**A Spatially Explicit Agent Based Model of Food Deserts**

HAROLD J. WALBERT

Department of Computational Social Science

George Mason University

Fairfax, VA 22030

hwalbert@gmu.edu

**Abstract**

Understanding the influences and factors that determine food accessibility and create food deserts is important if policymakers are to implement changes to improve overall public health outcomes. The aim of this study was to model the factors of space and distance, income, education, and food assistance to better understand individual food accessibility and the overall food accessibility of a region. The study uses a spatially explicit agent based model to model the population of Washington DC, their accessibility to food, and their health status. This is an empirical model with data collected from a variety of sources and used to construct a representation of the population of Washington DC, their unique socioeconomic status, their unique health status and their spatial location in relation of sources of food (supermarkets and convenience stores). This model generates results that are consistent with where food deserts are reported to exist in the District of Columbia. In addition the variable of accessibility to food sources is found to be statistically significant and to have a positive relationship to overall health perception for the population of Washington DC. Although there is more to be done regarding food accessibly research, this model is a proof of concept for modeling food accessibility using agent based models. The model provides tools for understanding how changes in policy can effect outcomes and for understanding where interventions could be most effective.

**Model – Overview**

*Overview – Purpose*

For over twenty years research on food deserts and food accessibility has increased and has been contributed to from a diverse set of disciplines including, but not limited to, agriculture, sociology, economics, public policy, sociology, and social epidemiology (Adams et. al., 2010). In 2010 the topic of food accessibility received special attention in the United States national spotlight when the White House made eliminating food deserts a part of the First Lady’s Let’s Move campaign. There are multiple definitions that have been proposed for what constitutes a food desert but they all tend to include some of the same themes of limited or no access to supermarkets, limited transportation options, and low income. Other literature also suggests that education plays a part in the types of food choices that are made (Moreira & Padrão, 2004). A large literature in the medical field has been established that shows that medical conditions such as heart disease, diabetes, high blood pressure and obesity are made worse by poor diets. In this paper I attempt to add a computational social science perspective to the discussion of food deserts that draws on the already existing interdisciplinary research and empirical data. Modeling where food deserts occur, the factors involved in their appearance, and developing ways to test policy interventions are all important if food deserts and to be eliminated and accessibility to quality healthy food is to be increased.

Understanding food deserts requires the researcher to consider the physical location, financial situation and internal sophistication of the people involved (Shaw, 2006). Depending on the availability of cars or public transportation a person may have very different abilities to get to a supermarket and be able to transport the food back to their home. Distance and the ease with which a person can transport themselves are important for understanding overall accessibility (Bitler & Haider, 2010).

A separate by intertwined issued with transportation is income and poverty. “The most basic determinants of the demand for healthy food are income, prices, and preferences. Economic theory suggests that the quantity of healthy food demanded is decreasing in its own price and increasing in the price of substitute foods. Assuming healthy food is a normal good, the demand for healthy food will increase with income levels. This observation implies that there will be more food stores with healthy food in high-income areas when compared to low-income areas, even if there were sufficient food stores with healthy food in both. Preferences determine the degree to which prices and income affect food consumption” (Bitler & Haider, 2010).

Even in those situations where healthy and affordable food is available, people may choose to still eat unhealthy food. Although most of the literature does not list education as a determinant of what makes a food desert, it is still important for understanding the food accessibility equation. In their 2004 paper, Moreira & Padrão report their findings from a cross-sectional survey looking at the effect of education level and income on the food consumption choice. They report that their study suggests that “education and income have distinct associations with food choice. The low and high income groups are or tend to be similar in regard to the majority of food choices, and access to education appears to be the key element to a better food pattern as indicated by higher frequency of milk, vegetable soup, vegetables, fruit, and fish consumption.”

This paper uses a spatially explicit, empirically grounded agent based model of the population of the District of Columbia in the United States to model the population’s access to sources of food (supermarkets and convenience stores). This model will investigate the phenomenon of food deserts and try to model and understand the varying level of access that people have to different types of food. This model will be a spatially explicit model of agents interacting with their environment and choosing a location where they will acquire food. It is a model that will focus on an urban setting (Washington D.C) where agents will have a choice of many, few, or no stores (based on their location) from which to acquire food. Some of these stores will be traditional grocery stores while others will be modeled as convenience stores. The assumption is that grocery stores are more likely to have a healthy set of food options and that smaller (but more plentiful) convenience stores will have a more limited and less healthy set of food options and is supported by Ghirardelli et.al (2010). It is also supported by Adams et.al. (2010), who write that people in impoverished neighborhoods often do not have access to stores that provide affordable and healthy foods and that these same communities also tend to have many convenience stores scattered throughout.

*Overview – Entities, State Variables, and Scales*

The agents and the environment in which they exist are created using empirical data collected from several different datasets. The agents in this model represent people and have heterogeneous attributes that are empirically derived from real world data. Agents have unique values for education, food assistance status (food stamps), poverty, the number of stores they have access to from their unique spatial position, and their individual health status. The agents are distributed according to population across the landscape. Each agent is represented as a black dot on the landscape.

Agents exist on a landscape that represents Washington, D.C. in the United States. The visualization of the model uses black lines to delineate census blocks and orange lines to outline the different wards that make up the city. The landscape is dotted with stores that are either supermarkets or something other than a supermarket but that still appeared in the supermarket data. In this paper I treat these “non-supermarkets” as the equivalent of a convenience store. Supermarkets are represented in the visualization of the model as green dots and the convenience stores are shown as red dots. A more detailed discussion of agent and environment initialization as well as the input data used will appear later in this paper.

The global parameters in the model are the radius size, d\*, and the education factor. Radius size is the distance around each agent that they are willing to travel in to get to a supermarket or convenience store. The number of stores they have access to from their unique spatial position is equal to the number of stores that appear in their radius. The global parameter d\* is an input to the accessibility score calculation that each agent makes when the simulation is run. It represents the distance from the agent where accessibility decreases most quickly. It is most usefully interpreted in this model as the ease by which an agent can move around the city and could be thought of as a proxy for how effective public transportation is or how costly it is to move around the city. Lower values of d\* correspond with more difficulty in moving around the city and higher values correspond with less difficulty moving around the city. The education factor is an input to the accessibility score calculation when the accessibility calculation is set to use socioeconomic data. A more detailed discussion of the accessibility score calculation will appear later in this paper.

*Overview – Process Overview and Scheduling*

Before the simulation begins each agent is instantiated with a random number that determines when they will perform the function where they determine their individual accessibility to food (i.e. perform the accessibility score calculation). One run of the model represents one week and in this week all the agents in the simulation make a decision (on different days) to determine their individual accessibility score. When an agent decides it is time to calculate their accessibility score they will do it in one of two ways depending on whether the model is set to run using the basic accessibility score equation or the more advanced accessibly score equation that incorporates the agents socioeconomic data (the agent attributes for poverty, education, and food assistance).

The basic accessibility score calculation uses a modification of the basic model set out by Guy (1983) that gives a measure of spatial accessibility to food resources. The accessibility for an individual store (Ai) is calculated using this equation where Sj is the size of the store (interpreted in the model as being larger when the store is a supermarket and smaller when the store is not a supermarket), dij is the distance between the agent and the store, and d\* is a global variable that is a proxy for how easy it is to move around the landscape:

Each agent assigns a score to each store that is in their unique radius and this score is unique to each agent-store pairing. An agent’s total accessibility score is equal to the sum of all the individual stores accessibility scores.

The more advanced accessibility score is a variation on the basic model that uses information about the individual’s socioeconomic status to augment the calculation that determines overall accessibility. The three variables that augment the basic calculation are poverty, education, and food assistance. The next three paragraphs will describe how these socioeconomic variables modify the basic accessibility equation.

If the agent is determined to have some form of food assistance that will add to the Sj term in the equation and cause the overall accessibility to increase for that agent regardless of what happens in the rest of the calculation. Increasing the affordability and overall ability of people to acquire food is a basic function of food assistance that is reflected in this modification of the Sj term in the accessibility equation.

Education is assumed to positively affect an agent’s ability to make good choices to eat healthy foods. Although the health of the food and the types of food available at supermarkets and non-supermarkets is not explicitly modeled in this paper, I make the assumption that more accessibility is desirable over lower accessibility and that, in general, more healthful foods are available at supermarkets as opposed to non-supermarkets. Education is assumed to positively affect accessibility in this model and lack of education is assumed to detract from accessibility in this model. In addition, education will affect the agent’s interaction with supermarkets and non-supermarkets differently. Education will increase the Sj term of supermarkets and will decrease the Sj term of non-supermarkets. Lack of education will increase the Sj term of non-supermarkets and will decrease the Sj term of supermarkets.

While food assistance and education work on the Sj term of the equation, poverty affects the d\* term. The d\* term can best be thought of in this model as the ease by which an agent is able to move about the city. Higher levels of d\* are associated with higher levels of mobility while lower levels of d\* are correlated with is being harder to move around the landscape. If the agent is determined to be in poverty then their d\* term will be lower and if they are determined to not be in poverty their d\* term will be higher. This assumes that people with more money have the means to transport themselves around the city more easily or possibly have more modes of transportation at their disposal that decreases their overall transportation burden.

**Model – Design Concepts**

*Design Concepts – Emergence*

Several macro level variables emerge from the combined decisions that the agents make. Total accessibility, the distribution of accessibility, and the ratio of agents choosing supermarkets vs. convenience stores all emerge from the interactions in the model. These variables are affected by the global parameters set at the start of the run as well as the decision of the modeler to include socioeconomic variables in the accessibility calculation. A discussion of the results will be included later in this paper.

*Design Concepts – Sensing, Fitness and Interaction*

Each agent is able to “see” in a specified radius around their location and only considers the stores that appear in this radius. All agents in the model can see the same distance around themselves and this parameter is set as a global constant in the model. The stores that the agent can consider are of two types, supermarkets or non-supermarkets. The agent is able to distinguish between these two types of stores and uses this information when calculating accessibility. The agents end goal is to pick a store with the highest accessibility score. The agent does not explicitly model the health of the food available at a certain store. Agents only interact with their environment in this model.

*Design Concepts – Stochasticity and Observation*

Agents are instantiated with empirically derived data for poverty, food assistance, education, health perception, heart disease, and obesity. Although these measures are taken from empirical data they are assigned to the agents based on a random normal distribution. The mean and standard deviations used in the random normal distribution are drawn from the empirical data.

Data about the individual agents and the macro level outcomes are collected from each run of the model. Data about the total and individual accessibility, average accessibility, and minimum and maximum accessibility is collected. Information about the distribution of accessibility and the percent of agents choosing different types of stores is also reported. An Accessibility-Index (a Gini-Index applied to accessibility) is calculated after each run of the model. A more in depth discussion of the results as well as a parameter sweep will be included in the next section of the paper.

**Model – Details**

*Details – Initialization and Input Data*

The population’s spatial distribution as well as the information about the agents education, food assistance, and poverty is taken from the United States Census Bureau’s TIGER (Topologically Integrated Geographic Encoding and Referencing) GIS data and is at the census block group level of granularity. A table showing the exact data used to instantiate the population is shown in Exhibit 1. This data is assigned to the agents at the beginning of each model run according to a standard normal distribution with a mean value equal to the value for the particular census block group and a standard deviation equal to the standard deviation in the empirical data for all census block groups. The global parameter called population denominator is used to scale down the total population of Washington DC to a level that can be simulated more easily. For all simulations the population denominator is set to 300 meaning that one agent is representative of approximately 300 people.

The map showing the outline of the census block groups is generated from the polygon shapefile also made available from the Census. The map showing the outline of the different wards in the District of Columbia was downloaded from the ARCGIS open data website which collected this data from the District of Columbia GIS data website. In order to increase the speed at which the model is set up, the polygons for both census block groups and DC wards were simplified using a tool called Mapshaper that smooths out the rough edges of the polygon data and reduces the size of the file being imported. For this reason some of the lines from the block group polygons and the wards polygons to not overlap exactly when overlaid on one another.

Supermarket data comes from the Supermarket Access Map made available from ARCGIS. Locations in this database are marked as either being supermarkets (denoted in the dataset with a “T” flag in the “Supermarke” column) while those that are not supermarkets are denoted with a “F” or “NULL”. Included as attributes of the supermarket data is information about the sales volume at each store (the unit and time scale for sales volume is not reported). I use this data as a proxy for the size of the store (Sj in the accessibility equation). The average sales volume reported for all stores not listed as supermarkets is 651 and the average sales volume reported for all stores listed as supermarkets is 25,146. The ratio of average sales volume for non-supermarkets to the average sales volume for supermarkets is 0.026 and this is the value used for Sj in the accessibility equation for non-supermarkets. A value of 1.0 is used as the Sj value for supermarkets. Exhibit 2 shows the map of Washington DC with census blocks outlines in black, wards outlined in orange, supermarkets marked with green circles, and convenience stores marked with red circles.

Measures of the health of the population were incorporated into the agents in the simulation. These measures include overall health perception, heart disease mortality, and obesity. This data was retrieved from the 2014 District of Columbia Community Health Needs Assessment which was prepared by the DC Department of Health. This data is for all eight wards in the Washington DC. This data is assigned to the agents at the beginning of each model run according to a standard normal distribution with a mean value equal to the values for the particular ward and a standard deviation equal to the standard deviation in the empirical data for all wards in Washington DC.

*Details – Submodels*

The model for accessibility has been explained in the previous section and it is responsible for the core agent decision making behavior. The accessibility outcome is unique for each agent based on their unique location and attributes. In order to connect food accessibility with health outcomes an ordinary least squares (OLS) linear regression model was used to estimate the effect of accessibility on the empirical data on overall health perception taken from the 2014 District of Columbia Community Health Needs Assessment which was prepared by the DC Department of Health.

The hypothesis is that accessibility to sources of food should be positively related to overall health outcomes (i.e. when you see higher accessibility you should see better health outcomes). In order to test this the following regression was specified:

A higher Overall Health Outcome is associated with better health just as higher accessibility is associated with better access to food. The experiments showed that after the running the model 100 times and sweeping through the parameters of d\* (ease of transportation) and the education factor that the coefficient on accessibility (β) was positive and statistically significant at the 99 percent confidence level.

**Experiments and Results**

This model was built using the agent based modeling language Netlogo and the statistical programming language R was used to run the experiments and analyze the results. The R package RNetLogo was used to connect R to Netlogo. A parameter sweep was run on both the d\* parameter and the Education Factor parameter. The sweep of the d\* parameter was done on the model twice with the only difference being that one sweep was done while using the original accessibility model and the other sweep was done using the augmented accessibility equation that takes into account socioeconomic factors. The parameter d\* was swept at the following levels: 0.3 (the low level where transportation was difficult), 1.1 (the medium level where transportation is more accessible), and 2.0 (where transportation around the landscape is easiest).

The education factor was swept at three different levels as well. It was swept at the 0.1 (low impact of education), 0.5 (medium impact of education), and 1.0 (high impact of education) levels. It is important to remember how these education factors work. Education is assumed to positively affect accessibility in this model and lack of education is assumed to detract from accessibility in this model. Education will affect the agent’s interaction with supermarkets and non-supermarkets differently by increasing the Sj term of supermarkets and decreasing the Sj term of non-supermarkets. Conversely, lack of education will increase the Sj term of non-supermarkets and decrease the Sj term of supermarkets. In this model education cuts both ways and is very tied to the percent of the population that is considered educated.

A sample of the code used to run the experiments on one of the levels of d\* and the education factor is included below as an aid to understand the structure of the experiments. This code includes commands to reset default values for the experiments, setting the number of repetitions for the experiments, setting the values for the parameter being swept, the regression model, and the collection and structuring of the resulting data.

Code for running 100 experiments with a d\* of 0.3 and using the socioeconomic data in the accessibility equation:

ResetDefaultValues <- function(){

NLCommand("set radius-size 6")

NLCommand("set d\* 1.1")

NLCommand("set populationDenominator 300")

NLCommand("set education-factor 0.5")

NLCommand("set Calculate-Accessibility-Using-Socioeconomic-Data? TRUE")

}

NumberOfRepetitions <- 100

#Reset default values and set d\* to Low value

ResetDefaultValues()

NLCommand("set d\* 0.3")

LowdResultsEstimate <- data.frame()

LowdResultsStErr <- data.frame()

LowdPValue <- data.frame()

LowdRSqr <- data.frame()

LowdSumAcc <- data.frame()

LowdMeanAcc <- data.frame()

LowdAccIndex <- data.frame()

LowdPerSupermarket <- data.frame()

LowdPernotSupermarket <- data.frame()

LowdRatio <- data.frame()

for (i in 1:NumberOfRepetitions)

{

NLCommand("setup")

NLCommand("draw")

NLCommand("make-pop")

NLDoCommand(7, "go")

peopleProperties <- NLGetAgentSet(c("accessibility","health","heart","overweight","edu","pov","fs"), "people", as.data.frame=TRUE)

reg <- lm(peopleProperties$health ~ peopleProperties$accessibility)

LowdResultsEstimate <- as.data.frame(append(LowdResultsEstimate, coef(summary(reg))["peopleProperties$accessibility","Estimate"]))

LowdResultsStErr <- as.data.frame(append(LowdResultsStErr, coef(summary(reg))["peopleProperties$accessibility","Std. Error"]))

LowdPValue <-as.data.frame(append(LowdPValue, coef(summary(reg))["peopleProperties$accessibility", "Pr(>|t|)"]))

LowdRSqr <- as.data.frame(append(LowdRSqr, summary(reg)$adj.r.squared))

LowdSumAcc <- as.data.frame(append(LowdSumAcc, NLReport("sum ([accessibility] of people)")))

LowdMeanAcc <- as.data.frame(append(LowdMeanAcc, NLReport("mean [accessibility] of people")))

LowdAccIndex <- as.data.frame(append(LowdAccIndex, NLReport("(accessibility-index-reserve / (count people)) / 0.5")))

LowdPerSupermarket <- as.data.frame(append(LowdPerSupermarket, NLReport("100 \* (count people-with-store with [[supermarket-here?] of store-choice = TRUE]) / (count people)")))

LowdPernotSupermarket <- as.data.frame(append(LowdPernotSupermarket, NLReport("100 \* (count people-with-store with [[notSupermarket-here?] of store-choice = TRUE]) / (count people)")))

LowdRatio <- as.data.frame(append(LowdRatio, NLReport("(100 \* (count people-with-store with [[supermarket-here?] of store-choice = TRUE]) / (count people)) / (100 \* (count people-with-store with [[notSupermarket-here?] of store-choice = TRUE]) / (count people))")))

}

LowdResultsEstimate <- t(LowdResultsEstimate)

LowdResultsStErr <- t(LowdResultsStErr)

LowdPValue <- t(LowdPValue)

LowdRSqr <- t(LowdRSqr)

LowdSumAcc <- t(LowdSumAcc)

LowdMeanAcc <- t(LowdMeanAcc)

LowdAccIndex <- t(LowdAccIndex)

LowdPerSupermarket <- t(LowdPerSupermarket)

LowdPernotSupermarket <- t(LowdPernotSupermarket)

LowdRatio <- t(LowdRatio)

LowdCombined <- data.frame(LowdResultsEstimate, LowdResultsStErr, LowdPValue, LowdRSqr, LowdSumAcc, LowdMeanAcc, LowdAccIndex, LowdPerSupermarket, LowdPernotSupermarket, LowdRatio)

Code for running 100 experiments with an Education Factor of 0.5 and using the socioeconomic data in the accessibility equation:

#Reset default values and set education-factor to Medium value

ResetDefaultValues()

NLCommand("set education-factor 0.5")

MediumEduResultsEstimate <- data.frame()

MediumEduResultsStErr <- data.frame()

MediumEduPValue <- data.frame()

MediumEduRSqr <- data.frame()

MediumEduSumAcc <- data.frame()

MediumEduMeanAcc <- data.frame()

MediumEduAccIndex <- data.frame()

MediumEduPerSupermarket <- data.frame()

MediumEduPernotSupermarket <- data.frame()

MediumEduRatio <- data.frame()

for (i in 1:NumberOfRepetitions)

{

NLCommand("setup")

NLCommand("draw")

NLCommand("make-pop")

NLDoCommand(7, "go")

peopleProperties <- NLGetAgentSet(c("accessibility","health","heart","overweight","edu","pov","fs"), "people", as.data.frame=TRUE)

reg <- lm(peopleProperties$health ~ peopleProperties$accessibility)

MediumEduResultsEstimate <- as.data.frame(append(MediumEduResultsEstimate, coef(summary(reg))["peopleProperties$accessibility","Estimate"]))

MediumEduResultsStErr <- as.data.frame(append(MediumEduResultsStErr, coef(summary(reg))["peopleProperties$accessibility","Std. Error"]))

MediumEduPValue <-as.data.frame(append(MediumEduPValue, coef(summary(reg))["peopleProperties$accessibility", "Pr(>|t|)"]))

MediumEduRSqr <- as.data.frame(append(MediumEduRSqr, summary(reg)$adj.r.squared))

MediumEduSumAcc <- as.data.frame(append(MediumEduSumAcc, NLReport("sum ([accessibility] of people)")))

MediumEduMeanAcc <- as.data.frame(append(MediumEduMeanAcc, NLReport("mean [accessibility] of people")))

MediumEduAccIndex <- as.data.frame(append(MediumEduAccIndex, NLReport("(accessibility-index-reserve / (count people)) / 0.5")))

MediumEduPerSupermarket <- as.data.frame(append(MediumEduPerSupermarket, NLReport("100 \* (count people-with-store with [[supermarket-here?] of store-choice = TRUE]) / (count people)")))

MediumEduPernotSupermarket <- as.data.frame(append(MediumEduPernotSupermarket, NLReport("100 \* (count people-with-store with [[notSupermarket-here?] of store-choice = TRUE]) / (count people)")))

MediumEduRatio <- as.data.frame(append(MediumEduRatio, NLReport("(100 \* (count people-with-store with [[supermarket-here?] of store-choice = TRUE]) / (count people)) / (100 \* (count people-with-store with [[notSupermarket-here?] of store-choice = TRUE]) / (count people))")))

}

MediumEduResultsEstimate <- t(MediumEduResultsEstimate)

MediumEduResultsStErr <- t(MediumEduResultsStErr)

MediumEduPValue <- t(MediumEduPValue)

MediumEduRSqr <- t(MediumEduRSqr)

MediumEduSumAcc <- t(MediumEduSumAcc)

MediumEduMeanAcc <- t(MediumEduMeanAcc)

MediumEduAccIndex <- t(MediumEduAccIndex)

MediumEduPerSupermarket <- t(MediumEduPerSupermarket)

MediumEduPernotSupermarket <- t(MediumEduPernotSupermarket)

MediumEduRatio <- t(MediumEduRatio)

MediumEduCombined <- data.frame(MediumEduResultsEstimate, MediumEduResultsStErr, MediumEduPValue, MediumEduRSqr, MediumEduSumAcc, MediumEduMeanAcc, MediumEduAccIndex, MediumEduPerSupermarket, MediumEduPernotSupermarket, MediumEduRatio)

The results table below gives the summary statistics for low d\*, medium d\*, and high d\* for the following outputs of the model: Accessibility Index which is a Gini-Index applied to accessibility (AccIndex), Percent of agents choosing supermarkets (PerSupermarket), percent of agents choosing convenience stores (PernotSupermarket), the ratio of agents choosing supermarkets to convenience stores (Ratio), Total accessibility in the model (SumAcc), and average accessibility in the model (MeanAcc). Results for all regressions are not included as a part of the summary and are discussed in the submodels section of the paper.

Results Table for sweep of d\* using the augmented accessibility equation for socioeconomic data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| LowdAccIndex | LowdPerSupermarket | LowdPernotSupermarket | LowdRatio | LowdSumAcc | LowdMeanAcc |
| Min. :0.6550 | Min. :28.95 | Min. :67.11 | Min. :0.4095 | Min. :543.6 | Min. :0.2861 |
| 1st Qu.:0.6630 | 1st Qu.:30.88 | 1st Qu.:68.13 | 1st Qu.:0.4468 | 1st Qu.:571.9 | 1st Qu.:0.3010 |
| Median :0.6659 | Median :31.29 | Median :68.68 | Median :0.4544 | Median :580.0 | Median :0.3053 |
| Mean :0.6662 | Mean :31.25 | Mean :68.69 | Mean :0.4551 | Mean :579.4 | Mean :0.3049 |
| 3rd Qu.:0.6698 | 3rd Qu.:31.79 | 3rd Qu.:69.11 | 3rd Qu.:0.4659 | 3rd Qu.:586.6 | 3rd Qu.:0.3087 |
| Max. :0.6806 | Max. :32.95 | Max. :70.68 | Max. :0.4910 | Max. :606.6 | Max. :0.3192 |
| MediumdAccIndex | MediumdPerSupermarket | MediumdPernotSupermarket | MediumdRatio | MediumdSumAcc | MediumdMeanAcc |
| Min. :0.5999 | Min. :29.21 | Min. :65.53 | Min. :0.4151 | Min. :2283 | Min. :1.201 |
| 1st Qu.:0.6099 | 1st Qu.:31.47 | 1st Qu.:67.58 | 1st Qu.:0.4585 | 1st Qu.:2345 | 1st Qu.:1.234 |
| Median :0.6135 | Median :31.95 | Median :68.18 | Median :0.4689 | Median :2370 | Median :1.248 |
| Mean :0.6135 | Mean :31.92 | Mean :68.16 | Mean :0.4686 | Mean :2370 | Mean :1.247 |
| 3rd Qu.:0.6170 | 3rd Qu.:32.42 | 3rd Qu.:68.63 | 3rd Qu.:0.4800 | 3rd Qu.:2394 | 3rd Qu.:1.260 |
| Max. :0.6256 | Max. :34.42 | Max. :70.42 | Max. :0.5253 | Max. :2489 | Max. :1.310 |
| HighdAccIndex | HighdPerSupermarket | HighdPernotSupermarket | HighdRatio | HighdSumAcc | HighdMeanAcc |
| Min. :0.5953 | Min. :30.63 | Min. :65.05 | Min. :0.4399 | Min. :3687 | Min. :1.940 |
| 1st Qu.:0.6049 | 1st Qu.:32.36 | 1st Qu.:66.58 | 1st Qu.:0.4792 | 1st Qu.:3877 | 1st Qu.:2.040 |
| Median :0.6088 | Median :32.87 | Median :67.21 | Median :0.4894 | Median :3926 | Median :2.066 |
| Mean :0.6087 | Mean :33.01 | Mean :67.17 | Mean :0.4917 | Mean :3920 | Mean :2.063 |
| 3rd Qu.:0.6124 | 3rd Qu.:33.58 | 3rd Qu.:67.74 | 3rd Qu.:0.5055 | 3rd Qu.:3969 | 3rd Qu.:2.089 |
| Max. :0.6270 | Max. :35.47 | Max. :69.63 | Max. :0.5435 | Max. :4055 | Max. :2.134 |

Results Table for sweep of d\* using the original accessibility equation without socioeconomic data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Lowd\_FMeanAcc | Lowd\_FAccIndex | Lowd\_FPerSupermarket | Lowd\_FPernotSupermarket | Lowd\_FSumAcc | Lowd\_FRatio |
| Min. :0.04118 | Min. :0.7726 | Min. :45.16 | Min. :49.42 | Min. :78.25 | Min. :0.8684 |
| 1st Qu.:0.04406 | 1st Qu.:0.7895 | 1st Qu.:46.37 | 1st Qu.:50.16 | 1st Qu.:83.72 | 1st Qu.:0.9098 |
| Median :0.04547 | Median :0.7938 | Median :46.66 | Median :50.58 | Median :86.40 | Median :0.9238 |
| Mean :0.04545 | Mean :0.7936 | Mean :46.70 | Mean :50.55 | Mean :86.36 | Mean :0.9240 |
| 3rd Qu.:0.04676 | 3rd Qu.:0.7987 | 3rd Qu.:47.05 | 3rd Qu.:50.89 | 3rd Qu.:88.85 | 3rd Qu.:0.9365 |
| Max. :0.05068 | Max. :0.8142 | Max. :48.05 | Max. :52.00 | Max. :96.29 | Max. :0.9713 |
| Mediumd\_FMeanAcc | Mediumd\_FAccIndex | Mediumd\_FPerSupermarket | Mediumd\_FPernotSupermarket | Mediumd\_FSumAcc | Mediumd\_FRatio |
| Min. :0.3599 | Min. :0.5353 | Min. :70.42 | Min. :23.84 | Min. :683.8 | Min. :2.671 |
| 1st Qu.:0.3623 | 1st Qu.:0.5381 | 1st Qu.:71.26 | 1st Qu.:24.78 | 1st Qu.:688.4 | 1st Qu.:2.813 |
| Median :0.3634 | Median :0.5398 | Median :71.58 | Median :25.05 | Median :690.5 | Median :2.856 |
| Mean :0.3635 | Mean :0.5400 | Mean :71.61 | Mean :25.07 | Mean :690.7 | Mean :2.858 |
| 3rd Qu.:0.3646 | 3rd Qu.:0.5420 | 3rd Qu.:71.89 | 3rd Qu.:25.37 | 3rd Qu.:692.8 | 3rd Qu.:2.909 |
| Max. :0.3678 | Max. :0.5460 | Max. :73.00 | Max. :26.37 | Max. :698.9 | Max. :3.057 |
| Highd\_FMeanAcc | Highd\_FAccIndex | Highd\_FPerSupermarket | Highd\_FPernotSupermark | Highd\_FSumAcc | et Highd\_FRatio |
| Min. :0.6679 | Min. :0.4898 | Min. :70.63 | Min. :23.74 | Min. :1269 | Min. :2.695 |
| 1st Qu.:0.6743 | 1st Qu.:0.4944 | 1st Qu.:71.16 | 1st Qu.:24.74 | 1st Qu.:1281 | 1st Qu.:2.805 |
| Median :0.6761 | Median :0.4959 | Median :71.53 | Median :25.11 | Median :1285 | Median :2.847 |
| Mean :0.6759 | Mean :0.4958 | Mean :71.59 | Mean :25.10 | Mean :1284 | Mean :2.854 |
| 3rd Qu.:0.6779 | 3rd Qu.:0.4976 | 3rd Qu.:71.95 | 3rd Qu.:25.42 | 3rd Qu.:1288 | 3rd Qu.:2.904 |
| Max. :0.6838 | Max. :0.5011 | Max. :72.95 | Max. :26.21 | Max. :1299 | Max. :3.067 |

Results Table for sweep of education factor:

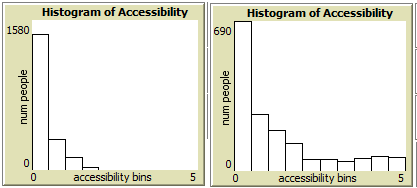
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| LowEduAccIndex | LowEduPerSupermarket | LowEduPernotSupermarket | LowEduRatio | LowEduSumAcc | LowEduMeanAcc |
| Min. :0.5433 | Min. :70.11 | Min. :33.58 | Min. :1.965 | Min. :1217 | Min. :0.6407 |
| 1st Qu.:0.5499 | 1st Qu.:71.09 | 1st Qu.:34.26 | 1st Qu.:2.039 | 1st Qu.:1236 | 1st Qu.:0.6504 |
| Median :0.5529 | Median :71.34 | Median :34.58 | Median :2.065 | Median :1245 | Median :0.6552 |
| Mean :0.5530 | Mean :71.32 | Mean :34.59 | Mean :2.063 | Mean :1244 | Mean :0.6548 |
| 3rd Qu.:0.5552 | 3rd Qu.:71.58 | 3rd Qu.:34.89 | 3rd Qu.:2.088 | 3rd Qu.:1251 | 3rd Qu.:0.6585 |
| Max. :0.5639 | Max. :72.42 | Max. :35.79 | Max. :2.152 | Max. :1274 | Max. :0.6705 |
| MediumEduAccIndex | MediumEduPerSupermarket | MediumEduPernotSupermarket | MediumEduRatio | MediumEduSumAcc | MediumEduMeanAcc |
| Min. :0.6011 | Min. :29.74 | Min. :66.11 | Min. :0.4201 | Min. :2284 | Min. :1.202 |
| 1st Qu.:0.6097 | 1st Qu.:31.53 | 1st Qu.:67.57 | 1st Qu.:0.4594 | 1st Qu.:2346 | 1st Qu.:1.235 |
| Median :0.6126 | Median :32.05 | Median :68.13 | Median :0.4713 | Median :2376 | Median :1.251 |
| Mean :0.6132 | Mean :32.00 | Mean :68.09 | Mean :0.4701 | Mean :2372 | Mean :1.249 |
| 3rd Qu.:0.6164 | 3rd Qu.:32.58 | 3rd Qu.:68.63 | 3rd Qu.:0.4818 | 3rd Qu.:2396 | 3rd Qu.:1.261 |
| Max. :0.6254 | Max. :33.89 | Max. :70.79 | Max. :0.5123 | Max. :2453 | Max. :1.291 |
| HighEduAccIndex | HighEduPerSupermarket | HighEduPernotSupermarket | HighEduRatio | HighEduSumAcc | HighEduMeanAcc |
| Min. :0.6193 | Min. :20.58 | Min. :75.11 | Min. :0.2625 | Min. :3721 | Min. :1.958 |
| 1st Qu.:0.6286 | 1st Qu.:21.72 | 1st Qu.:76.41 | 1st Qu.:0.2797 | 1st Qu.:3870 | 1st Qu.:2.037 |
| Median :0.6318 | Median :22.26 | Median :77.03 | Median :0.2881 | Median :3914 | Median :2.060 |
| Mean :0.6323 | Mean :22.31 | Mean :77.00 | Mean :0.2899 | Mean :3909 | Mean :2.057 |
| 3rd Qu.:0.6356 | 3rd Qu.:22.91 | 3rd Qu.:77.53 | 3rd Qu.:0.2994 | 3rd Qu.:3952 | 3rd Qu.:2.080 |
| Max. :0.6509 | Max. :24.58 | Max. :79.00 | Max. :0.3266 | Max. :4044 | Max. :2.129 |

**Discussion & Further Work**

This paper has provided the outline for a model that is a proof of concept for using agent based modeling to understand food accessibility and to model food deserts. This model has several features built in that could allow a policymaker to experiment with different policy interventions designed to effect accessibility.

The parameter sweep of d\* (ease of transportation) shows that when transportation is easier overall equality of accessibility is higher. When the level for d\* is low we see many more people with low accessibility and very few people with high accessibility. The visualizations in Graph 1 show the distribution of accessibility at low levels of d\* (d\* = 0.3 on left) and high levels of d\* (d\* = 2.0 on right).

Graph 1:



The sweep of the education factor shows that education is a double edged sword as mentioned earlier in the paper. If an agent is education it increases the attractiveness of supermarkets and decreases the attractiveness of convenience stores while doing the inverse if the agent is not educated. This means that the overall percent of people that are considered educated is important to the outcomes in the model. Lower education levels actually have more people choosing supermarkets than higher education levels. This is because only about 30% of the population is considered to be “educated”. A tool for changing the level of education in the population is discussed later in this section.

The first tool is the voronoi tessellation of the Washington DC landscape. The model uses the supermarkets as the starting set of points from which the tessellation is created. Every polygons created from this has the unique property of having every bit of area inside it closer to the supermarket inside it than to any other supermarket on the landscape. This provides a policymaker with the ability to strategically place new sources of food at optimal locations. For example, a policymaker could pick a boundary between two of the polygons that is also situated in an area that is low on the accessibility scale. It is important to note that doing this might not serve the most number of people as population density could be low in the area. Overall accessibility might be better served by placing a new source of food in a densely population area where many people could go but these areas (at least for the Washington DC model) tend to be in areas that do not have a problem with food accessibility. The model comes stocked with a pre-populated addition of nine additional supermarkets that are placed in locations where voronoi polygons have their boundaries. Using this allocation of new sources of food increase overall accessibility in the model. Although more needs to be done in order to understand how best to serve the underserved urban areas this is a starting points for conversation that uses empirical data to ground the discussion.

The second tool explored in this model is increasing the percent of the population that is “educated”. The education level census data used for this model is for education attainment for the population 25 years and over. If we think about education more broadly as opposed to formal education and consider specific public health education aimed at targeted populations then it becomes reasonable for the policymaker to think about how overall accessibility might change if the percent of the population that is educated about healthy eating choices changes. The “Educate” function in the model serves as a way to take twenty percent of the population that is not “educated” and make them “educated”. In practice this might involve public education initiatives, school programs, advertisements, or some combination of other education schemes.

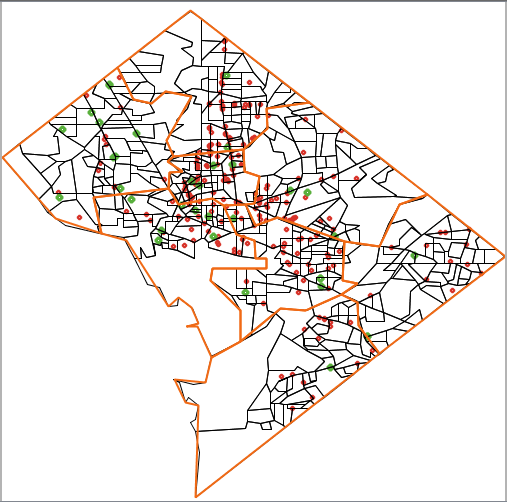
**Conclusion**

This model provides a proof on concept for using agent based models to model food accessibility and food deserts. It provides qualitative agreement with already existing models of food deserts (Graph 2). It also shows that higher accessibility is significant and correlated positively with better health perceptions.

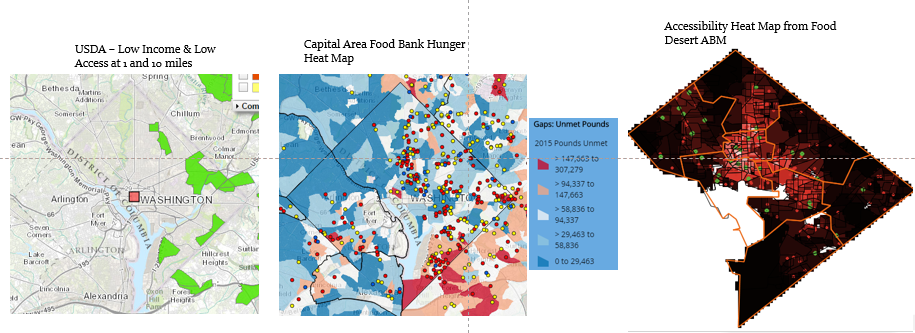
Exhibit 1

|  |  |  |
| --- | --- | --- |
| Used to Instantiate: | Census Data ID | Census Description |
| Population | B01001e1 | SEX BY AGE: Total: Total population -- (Estimate) |
| Education | B15003e1 | EDUCATIONAL ATTAINMENT FOR THE POPULATION 25 YEARS AND OVER: Total: Population 25 years and over -- (Estimate) |
| Poverty | B17010e1 | POVERTY STATUS IN THE PAST 12 MONTHS OF FAMILIES BY FAMILY TYPE BY PRESENCE OF RELATED CHILDREN UNDER 18 YEARS BY AGE OF RELATED CHILDREN: Total: Families -- (Estimate) |
| Food Stamps | B22010e2 | RECEIPT OF FOOD STAMPS/SNAP IN THE PAST 12 MONTHS BY DISABILITY STATUS FOR HOUSEHOLDS: Household received Food Stamps/SNAP in the past 12 months: Households -- (Estimate) |
| All values at Census Block level of granularity | | |

EXHIBIT 2



**Graph 2**

****

**Works Cited**

Larson, J., & Moseley, W. G.. (2012). Reaching the limits: a geographic approach for understanding food insecurity and household hunger mitigation strategies in Minneapolis-Saint Paul, USA. *Geojournal*,*77*(1), 1–12. Retrieved from <http://www.jstor.org/stable/41431187>

Adams, A. T., Ulrich, M. J., & Coleman, A.. (2010). Food Deserts. *Journal of Applied Social Science*,*4*(2), 58–62. Retrieved from <http://www.jstor.org/stable/23548938>

Shaw, H. J.. (2006). Food Deserts: Towards the Development of a Classification. *Geografiska Annaler. Series B, Human Geography*, *88*(2), 231–247. Retrieved from <http://www.jstor.org/stable/3878390>

Mead, M. N.. (2008). Urban Issues: The Sprawl of Food Deserts. *Environmental Health Perspectives*,*116*(8), A335–A335. Retrieved from <http://www.jstor.org/stable/25071122>

Bitler, M., & Haider, S. J.. (2010). AN ECONOMIC VIEW OF FOOD DESERTS IN THE UNITED STATES. *Journal of Policy Analysis and Management*, *30*(1), 153–176. Retrieved from <http://www.jstor.org/stable/40961588>

Food Systems, Food Movements, Food Justice. (2016). Food Systems, Food Movements, Food Justice. In *More Than Just Food: Food Justice and Community Change* (1st ed., pp. 33–59). University of California Press. Retrieved from <http://www.jstor.org/stable/10.1525/j.ctt19b9jf9.7>

Bader, M. D. M., Purciel, M., Yousefzadeh, P., & Neckerman, K. M.. (2010). Disparities in Neighborhood Food Environments: Implications of Measurement Strategies. *Economic Geography*,*86*(4), 409–430. Retrieved from <http://www.jstor.org/stable/40929682>

Crooks, A. T., & Wise, S. (2013). GIS and agent-based models for humanitarian assistance. Computers, Environment and Urban Systems, 41, 100-111.

Ghirardelli, A., Quinn, V., & Foerster, S. B. (2010). Using Geographic Information Systems and Local Food Store Data in California’s Low-Income Neighborhoods to Inform Community Initiatives and Resources. American Journal of Public Health, 100(11), 2156–2162. <http://doi.org/10.2105/AJPH.2010.192757>

Moreira, P. A., & Padrão, P. D. (2004). Educational and economic determinants of food intake in Portuguese adults: a cross-sectional survey. BMC Public Health, 4(1), 1.

Leung, C.W., Blumenthal, S.J., Hoffnagle, E.E., Jensen, H.H., Foerster, S.B., Nestle, M., Cheung, L.W., Mozaffarian, D. and Willett, W.C., 2013. Associations of food stamp participation with dietary quality and obesity in children. Pediatrics, 131(3), pp.463-472.

Guthrie, J., Andrews, M., Frazão, E., Leibtag, E., Lin, B., Mancino, L., Nord, M., Prell, M., Smallwood, D., Variyam, J., & Ver Ploeg, M. (2007). Can Food Stamps Do More To Improve Food Choices? An Economic Perspective. USDA Economic Research Service, Economic Information Bulletin Number 29-1.