Where's dinner coming from? A utility-based investigation of access to nutrition in Utah.

Gregory S. Macfarlane[a,*], Emma Stucki[a], Alisha H. Redelfs[b], Lori A. Spruance[b]

[a] *Civil and Construction Engineering Department, Brigham Young University, Provo, Utah USA, 84602*

[b] *Public Health Department, Brigham Young University, Provo, Utah USA, 84602*

Abstract

Convenient access to high-quality nutrition is a critical element of public health as well as an important interface between communities and the transportation system. In this research, we seek to construct a detailed picture of the nutrition environment in three communities in Utah, alongside the community members' ability to access that environment through multiple transportation modes. In doing so we construct a utility-based accessiblity model enabled by modern mobility device data. This model reveals the tradeoffs between the quality and price of goods on one hand and the distance traveled to reach them on the other. We then apply this model to a series of potential access-improving policies: building a new store, improving an existing store, and improving the non-automobile transport network between residents and existing stores. The results show that new or improved store locations bring substantially higher benefits than improvements to the transportation system, at likely lower costs. The findings suggest that transportation agencies work to increase the availability of community-sized grocery stores in low-access areas, and consider activity-based methods of measuring resource access.

Keywords: Accessibility, Utility-based access, Access to nutrition, Passive location data

1. Introduction

The ability of people to access quality nutrition has been studied at length in public health and urban geography for decades (Beaulac et al., 2009; Walker et al., 2010). This interest is motivated in large part by an observed spatial disparity in nutrition access in many communities — though this issue may be particularly pronounced in the United States (Beaulac et al., 2009). The spatial disparity has been linked at an aggregate level with negative public health outcomes (Chen et al., 2016; Cooksey-Stowers et al., 2017), though other complicating factors including prices and habits may be present as well (Ghosh-Dastidar et al., 2014).

*Corresponding author
Email address: gregmacfarlane@byu.edu (Gregory S. Macfarlane)

At the same time, access to nutrition and to community resources in general is frustratingly hard to define. "Accessibility" is an abstract concept without a specific quantitative definition (Handy & Niemeier, 1997). However, using accessibility as a policy measure requires comparative quantification, and transportation and public health researchers have constructed several quantitative measures, such as presence of a store within a travel time buffer, or the distance to the nearest store. These measures are relatively easy to calculate using readily available GIS software, but elide much useful information (Dong et al., 2006; Logan et al., 2019). These types of measures require the researcher to make a series of assumptions and assertions: why is 30 minutes chosen instead of 40? Is that time by transit or highway or walking? Should these definitions change for individuals in different socioeconomic groups? And people do not always go to the closest grocery store to begin with (Clifton, 2004; Hillier et al., 2011); how much further are people willing to travel to go to a store that is cheaper or that has a wider variety of goods? And perhaps the home location isn't the only spatial point of reference (Liu et al., 2022). A measure that potentially combines many of these different considerations is desirable.

In this research, we develop and explore an accessibility measure based on destination choice models estimated for three distinct communities in Utah. This methodology is based on a unique dataset made by linking between three extensive data sources:

1. A detailed survey of the nutrition market in three Utah communities.
2. Location-based services data derived from mobile phone records revealing which grocery stores are frequented by residents of different neighborhoods.
3. Multi-modal network data providing detailed mobility data by car, walking, and public transit.

These data will be combined in order to develop accurate logit models that demonstrate the variables that are significant to grocery store choice in Utah. These models could then be used to find accessibility to stores and impact transportation policy to improve quality of life for all communities in Utah.

This paper is organized in a typical manner. A methodology for data collection and modeling is described in Section 2 and a description of the nutrition environment and choice models estimates follows in Section 3. Section 4 presents a series of scenarios to which we apply the models estimated in Section 3, illustrating the interrelated elements of nutrition quality and transportation infrastructure in developing more complete access to nutrition. Section 5 places the findings of this research in context, while discussing limitations to the methodology and associated future research opportunities.

2. Methods

The nutrition access literature is long and has been approached from numerous angles including public health, urban science and economics, and social justice. In general, researchers have sought to link spatial access to nutrition with health outcomes including obesity, caloric intake, and the like. Complete — though somewhat dated — reviews of this literature can be had from (Beaulac et al., 2009) and (Walker et al., 2010). More recent work has extended the description and refinement of the measures used to evaluate food access and control for confounding variables. (Widener & Shannon, 2014) considered that temporal access to quality food is as important as spatial access. (Aggarwal et al., 2014) suggested that spatial access was not as important as store choice, given that most people were not observed to shop at the nearest vendor. By contrast, (Chen et al., 2016) compared spatial access to quality food vendors with observed food expenditures and showed poor access explained obesity even when controlling for consumption. (Cooksey-Stowers et al., 2017) jointly pursued spatial proximity with quality of offerings and showed the latter might be more predictive of obesity rates.

What has not been frequently attempted in the nutrition access literature, however, is a serious comparison of multiple alternative policies to address the problem, which would require a multi-dimensional analysis of spatial access, store quality, and observed tradeoffs between the two. (Macfarlane et al., 2021) illustrated the potential for a utility-based model of access to establish relationships between urban green space access and health, and then continued that methodology into a policy analysis of park space during Covid-19 (**macfarlane2022a?**). The potential for application of this methodology to the nutrition literature is well-motivated by the previous attempts as well as the lack of clear policy solutions (Wright et al., 2016).

This section describes how we construct a model of access to grocery stores in communities in Utah. We first describe the theoretical model, and then describe data collection efforts to estimate this model and apply it.

2.1. Model

A typical model of destination choice (Recker & Kostyniuk, 1978) can be described as a random utility maximization model where the utility of an individual i choosing a particular destination j is

$$U_{ij} = \beta_s f(k_{ij}) + \beta_x(X_j) \tag{1}$$

where $f(k_{ij})$ is a function of the travel impedance or costs from i to j and X_j represents the location attributes of j. The coefficients β can be estimated given sufficient data revealing the choices of individuals. The probability that individual at location i will choose alternative j from a choice set J can be estimated with a multinomial logit model (MNL) (McFadden, 1974),

$$P_i(j) = \frac{\exp(U_{ij})}{\sum_{j' \in J} \exp(U_{ij'})} \tag{2}$$

The overall fit of the model can be described with the Akaike Information Criterion (AIC) — which should be minimized — or by the McFadden likelihood ratio $\rho_0^2 = 1 - \ln \mathcal{L} / \ln \mathcal{L}_0$. In this ratio $\ln \mathcal{L}$ is the model log-likelihood and $\ln \mathcal{L}_0$ the log-likelihood of an alternative model where all destinations are equally likely; a higher ρ_0^2 value indicates more explanatory power relative to this null, random chance only model.

The idea of using destination choice logsums as accessibility terms is not new, and the theory for doing so is described in (Ben-Akiva & Lerman, 1985, p. 301). Effectively, the natural logarithm of the denominator in Equation 2 represents the consumer surplus — or total benefit — available to person i:

$$CS_i = \ln \left(\sum_{j \in J} \exp(U_{ij}) \right) \tag{3}$$

A difference in logsum measures may exist for a number of reasons that affect the utility functions described in Equation 1. For example, individuals at different locations or with different mobility will see different impedance values k_{ij} and therefore affected utility. Changes to the attributes of the destinations X_j will likewise affect the utility.

Despite the relative maturity of this theory, applications of utility-based access in the literature are still rare, outside of public transport forecasting analyses (Geurs et al., 2010). The rarity is likely explained by an unfamiliarity with destination choice models and the ready availability of simpler methods on one hand (Logan et al., 2019), and the difficulty in obtaining a suitable estimation dataset for particular land uses on the other (Kaczynski et al., 2016). This second limitation has been somewhat improved by a new methodology developed by (**macfarlane2022a?**), relying on commercial location-based services data to estimate the affinity for simulated agents to visit destinations of varying attributes and distances.

2.2. Data

In this research, we develop a unique dataset to estimate the destination choice utility coefficients for grocery store choice in three different communities in Utah. The three communities were selected to maximize potential observed differences in utility between community residents. The three communities are Utah County, West Salt Lake County, and San Juan County. Note that in this document we refer to the second community as "Salt Lake" even though this does not refer to the entire Salt Lake County nor to Salt Lake City, rather, we focus on communities in the western part of the valley, such as Magna, Kearns, and West Valley City. The communities are shown in a wider context in Figure 1.

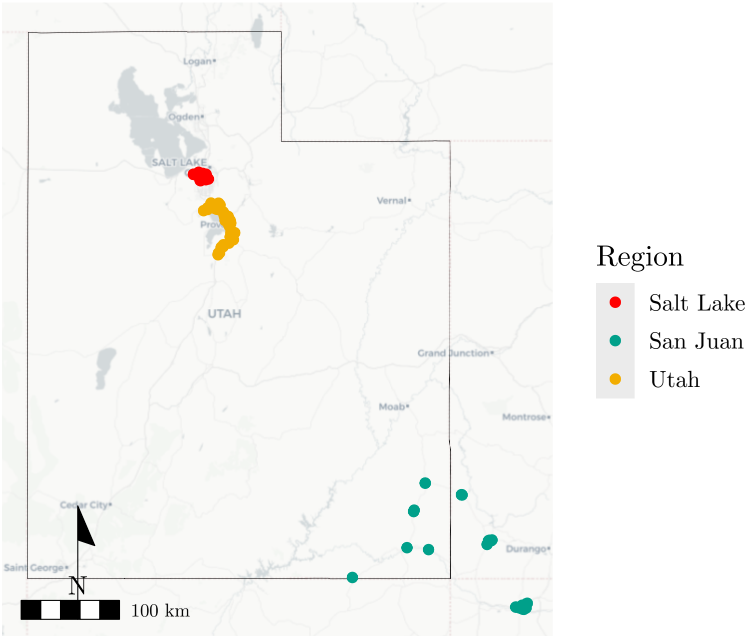


Figure 1: Location of study regions in Utah.

Table 1: Demographic Statistics of Study Regions

	Utah	Salt Lake	San Juan
Total population	627,098	655,830	7,091
Total households	171,538	216,731	2,090
Housing units per sq. km	599	831	103
Median income	79,453	64,868	58,586
Percent minority individuals	18	36	26

Table 1 shows several key population statistics based on 2021 American Community Survey data for block groups in the three communities of interest. Utah County is a largely suburban county with high incomes and a low percentage of minority individuals. The Salt Lake region is more dense with somewhat lower incomes and household sizes but a high share of minority individuals. San Juan County is primarily rural, with a few small communities and a large reservation for the Navajo Tribe.

Estimating the utility model described in Equation 2 for grocery stores requires three interrelated data elements:

1. An inventory of grocery store attributes X_j;
2. A representative travel impedance matrix K composed of all combinations of origin i and destination j;

3. A database of observed person flows between i and j by which to estimate the β coefficients. We describe each of these elements in turn in the following sections.

2.2.1. Store Attributes

The store attributes were collected using the Nutritional Environment Measures Survey — Stores (NEMS-S) tool (Glanz et al., 2007). This tool was developed to reveal significant differences in the availability and cost of healthy foods in an environment, and has been validated for this purpose. Beyond superficial attributes such as the store category (dollar store, convenience store, ethnic market, etc.) and the number of registers, the NEMS-S collects detailed information about numerous store offerings such as the availability of produce, dairy products, lean meats, juices, and canned and dry goods of various specific types. Of particular interest to the survey are availability and price differentials of lower-fat alternatives: for example, the survey instrument requests the shelf space allocated to milk products of various fat levels and the price of each product.

Student research assistants collected the store attributes by visiting grocery stores, dollar stores, ethnic markets, and other food markets in the three communities of interest described above. Stores were identified using internet-based maps combined with in-person validation and observation. The student researchers completed the NEMS-S instrument with the aid of a digital survey and a tablet computer. Each researcher who collected data was trained to use the survey at a control store in Provo, and the training data helped to eliminate the risk of surveyor bias. The store attributes were collected in the spring of 2021 for Utah County and spring of 2022 for Salt Lake and San Juan Counties. In Utah and Salt Lake Counties, we included dollar stores and grocery stores but did not include convenience stores. Given the rural nature of San Juan County, we made two adjustments to capture the entirety of the nutrition environment. First, we included convenience stores and trading posts if they were the only food market in a community. We also included full-service grocery stores in Cortez, Colorado, and Farmington, New Mexico in the San Juan data collection, as community conversations made it clear that many residents will drive these long distances for periodic shopping with greater availability and lower prices.

Using the information in the NEMS-S survey, two measures of a store can be calculated: an availability score based on whether stores stock particular items as well as lower-calorie options; and a cost score describing the spread between prices of these options. These score values are described in (Lunsford et al., 2021), and we developed an R package to compute the scores; this package is available at https://github.com/byu-transpolab/nemsr. In the availability score, each store is given a value for whether or not there are more healthful options available in the store, such as low-calorie chips, or low-fat milk. If the store does not have a more healthful option in a category it receives a lower score, so stores with more availability of healthful food options will receive a higher availability score. For the cost score, the measure is the price spread between healthful and less healthful options: if the price of whole wheat bread is cheaper than white bread, the store receives positive points for the cost option, if the price is the same then zero points are awarded, and if the wheat bread is more expensive then the store receives negative points. Thus a store with a higher availability and cost score will have both more healthful options, and a more advantageous pricing scheme towards those options.

One important store attribute that the NEMS-S instrument does not collect or compute directly is a measure of the cost of common goods that can be compared across stores. We therefore used the data collected from the NEMS-S instrument to construct a market basket-based affordability measure that could be compared across stores, following the approach of (Hedrick et al., 2022). This market basket score is based on the US Department of Agriculture (USDA) 2021 Thrifty Food Plan (FNS, 2021), which calculates a reference market basket for a family of four. Because this market basket contains more (and sometimes different) items than what the NEMS-S instrument requests, we chose relevant items from our NEMS-S data as replacements. For example, the USDA market basket contains a certain amount of poultry, but the NEMS-S score collects the per-pound cost of ground beef at various fat contents. For any stores that were missing any of the elements in the market basket, we first substituted with another ingredient that would fit the nutritional requirements. If no substitute was available, we included the average price of the missing good at other stores in that community multiplied by 1.5 as a penalty for not containing the product. The final market basket score is the total cost of all foods in the market basket. These costs can then be compared from store to store to understand general affordability comparisons between stores.

Table 2 presents the store attribute data collected for each community. Utah County generally has the largest average store size (as measured by the number of checkout registers) while having the lowest market basket cost, the highest availability of healthful food (measured by the NEMS-S availability score) and the lowest difference between healthy and unhealthy food (the NEMS-S cost score). San Juan County has the smallest average stores, highest costs, and the lowest availability of healthy options, and Salt Lake falls in between.

2.2.1.1. Imputation of Missing Store Data.

We collected detailed store attributes for a complete census of stores in Utah County, San Juan County, and a portion of Salt Lake County using the NEMS-S survey instrument. These attributes form the basis of the choice models used to determine access and provide a complete picture of access in those communities, assuming people do not leave the communities for grocery trips. But understanding access in other parts of Salt Lake County – including how stores outside of the West Salt Lake County area might shape access inside that community — requires us to impute the measured attributes onto the stores that we did not directly measure.

Table 2: Grocery Store Attributes

		Utah (N=63)		Salt Lake (N=39)		San Juan (N=50)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Registers (incl. self checkout)		12.5	11.7	9.9	8.9	6.1	8.8
NEMS-S availability score		18.7	8.4	16.2	8.1	13.2	7.6
NEMS-S cost score		1.9	2.3	2.3	2.2	1.9	1.9
Market basket cost		126.1	21.5	141.6	19.2	157.6	16.8
		N	Pct.	N	Pct.	N	Pct.
Type	Convenience Store	2	3.2	0	0.0	10	20.0
	Dollar Store	5	7.9	11	28.2	15	30.0
	Grocery Store	50	79.4	27	69.2	19	38.0
	Other	6	9.5	1	2.6	6	12.0
Pharmacy	FALSE	42	66.7	32	82.1	43	86.0
	TRUE	21	33.3	7	17.9	7	14.0
Ethnic market	FALSE	55	87.3	30	76.9	47	94.0
	TRUE	8	12.7	9	23.1	3	6.0
Other merchandise sold	FALSE	52	82.5	35	89.7	47	94.0
	TRUE	11	17.5	4	10.3	3	6.0

To do this, we used web-based mapping databases (including OpenStreetMap and Google Maps) to obtain a list of grocery stores, dollar stores, and appropriate convenience stores throughout the state. From this search, we were able to determine each store's location, brand name, and store type, which we also collected in the manual data assembly efforts. Using this information, we built a multiple imputation model using the `mice` package for R (van Buuren & Groothuis-Oudshoorn, 2011). The predictor variables in the imputation included the store brand and type, as well as the average income and housing density in the nine closest block groups to the store location (based on population-weighted block group centroids and Euclidean distances).

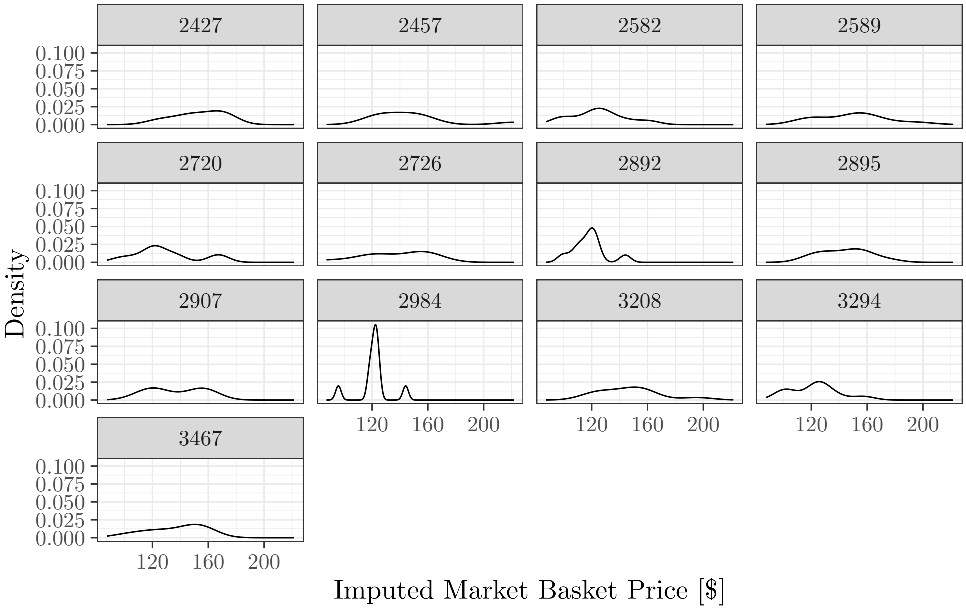


Figure 2: Imputed market price values for 12 random grocery stores.

Thirty iterations of the multiple imputation algorithm were run for each of ten independent imputations. Figure 2 shows the density of the ten imputed market basket prices for a randomly selected set of 12 stores. As the figure reveals, there is some general peaking in the predicted market price for most stores, but the imputation model still predicts a wide range of possible prices for most stores. When using the imputed data for analysis, we take the mean of the ten predictions for continuous values, and the mode for discrete values.

2.2.2. Travel Impedances

The second element of the utility equation in Equation 1 is the travel impedance between i and j. Many possibilities for representing this impedance exist, from basic euclidean distance to complex network paths. A primary purpose of the model we are developing in this research is to study comparative tradeoffs between infrastructure-focused and environment-focused improvements to the nutrition access of households. It is therefore essential that we use a travel impedance measure that can combine and compare the cost of traveling by multiple modes so that highway improvements and transit / active transport improvements can be compared in the same basic model.

Just as the log-sum of a destination choice model is a measure that sums the utility of multiple destination attributes and costs in a rigorous manner, the log-sum of a mode choice model combines the utilities of all available travel modes. In this study we assert the following mode choice utility equations:

$$\begin{aligned}
V_{\text{auto},ij} &= -0.028(t_{\text{auto},ij}) \\
V_{\text{bus},ij} &= -4 - 0.028(t_{\text{bus},ij}) - 0.056(t_{\text{wait},ij}) - 0.056(t_{\text{access},ij}) \\
V_{\text{walk,ij}} &= -5 - 0.028(t_{\text{walk},ij}) - 1.116(d_{ij<1.5}) - 5.58(d_{ij>1.5})
\end{aligned}$$

where t is the in-vehicle travel time in minutes for each mode between i and j. The transit utility function additionally includes the wait time for transit as well as the time necessary to access the transit mode on both ends by walking. The walk utility includes a per-mile distance disutility that increases for distances greater than 1.5 miles. These equations and coefficients are adapted from a statewide mode choice model for home-based non-work trips in urban and rural regions developed for UDOT research (Barnes, 2021).

The log-sum, or total weighted impedance by all modes is therefore

$$k_{ij} = \ln(e^{V_{\text{auto},ij}} + e^{V_{\text{bus},ij}} + e^{V_{\text{walk},ij}}) \tag{4}$$

In this implementation, i is the population-weighted centroid of a 2020 Census block group, and j is an individual grocery store. We measure the travel times from each i to each j using the `r5r` implementation of the R5 routing engine (Conway et al., 2017, 2018; Conway & Stewart, 2019; Pereira et al., 2021). This algorithm uses common data elements — OpenStreetMap roadway and active transport networks alongside General Transit Feed Specification (GTFS) transit service files — to simulate multiple realistic route options by all requested modes. We obtained OpenStreetMap networks and the Utah Transit Authority GTFS file valid for May 2023 and requested the minimum total travel time by each mode of auto, transit, and walking for a departure between 8 AM and 9 AM on May 10, 2023. The total allowable trip time by any mode was set to 120 minutes, and the walk distance was capped at 10 kilometers; if a particular i, j pair exceeded these parameters then the mode was presumed to not be available and contributes no utility to the log-sum.

2.2.3. Mobile Device Data

The final element of destination utility presented in Equation 1 is the set of coefficients, which are often estimated from household travel surveys in a travel demand context. It is unlikely, however, that typical household diaries would include enough trips to grocery stores and similar destinations to create a representative sample.

Emerging mobile device data, however, could reveal the typical home locations for people who are observed in the space of a particular store. (**macfarlane2022a?**) present a method for estimating destination choice models from such data, which we repeat in this study. We provided a set of geometric polygons for the grocery stores of interest to StreetLight Data, Inc., a commercial location-based services aggregator and reseller. StreetLight Data in turn provided data on the number of mobile devices observed in each polygon grouped by the inferred residence block group of those devices during summer 2022. We then created a simulated destination choice estimation dataset for each community resource by sampling 10,000 block group - grocery store "trips" from the StreetLight dataset. This created a "chosen" alternative; we then sampled ten additional stores from the same community at random (each simulated trip was paired with a different sampled store) to serve as the non-chosen alternatives. Random sampling of alternatives is a common practice that results in unbiased estimates, though the standard errors of the estimates might be larger than could be obtained through a more carefully designed sampling scheme (Train, 2009).

3. Results

This section presents results on the nutrition environment in each of the three communities of Utah County, West Salt Lake County, and San Juan County, along with destination choice model estimates and their application to creating accessibility maps of each community and the entire state of Utah.

3.1. Nutrition Environment

Though some basic descriptive statistics of the grocery store attributes were presented in Table 2, some additional exploration of these attributes is valuable to understand the nutrition environment in these three communities.

Figure 3 presents the relationship between the recorded NEMS availability score and the USDA market basket cost at the stores by community and store type. In all three communities, the relationship is strongly negative, with stores that stock more varieties of goods also having overall lower prices for those goods. This is emphasized by the bottom-right quadrants of these plots (high availability, low-cost) being dominated by full-service grocery stores, which have more availability and lower prices than convenience stores or dollar stores, but require higher traffic and demand to make up for their lower profit margins. Average prices in Utah County are lower than prices in the other two communities across the availability spectrum; this is true even after adjusting for 9.4% annual inflation between March 2021 and March 2022 in food products (Bureau of Labor Statistics, 2023).

Figure 4 shows the relationship between the NEMS availability and cost scores. In this case the relationship is generally positive, with stores that stock more healthful options also placing these options at competitive

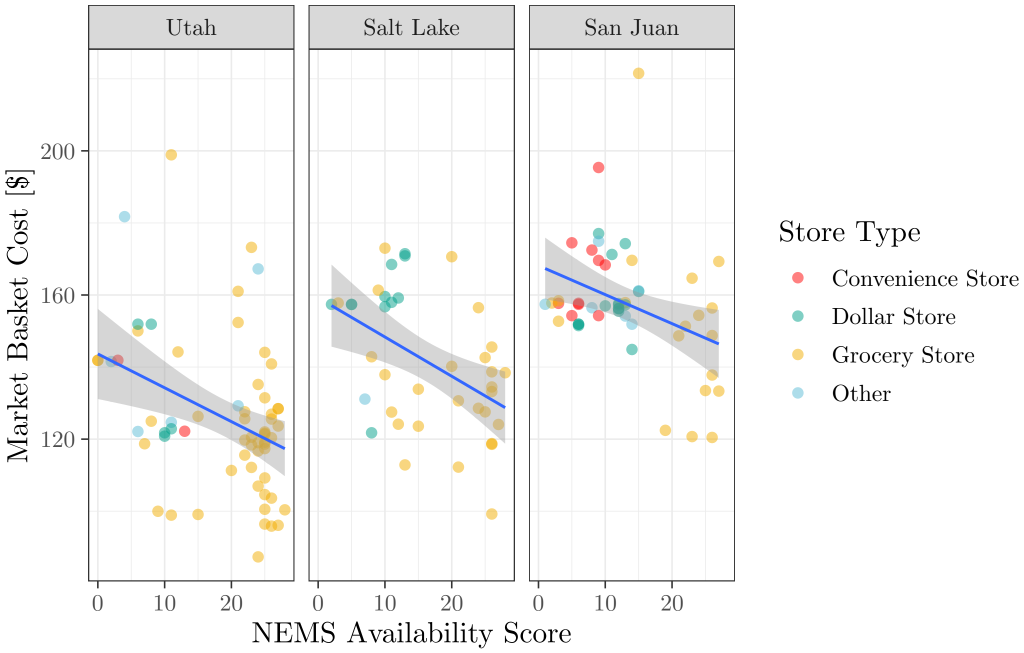


Figure 3: Relationship between NEMS availability score and market basket score in each study community. Utah county prices adjusted for 2021-2022 annual inflation.

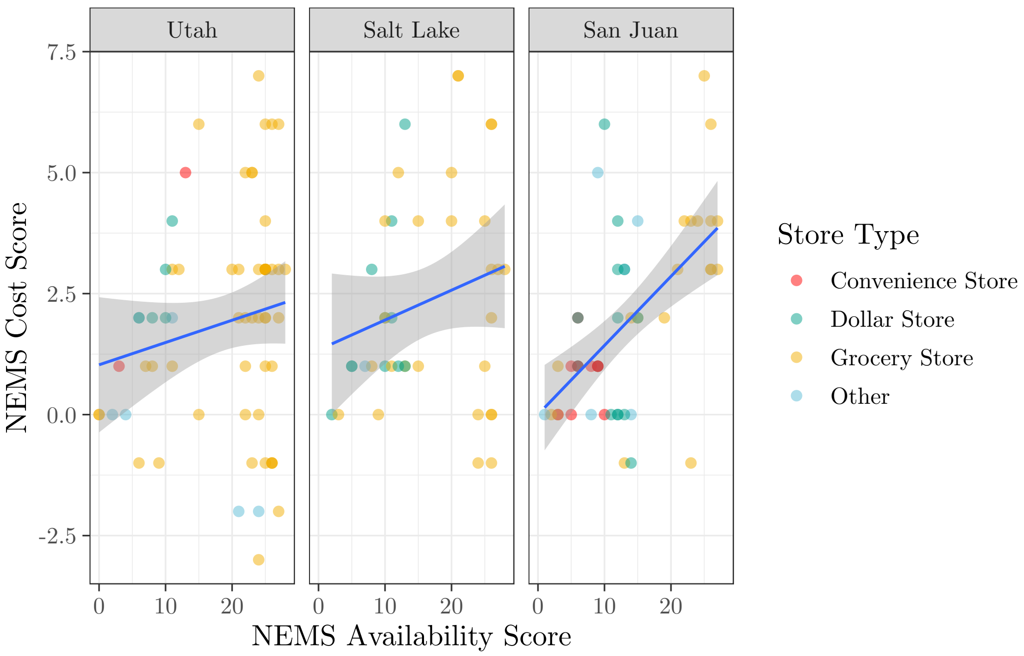


Figure 4: Relationship between NEMS availability score and cost score in each study community.

prices. Conversely, stores with fewer options tend to place the options they do stock at a higher price point. This relationship between availability and cost of healthful goods is strongest in San Juan County, with convenience stores anchoring the low-availability, high-premium quadrant for healthy food. It should be noted that these convenience stores also exist in the Utah County community, but we explicitly included them in the San Juan data collection as they are the only food markets of any kind in multiple towns, with dozens of miles separating towns from each other.

3.2. Destination Choice

Using the data collected and MNL destination choice model as described in Section 2, we estimate a series of model specifications in each community with the `mlogit` package for R (Croissant, 2020). To illustrate the role of different data elements on destination choice, we develop and estimate four different utility equations:

$$\begin{aligned}
\text{Access} &= \beta_{MCLS}(k_{ij}) \\
\text{NEMS} &= \beta_{n-a}(\text{NEMS} - \text{Availability}) + \beta_{n-c}(NEMS - Cost) \\
\text{Attributes} &= \beta_{mkt}(\text{MarketBasket}) + \beta_{reg}(\text{Registers}) + \beta_{type}(\text{Type}) \\
\text{All} &= \text{Access} + \text{NEMS} + \text{Attributes}
\end{aligned}$$

The Access model includes only the mode choice logsum described in Equation 4. The NEMS model includes the NEMS cost and availability scores describing the goods the store offers, while the Attributes model contains information that might be more conventionally available to shoppers including the size, type, and average prices at the store. As the nutrition environment in each community contains different types of stores, the specific type coefficients differ by community. The All model contains all of the other three sets of estimated coefficients.

Table 3 presents the estimated coefficients in the Utah County community. In general, the utility coefficients are statistically significant and in a direction that would be expected by informed hypothesis. The Access model has a positive coefficient on its mode choice log-sum term, which indicates that as the mode choice logsum between a block group and a store increases — indicating lower travel costs between Census block groups and the store, because travel times in Equation 4 have a negative relationship with utility — a higher proportion of mobile devices residing in that block group are observed to travel to that store. The NEMS model shows a positive relationship between both environment variables and utility, indicating that people are more likely to choose stores with higher availability of healthy goods and more advantageous prices for those goods, all else equal. The Attributes model suggests that people are less willing to visit stores with higher prices, fewer registers, and convenience stores or other non-standard grocery stores with the exception of dollar stores, which they are *more* attracted to. Combining all of these variables in the All model retain the

Table 3: Estimated Models of Utah County

	Access	NEMS	Attributes	All
Mode Choice Log-sum	8.063**			8.686**
	(95.356)			(87.770)
NEMS Availability Score		0.032**		-0.024**
		(22.020)		(-7.369)
NEMS Cost Score		0.035**		0.041**
		(7.688)		(5.946)
USDA Market Basket			-0.007**	-0.008**
			(-9.814)	(-8.855)
Registers			0.056**	0.065**
			(54.586)	(41.357)
Store Type: Dollar Store			1.948**	2.086**
			(56.866)	(41.000)
Store Type: Convenience Store			-1.990**	-2.422**
			(-7.400)	(-8.495)
Store Type: Other			-1.785**	-1.813**
			(-12.641)	(-11.991)
AIC	31,874.44	49,051.55	41,985.6	25,402.43
ρ_0^2	0.359	0.013	0.155	0.489

* $p < 0.05$, ** $p < 0.01$

t-statistics in parentheses

Table 4: Estimated Models of West Salt Lake Valley

	Access	NEMS	Attributes	All
Mode Choice Log-sum	9.870**			12.044**
	(74.095)			(74.423)
NEMS Availability Score		0.129**		0.002
		(71.898)		(0.580)
NEMS Cost Score		-0.036**		0.057**
		(-7.857)		(9.607)
USDA Market Basket			-0.010**	-0.006**
			(-13.065)	(-6.953)
Registers			0.104**	0.129**
			(70.019)	(52.110)
Store Type: Dollar Store			0.325**	0.494**
			(7.250)	(8.943)
Store Type: Other			0.248*	0.556**
			(2.211)	(4.614)
AIC	42,820.66	42,541.44	40,195.01	31,863.31
ρ_0^2	0.138	0.144	0.191	0.359

* $p < 0.05$, ** $p < 0.01$

t-statistics in parentheses

significance, direction, and basic scale of all previous estimates with the exception of the NEMS availability variable. In this case, it seems that the previous positive relationship may have been a result of correlation between NEMS availability and other variables such as cost or the number of registers. And when controlling for all other variables, the role of transportation access becomes somewhat more important than considering only distance alone, implying that people are willing to travel somewhat further for stores with attributes they value.

The overall fit of the four models in Table 3 is also revealing: the model with only NEMS variables against almost no predictive power over randomly selecting any store in the community (as revealed by the ρ_0^2 statistic). Though all sets of variables contribute to the overall fit, it is apparent that the bulk of model explanatory power is due to transportation proximity.

Table 4 presents the estimated coefficients in the west Salt Lake County community, and Table 5 presents

Table 5: Estimated Models of San Juan County

	Access	NEMS	Attributes	All
Mode Choice Log-sum	0.709**			1.205**
	(81.466)			(73.616)
NEMS Availability Score		0.139**		0.065**
		(63.869)		(12.857)
NEMS Cost Score		0.227**		0.055**
		(35.386)		(6.721)
USDA Market Basket			-0.011**	-0.033**
			(-15.376)	(-24.850)
Registers			0.022**	0.049**
			(20.517)	(25.969)
Store Type: Dollar Store			-2.222**	-1.377**
			(-46.835)	(-18.903)
Store Type: Convenience Store			-3.597**	-1.080**
			(-31.496)	(-7.747)
Store Type: Other			-1.451**	-1.331**
			(-28.680)	(-19.097)
AIC	40,811.4	35,119.64	37,230.81	23,556.77
ρ_0^2	0.179	0.293	0.251	0.526

* $p < 0.05$, ** $p < 0.01$

t-statistics in parentheses

the estimated coefficients in San Juan County. The same general story about coefficient direction and hypotheses applies in both of these communities, except in regards to the NEMS variables. In Salt Lake, the NEMS cost score appears negative when estimated alone but becomes positive when other variables are included. In San Juan, these variables are consistently positive. Additionally, the story of model fit is reversed: in both Salt Lake and San Juan, the attributes of the store explain more of the model fit than the transportation impedance term.

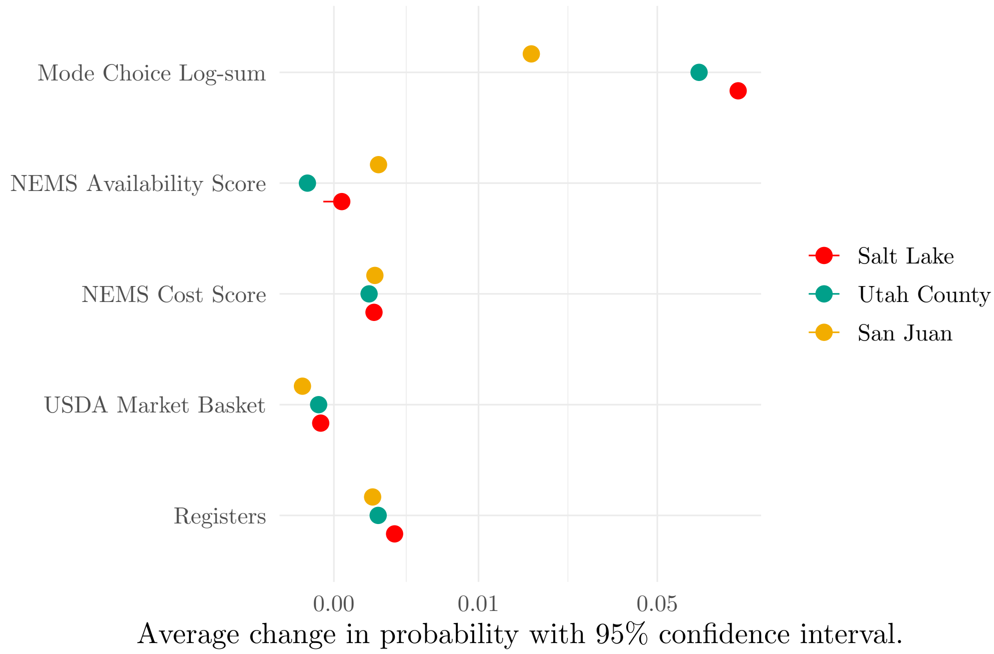


Figure 5: Average change in the probability of the chosen alternative with respect to a 100% change in model variables.

To better visualize how the preferences in the three communities differ from each other, Figure 5 plots the average marginal elasticity on the choice probability of the chosen store with respect to many variables in the "All" model for each county. Specifically, we increased the variable of interest by 100% for all synthetic choice makers, and then calculated the average change in the probability of the chosen alternative for all choice makers with its sample confidence interval. The mode choice log-sum elasticity is strongly significant in all three communities (and is the most influential variable on choice), but it has its smallest value in San Juan County where people often must travel long distances to reach any stores. The highest mode choice log-sum value is in Salt Lake, but this explains a smaller proportion of the model outcomes than the lower value in Utah County; a possible hypothesis for this observation may include the higher density of stores in Salt Lake — attributes are more important when so many stores are close together — paired with the somewhat lower vehicle ownership in that community driving up the coefficient value.

3.3. *Accessibility*

With the models estimated in Section 3.2, we can evaluate the spatial access of each community. Figure 6 shows the value of the grocery store destination choice log-sum for block groups in Utah County. Unsurprisingly, the block groups in the core of the urban areas of the region have the highest access to grocery stores, because this is where the stores are located and also where the transportation access to multiple destinations is highest. This map also contains somewhat interesting implications for the equity of access. A perhaps unique feature of Utah County's demographic geography is that the wealthiest neighborhoods tend to be located on the mountain benches east of the main urban areas. This means that in Utah County, at least, the neighborhoods with the lowest access to grocery stores are actually some of the wealthiest neighborhoods with the lowest concentrations of ethnic minorities in the region.

Of course, much of this high access in the urban core of Utah County is achieved by cheap and available automobile transportation. We can consider what access looks like for those without cars by re-computing the mode choice log-sum described in Section 2.2.2 between all block group / store pairs but eliding the automobile mode, and examining the resulting impact to destination choice utility. Figure 7 shows the results of this analysis: whereas the total access (with car included) is a smooth gradient across the valley, the access for individuals without vehicles is blocky and discontinuous, with neighborhoods of relatively good access immediately next to neighborhoods with bad or non-existent access. This may reflect the discontinuous nature of active transport and public transit facilities in the region, as well as the auto-dominated locations of many grocery stores. Note also that even for neighborhoods of relatively good non-vehicle access, the destination choice log-sum value is substantially lower than the logsum with vehicle access; the minimum value on with vehicles is just below 0, whereas the *maximum* log-sum without vehicles is around -100. Because the log-sum occurs on the same scale in both cases, this represents a serious additional cost for non-vehicle users.

4. Application

In this section, we develop a series of scenarios to which we apply the models estimated in Section 3. These scenarios are constructed to ascertain what may be the best strategy to improve nutrition access in a community. We first describe how each scenario was constructed, and then discuss the results together.

The accessibility of each scenario is determined using the approximated utility coefficients of the All model estimated in Section 3,

$$U_{ij} = \beta_{MCLS}(k_{ij}) + \beta_{n-a}(\text{NEMS} - \text{Availability}) + \beta_{n-c}(NEMS - Cost) + \beta_{mkt}(\text{MarketBasket}) + \boldsymbol{\beta}_{type}(\text{Type}) \tag{5}$$

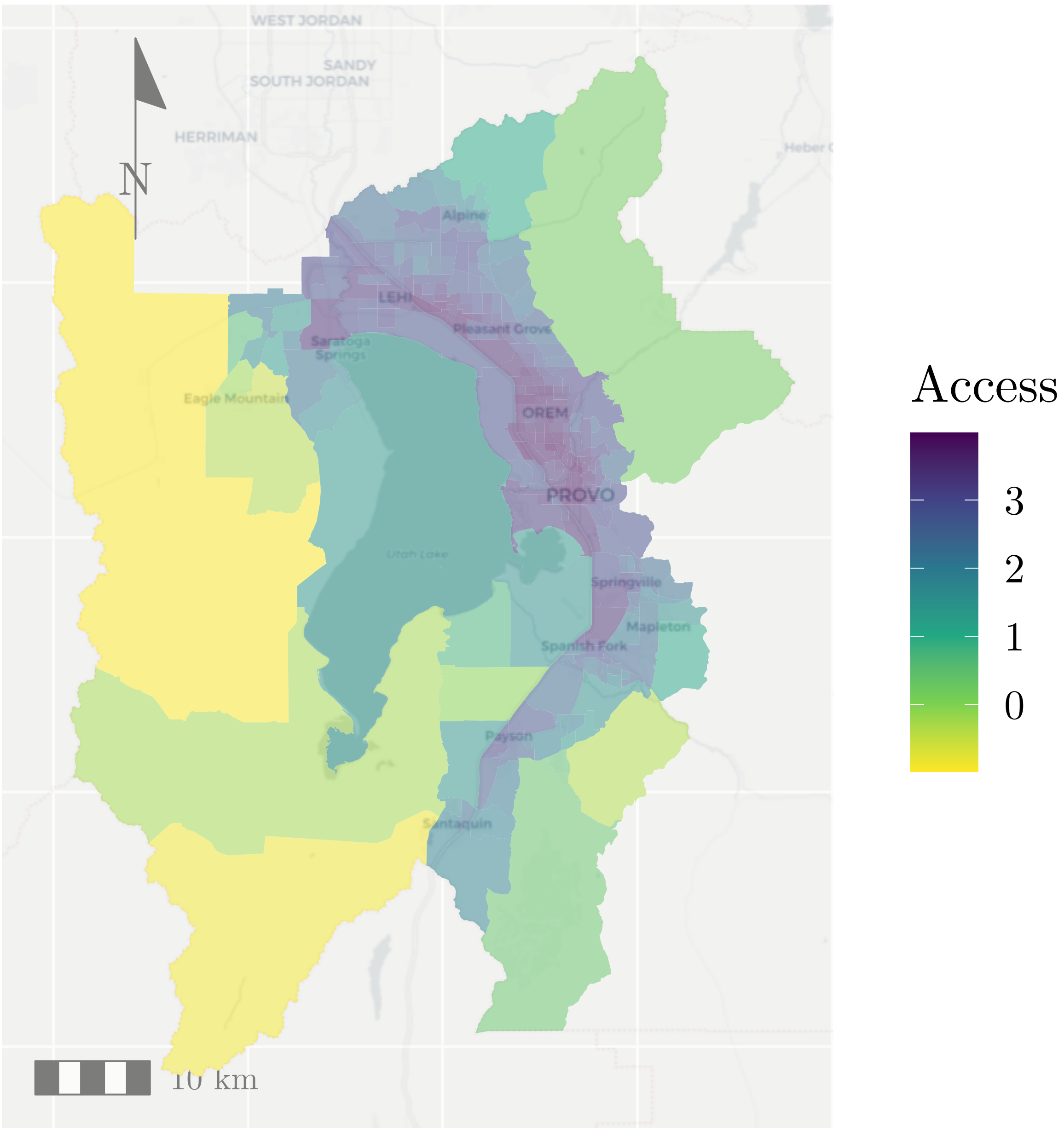
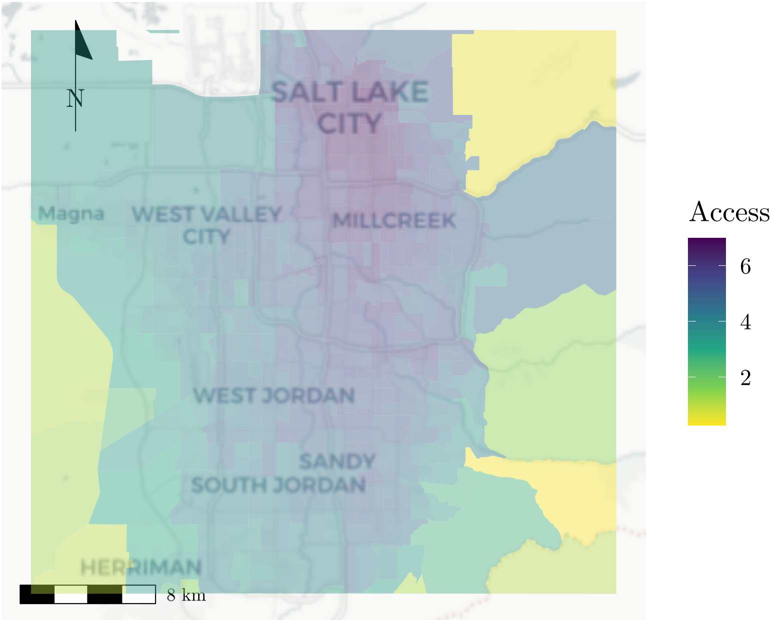
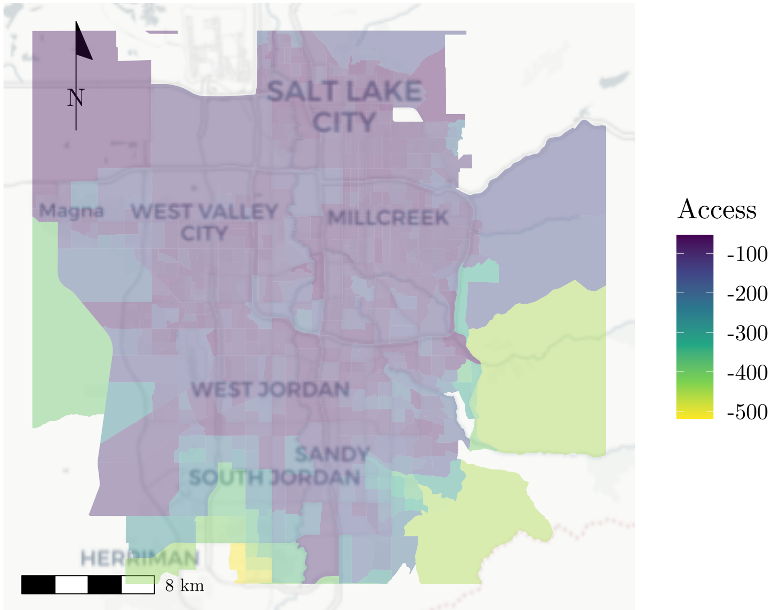


Figure 6: Modeled access to grocery stores in Utah County



(a) With vehicles



(b) Without vehicles

Figure 7: Access to groceries in Salt Lake County with and without a vehicle.

and the total access benefit is defined as the consumer surplus given in Equation 3. Note that this benefit is denominated in units of utility (Ben-Akiva & Lerman, 1985), and the coefficients in Equation 5 serve to convert between the units of the variable and utility. Specifically, the β_{mkt} coefficient represents how many dollars of grocery cost a person is willing to spend to increase their utility by one unit.

In actuality, the utility formula in Equation 5 is a relative utility added by an unknown constant, $U_{ij} = f(\beta, X_{ij}) + C$, but this C term is included in all the alternatives and therefore cancels out (Train, 2009) in estimation. This means we cannot assess the *absolute* value of utility, but we can assess the *relative* monetary benefit of the difference between two scenarios as

$$\text{Benefit} = \sum_i \left(-\frac{\omega_i}{\beta_{mkt}} (CS'_i - CS_i) \right)$$

with the consumer surplus of the "improved" scenario in each origin zone i indicated as CS'_i and the unimproved counterpart as CS_i, a weight ω accounts for the population in the zone, and the β_{mkt} converts the difference in utility into a dollar amount.

4.1. Scenario Descriptions

There are three general strategies we develop scenarios around:

1. Erect a new grocery store in the community, in a place where one does not already exist.
2. Improve an existing convenience store or dollar store so that it has the attributes of a full-service grocery store.
3. Improve the transit and non-motorized access to stores in a region.
4. Increase the availability of grocery delivery in the region.

We implement all three strategies in scenarios on the west Salt Lake valley community. We also implement strategy 2 (an improved store) in the San Juan and Utah County communities for comparison.

4.1.1. Erect a new store

This strategy assumes that the nutrition environment would benefit from a new store located in a place that presently has low grocery store access. To examine the potential for this strategy to improve access to nutrition in each community, we calculate the change in destination choice log-sum when a new store is added to the region in a location where access is currently poor. The new store is a full-service grocery store with a number of registers equal to the mean of other grocery stores in the community, and NEMS availability score, NEMS cost score, and market basket cost equivalent to the 75^{th} percentile for the community. Thus the store is expected to be better-than-average quality as perceived by the residents of the community. The location for this new store is at 4100 S and 2700 W in West Valley City.

4.1.2. Improve an existing store

This strategy assumes that existing stores are in locations that the community values and can access, but that those stores may not have high availability of quality goods. To examine the potential for this strategy to improve access to nutrition, we improve the attributes of an existing dollar store in the community so that it has the size, prices, and availability of goods as a full-service grocery store. As above, we create a full-service grocery store with a number of registers equal to the mean of other grocery stores in the community, and NEMS availability score, NEMS cost score, and market basket cost equivalent to the 75^{th} percentile for the community. Thus the store is expected to be better-than-average quality as perceived by the residents of the community; the difference from the previous scenario is that the improved store takes the place of an existing convenience store or dollar store.

The improved stores are at the following locations in each community:

- An ethnic store near 2700 W 3500 S in West Valley City (Salt Lake)
- A small grocery store in Santaquin (Utah County)
- A dollar store in Blanding (San Juan County)

4.1.3. Improve transit and non-motorized transport

This strategy assumes that people cannot easily travel to existing stores because they cannot or do not drive for a variety of reasons, and that the public and active transport networks provide an insufficient level of service. To examine the potential for this strategy to improve access to nutrition, we improve the travel time costs in the Salt Lake community for non-motorized and public transportation in the region and calculate the change in destination choice log-sum.

For active transportation, the lack of pedestrian facilities across and alongside roads both in reality and in the OpenStreetMap dataset may substantially increase measured walk distances and times. In this scenario, we replace the times measured from OpenStreetMap using R5 with an idealized distance function,

$$t_{\text{walk}} = \frac{\sqrt{2} * d'_{ij}}{v_{\text{walk}}} \tag{6}$$

where d'_{ij} is the Euclidean (straight-line) distance between i and j and v_{ij} is an average walking speed equivalent to 3.5 feet per second (Fitzpatrick et al., 2006). The distance is multiplied by the square root of 2 to reflect the Manhattan distance (along a gridded street system). We retain the cap on walking distance at 10 kilometers. Though this distance may radically understate the real walking distance, we are trying to create an idealized scenario of effectively frictionless active transport. For public transit, we assume that the frequency of service is such that all transfer and initial wait times are at most 5 minutes, and that no person must walk more than 10 minutes to access their first public transport service. Presently, the *mean* walk access time in the scenario region is over 20 minutes, and the wait time over 10 minutes.

Travel times are improved in this way for all block group — store pairs in the west Salt Lake Valley community.

4.1.4. Grocery delivery services

This strategy assumes that the most effective way to bring quality nutrition to an inaccessible location is not to change the distribution of stores or the quality of transport, but to support and enable delivery services that will bring the nutrition to individuals. To model this scenario, we assume that 20% of the full-service grocery stores in the west Salt Lake Valley community offer a delivery service. This raises the price of the market basket by $10, but reduces the travel time on all modes to 0 if the store is within 3 kilometers. All of these assumptions are imperfect: a local grocery chain presently offers services at a $7 delivery fee, which is not paid by all customers; additionally, delivery services require that customers expend time placing online orders; and service providers in the region typically place a long-distance charge rather than a service area limit. Regardless, these assumptions will work for the purpose of this exercise.

4.2. Scenario Results

Using the methodology described above, we recalculated the destination choice log-sum value for each block group under each scenario, and compared the change in accessibility resulting from the improvement.

Figure 8 shows the geographic distribution of benefits associated with locating a new store at a site in the Salt Lake community. The benefits are largest immediately next to the new store, where they exceed 3 for each household each time the household makes a trip to a grocery store.

Figure 9 shows the results of the scenario improving an existing store in the Salt Lake community. Compared to the results of the new store scenario, the scale of the benefits are not as substantial (a maximum per-household-trip benefit of less than $1), and seem to not cover quite as large a geographic region. Figure 10 shows the results of improving a store in Utah and San Juan Counties. As in Salt Lake, the benefits are most strongly concentrated in the immediate vicinity of the improved store. One interesting observation — especially in Utah County — is that the improvements are felt more strongly in the block groups near the improved store that have lower availability of other options. The block group in Utah County directly containing the improvement sees a per-household-trip benefit well over $2, considerably more than the maximum benefit in Salt Lake. This is intuitive, as the improvement of a store matters less if the stores close to you are already sufficient.

The results of the third scenario, improving the access of non-motorized and public transit access to stores, are shown in Figure 11. This benefit is spread over a larger area, and is concentrated on the 35 MAX bus rapid transit corridor where the improvement in walk access time to transit couples with high frequency

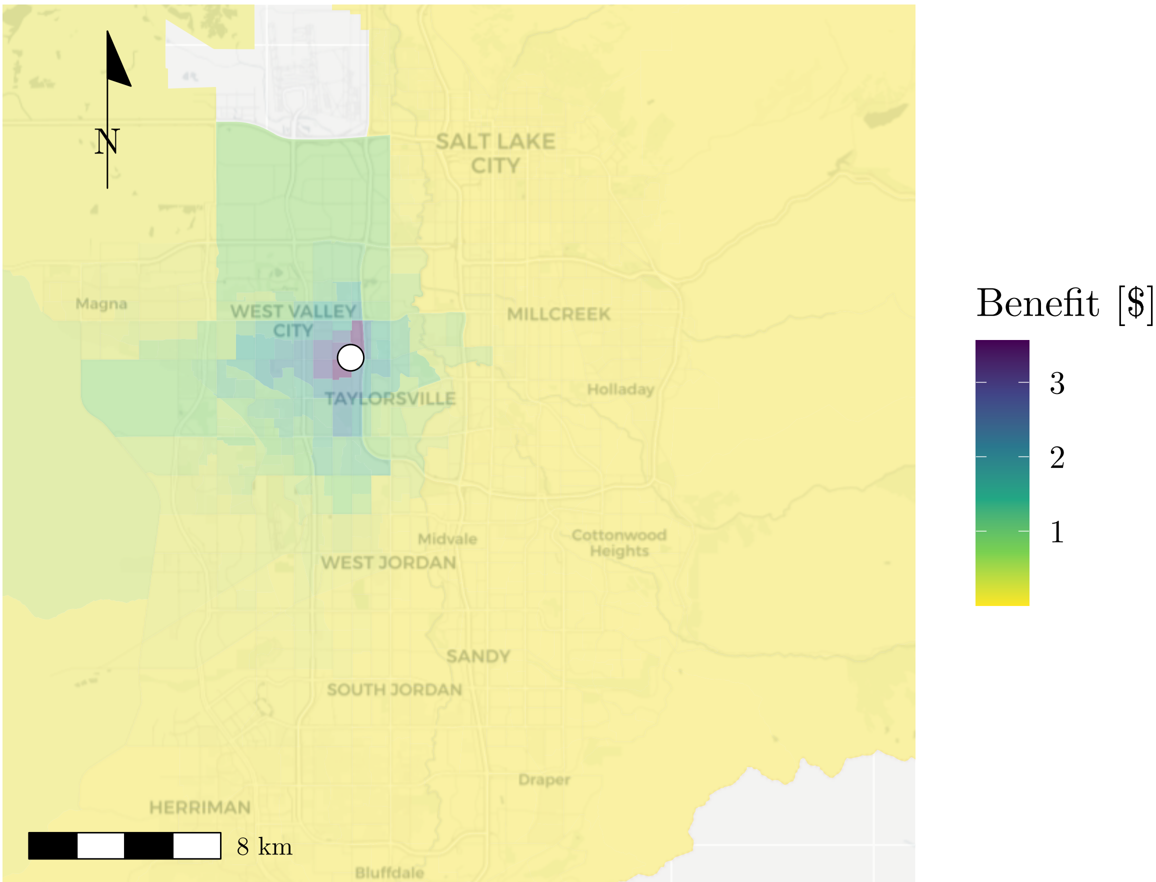


Figure 8: Estimated per-household benefits of adding new full-service store in west Salt Lake Valley

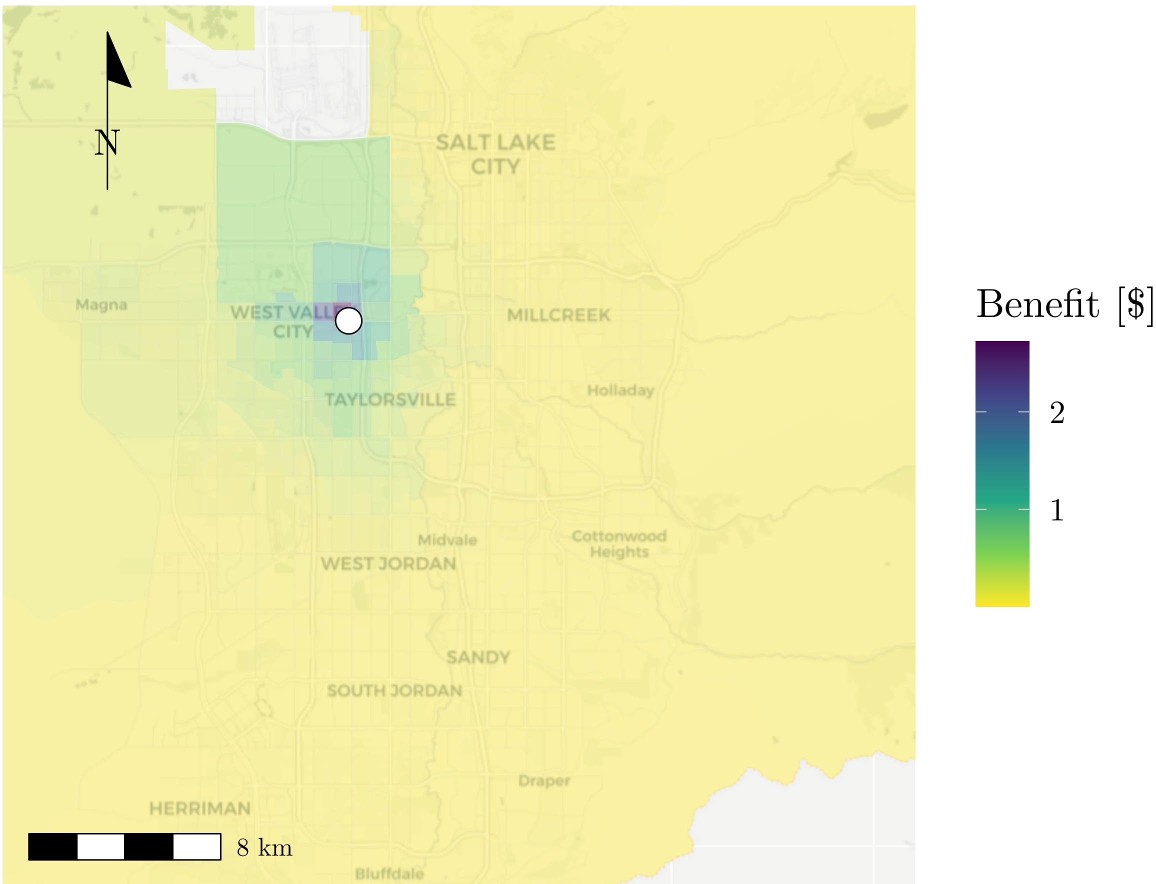
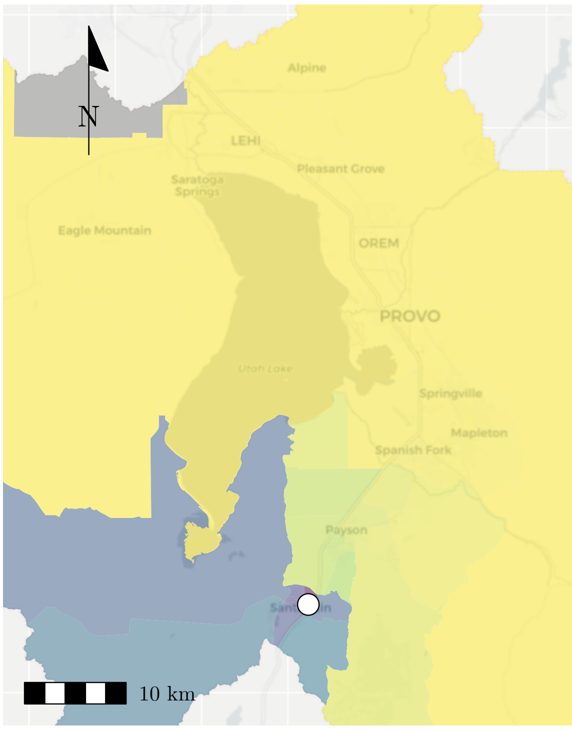
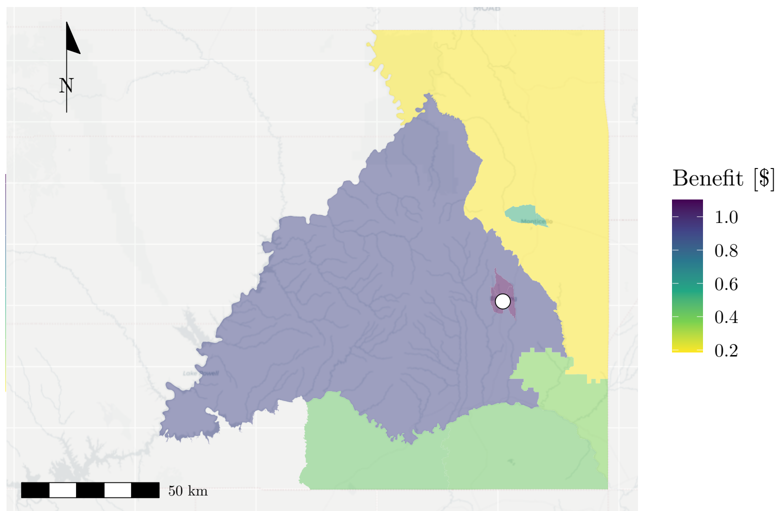


Figure 9: Estimated per-household benefits of improving an existing store in west Salt Lake Valley



(a) Utah County (in Santaquin)



(b) San Juan County (in Blanding)

Figure 10: Estimated per-household benefits of improving an existing store in Utah and San Juan Counties.

transit service to large grocery stores on the corridor. The per-household-trip benefit is very small however, with a maximum benefit on the order of $0.25.

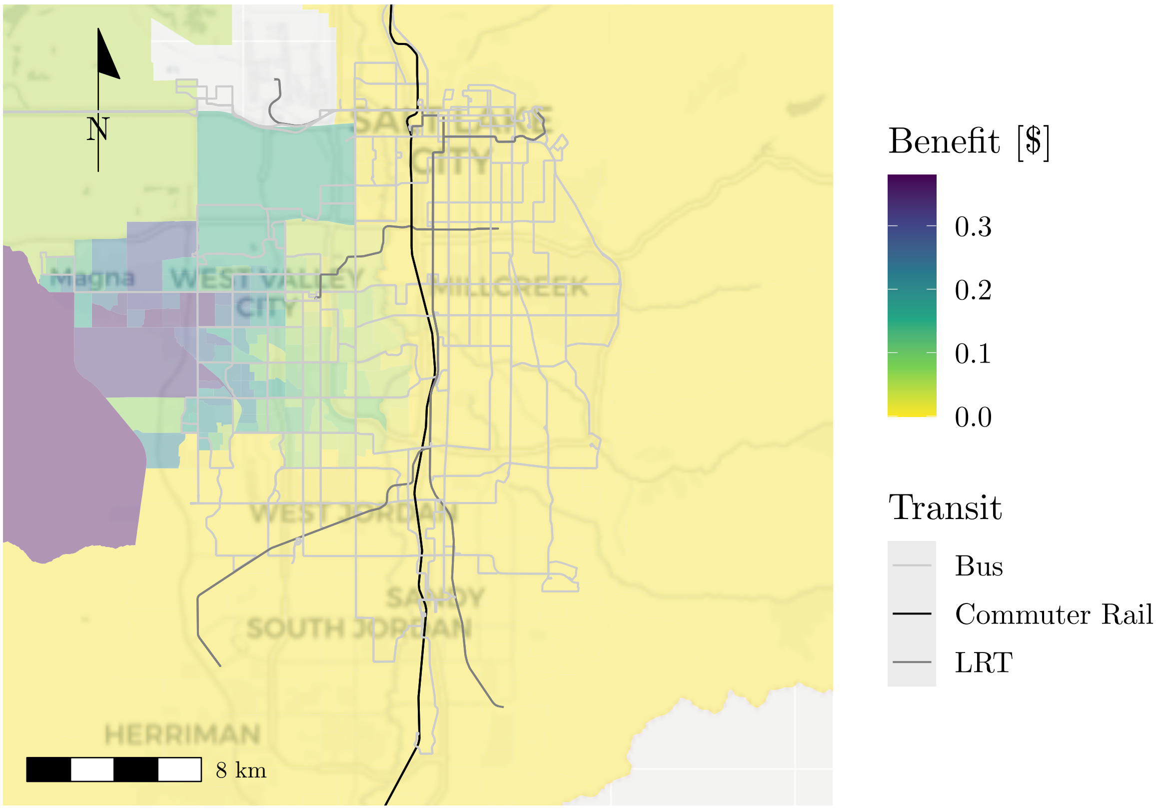


Figure 11: Estimated per-household benefits of improving non-motorized and public transport access to groceries in Salt Lake Valley.

Although comparing the geographic distribution of benefits is helpful, the aggregate benefit is more likely to guide policy. Additionally, the aggregate benefits can be weighted in different ways to understand the effects of the various policies on different populations. Table 6 presents the aggregate benefit from each of the three scenarios (and the result of the second scenario in all three communities). The Households column multiplies the difference in destination choice log-sum at each block group by the number of households in that block group, while the non-white and low-income columns weight the difference by the share of non-white individuals and low-income and zero-vehicle zero-vehicle respectively. All demographic data comes from the American Community Survey (ACS) 5-year aggregations; the ACS discloses households by vehicles available at the tract level, while all other household characteristics are at the more spatially refined block group level.

These alternate weighting schemes help to illustrate the potential equity of the benefit distribution should each scenario be pursued. The new store scenario in the Salt Lake community, for example, has a total

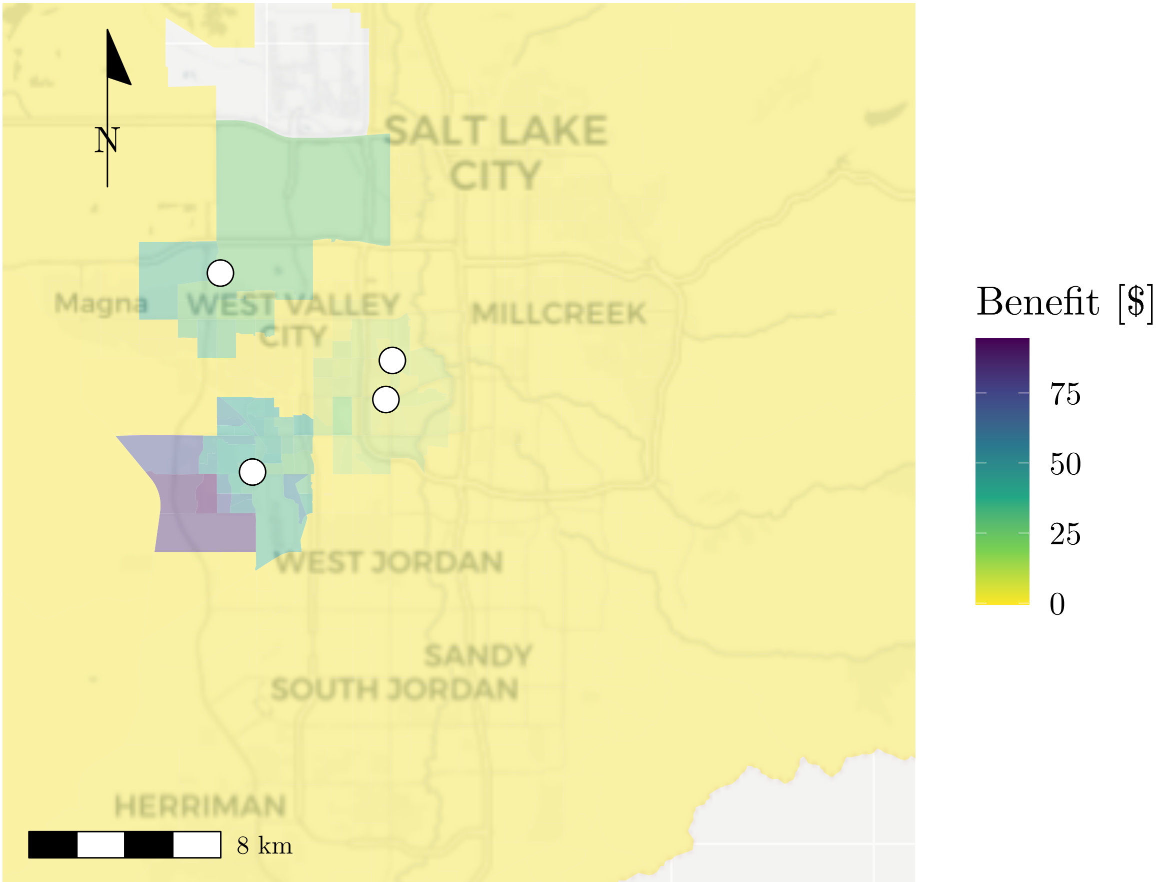


Figure 12: Estimated per-household benefits of grocery deliveries in Salt Lake Valley.

Table 6: Scenario Benefits

	Weighted by			
Scenario	Households	Non-white	Low-Income	Zero-Vehicle
New Store	$12,860,857	$5,612,293	$2,217,757	$555,185
Improved Store				
Salt Lake Valley	$6,273,501	$3,019,389	$1,255,096	$333,377
Utah County	$8,413,589	$1,340,090	$1,319,575	$163,353
San Juan County	$79,343	$35,482	$25,767	$5,383
Improved Transport	$1,686,717	$767,224	$234,676	$52,085
Delivery Services	$228,192,080	$88,977,055	$26,216,631	$5,474,393

benefit of approximately $13 million. Of this amount, somewhat less than half the benefits go to non-white individuals and less than one-fifth to low-income households. These ratios are more or less the same for the store improvement scenario and the improved transport scenario in the Salt Lake Valley. The Utah County store improvement, on the other hand, has a somewhat higher proportion of benefits going to low-income households relative to non-white households. This is a reflection of the lower minority population in southern Utah County vis a vis west Salt Lake valley, but is nonetheless a metric that program evaluators might pay attention to.

Overall, the improved store brings more than twice the benefits of improving non-automobile transportation, and the new store more than five times the benefits. Understanding the costs of these various alternatives is outside the scope of this research, but the level of infrastructure investment required to increase transportation level of service to that constructed in the simulation is likely an order of magnitude higher than the cost of a single new grocery store. It should also be acknowledged that improving non-automobile infrastructure and services would have benefits beyond just grocery store trips that we do not attempt to enumerate here. It is also not clear whether the level of improvement simulated in the second scenario could be accomplished within the envelope of the existing store; it may be that such improvements would meet or exceed the cost of building a new store on a new site, along with stocking, staffing, and operating the store.

By far the highest benefit observed in these scenarios is the increased availability of delivery services, but this scenario must be treated with some skepticism. Delivery services place costs on shoppers that are not well-quantified in the existing model: potential increases in shopping time, uncertainty over selection, inability to inspect goods prior to purchase, delays in delivery, and so forth. Additionally, low extant delivery prices may be a function of predatory pricing strategies by application developers seeking to capture market share first, and achieve profitability later (Moore, 2022). Resolving all of these issues cannot be handled in the present analysis, but it must be at least considered that grocery delivery may have the greatest consumer surplus of these four strategies by a wide margin.

5. Conclusions and Recommendations

Access to nutrition is a critical topic that has been a frequent focus of academic literature in public health, community planning, and engineering, but the varying and incomplete quantitative definitions of access have perhaps limited efforts at developing solutions.

In this research, the model we developed and informed with location-based services data suggested that policies that increased the availability and quality of grocery stores would do more to enhance access to nutrition than strategies that increased the transit level of service and lowered active transport travel times in three different Utah communities. These findings may not be general, but they largely support the review of food desert literature in (Beaulac et al., 2009) as well as critical evaluations by (Shannon, 2014) and (Wright et al., 2016).

5.1. Limitations

A number of assumptions made in Section 2 lead to limitations that other reasonable researchers might pursue differently, and thereby obtain marginally different outcomes.

In selecting a survey instrument with which to collect the store attribute data, we selected an existing and validated instrument from the nutrition environment literature. The NEMS-S focus on low-calorie and low-fat alternatives may be somewhat outdated in view of modern nutritional guidelines. For example, the NEMS-S does not track the availability and price of poultry, as "lean" poultry is not a goods category in the way that ground beef comes with multiple fat contents. Thus a major source of protein in typical American diets (FNS, 2021) was not traced across stores in each community. On a more basic level, the NEMS-S attempts to measure a store's stocking of goods that researchers believe are beneficial, and does not measure either what people wish to buy, or what they are actually buying at a store. Future research might attempt to survey shoppers on what they actually purchased at each store — or collect receipts of their purchases — though this would substantially raise the difficulty of collecting data.

Most households do not obtain all their groceries at a single store, though this research of necessity assumed that a simulated person chose exactly one store from stores available to them. Similarly, the location-based services data provided by StreetLight and used to identify which stores people traveled to only reveal whether a device was identified inside a geographic polygon, and not what they were actually doing in that polygon. This research had no way of distinguishing, for example, whether an individual device observed at a dollar store or a super market (e.g. Wal-Mart) was there to purchase groceries or some other household goods that might not be offered at more traditional food markets.

A number of simplifying assumptions concern the socioeconomic and spatio-temporal detail supplied to the choice models and accessibility calculators. The StreetLight data do not contain any demographic information on the individuals making trips, beyond the inferences made possible by the residence block group. This makes it difficult to estimate whether lower income households are more or less sensitive to travel distances or prices. Additionally, the research assumed that every trip was from the population-weighted block group centroid, which may vary substantially from the actual distance traveled, especially in large block groups. The methodology also used travel times calculated in the AM peak hour; though this time maximizes the availability of transit options, it is not a typical peak time for grocery shopping. The mode choice model was also selected for general convenience, and not fully calibrated to grocery trips in the specific regions. All of these limitations could potentially be relaxed by using a synthetic population with detailed socioeconomic data and parcel-level location choices determined by an activity-based model, as proposed by (Dong et al., 2006). In this exercise, which we leave to future research, the grocery maintenance trips could be explicitly modeled, with synthetic individuals of unique characteristics choosing destinations that are available on the course of their other daily activities, using their chosen travel modes. This proposal would effectively extend the methods of (Widener & Shannon, 2014) within the framework of explicit activity-travel modeling.

Finally, the scenarios presented in Section 4 are designed to illustrate potential applications of this accessibility methodology, with a comparative analysis of strategies to improve access to nutrition. Selecting different sites, attribute levels, or transport policies might substantially change the scale or rank-ordering of the estimated benefits. A comprehensive search for the location that would maximize benefits would be an interesting exercise, which we also leave to future research.

Acknowledgments

The authors acknowledge the Utah Department of Transportation (UDOT) for funding this research. The authors alone are responsible for the preparation and accuracy of the information, data, analysis, discussions, recommendations, and conclusions presented herein. The contents do not necessarily reflect the views, opinions, endorsements, or policies of the Utah Department of Transportation or the U.S. Department of Transportation. The Utah Department of Transportation makes no representation or warranty of any kind, and assumes no liability therefore.

Student research assistants who collected the store data include Kaeli Monahan, Brooke Jones, Emily Aleson, Andrea Barney, Mattie Earnest, Tare Ekpagha, Mackay Graper, Olivia Harrison, Mary Jacobsen, Gillian Martin, Maggie Plessinger, and Jenna Smith. Madeleine Smith coordinated the data collection efforts.

Graphics in the document are produced with multiple R packages (Arel-Bundock, 2022; Dunnington, 2023; Ram & Wickham, 2018; Wickham, 2016).

Author Contribution Statement

Gregory S. Macfarlane: Conceptualization, Methodology, Software, Resources, Writing - original draft, Visualization, Supervision **Emma Stucki:** Software, Investigation, Data curation, Writing - original draft **Myrranda Salmon:** Investigation, Data curation **Alisha H. Redelfs:** Methodology, Resources, Writing - review & editing **Lori Spruance:** Methodology, Resources, Writing - review & editing

References

Aggarwal, A., Cook, A. J., Jiao, J., Seguin, R. A., Vernez Moudon, A., Hurvitz, P. M., & Drewnowski, A. (2014). Access to Supermarkets and Fruit and Vegetable Consumption. *American Journal of Public*

Health, *104*(5), 917–923. https://doi.org/10.2105/AJPH.2013.301763

Arel-Bundock, V. (2022). modelsummary: Data and model summaries in R. *Journal of Statistical Software*, *103*(1), 1–23. https://doi.org/10.18637/jss.v103.i01

Barnes, M. (2021). *Resiliency of Utah's Road Network: A Logit-Based Approach* [{{MS Thesis}}]. Brigham Young University.

Beaulac, J., Kristjansson, E., & Cummins, S. (2009). A Systematic Review of Food Deserts, 1966-2007. *Preventing Chronic Disease*, *6*(3), A105.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Applications to Travel Demand*. MIT Press. https://www.jstor.org/stable/1391567

Bureau of Labor Statistics. (2023). 12-month percentage change, Consumer Price Index, selected categories. In *Consumer Price Index*. https://www.bls.gov/charts/consumer-price-index/consumer-price-index-by-category-line-chart.htm.

Chen, D., Jaenicke, E. C., & Volpe, R. J. (2016). Food Environments and Obesity: Household Diet Expenditure Versus Food Deserts. *American Journal of Public Health*, *106*(5), 881–888. https://doi.org/10.2105/AJPH.2016.303048

Clifton, K. J. (2004). Mobility Strategies and Food Shopping for Low-Income Families: A Case Study. *Journal of Planning Education and Research*, *23*(4), 402–413. https://doi.org/10.1177/0739456X04264919

Conway, M. W., Byrd, A., & Eggermond, M. van. (2018). Accounting for uncertainty and variation in accessibility metrics for public transport sketch planning. *Journal of Transport and Land Use*, *11*(1). https://doi.org/10.5198/jtlu.2018.1074

Conway, M. W., Byrd, A., & van der Linden, M. (2017). Evidence-Based Transit and Land Use Sketch Planning Using Interactive Accessibility Methods on Combined Schedule and Headway-Based Networks. *Transportation Research Record*, *2653*(1), 45–53. https://doi.org/10.3141/2653-06

Conway, M. W., & Stewart, A. F. (2019). Getting Charlie off the MTA: A multiobjective optimization method to account for cost constraints in public transit accessibility metrics. *International Journal of Geographical Information Science*, *33*(9), 1759–1787. https://doi.org/10.1080/13658816.2019.1605075

Cooksey-Stowers, K., Schwartz, M. B., & Brownell, K. D. (2017). Food Swamps Predict Obesity Rates Better Than Food Deserts in the United States. *International Journal of Environmental Research and Public Health*, *14*(11), 1366. https://doi.org/10.3390/ijerph14111366

Croissant, Y. (2020). Estimation of random utility models in R: The mlogit package. *Journal of Statistical Software*, *95*(11), 1–41. https://doi.org/10.18637/jss.v095.i11

Dong, X., Ben-Akiva, M. E., Bowman, J. L., & Walker, J. L. (2006). Moving from trip-based to activity-based measures of accessibility. *Transportation Research Part A: Policy and Practice*, *40*(2), 163–180. https://doi.org/10.1016/j.tra.2005.05.002

Dunnington, D. (2023). *Ggspatial: Spatial data framework for ggplot2*. https://CRAN.R-project.org/packag e=ggspatial

Fitzpatrick, K., Brewer, M. A., & Turner, S. (2006). Another Look at Pedestrian Walking Speed. *Transportation Research Record*, *1982*(1), 21–29. https://doi.org/10.1177/0361198106198200104

FNS. (2021). *Thrifty Food Plan*. Food and Nutrition Service, United States Department of Agriculture.

Geurs, K., Zondag, B., de Jong, G., & de Bok, M. (2010). Accessibility appraisal of land-use/transport policy strategies: More than just adding up travel-time savings. *Transportation Research Part D: Transport and Environment*, *15*(7), 382–393. https://doi.org/10.1016/j.trd.2010.04.006

Ghosh-Dastidar, B., Cohen, D., Hunter, G., Zenk, S. N., Huang, C., Beckman, R., & Dubowitz, T. (2014). Distance to Store, Food Prices, and Obesity in Urban Food Deserts. *American Journal of Preventive Medicine*, *47*(5), 587–595. https://doi.org/10.1016/j.amepre.2014.07.005

Glanz, K., Sallis, J. F., Saelens, B. E., & Frank, L. D. (2007). Nutrition Environment Measures Survey in Stores (NEMS-S): Development and Evaluation. *American Journal of Preventive Medicine*, *32*(4), 282–289. https://doi.org/10.1016/j.amepre.2006.12.019

Handy, S. L., & Niemeier, D. A. (1997). Measuring accessibility: An exploration of issues and alternatives. *Environment and Planning A: Economy and Space*, *29*(7), 1175–1194. https://doi.org/10.1068/a291175

Hedrick, V. E., Farris, A. R., Houghtaling, B., Mann, G., & Misyak, S. A. (2022). Validity of a Market Basket Assessment Tool for Use in Supplemental Nutrition Assistance Program Education Healthy Retail Initiatives. *Journal of Nutrition Education and Behavior*, *54*(8), 776–783. https://doi.org/10.1016/j.jneb.2022.02.018

Hillier, A., Cannuscio, C. C., Karpyn, A., McLaughlin, J., Chilton, M., & Glanz, K. (2011). How Far Do Low-Income Parents Travel to Shop for Food? Empirical Evidence from Two Urban Neighborhoods. *Urban Geography*, *32*(5), 712–729. https://doi.org/10.2747/0272-3638.32.5.712

Kaczynski, A. T., Schipperijn, J., Hipp, J. A., Besenyi, G. M., Wilhelm Stanis, S. A., Hughey, S. M., & Wilcox, S. (2016). ParkIndex: Development of a standardized metric of park access for research and planning. *Preventive Medicine*, *87*, 110–114. https://doi.org/10.1016/j.ypmed.2016.02.012

Liu, B., Widener, M. J., Smith, L. G., Farber, S., Minaker, L. M., Patterson, Z., Larsen, K., & Gilliland, J. (2022). Disentangling Time Use, Food Environment, and Food Behaviors Using Multi-Channel Sequence Analysis. *Geographical Analysis*, *54*(4), 881–917. https://doi.org/10.1111/gean.12305

Logan, T., Williams, T., Nisbet, A., Liberman, K., Zuo, C., & Guikema, S. (2019). Evaluating urban accessibility: Leveraging open-source data and analytics to overcome existing limitations. *Environment and Planning B: Urban Analytics and City Science*, *46*(5), 897–913. https://doi.org/10.1177/2399808317736528

Lunsford, J., Brunt, A., Foote, J., Rhee, Y., Strand, M., & Segall, M. (2021). Analysis of Availability, Quality, and Price of Food Options in Denver, CO Grocery Stores. *Journal of Hunger & Environmental Nutrition*, *16*(3), 297–303. https://doi.org/10.1080/19320248.2020.1741482

Macfarlane, G. S., Boyd, N., Taylor, J. E., & Watkins, K. (2021). Modeling the impacts of park access on health outcomes: A utility-based accessibility approach. *Environment and Planning B: Urban Analytics and City Science*, *48*(8), 2289–2306. https://doi.org/10.1177/2399808320974027

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press.

Moore, D. V. (2022). Predatory Pricing, Network Effects, and the Food Delivery Industry: A Case Study. *Northern Kentucky Law Review*, *49*(1), 93–114.

Pereira, R. H. M., Saraiva, M., Herszenhut, D., Braga, C. K. V., & Conway, M. W. (2021). R5r: Rapid Realistic Routing on Multimodal Transport Networks with R^5 in R. *Findings*. https://doi.org/10.32866/001c.21262

Ram, K., & Wickham, H. (2018). *Wesanderson: A wes anderson palette generator*. https://CRAN.R-project.org/package=wesanderson

Recker, W. W., & Kostyniuk, L. P. (1978). Factors influencing destination choice for the urban grocery shopping trip. *Transportation*, *7*(1), 19–33. https://doi.org/10.1007/BF00148369

Shannon, J. (2014). Food deserts: Governing obesity in the neoliberal city. *Progress in Human Geography*, *38*(2), 248–266. https://doi.org/10.1177/0309132513484378

Train, K. E. (2009). *Discrete Choice Methods with Simulation* (2nd edition). Cambridge University Press.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Walker, R. E., Keane, C. R., & Burke, J. G. (2010). Disparities and access to healthy food in the United States: A review of food deserts literature. *Health & Place*, *16*(5), 876–884. https://doi.org/10.1016/j.healthplace.2010.04.013

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Widener, M. J., & Shannon, J. (2014). When are food deserts? Integrating time into research on food accessibility. *Health & Place*, *30*, 1–3. https://doi.org/10.1016/j.healthplace.2014.07.011

Wright, J. D., Donley, A. M., Gualtieri, M. C., & Strickhouser, S. M. (2016). Food Deserts: What is the Problem? What is the Solution? *Society*, *53*(2), 171–181. https://doi.org/10.1007/s12115-016-9993-8