A Comprehensive Method for Identifying Optimal Areas for Supermarket Development

TRF Policy Solutions

July 14, 2011

Capital at the point of impact.

Profile of TRF

The Reinvestment Fund builds wealth and opportunity for low-wealth communities and low and moderate income individuals through the promotion of socially and environmentally responsible development.

We achieve our mission through:

Capital

- **Grants, loans and equity investments**

Knowledge

- **Information and policy analysis; Policy Solutions & PolicyMap**

Innovation

- **Products, markets and strategic partnerships**

The Problem

- A growing body of research suggests that there are areas in the United States that suffer from poor access to healthy and affordable food.
- U.S. Department of Agriculture's Economic Research Service reports that:
  - People without access to full service grocery stores often depend on small grocery or convenience stores that may not carry all the foods needed for a healthy diet.
  - Convenience stores often charge more than grocery stores for the same items, keeping people from eating a balanced diet. *
  - 23.5 million people live in low income areas (below 200% poverty) that are 1 mile from a supermarket.

* "Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences" (June, 2009)

Limitations of Prior Work

Methodological limitations of previous research:

- Imprecise definitions (e.g., the many and varied definitions of the term “food desert.”)
- Data quality / coverage
- Use fixed distances for urban, suburban, and rural areas throughout entire counties, metro areas, and even states (does not take account of population density, access to transportation, etc.).
- Use large geographies (zip codes, counties) as unit of analysis.
- Do not establish meaningful benchmarks for defining “adequate access.”
- Based on presence/absence of food retail, not travel distance or other physical or economic impediments to access that retail.

Assumption / Goals

Assumption: The degree to which food retail is available in middle and upper income places is the standard against which other locations are measured. Stated differently, *we assume that "the market" is operating well in the higher income areas.*

1. Locate small geographic areas showing the strongest need for additional supermarket development **and** a sufficiency of demand to support additional development.
2. Create a demographic profile of low access areas.
3. Calculate the food at home expenditures leaked from the low access areas.
4. Establish a valid and reliable method for measuring low grocery access nationwide.
5. Provide CDFIs and policymakers operating at the city, metro, state, regional or national level summaries of inadequate supermarket access problem.

Methods

1. Software Applications:
   i. ARCMAP 10, Spatial Analyst, Network Analyst, Model Builder
   ii. SQL Server Management Studio
   iii. R – Version 2.11.1 64 bit
   iv. OPENGEODA

2. Datasets
   i. Trade Dimensions
      - Supermarket Locations, Sales, Square Feet
   ii. Street Map USA Data 9.3
      - Road Network
   iii. America Community Survey (2005-2009)
      - Car Ownership Rate
      - Median Household Income
   iv. 2010 Census
      - Population
      - Land Area
      - Number of Households
   v. Bureau of Labor Statistics Consumer Expenditure Data

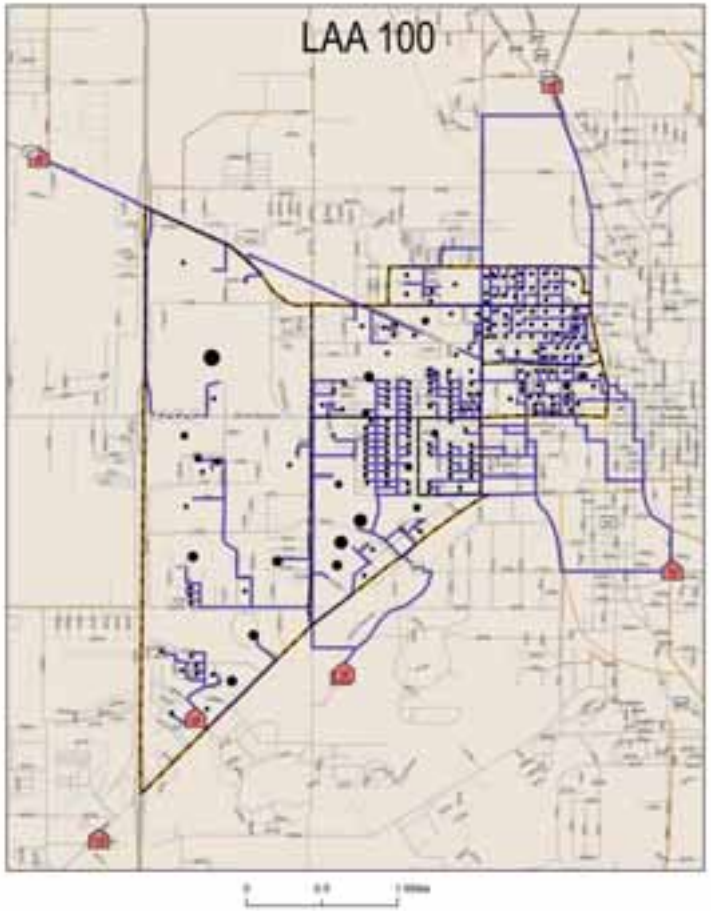
Block Group Population Density and Car Ownership

- Obtain location data for grocery stores in all lower 48 states and DC from Trade Dimensions (~ 70,000) stores, plus chain pharmacies.

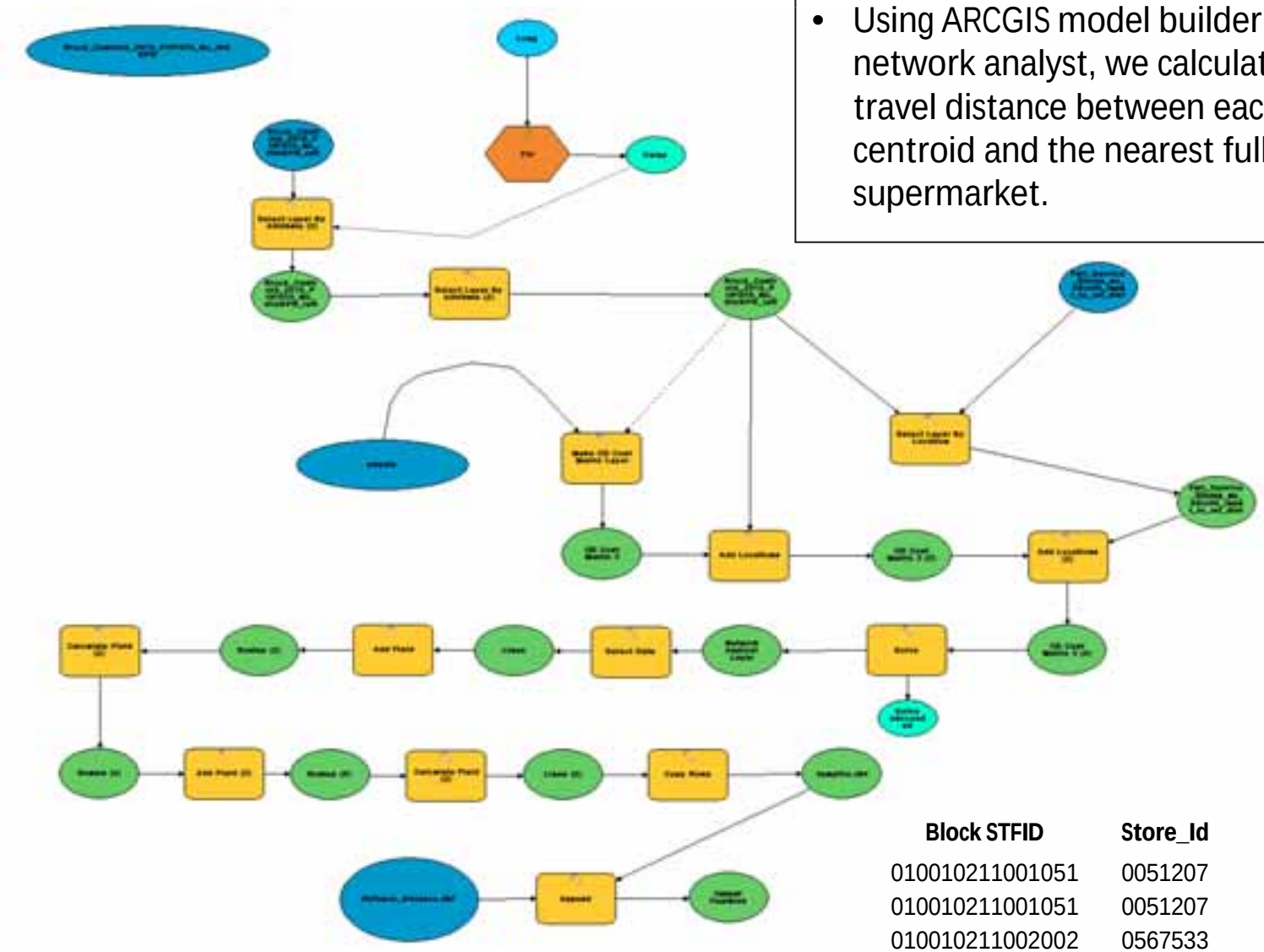
- Geocode grocery stores categorized as supermarkets, supercenters, natural/ gourmet foods or wholesale clubs (i.e., stores with a full selection of foods ~ 32,000 stores).

- Calculate the network distance from each 2010 census block centroid to the nearest supermarket, supercenter, Warehouse Grocery and , natural foods or wholesale club.

- We processed the network dataset as a sparse matrix using ARCGIS model builder to iterate through 5,000 block centroids at a time searching only for stores that were within 100 Euclidian miles.



Methods



- Using ARCGIS model builder and network analyst, we calculated the travel distance between each block centroid and the nearest full service supermarket.

Block STFID	Store_Id	Total Time	Total Length
010010211001051	0051207	15.1321	10.0035
010010211001051	0051207	15.1321	10.0035
010010211002002	0567533	33.2163	18.1791
010010211002009	0567533	26.2290	14.0377
010010211002041	0567533	24.4221	14.2995

Methods

- Calculated residential population density for all block groups with population > 250 population, netting out blocks with zero population.
- Apportioned 2010 census block data to the 2000 block group using a spatial bridge table created using python and arcpy.Union_analysis.
- Adjusted each 2000 block group's land area using the 2010 census block boundaries – blocks with zero population were excluded as well as areas of water.
- Classified 2000 block groups into 7 population density categories using a CART Model (Cartographic and Regression Tree) model in R.
- Stratified density categories based on car ownership rates for the higher density areas, also using a CART Model.
- Calculated reference distances based on mean travel distance to full service supermarkets for blocks with median household incomes above 120% of their Area Median Incomes (AMI)

Groups	Distance
7LC	0.15
7MC	0.41
7HC	0.73
7VHC	0.90
6MC	1.19
6HC	1.49
5M	2.30
5H	3.00
4H	5.51
3H	8.46
2H	11.34
1H	19.33

Methods

- Calculated an Access Score for each block group based on the difference between the minimum travel distance to the nearest full service store and its reference group distance.

Market Distance: 1.25 miles

Reference Distance (7HC): 0.73miles

Access Score: (1.25 -0.73)/1.25= **.416**

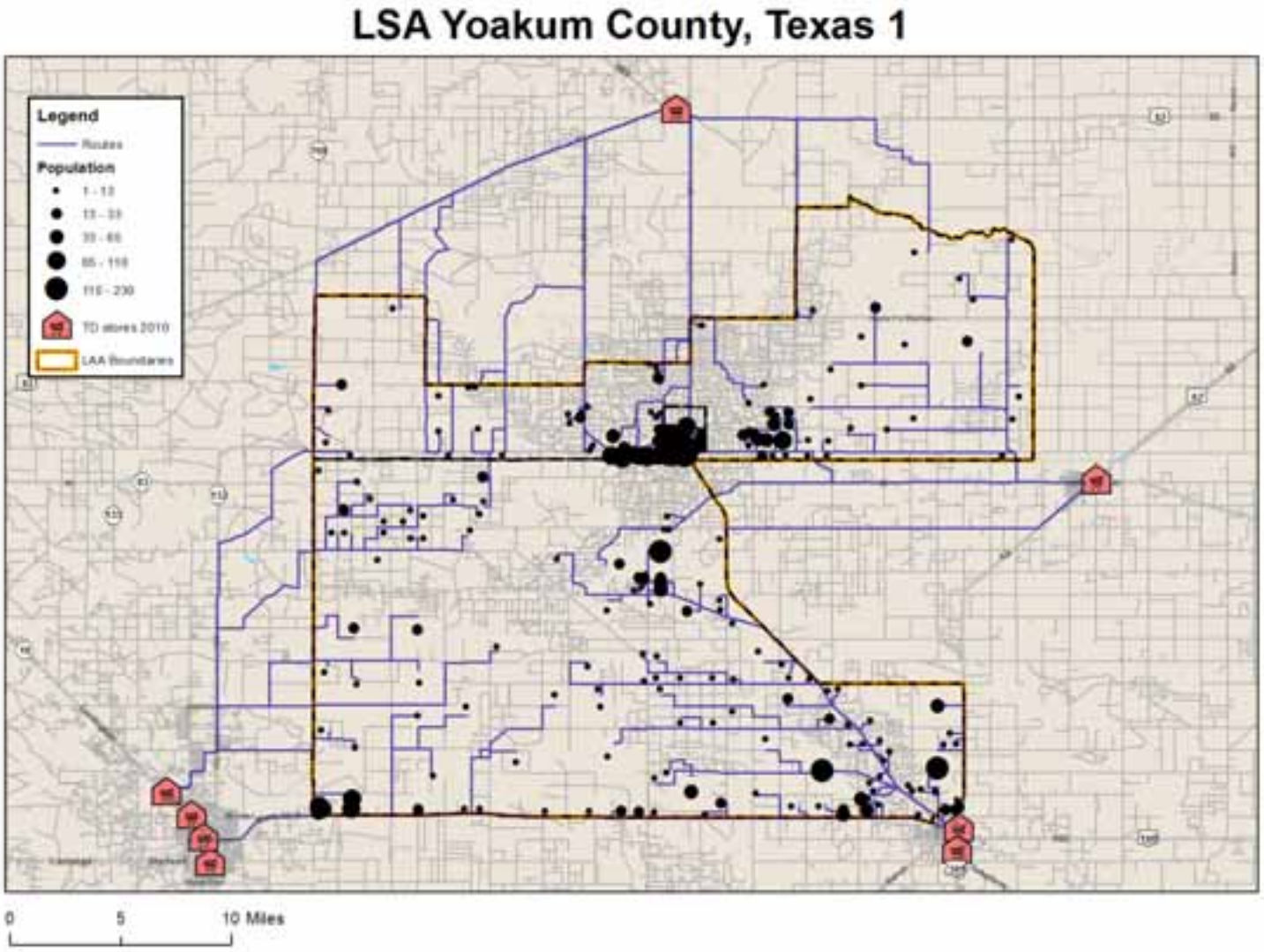
This block group's travel distance would have to be decreased by 41.6% to equal the distance traveled by its non-LMI counterparts.

- Conducted a Local Indicator of Spatial Association (LISA) using open GEODA to identify block groups with positive access scores that are contiguous to neighboring block groups

- Spatial neighbors defined as all block groups that are contiguous and share a road.

- Cluster membership is based on spatial correlation.

Results


LSA Yoakum County, Texas 1
Legend
Routes
Population
1 - 12
13 - 33
33 - 66
66 - 115
115 - 230
TD stores 2010
LAA Boundaries
0
5
10 Miles

Grocery Retail Leakage

- A measure of how much money is spent on food prepared at home at grocery stores located outside a block group's reference distance – this money is "leaking" from the local community into adjacent communities.

- To measure leakage we ran a secondary Network analysis to identify all stores (~ 72,000) within 20 miles of each census block centroid.

- We then partitioned each store's annual food sales to the population in blocks that encompass the store(s) within their reference distances (e.g. if a store has 4 million dollars a year in food sales and 10,000 people live within blocks whose reference distance encompasses that store, then each person is given 400 dollars of food sales a year from that store, plus any other stores within their reference distance).

- Next we summed the total amount of sales available from all stores within each block group's reference distance.

- Lastly we subtracted each block group's total demand for food prepared at home, based on the Bureau of Labor Statistics (BLS) Consumer Expenditure Survey Micro-Data.

- The Difference between the two (demand minus supply) represents leakage, which can be positive or negative.