Detroit Project
Technical Appendix

Companion to

Examining the Impact of Food Deserts on Public Health in Detroit

Sponsored by

LaSalle Bank

June 2007

MARI GALLAGHER
RESEARCH & CONSULTING GROUP

Methodology Only

Methodology

In addition to our group's personal commentary and key acknowledgements, the purpose of the Detroit Project Technical Appendix is to provide additional technical details for which we did not have room in the project's Executive Briefing. Much of what we provide here is for other researchers and students of urban planning, public health, and other disciplines. We also welcome Detroit community groups and leaders to put it to use!

Our methodology begins with a very simple premise: that the vitality and health of any urban community is a block-by-block phenomenon. When we think of concepts such as "home" and "community," we typically think of the very block where we live, and whether or not it is a well-balanced, life-supporting environment. We begin our study of food access and health outcomes at the block level for this reason. We also analyze patterns by Census tract, official Community Areas, and Zip Codes. Our research objective is to compare food access and diet-related health outcomes, holding other influencers such as income, education, and race constant to the degree possible given time and resource constraints.

After documenting neighborhood differences in the availability of grocery stores and fringe food outlets, we test the theory that a *balanced* food environment is an important determinant of community health. In other words, do food deserts (areas that cluster with no or distant grocery stores) or out-of-balance areas (food deserts with nearby concentrations of fringe food options) face nutritional challenges evident in poorer diet-related health outcomes, and do those outcomes also worsen when the food desert has high concentrations of nearby fast food alternatives?

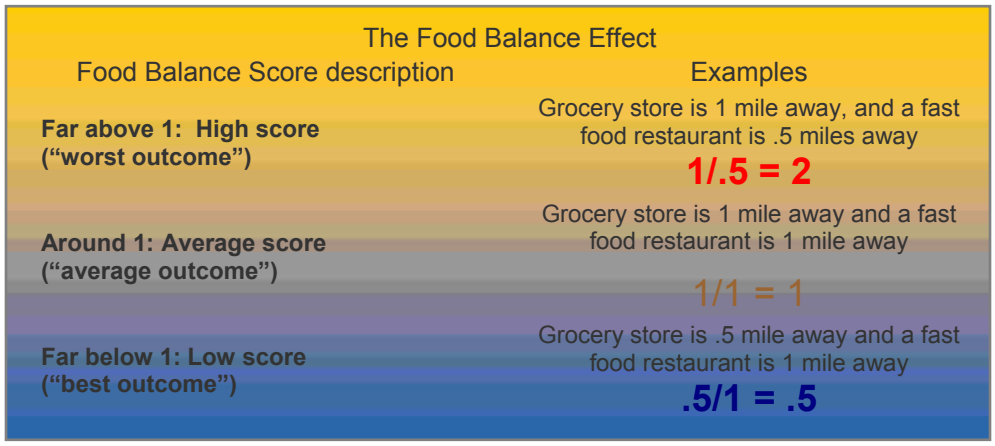
To calculate the average distance to a food venue (such as a grocery store or fast food restaurant), we measured the distance between the geographic centers of each block and the locations of each food venue in the Detroit region. The distance from the center of each block to each food venue was calculated, using the latitude and the longitude of each food venue and of each block center. Of these distances, the minimum distance was calculated for each block, representing the distance from that block to the nearest food venue by category. For each block, a weight was created to reflect the share of the city's population living in that block.

The average distance for an area larger than a block (a tract, community area, zip code, etc.) is the weighted average of the distance from each block within that area to the nearest food venue, with greater weights given to blocks with larger numbers of residents. Therefore, we account for density, and we exclude areas with no population. The distance score, calculated in miles, is the distance the average person in that area would need to travel to reach a food venue.

To test our core theory that food venue balance matters for health outcomes, we developed a ratio score: the distance to any grocer divided by the distance to any fringe food venue. The ratio for the entire city, a ZIP Code, a Community Area, or a census

tract, is the ratio of the average distance to a grocer to the average distance to a fringe food venue, where these distances are averaged across all blocks in the areas using block-level populations as weights. So greater weight is given to blocks with larger numbers of residents.

We call this ratio the **Food Balance Score** and we call its impact the **Food Balance Effect**. These measures are unique to our group; we developed them as part of our work on food access and its impact on public health.



Food venue data was collected for Detroit and the 5 counties including and surrounding Detroit. This allowed us to calculate the actual distance to the nearest food venue, rather than be constrained by artificial boundaries such as concentric rings, Community Areas, and so on, that might not reflect the realities of food purchasing. We believe the list is current (2007).

Food venue data consists of grocery stores, fast food restaurants, and other fringe locations that sell groceries. Generally, the definition of fast food is take-out or self-carry to tables within the restaurant. Primarily we focused on burger, chicken, taco, and hot dog places.

Death data was obtained for a three year period (2004, 2005, and 2006) for the entire state, and culled for the 5 counties including and surrounding Detroit. From that we culled all diet-related deaths. Cardiovascular disease rates and deaths were segregated by those that are diet-related, namely: essential hypertension, hypertensive heart disease, hypertensive renal disease, hypertensive heart and renal disease, acute myocardial infarction, subsequent myocardial infarction, certain current complications following acute myocardial infarction, other acute ischaemic heart diseases, chronic ischaemic heart disease, all cerebrovascular diseases (stroke) and artheroscelorosis. Excluded cardiovascular disease rates and deaths were acute rheumatic fever, chronic rheumatic heart diseases, and pulmonary heart disease.

The City of Detroit has 313 census-defined tracts and 11,373 census-defined blocks with non-zero populations (total population 951,270). The area of Metro Detroit (with

Detroit excluded) for which we calculated distances to food retailers has 32,419 census-defined blocks with non-zero populations (total population 2,624,204). This is slightly smaller than the whole 5 counties because of our need to draw a buffer zone. See the *Additional Maps* section.

In our analysis of the relationship between food balance and health outcomes, we used 41 Detroit community areas (total population 862,934) for all diet-related deaths and 29 Detroit community areas (total population 692,386) for all cardiovascular diet-related deaths.

In Metro Detroit (with Detroit excluded), we used 507 census tracts (total population 2,133,446) for all diet-related deaths and 289 census tracts (total population 1,285,445) for all cardiovascular diet-related deaths.

The number of places for which we calculated distances to food retailers is larger than the number for which we calculated the relationship between distances and health outcomes because we limited the analysis to places (tracts for the non-Detroit area, community areas for Detroit) that had 20 or more diet-related deaths (or 20 or more cardiovascular diet-related deaths), in order to produce more reliable estimates of Years of Potential Life Lost (YPLL).

Very detailed information was developed on distances to particular food venues, including the distribution of USDA Food Stamp Retailers, not only for Detroit, but also for Metro Detroit and other Michigan locations. Some data were sorted by majority race of the block or tract. **See Additional Tables** section for more details.

No data set can provide a complete and fully accurate picture of the locations of commercial venues across large, diverse, and dynamic urban geographies. Nonetheless, we believe that our overall patterns are accurate and that they reflect the current realities of food access in Detroit today.

From that data set we use regression analysis to compare food ratio scores and their impact on YPLL, holding education, income, and race constant. Here are the details without controls:

Community Area Level (Detroit only)

	Dependent Variable							
Independent Variable	YPLL (diet-related)		YPLL (diet-related)		YPLL (cardiovascular)		YPLL (cardiovascular)	
	beta	prob(beta≠0)	beta	prob(beta≠0)	beta	prob(beta≠0)	beta	prob(beta≠0)
Food Balance	1.54	0.84			1.71	0.88		
Mean Dist Grocer			10.29	0.90			8.19	0.85
Mean Dist Fast Food			-34.48	0.89			-19.12	0.72
Intercept	18.85	0.99	24.49	0.99	8.48	0.99		
Observations	41		41		29		29	

Tract Level (non-Detroit only)

	Dependent Variable							
	YPLL (diet-related)		YPLL (diet-related)		YPLL (cardiovascular)		YPLL (cardiovascular)	
Independent Variable	beta	prob(beta≠0)	beta	prob(beta≠0)	beta	prob(beta≠0)	beta	prob(beta≠0)
Food Balance	3.65	0.99			21.88	0.99		
Mean Dist Grocer			4.57	0.90			5.19	0.88
Mean Dist Fast Food			-17.50	0.99			-19.83	0.99
Intercept	37.20	0.99	48.01	0.99	21.88	0.99		
Observations	752							

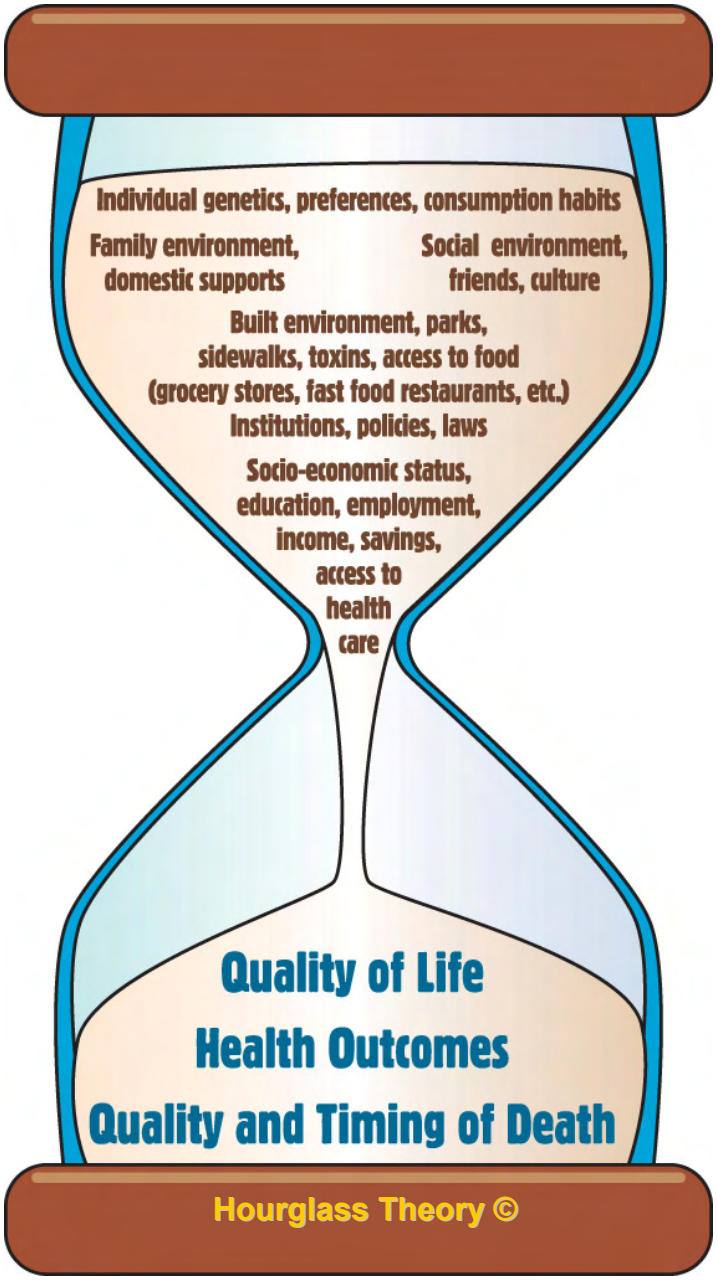
With controls:

Community Area Level (Detroit only)			Red = Food Balance or distance coefficient goes in the "wrong" direction					
	Dependent Variable							
	YPLL (diet-related)		YPLL (diet-related)		YPLL (cardiovascular)		YPLL (cardiovascular)	
Independent variable	beta	prob(beta≠0)	beta	prob(beta≠0)	beta	prob(beta≠0)	beta	prob(beta≠0)
Food Balance	-0.10	0.06			1.65	0.78		
Mean Dist Grocer			2.09	0.24			8.05	0.79
Mean Dist Fast Food			4.24	0.11			-27.40	0.66
Pct Black	0.09	0.97	0.08	0.96	-0.02	0.38	-0.02	0.39
Pct H.S. Education or less	0.27	0.94	0.25	0.91	0.10	0.53	0.09	0.47
Median H.H. Income								
Pct Below H.H. Inc. $25,000	-0.03	0.22	-0.01	0.12	-0.07	0.63	-0.09	0.68
Intercept	8.71	0.94	6.81	0.63	10.68	0.98	17.24	0.95
Observations	41		41		29		29	

Tract Level (non-Detroit only)								
	Dependent Variable							
	YPLL (diet-related)		YPLL (diet-related)		YPLL (cardiovascular)		YPLL (cardiovascular)	
Independent variable	beta	prob(beta≠0)	beta	prob(beta≠0)	beta	prob(beta≠0)	beta	prob(beta≠0)
Food Balance	2.81	0.99			2.80	0.99		
Mean Dist Grocer			5.62	0.98			6.83	0.98
Mean Dist Fast Food			-1.75	0.35			-5.85	0.74
Pct Black	0.09	0.94	0.08	0.93	0.09	0.92	0.09	0.92
Pct H.S. Education or less	0.87	0.99	0.88	0.99	0.67	0.99	0.68	0.99
Median H.H. Income	-0.15	0.98	-0.20	0.99	-0.12	0.95	-0.13	0.96
Pct Below H.H. Inc. $25,000								
Intercept					16.00	0.99	19.14	0.99
Observations	425		425		246		246	

2006 driver's license data, which reports height and weight, was used to construct measures of body mass index, which is an accepted measure for obesity. BMI calculations were at the Zip Code level. There was too much noise in the data; we could not find a meaningful pattern. More work is needed here.

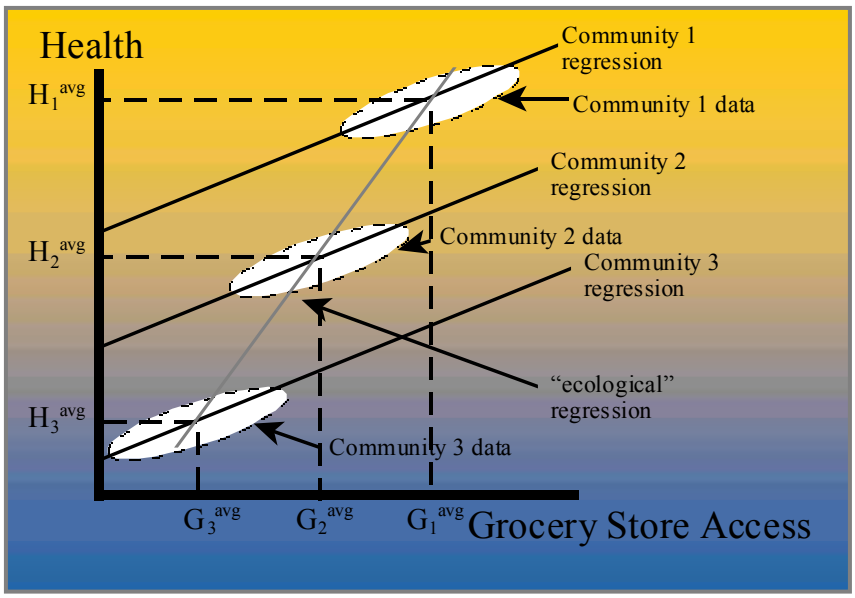
We recognize the "time factor" challenge in studying food access and health outcomes. How much time needs to pass for the lack of nutritious food access to have an effect on community health? For example, lack of adequate nutrition in childhood might not be evident until later in life. Or it could be evident in childhood obesity patterns. This is not a challenge we can control for in this study, nor do we speculate on the theoretical possibilities of time factors on health impacts, other than in our Hourglass Theory where we illustrate the many non-linear, highly dynamic influencers of community health. We believe our analysis is a valid approach. Health conditions do not develop overnight, nor do new grocery stores or fast food restaurants.



A bigger challenge in correlating neighborhood characteristics and health is that neighborhood effects are, by definition, endogenous to the compositional characteristics of neighborhoods. The project recognizes the many complex methodological challenges in isolating cause and effect and holding constant potential statistical confounders, such as income, race, genetics, culture, food preferences, and self-selection into specific types of communities. We utilize national and local best methodological practices, taking deliberate measures to maximize the robustness and accuracy of our data and models,

and exercising caution and care in stating our findings. We are particularly mindful of what has been called the "Ecological Fallacy."

Here we present a scenario, not based on actual individual level data, but one that should be kept in mind when describing outcomes and findings. Within each community of our scenario, the relationship between access to grocery stores and health is positive (better access ► better health). The relationship is also positive across the entire population (if we pool observations from all three communities, the regression line would coincide with the Community 2 regression line).



Ecological Regression Illustration

If we only have averages by community, we could still conclude that the relationship is positive: the ecological regression line that best fits the *average* values for the three communities – (H_1^{avg}, G_1^{avg}), (H_2^{avg}, G_2^{avg}), (H_3^{avg}, G_3^{avg}) – slopes upward. But the relationship is stronger when we use community averages than when we use data on individuals (either analyzing all individuals pooled or analyzing individuals separately by community). This is because the communities have very different compositions (Community 1: high health/high access; Community 2: moderate health/moderate access; Community 3: low health/low access).

Based on the ecological regression, the most we can say is "communities with better average access to grocery stores have better average health" – we *cannot* say that "an individual person who has better access to grocery stores will have better health" (i.e. we need to limit our generalizations to the unit of observation – the community in this case – that we are using, and not generalize to smaller units of observation – the individual).

Therefore, while we demonstrate in this study the positive association between access to better foods and better diet-related health outcomes and conditions, particularly concerning certain diseases and races, we must set our findings in the context of the challenges and limitations of linking cause and effect and of predicting, with certainty, the exact statistical magnitude of the relationship. Nonetheless, we stand by our conclusion that food deserts – particularly those with high concentrations of fast food restaurants – pose serious health and wellness challenges to the residents who live within them and to Detroit and the region as a whole.