expenditure	
Variable Name	**Rural**	**Urban**	**Rural**	**Urban**
Difference in distance (Primary vs. Nearest)	-0.268(0.08)***	-0.714(0.21)**	-0.261(0.08)**	-0.691(0.24)**
Primary Store Choice Reasons				
Good prices	-0.017(0.17)	-0.285(0.09)**	-0.148(0.23)	-0.304(0.12)*
Good quality	-0.251(0.24)	-0.076(0.12)	-0.130(0.27)	-0.260(0.15)
Close to home	0.436(0.18)*	0.242(0.09)*	0.541(0.28)	0.374(0.13)**
Local Food Environment: 10-mile availability				
Supermarket/Superstore	-0.037(0.02)*	0.001(0.002)	-0.041(0.02)	-0.001(0.002)
Grocery store	0.011(0.01)	-0.0003(0.0005)	0.006(0.01)	-0.00002(0.001)
Convenience store	-0.001(0.01)	-0.0001(0.0004)	-0.008(0.01)	0.0001(0.001)
Fast-food restaurant	0.002(0.01)	-0.0005(0.001)	0.005(0.01)	-0.001(0.001)
Non-fast food restaurant	0.001(0.002)	0.0001(0.0001)	0.003(0.002)	0.0002(0.0002)
Household Level Characteristics				
Household size	-0.067(0.08)	0.049(0.04)	-0.091(0.09)	0.020(0.04)
Income ($1,000/month)	0.050(0.02)*	-0.009(0.009)	0.056(0.03)*	-0.011(0.01)
At lease one child	-0.05(0.23)	-0.171(0.12)	0.172(0.26)	0.004(0.15)
Own/lease car	-0.045(0.27)	-0.062(0.11)	0.025(0.32)	0.051(0.14)
Participate in SNAP	0.307(0.20)	0.197(0.10)*	0.471(0.24)	0.163(0.12)
Primary Respondent Characteristics				
Female	-0.443(0.18)*	-0.097(0.11)	-0.488(0.21)*	-0.196(0.14)
Caucasian	0.11(0.34)	0.117(0.12)	0.143(0.38)	0.180(0.14)
African American	-0.200(0.41)	0.102(0.16)	-0.309(0.44	0.047(0.20)
College Education or more	-0.013(0.19)	0.219(0.11)*	0.018(0.21)	0.123(0.13)
Model Statistics				
Correlation coefficient (Selection_PNN)	0.018(0.30)	-0.052(0.10)	-0.312(0.55)	0.043(0.10)
Correlation coefficient (Selection_PN)	0.237(0.24)	-0.813(0.10)*	0.131(0.28)	-0.008(0.32)

Note: * denotes 5% significance level, ** denotes 1% significance level, *** denotes 0.1% significance level.