PREDICTIVE FOOD DESERT SIMULATION MODELING TO INCREASE FOOD ACCESS IN UNDERSERVED COMMUNITIES

Abhilash C. Singh
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Tel: 512-471-4535; Email: acsingh@utexas.edu

Kelsey C. Abel
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
E-mail: kelsey.abel@utexas.edu

Joseph W. Hutchinson
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Tel: 512-471-4535; Email: jhutchinson@utexas.edu

Kasey M. Faust
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
E-mail: faustk@utexas.edu

and
Chandra R. Bhat (corresponding author)
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Tel: 512-471-4535; Email: bhat@mail.utexas.edu
and
The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

December 2018
Keywords: Food Access, Food Deserts, Location-choice model.

1. INTRODUCTION

Land use consolidation in United States had a profound impact on food-shopping activities. A small number of large grocery stores that offered a wide range of products and facilitated *one-stop shopping* replaced the plethora of small specialty-markets serving single neighborhoods. While the selection offered in these stores meant fewer total trips for shoppers, farther distances needed to be traveled to purchase groceries. The shift in scenario diminished the ability to access healthy food for those without sufficient mobility. This inevitably led to the deterioration of *urban food environments*, defined as the prevalence and proximity of grocery stores in a given neighborhood. Poor urban food environments with impeded access to healthy food have been correlated with higher rates of chronic, diet-related diseases and poor academic performance of school-age children (Shonkoff, 2000; WHO, 2003; UCLA, 2008). The effects of this transformation are critical to analyze, as studies show that one's food environment plays an important role in determining one's overall health (Baker, 2006; UCLA, 2008; Morland *et al.*, 2002; Morland *et al.*, 2006; WHO, 2003).

Inarguably, the deterioration of urban food environments tend to disproportionately affect low-income populations. Many studies (WHO, 2003; Morland *et al.*, 2006; Dixon *et al.*, 2007) attribute increased incidence of diet-related diseases to the consumption of over-processed and sugary foods, which were found to outnumber healthy food choices in low-income communities. Several studies (Algert *et al.*, 2006; Dai and Wang, 2011; Eckert and Vojnovic, 2017; Galvez *et al.*, 2008; Helling and Sawicki, 2003) have found that the prevalence of stores carrying healthy food options is lower in low-income neighborhoods than higher-income neighborhoods, leaving low-income residents with fewer options to purchase food.

The study of inaccessibility of healthy food options in low-income neighborhoods and its negative impacts has contributed to the concept of food desert (FD), defined by Cummins and Macintyre (2002) as "poor urban areas, where residents cannot buy affordable, healthy food." The USDA (2017) defines FDs more precisely as low-income census tracts in which either 500 people or 33 percent of the population reside more than one mile (measured radially) from the nearest supermarket or large grocery store. According to the data released by USDA, 6,529 census tracts (or approximately 9% of all U.S. census tracts) qualify as FDs, consisting of an estimated 19 million people (roughly 6% of the U.S. population). Remarkably, three fourths of which are urban, while, the rest are rural. These statistics by the USDA are conservative compared to a study by Ver Ploeg (2012), which suggests that nearly 10% of the U.S. population (29.7 million people) lives in low-income areas more than one mile from a supermarket. This discrepancy indicates that the USDA's definition of a FD may be inadequate and in need of further study.

The definition of FDs is unrealistic for multiple reasons. First, the classification of a census tract as low-income is not clear. Standard measures to define a low-income census tract classify it as such when (1) 20% or more of its residents are below the federal poverty line (2) if the census tract's median household income is less than or equal to 80% of the median state-wide household income, or (3) if the census tract is in a metro area and has a median income less than or equal to 80% of the metropolitan area's median income. Second, even when properly defined, income is not accurately reported by both high- and low-income individuals, resulting in unreliable data. Third, the definition seems to imply that most people shop for groceries within the census tract where they live, but in reality, fewer than one in five do (Aggarwal *et al*., 2014). Fourth, this definition is based strictly on geographic proximity, ignoring other key factors that contribute to food access, such as vehicle ownership. Fifth, the distance used to determine access is measured radially rather than through-the-network. Radial distances can be substantially smaller than through-the-network distances, and hence, contribute towards an overly optimistic measurement of the extent of FDs. The impact of mismatch between radial and through-the-network distances has been demonstrated in prior studies, including U.S. Veteran's ability to access VA hospitals (CBS, 2015).

Because of the limitations in the USDA FD definition, many studies, including this one, have sought alternative metrics that better capture the diminished food-accessibility. In addition to geographic proximity, food accessibility metrics have included the measures of physical and economic barriers, the quality of food available to purchase in a given region, the amount of time that a time-impoverished person allocates to grocery shopping, and the affordability of healthy foods normalized for the socio-economic condition in the community. Studying low-income neighborhoods may lend a more nuanced look into regions classified as FDs.

This study employs a location-choice model and proposes a metric of food accessibility to refine the definition of a FD. In addition to identifying a metric that correlates with food insecurity, the results of this study will uncover how lack of transportation is a barrier to accessing food for low-income residents who do not own a car or cannot afford the cost of taking public transportation. Rather than the traditional way of measuring at an aggregated level of a census tract, we are applying the concept of FD at the most disaggregate level possible – by analyzing food shopping trips for the each individual. By identifying the factors that impact the location of food shopping for low income individuals, this study will inform decision makers and researchers about potential ways to isolate and investigate food-shopping behavior of low-income populations, and to help tailor food assistance programs to better fit the needs of low-income residents in FDs.

2. LITERATURE REVIEW

Traditionally, studies (such as Chung and Myers, 1999; Eckert and Shetty, 2011; Giang *et al.*, 2008; Wrigley, 2002; Wrigley *et al.*, 2003; Widener *et al.*, 2011; Widener *et al.*, 2013; Wilkins, 2017) have looked at how temporal and spatial parameters influence food access, and how policy levers can be used to potentially change food access scenario. Additionally, the distance from residential locations is employed to represent food accessibility (Chaix, 2009; Lee and Lim, 2009; Bertrand *et al.*, 2008), while neglecting people's mobility behaviors. Furthermore, most previous studies do not consider the potential impact of people's time-budget and mode of transportation on food accessibility. However, recent efforts (Auchincloss *et al.*, 2011; LeDoux and Vojnovic, 2013 & 2014; Rose, 2011; Vojnovic *et al.*, 2013) have looked beyond assessing proximity to food retail opportunities by incorporating data on actual travel behavior to understand how residents navigate these food landscapes. In reality, people travel beyond their residential neighborhoods in their daily activities. Specifically, adolescents are likely to visit food stores beyond their residential neighborhoods (Shearer *et al.*, 2015). To overcome the inadequacies of the food accessibility measurements, it is necessary to consider mobility and time as limiting resources (Kwan, 2013; Widener and Shannon, 2014).

Preliminary work on the FD definition that takes socio-demographic factors into consideration confirms the potential of the USDA definition to overestimate food access levels (Raj and Lentz, 2016) and potentially impede access (Abel and Faust, 2017; Ver Ploeg *et.al.*, 2014). Studies, such as Walker *et al*. (2010), draw attention to individuals' health conditions and demographic and socio-economic characteristics, while analyzing diminishing access to food stores. It is also evident that there are racial, socio-economic, and rural/urban disparities, with food deserts being more commonly located in predominantly African-American, low-income, and rural communities (Walker *et al.*, 2010; McKinnon *et al.*, 2009; Sallis *et al.*, 2009). Theoretically, high-income census tracts can also be geographic food deserts, however, high-income households typically have sufficient mobility to reach stores or can afford alternatives such as home-delivery of food and groceries (Eckert and Vojnovic, 2017). However, there have been inconsistent results in previous studies. Such as the relationship between income and access to healthy foods has been found to be positive (Morland *et al.* 2006; Larsen and Gilliland 2008), null (Bertrand *et.al.* 2008), or negative (Black *et al.* 2011). These mixed results might be caused by differences in the scale and measurement approach between these studies.

An interesting study done by Shannon and Christian (2017) uses network distance from home to all visited food sources to understand food-shopping behavior. The main highlights of their findings are: the residential food environment plays a limited role in households' food provisioning practices; most individuals travelled significantly farther than the nearest grocery store to shop for food; automobile use was significantly lower and use of both transit and walking were more frequent for households with incomes under \$25,000; both travel and stay times on food related tours were longest for households with incomes under \$25,000 and shortest for higher income groups, suggesting that the constraints of lower incomes—such as limited access to reliable transportation, thereby increasing the time and opportunity costs of food shopping. These results do show that interventions to improve healthy food options in low-income areas must look beyond the immediate residential neighborhood. Determining what drives or impedes residents' store selection and frequency of food shopping trips for low-income households is the first step toward creating equitable access to food.

This study seeks to develop an improved metric of food accessibility that can help us better understand food accessibility issues at the individual level. We propose to do so by developing a model of food-shopping location for low-income individuals. The National Household Travel Survey (NHTS, 2017) in the NCTCOG Region provides information on the trip destination at a Traffic Survey Zone (TSZ) level and we extract the food-shopping trips by cross-referencing trips where the destination zone contained a food shopping location with trips where the purpose is shopping. The socio-economic information in the NHTS dataset includes information allowing for the examination of households in low and high-density census tracts. This data will then be supplemented with land-use and transportation network GIS data for the region. Based on this information, a location-choice model will be developed at the most disaggregate spatial level using the Dallas-Fort Worth region sub-sample. Using the revealed choice of low-income individuals' food-shopping locations, the influence of demographic, mobility, and food-environment factors will be teased out. The factors of importance identified in this process will be considered in terms of their tempering or enhancing effects on the attractiveness of food shopping locations. The authors propose the use of available food shopping locations at a zonal level as a part of the consideration set for alternative destination choices to estimate a location-choice model framework. Bhat *et al.* (2003) describe the mathematical procedure used to apply the spatial location-choice model. The methodology employs a choice-set generation method in the determination of the candidate locations for the stop. The exogenous variables for this model framework include socio-economic factors, demographic factors and built environment attributes. Subsequently, a Multinomial Logit (MNL) model is estimated to predict the spatial location-choice among the candidate locations in the choice-set. Finally, a gravity accessibility measure based on utility theory is suggested for characterizing accessibility in low-income households. This metric will fill a gap by bringing together the factors that either provide attraction towards shopping destinations or impede the access to them.

3. DATA DESCRIPTION

The following sub-sections describe data sources, the process of cleaning the data and how the data was used for estimation purposes.

3.1 Data Sources

The activity-travel data used for this study is the 2017 NHTS data from the North Central Texas Council of Governments (NCTCOG) complete with location data at the TSZ-level. TSZs are microanalysis zones with boundaries defined by NCTCOG for traffic analysis studies in the Dallas-Fort Worth-Arlington region. The study area consists of 5,252 TSZs for which variables representing attractiveness of the TSZ (e.g., household, employment, and population density per square mile) have been computed. Other computed measures of attractiveness include the number of different types of developments (such as commercial, industrial, and governmental) whose proximity to food-shopping locations may attract low-income individuals for other activities before or after food shopping. The shopping locations were aggregated by TSZs to characterize zones where food-shopping activities take place. These locations include grocery stores, supercenters, and wholesale stores.

3.2 Data Cleaning and Sample Formation

To ensure that the data was suitable for modeling, we removed the observations missing the following information: age, gender, driver status, relationship to survey respondent, Hispanic status, race, worker status, full-time or part time worker status, occupation type, income, trip mileage and educational achievement. The cleaning process yielded 475 food-shopping locations, located in 409 TSZs in the study area. This includes 309 grocery stores (example Krogers, H-E-B, Whole Foods), 136 supercenters (example Walmart and Target), and 30 wholesale stores (example Sam's Club and Costco).

We identified low-income households in the 2017 NHTS DFW-Arlington sub-sample using the guidelines of the United States Department of Agriculture's Supplemental Nutrition Assistance Program (USDA SNAP) income eligibility limits. We classified households as *low-income* that were just above the USDA SNAP income eligibility limits to capture households that may be considered low-income given a large household size. However, the distribution of household income from the DFW-Arlington sub-sample of the 2017 NHTS compared to the breakdown from the 2016 American Community Survey (ACS) presents a few discrepancies. In the 2017 NHTS DFW sample, 33.7% of households report making less than $50,000 per annum compared to 41.7% making less than $50,000 per annum in the 2016 ACS. The sample-share of low-income households may be under reported in the DFW-Arlington sub-sample of the 2017 NHTS, as compared to the actual share of low-income households in the region. The final sample for model estimation consisted of all the low-income households' trips that were classified in the NHTS dataset as "shopping/errands" trips whose destination was within the TSZs with known food shopping locations. This process yielded a final sample of 1,005 food-shopping trips made by 683 low-income households.

3.3 Data Description

Table 1 provides the average household (HH) characteristics for low and medium/high income households who made food-shopping trips in the DFW subsample of the 2017 NHTS. Medium/high income households on average take 1.74 more household trips per day than low-income households. Fewer trips in a day might indicate fewer opportunities to shop for food for low-income individuals. Population density per square mile and the percentage of renter housing are lower in Census Block Groups (CBG) where wealthier households reside. Low-income households are more likely to be African-American than medium/high income households, as evident by the fact that African-Americans constitute 8.2% of medium/high income households, as compared to 19.0% of low-income households. Comparatively, 84.7% of medium/high income households are Caucasian, while 69.4% of low-income households are Caucasian. The difference in the proportion of African-American households between low and medium/high income categories indicates that household race may have an indirect impact on food accessibility. Similarly, 12.0% of low-income households are Hispanic compared to 8.0% of medium/high income households.

The characteristics of food-shopping locations, frequented by individuals from low-income households, might have an impact on food shopping location choices. These include the number of wholesale stores, supercenters and restaurants in the location where an individual shops for food. Possibility of adding another stop for an activity near the food-shopping destination may add to the utility of shopping at any particular destination. The average distance between a low-income household's home TSZ and the TSZ of the food shopping location is 6.64 miles. For the low-income food shopping trips, the mode-share by car, SUV, truck, or van, is 93.5%. When measuring food accessibility, the mode of travel may not be a reliable indicator of food access, because almost all low-income individuals in DFW rely on private vehicles for their food shopping trips. However, increasing food accessibility for low-income individuals in a sustainable way may require improving public transportation options, which has a mode-share of less than 3% of food-shopping occasions in our sample.

4. METHODOLOGY

In this section, the methodology employed for estimation of the location-choice model is described, followed by the formulation of the metric used to identify attractiveness of the food-shopping destinations.

4.1 Choice-set generation

Bhat *et al.* (2003) describe the mathematical procedure used to apply the spatial location-choice model. The methodology employs a probabilistic choice-set generation method that uses the travel distance to the destination TSZ (from the origin TSZ) in the determination of the candidate locations. Subsequently, MNL prediction procedure is used to predict the spatial location-choice among the candidate locations in the choice-set. At this stage, the utilities of the alternatives in the choice-set are compared directly with each other in a utility maximizing process. The difference in the process at the choice-set generation and choice stages enables an attribute associated with a choice to have two separate effects: a consideration effect (*i.e.*, the impact on the consideration set of locations) and a choice effect (*i.e.*, the impact on the choice of a location, given that the location is considered by the individual).

It was found that the probabilistic choice-set generation method was giving rise to unreasonably far (from the origin zone) spatial location-choice considerations. Hence, a deterministic choice-set generation method was utilized to ensure the spatial consistency of the choice set (or consideration set). The deterministic choice-set generation method uses the travel distance to the destination TSZ (from the origin TSZ) in the determination of the candidate locations for the activity. The rationale behind using the travel distance to the destination TSZ in generating the location-choice-set is that the TSZ location to be predicted should be within a certain range of the given travel distance of that trip. Hence, the location-choice-set for destination TSZ consists of the zones that fall within a certain range of travel distances from the origin TSZ. If the travel distance is greater than the intra-zonal travel distance, then half of the candidate zones selected into the location-choice-set have shorter travel distances (from the origin TSZ) than the travel distance, while the other half have travel distances greater than or equal to the travel distance. The detailed step-by-step process to formulate the choice set is described below.

The steps involved in the deterministic choice-set generation are:

a. If the travel distance is greater than the intra-zonal travel distance, follow the steps below:
   i. Arrange all the zonal locations in the ascending order of travel distance from the previous stop.
   ii. Select the first spatial zone *Z*, whose travel distance from the origin TSZ (d_z) is greater than (or equal to) the travel distance.
   iii. Select twenty-five zones with travel distance (from the origin TSZ) less than d_z and twenty-five zones with travel distance greater than d_z. If twenty-five zones are not available on one or both sides of d_z, select the minimum number of zones available on both sides to maintain symmetry of travel distances of the candidate zones in the choice-set.

b. If the travel distance is less than the intra-zonal travel distance from the origin TSZ, then in addition to the chosen alternative, select five zones with travel distance greater than d_z as candidate zones in the choice-set. There is also a dummy variable created based on whether the destination TSZ is the same as origin TSZ.

4.2 Estimating the spatial location-choice

Since zonal size attributes indirectly represent the elemental alternatives within a zone, it is important to introduce them in a form that ensures that a large zone will have a higher probability of being chosen than a small zone. For this purpose, a non-linear composite size measure is introduced, which is defined as follows (introduction of the composite size term is similar to Pozsgay and Bhat, 2002). A nonlinear-in-parameters MNL (NLMNL) model with the following utility expression was estimated and the best specification was identified for the composite size term:

$$\textit{Composite Size} = \alpha * distance + \beta * \ln(size - attributes) + \varepsilon$$

where, distance is measured from origin zone of the trip and natural log of size attributes is introduced, normalized using the overall area of the TAZ for which the composite size is estimated.

There are several zonal size attributes in the assembled dataset that capture the attractiveness of a zone. These include: (i) zonal area, (ii) number of wholesale stores, (iii) number of supercenters, (iv) number of grocery stores, (v) number of commercial/industrial developments, (vi) number of residential developments, (vii) number of educational developments, (viii) number of service developments, (ix) number of government developments, (x) number of transportation developments, (xi) number of entertainment developments, and (xii) number of retail developments. Using this composite size term as an exogenous variable, a Multinomial Logit (MNL) model was estimated for the location-choice. The following steps were performed to estimate the chosen alternative:

a) Compute the conditional probability ($P_1, P_2 \ldots P_K$) for each of the different K *(K = 50 or less)* candidate locations using the calibrated model parameters and the values of exogenous variables specific to the decision maker under consideration.
b) Generate a uniformly distributed random number (U) between 0 and 1.
c) The chosen alternative is determined using the computed choice probabilities and the uniform random number drawn as follows:
    i. If $0 <= U < P_1$, chosen alternative is A_1.
    ii. If $P_1 <= U < P_1+P_2$, chosen alternative is A_2.
    iii. If $P_1+P_2+\ldots P_{J-1} <= U < P_1+P_2+\ldots P_J$, chosen alternative is A_J.
    iv. If $P_1+P_2+\ldots P_{K-1} <= U < 1$, chosen alternative is A_K.

5. ESTIMATION RESULTS

The subsequent sub-sections provide results for the NLMNL estimation of the composite size variable and the MNL estimation of the location-choice using the composite size as an exogenous variable.

5.1 Composite Size Variable Calculation results

We provide the *non-adjusted* and the *adjusted* composite-size measure estimation results in Table 2 and 3. The *non-adjusted* estimation results (Table 2) are based on the fact that "composite-size" is not an observable and thus for its computation, we included the "area of alternative zone" variable, and we held the parameter on this variable constant in our estimation. Additionally, it is econometrically impossible to identify the composite size variable without holding a size-dependent variable constant, and therefore, area of alternative zone was held as equal to one. This non-linear in parameters result is *adjusted* in the final step of the composite-size estimation to reflect the true parameters which are utilized to calculate the composite-size variable. The best specification for the adjusted composite-size measure (Table 3) includes the number of recreational developments and the number of retail developments in the zone.

Results indicate that a zone with a greater number of retail and recreational opportunities is assigned a larger composite size. The other zonal size measures did not turn out to be significant in this specification, largely due to the correlations between these size measures. The coefficient on the composite size variable is significantly smaller than one, indicating that there are unobserved zonal attributes affecting the utility of elemental recreational and retail destinations within the zone. The number of retail and recreational developments have positive relationship with the food shopping location, indicating that shoppers will go to areas with a variety of stores, perhaps to make convenient stops at other shops and add their various activities as a part of one tour. In the MNL model subsequently estimated in this study, the composite size computed with the parameters identified in Table 3 is used as the only zonal size measure.

5.2 MNL Location-choice model

The final specification result of the home-based food-shopping location-choice model is presented in Table 3. The parameter signs on the variables are as expected. Among the zonal attributes, the coefficient on the composite size variable is positive, indicating that a larger *composite-size* zone is preferred for food-shopping choice by low-income households. The coefficients on the numbers of wholesale stores and supercenters are positive, indicating a high positive correlation between the choice of location and the number of shopping opportunities it provides. The number of grocery stores does not have a significant impact on the attractiveness of an alternative zone. A negative coefficient on population density indicates that zones with lower density are more attractive for food shopping. This could also indicate self-selection, as food-stores are more likely to have a large land take and reduce the area where housing can be built in a zone. A negative coefficient on distance from home indicates that the locations closer to home are preferred for food shopping. This inference is also supported by the positive parameters on the variables indicating food-shopping location was in an adjacent zone or the same zone as the home zone. The choice to shop for food at a location close to home is very attractive for low-income individuals. This is contrasting to the results provided by Aggarwal *et al*., 2014 where the authors say that fewer than one in five individuals shop in the census tract they live in. This difference could be attributed to the fact that this study focusses on low-income households only and employs a distinctive composite-area measure to account for some of the attractiveness of the food-shopping location.

The empirical results provide evidence of indirect effects through significant socio-demographic interactions. The sensitivity to distance is lower for individuals with many cars in the household. The explanation for this effect may be two-fold. First, individuals in low-income households may be willing to travel farther to explore inexpensive locations for food shopping. Second, low-income households may be in areas that are not near food facilities, especially supercenters and wholesale stores. The sensitivity to distance is higher for households with higher number of adults, which is counter-intuitive, as for a higher number of adults in the household, it may not be logistically possible to travel to greater distances if everyone in the household wishes to travel together. The sensitivity to distance is lower for high-density residential location households as they may be willing to travel to farther locations for inexpensive or larger supercenters that are usually not available in high-density residential neighborhoods. This suggests that after controlling for the low-income factor, sensitivity to distance is different for individuals with zero or one car in the household as compared to those with two or more cars.

The sensitivity to distance is lower for households with the presence of a child of less than five years of age, which is counter-intuitive. Although, it may be plausible that the presence of an infant may not be a hurdle to travel farther locations as more often, people choose to travel without young infants. On the other hand, low-income individuals may not be able to afford childcare and may have to take their young child shopping for food with them whether they want to or not. They may be less sensitive to distance in order to find a food shopping location that is safe and comfortable for their young child. For example, they may be willing to travel further away to shop for food at a location that has grocery carts with built-in seats to hold their small child. The sensitivity to distance is lower for full-time employed individuals as employment brings financial stability, which may catalyze longer trips for food shopping. Full-time employment may also make individuals more familiar with areas that are further away from home and closer to their work location or between home and work in an area that they would not frequent if they were not employed. The sensitivity to distance is lower for food-shopping destinations which are neither in the same zone nor in the adjacent zone to the home. This result suggests that once the threshold of residing neighborhood is overcome, individuals are willing travel much larger distances to shop for food. The sensitivity to distance is lower for number of wholesale stores and number of supercenters, reason being that as the number of stores increase in a certain zone, the attractiveness of the zone increases, and an individual can perform multiple activities.

The interaction between distance from an alternative to home zone and composite size indicates that individuals are willing to travel to larger zones that are further away. This reflects more utility for zones that have opportunities besides only stores where food can be purchased. Low-income individuals may be interested in completing other shopping trips for clothes or other goods, or complete recreational activities before shopping for food to maximize the efficiency of their travel-related time use.

5.3 Model fit statistics

Table 5 provides the model fit statistics for the analysis. The log-likelihood value at convergence is -2,515.59 and the log-likelihood value with zero coefficients is -3,931.56. The log-likelihood ratio test value for comparing the attraction-end choice model with the zero-coefficients model is 2,895.35. This value far exceeds the corresponding chi-squared value with 17 degrees of freedom at any reasonable level of significance. Thus, the test rejects the zero-coefficients model in favor of a location-choice model.

6. Food Accessibility Metric

Studies have conceptualized various methods based on spatial access and demographic characteristics at the neighborhood level. Widener *et al*. (2013) employ an interaction potential metric that uses inter-zonal commuting patterns to generate a healthy food accessibility score based on time–space prims. This accessibility measure incorporates aggregate travel patterns such as commuter flows and activity constraints. Farber *et al*. (2014) have also described a similar *Supermarket* Interaction Potential (SMIP) which considers the potential time available to commuters for shopping (interacting) at supermarkets given a time budget, their home location, and their work location. Each of these studies includes a measure of distance from a neighborhood to the nearest grocery store or supermarket, as well as the number of stores within 1000 meters of a neighborhood among the access measures calculated. Therefore, a way to merge the findings and get an inspired metric is to combine the characteristics of population being served with those of the food-shopping locations.

The main characteristics of a "food-access metric" must to include the following:

1. Attributes of the destination alternative must be included to measure attractiveness of a zone. These attributes can include number of retail-opportunity availability, food-shopping destinations, proximity dummy for destination TSZ, and measure of composite size of the destination TSZ.
2. Impedance to access the food-shopping destinations must be included in a tangible form.
3. Metric should provide an objective value to categorize the attractiveness of food shopping locations instead of a relative measure, thereby making it easier for policy analysis.

6.1 Proposed food-accessibility Metric

After studying the gamut of accessibility measures, we propose an accessibility measure based on the combination of the transportation system and land use patterns. Prior research (see, Bhat *et al.*, 2002) suggests that a gravity accessibility measure based on utility theory is best suited for characterizing accessibility in low-income households. This accessibility measure was evaluated using the actual dataset to derive the parameter values (see, Bhat *et al.*, 2001 and 2002 for estimation procedure details). Bhat *et al.* (2002) provide the region-specific default values of these parameters, we can directly use the ones estimated for Dallas-Fort Worth metropolitan region. Accessibility is dimensionless and can only be used in relative comparisons. The equation used to compute accessibility at the most disaggregate level is:

$$Acc_i = \ln\left[\sum_{j=1}^{J}\left(\frac{1}{J}\right)\left(\frac{O_i^{\alpha}}{C_{ji}^{\beta}}\right)\right]$$

In this equation,

α, β = parameters estimated from destination mode choice models for the region under consideration

O_i = sum of all measures of attractiveness for the TSZ *i* [Here, O_i = γ_1 (Number of wholesale stores) + γ_2 (Number of supercentres) + γ_3 (population density) + γ_4 (Composite Size of destination TSZ)]

$C_{ji} = \lambda * Distance$. This is an impedance measure between origin zones *j* and destination zone *i* based on distance (Bhat *et al.*, 2002 provide the region-specific default value of this parameter, we can directly use the ones estimated for Dallas-Fort Worth metropolitan region).

Therefore, for food-shopping purposes of low-income households: $\alpha_{shopping} = 0.2868$, $\beta_{shopping} = 3.0780$, $\lambda_{shopping} = 0.5992$, $\gamma_1 = 0.3590$, $\gamma_2 = 0.5000$, $\gamma_3 = 0.8490$, $\gamma_4 = 0.2501$

6.2 Comparison with distance-based metrics

Distance based metrics are abundant in food access studies. Fan et al. (2009) provide a measure of the presence of grocery stores within 1000 meters of a neighborhood centroid. Earlier studies either include a measure of distance from a neighborhood to the nearest grocery store or supermarket, or the number of grocery stores within 1000 meters of a neighborhood, and shall be called *grocery-store metric* to ease comparison. None of these metrics combine the factors that make a shopping-destination attractive with those that impede the access to those destinations. As we discussed in literature review, more individuals are travelling beyond their own neighborhood to access food stores. Thus, the idea of a measure based on distance of 1000 meters is inadequate. Therefore, the proposed metric fills a gap by bringing together the factors that either provide attraction towards shopping destinations or impede the access to them.

It is assumed that it is desirable for a measure to show variation in the region. There should not be a preponderance of zones of a very high or a very low level. One way to evaluate this aspect of a measure's performance is by looking at a frequency diagram. These frequencies were calculated for both the proposed formulation and the grocery-store metric which considers only the presence or absence of grocery stores within 1,000 meters of the neighborhood. To achieve this, the accessibility values (without the natural logarithm of the formulation) computed for each measure were organized into bins of range less than 1, 1 to 100, 100 to 1,000, more than 1,000. Also, the natural logarithm of the measure (as is given in the formulation) was taken and then the accessibility values computed for each measure were organized into bins of range less than 1, 1 to 10, 10 to 20, more than 20. This is presented in Figures 1 and 2. All the diagrams are presented at the same scale. Since all these results are normalized to a scale of 100 based on the highest and lowest values, it only takes one outlier to skew the normalized values to the other end of the scale.

The figures 1 and 2 give a clear indication that grocery-store metric provide an overly optimistic value to the attractiveness of the TAZs and thus, we must take into account all the factors considered in the proposed formulation to assist policy development for a more realistic approach to the food-access problem. The proposed metric can be modified to include built-environment factors and other measures of attractiveness of a food-shopping location. Taking a closer look at Figure 2 gives a clear picture of preponderance of zones of very low metric values. This skewness for the given low-income households denotes inaccessibility to food-shopping destinations for a large percentage of the households in consideration. Evidence provided by the proposed metric should be taken into consideration for policy development to increase food-access for low-income individuals.

7. CONCLUSION

A change in food-shopping scenario in the United States disproportionately affected the low-income population, thereby diminishing the ability to access healthy food. Long-term effects of such a transformation included higher rates of chronic and diet-related diseases. Adding to less prevalence of grocery stores, time poverty and fewer mobility options compound towards low-income residents' inability to access healthy food by limiting their ability to venture outside their neighborhoods to shop for food. Studies to understand the effect of this transformation coined the term food desert (FD). Not only are the standard definitions of FD conservative, but also ignore key factors that contribute to food access, such as vehicle ownership and network-distance to food-shopping locations. This study aims to uncover the effects of lack of transportation is a barrier to accessing food for low-income individuals.

This study presents a location-choice model to isolate and better understand the food accessibility for low-income households in the DFW-Arlington metro area. The model structure takes the form of a MNL formulation, and introduces a composite size measure, which is non-linear in parameters. Through the revealed choice of low-income individuals in the 2017 NHTS on where they undertake their food shopping from the many available locations around them, we determined the influence of socio-demographic, mobility, and built-environment factors in terms of their tempering or enhancing effects on the attractiveness of food shopping locations. Important implications from our empirical analysis are:

1. The number of retail and recreational developments have positive relationship with the food shopping location, indicating that shoppers will go to areas with other stores, perhaps to make convenient stops at other shops, providing evidence for trip-chaining behavior.
2. Composite Size of the zone should be incorporated in a specification, which captures its endogenous effect location of food shopping.
3. Households may not be influenced in the same way by food store proximity based on a wide range of factors besides household income. Therefore, the interactions of socio-demographic factors with built environment attributes give us an insight of how households value the choice of food shopping locations and how this shapes a low-income household's food environment.
4. Trip decision may not always be influenced by trip related variables such as travel-time.

Our results indicate that low-income households prefer to shop for food at wholesale stores and supercenters. Rather than increasing the access and attractiveness of grocery stores to low-income individuals, their food environments may be best improved if they have better access to supercenters and wholesale stores. Supercenters such as Walmart and wholesale stores such as Costco could enact programs for low-income individuals that give discounts on the total grocery purchase if it includes a high level of healthy foods such as vegetables and fruits. Additionally, stores like Costco could encourage low-income families to choose to shop there by waiving the membership fee or providing membership discounts.

Though we see that low-income families prefer to shop for food at locations closer to home, many low-income families also choose to travel to farther locations. The average distance from home to food shopping location in the current sample is 6.64 miles. This is also evident in the literature as households may be differentially affected by food store proximity (or lack thereof) based on a wide range of factors besides household income, such as access to a vehicle and through the network distances (Glanz *et al*. 2005). Few studies have indicated counter-intuitive results as stores used for food shopping can be two to three times as far from home as the closest supermarket (Ver Ploeg *et al*. 2015). The burden of traveling far for food for low-income families might be lifted if these low-income families are given access to services such as grocery delivery. The sustainable solution to food access would involve the increase of the use of public transportation or active transportation to complete food-shopping trips. However, our descriptive analysis of the sample households indicates that low-income individuals prefer a personal vehicle as compared to public transit for their food-shopping travel. Unless the public transportation system drastically improves its ability to connect low-income individuals with food-shopping locations, the prospect of transit-use for food shopping trips by low-income households does not seem promising.

A key contribution of our study is that it does not measure the overall food accessibility of a geographic region, which under the typical FD definition may include a large share of rich households with excellent food environments. We focus on the *actual* food shopping behavior of low-income households to understand what influences their decisions to undertake food-shopping trips at different locations. The departure from the typical food desert concept leads to a more disaggregate approach that can be applied to *any* low-income household or individual shopping for food. By doing this, we can help food researchers, policymakers, and planners to better formulate solutions to improving the food environment of low-income individuals.

A natural extension of our research is to apply a location-choice framework to the choices that low-income individuals make when purchasing food inside of the store. This may include a metric of the relative healthy choices of the total purchases made at the grocery store as the outcome, with the goal of determining what factors influence the choices to buy healthy or processed foods. Uncovering these effects might help food researchers to better formulate strategies that encourage low-income individuals to purchase healthier food. Food environment research has been dominated by studies aiming to characterize and identify food deserts, often neglecting other harder to measure but important factors such as food cost, convenience, quality, and store accommodations. Public health policy and practice should also turn its attention to these in-store characteristics of the food environment.

8. ACKNOWLEDGEMENT

The authors would like to thank North Central Texas Council of Governments for sharing National Household Travel Survey (NHTS, 2017) in the NCTCOG Region data set. This research was partially supported by the Center for Teaching Old Models New Tricks (TOMNET) (Grant No. 69A3551747116) as well as the Data-Supported Transportation Operations and Planning (DSTOP) Center (Grant No. DTRT13GUTC58), both of which are Tier 1 University Transportation Centers sponsored by the US Department of Transportation. The authors are grateful to Lisa Macias for her help in formatting this document.

9. REFERENCES

1. Abel, K.C.*, and Faust, K.M. (2017). Food Desert Modeling: Increasing Food Access by Understanding and Disrupting the Factors which Contribute to Food Desertification. Graduate and Industry Networking Event, University of Texas at Austin, Austin, TX, January 31, 2017
2. Aggarwal, A., Cook, A. J., Jiao, J., Seguin, R. A., Moudon, A. V., Hurvitz, P. M., & Drewnowski, A. 2014. Access to supermarkets and fruit and vegetable consumption. American Journal ofPublic Health, 104,917–923.
3. Algert, S. J., Agrawal, A., & Lewis, D. S. (2006). Disparities in access to fresh produce in low-income neighborhoods in Los Angeles. American Journal of Preventive Medicine, 30(5), 365–370.
4. Auchincloss, A., Riolo, R., Brown, D., Cook, J., & Diez Roux, A. V. (2011). An agent-based model of income inequalities in diet in the context of residential segregation. American Journal of Preventive Medicine, 40(3), 303–311.
5. Baker, E. A., Schootman, M. Barnidge, E. Kelly, C. (2006) The role of race and poverty in access to foods that enable individuals to adhere to dietary guidelines. Preventing Chronic Disease. 3(3) 1-11.
6. Bertrand, L., F. Therien, and M. S. Cloutier. 2008. Measuring and mapping disparities in access to fresh fruits and vegetables in Montreal. Canadian Journal of Public Health/Revue Canadienne de Sante'ePublique 99 (1): 6–11.
7. Bhat, C.R., Handy, S., Kockelman, K., Mahmassani, H.S., Chen, Q., Srour, I., and Weston, L., "Assessment of Accessibility Measures," Report 7-4938-3, prepared for the Texas Department of Transportation, August 2001.
8. Bhat, C.R., Handy, S., Kockelman, K., Mahmassani, H.S., Gopal, A., Srour, I., and Weston, L., "Development of an Urban Accessibility Index: Formulations, Aggregation, and Application," Report 7-4938-4, prepared for the Texas Department of Transportation, October 2002
9. Bhat, C.R., Srinivasan, S., Guo, J., and Sivakumar, A. (2003). Activity Based Travel Demand Analysis for Metropolitan Areas in Texas: A Micro-Simulation Framework for Forecasting, Report 4080-4, prepared for the Texas Department of Transportation.
10. Black, J. L., R. M. Carpiano, S. Fleming, and N. Lauster. 2011. Exploring the distribution of food stores in British Columbia: Associations with neighbourhood socio-demographic factors and urban form. Health&Place 17 (4): 961–70.
11. CBS (2015). VA rule change could double number of vets eligible for health care. CBS News. Retrieved April 13, 2018. https://www.cbsnews.com/news/va-rule-change-could-double-number-of-veterans-eligible-for-health-care/
12. Chaix, B. 2009. Geographic life environments and coronary heart disease: A literature review, theoretical contributions, methodological updates, and a research agenda. Annual Review of Public Health 30:81–105.
13. Chung, C., & Myers, S. (1999). Do the poor pay more for food? An analysis of grocery store availability and food price disparities. The Journal of Consumer Affairs, 33(2), 276–296.
14. Cummins, S., & Macintyre, S. (2002). Food deserts—evidence and assumption in health policy making. The BMJ, 325, 436–438.
15. Dai, D., & Wang, F. (2011). Geographic disparities in accessibility to food stores in southwest Mississippi. Environment and Planning B: Planning and Design, 38(4), 659–677.

16. Dixon, J., Omwega, A., Friel, S., Burns, C., Donati, K., & Carlile, R. (2007). The health equity dimensions of urban food systems. Journal of Urban Health, 84(1), 118–129.
17. Eckert, J. & Vojnovic I. (2017). Fast food landscapes: Exploring restaurant choice and travel behavior for residents living in lower eastside Detroit neighborhoods. Applied Geography, SN -0143
18. Eckert, J. & Vojnovic I. (2017). Fast food landscapes: Exploring restaurant choice and travel behavior for residents living in lower eastside Detroit neighborhoods. Applied Geography, SN -0143
19. Eckert, J., & Shetty, S. (2011). Food systems, planning and quantifying access: Using GIS to plan for food retail. Applied Geography, 31(4), 1216–1223.
20. Fan, J., Kowaleski-Jones, L., Zick, C. D.. (2009) Alternative measures of food environments. Paper presented at the conference Understanding the Economic Concepts and Characteristics of Food Access, Economic Research Service, USDA, Washington, DC, January.
21. Farber, S., M. Z. Morang, and M. J. Widener. 2014. Temporal variability in transit-based accessibility to supermarkets. Applied Geography 53:149–59.
22. Galvez, M., Morland, K., Raines, C., Kobil, J., Siskind, J., Godbold, J., et al. (2008). Race and food store availability in an inner-city neighbourhood. Public Health Nutrition, 11, 624–631.
23. Giang, T., Karpyn, A., Laurison, H. B., Hillier, A., & Perry, R. D. (2008). Closing the grocery gap in underserved communities: The creation of the Pennsylvania fresh food financing initiative. Journal of Public Health Management and Practice, 14(3), 272–279.
24. Glanz, K., Sallis, J. F., Saelens, B. E., & Frank, L. D. (2005). Healthy nutrition environments: Concepts and measures. American Journal of Health Promotion, 19(5), 330–333.
25. Helling, A., & Sawicki, D. (2003). Race and residential accessibility to shopping and services. Housing Policy Debate, 14(1-2), 69–101.
26. Kwan, M. P. 2013. Beyond space (as we knew it): Toward temporally integrated geographies of segregation, health, and accessibility: Space–time integration in geography and GIScience. Annals ofthe Association ofAmerican Geographers 103 (5): 1078–86.
27. Larsen, K., and J. Gilliland. 2008. Mapping the evolution of food deserts' in a Canadian city: Supermarket accessibility in London, Ontario, 1961–2005. International Journal of Health Geographics 7:16.
28. LeDoux, T., & Vojnovic, I. (2013). Going outside the neighborhood: The shopping patterns and adaptations of disadvantaged consumers living in the lower eastside neighborhoods of Detroit, Michigan. Health & Place, 19, 1–14.
29. LeDoux, T., & Vojnovic, I. (2014). Examining the role between the residential neighborhood food environment and diet among low-income households in Detroit, Michigan. Applied Geography, 55, 9–18.
30. Lee, G., and H. Lim. 2009. A spatial statistical approach to identifying areas with poor access to grocery foods in the city of Buffalo, New York. Urban Studies 46 (7): 1299–1315.
31. McKinnon RA, Reedy J, Morrissette MA, et al. Measures of the food environment: a compilation of the literature, 1990-2007. Am J Prev Med. 2009;36(4 suppl):S124-S133.
32. Morland, K., Roux, A. D., Wing, S. (2002) Supermarkets, Other Food Stores, and Obesity: The Atherosclerosis Risk in Communities Study. American Journal of Public Health. 30(4) 333-339.
33. Morland, K., Wing, S., Roux, A. D. (2006) The Contextual Effect of the Local Food Environment on Residents' Diets: The Atherosclerosis Risk in Communities Study. American Journal of Public Health. 92(11) 1761-1767.
34. Pozsgay, M.A., and C.R. Bhat (2002) Destination Choice Modeling for Home-Based Recreational Trips: Analysis and Implications for Land-Use, Transportation, and Air Quality Planning, Transportation Research Record, Vol. 1777, pp. 47-54.
35. Rose, D. (2011). Captive audience? Strategies for acquiring food in two Detroit neighborhoods. Qualitative Health Research, 21(5), 642–651.
36. Sallis JF, Glanz K. Physical activity and food environments: solutions to the obesity epidemic. Milbank Q. 2009;87:123-154.

37. "Shannon, J., and W. J. Christian. 2017. What is the relationship between food shopping and daily mobility? A relational approach to analysis of food access. GeoJournal. doi:10.1007/s10708-016-9716-0"
38. Shearer, C., D. Rainham, C. Blanchard, T. Dummer, R. Lyons, and S. Kirk. 2015. Measuring food availability and accessibility among adolescents: Moving beyond the neighbourhood boundary. Social Science & Medicine 133:322–30.
39. Shonkoff, J. P., Phillips, D. A. (Eds.) (2000) From Neurons to Neighborhoods: The Science of Early Childhood Development. Washington, D.C. National Academy Press.
40. UCLA (2008) Designed for disease: The link between local food environments and obesity and diabetes. California Center for Public Health Advocacy, Policy Link, and the UCLA Center for Health Policy Research.
41. USDA Economic Research Service. (2017). Food Access Research Atlas—Documentation. Retrieved from http:// www.ers.usda.gov/data-products/food-access-research-atlas/documentation.aspx#definitions.
42. Ver Ploeg, M. (Ed.). (2012). Access to affordable and nutritious food: Measuring and understanding food deserts and their consequences: Report to Congress. Washington, DC: US Department of Agriculture, Economic Research Service.
43. Ver Ploeg, M., Dutko, P., Brenerman, V. (2014) Measuring Food Access and Food Deserts for Policy Purposes. Applied Economic Perspectives and Policy. 37(2) 205-225.
44. Ver Ploeg, M., Mancino, L., Todd, J. E., Clay, D. M., & Scharadin, B. (2015). Where do Americans usually shop for food and how do they travel to get there? Initial findings from the national household food acquisition and purchase survey. Washington, DC: USDA Economic Research Service
45. Vojnovic, I., Lee, J., Kotval-K, Z., Podagrosi, A., Varnakovida, P., Ledoux, T., et al. (2012). The burdens of place: A socio- economic and ethnic/racial exploration into urban form, accessibility and travel behaviour in the lansing capital region, Michigan. Journal of Urban Design, 18(1), 1–35. doi:10.1080/13574809.2012.683403.
46. Walker, R. E., C. R. Keane, and J. G. Burke. 2010. Disparities and access to healthy food in the United States: A review of food deserts literature. Health & Place 16 (5): 876–84.
47. WHO (2003). Diet, Nutrition, and the Prevention of Chronic Diseases. World Health Organization.
48. Widener, M. J., & Shannon, J. (2014). When are food deserts? Integrating time into research on food accessibility. Health and Place, 30, 1–3. doi:10.1016/j.healthplace.2014.07.011.
49. Widener, M. J., Farber, S., Neutens, T., Horner, M.W. (2013) Using urban commuting data to calculate a spatiotemporal accessibility measure for food environment studies. Health and Place 21, 1-9.
50. Widener, M. J., Farber, S., Neutens, T., Horner, M.W. (2013) Using urban commuting data to calculate a spatiotemporal accessibility measure for food environment studies. Health and Place 21, 1-9.
51. Widener, M. J., Metcalf, S.S., Bar-Yam, Y. (2011) Dynamic Urban Food Environments: A Temporal Analysis of Access to Healthy Foods. American Journal of Preventative Medicine 41(4) 439-441.
52. Widener, M.J., Metcalf, S.S., Bar-Yam, Y. (2013) Agent-Based modeling of policies to improve food access for low-income populations. Applied Geography 40 (2013) 1-10.
53. Wilkins, E. L., Morris, M. A., Radley, D., Griffiths, C. (2017) Using Geographic Information Systems to measure retail food environments: Discussion of methodological considerations and a proposed reporting checklist (Geo-FERN). Health & Place. 44(2017) 110-117.
54. Wrigley, N. (2002). “Food deserts” in British cities: Policy context and research priorities. Urban Studies, 39(11), 2029–2040.
55. Wrigley, N., Warm, D., & Margetts, B. (2003). Deprivation, diet, and food-retail access: Findings from the leeds “food deserts” study. Environment and Planning A, 35(1), 151–188.

Table 1. Average household characteristics comparison

Characteristic	Low-Income Households	Medium/High-Income Households
Household size	2.13	2.32
Number of drivers	1.47	1.90
Number of Household vehicles	1.57	2.19
Number of adults	1.66	1.93
Number of young children	0.15	0.09
Number of workers	0.72	1.19
Percent Renter housing in HH CBG	42.14	28.32
People per sq. mi. in HH CBG	4,974.68	3,795.00
Housing units per sq. mi. in HH CBG	2,163.80	1,623.40

Table 2 *Non-Adjusted* Composite Size measure estimation

Variables	NLMNL	
	Coeff.	t-stat.
Distance from home	-0.4024	-47.45
Composite Size	0.3085	6.56
Area	1.0000	--
number of retail developments	-0.1322	-0.25
number of recreational developments	0.3921	0.87
Number of Observations	50,250	
Log-likelihood at convergence	-2567.20	

Table 3 *Adjusted* Composite Size measure estimation

Variables	NLMNL Results	
	Coeff.	t-stat.
Distance from home	-0.4024	-47.45
Composite Size	0.3085	6.56
number of retail developments	0.8762	1.89
number of recreational developments	1.4801	2.21
Number of Observations	50,250	
Log-likelihood at convergence	-2567.20	

Table 4 Location-choice estimation results

Variables	MNL model	
	Coefficient	t-statistic
Attributes of alternative		
Number of wholesale stores	0.2990	2.07
Number of supercentres	0.4152	4.69
TSZ is adjacent to home zone	0.8966	5.06
TSZ is the same as home zone	0.7995	4.03
Population Density (persons/sq. mi.)	-0.0010	-2.27
Composite Size measure	0.2240	1.97
Distance from home (miles)	-0.5491	-6.25
Socio-demographic interactions		
with Income >$50K		
TSZ neither same nor adjacent to home	0.6196	2.14
with Distance from home (miles)		
Number of vehicles in household	0.0396	3.66
Number of adults in the household	-0.0643	-4.01
High-density residential location dummy	0.0172	2.45
Presence of child under 5 years of age	0.0906	5.08
Full time employment dummy	0.1315	7.86
TSZ neither same nor adjacent to home	0.1507	1.99
Number of wholesale stores	0.0228	2.22
Number of supercentres	0.0248	2.56
Composite size measure	0.0262	2.55

Table 5 Model fit statistics

Summary Statistics	**Value**
Log likelihood at convergence	-2,515.59
Log likelihood with zero coefficients	-3,931.56
Log-likelihood ratio test	2,895.35
Number of parameters	17
Rho-Squared w.r.t. Zero	0.3602
Adjusted Rho-Squared w.r.t. Zero	35.58%

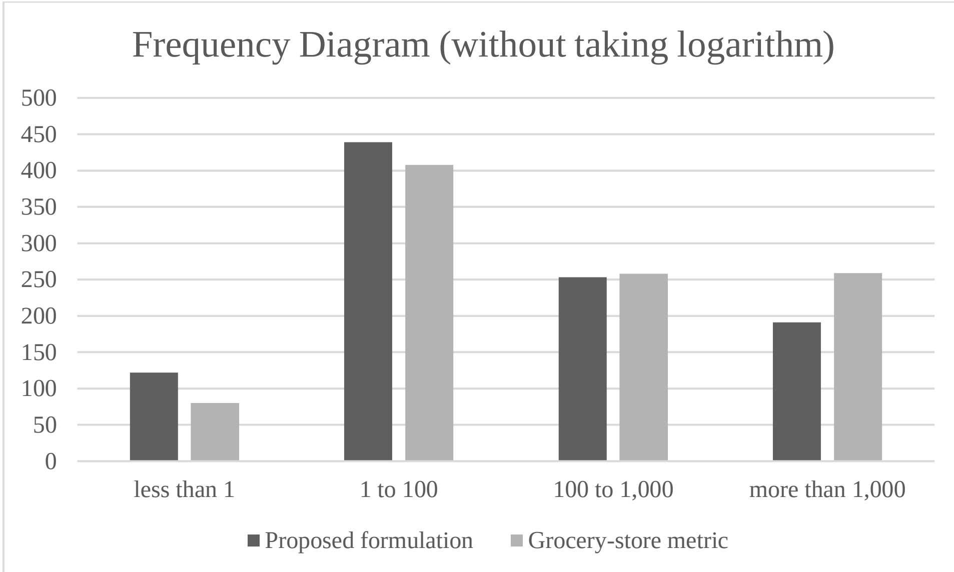

Figure 1 Frequency Diagram (without taking logarithm) to compare the proposed formulation with grocery-store metric

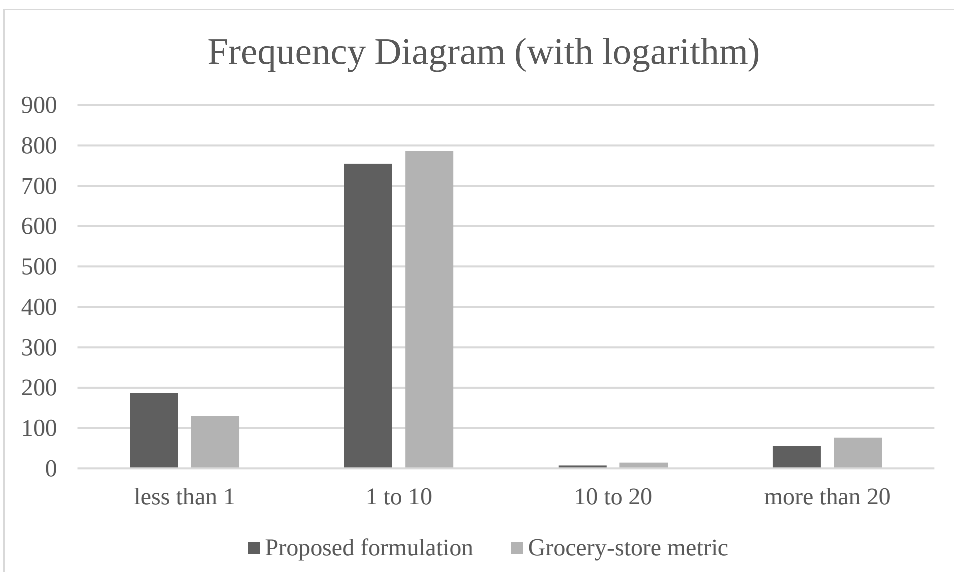

Figure 2 Frequency Diagram (with logarithm) to compare the proposed formulation with grocery-store metric