Mari Gallagher
Research & Consulting Group

312-339-0640

mari@marigallagher.com
www.marigallagher.com

TECHNICAL ADDENDUM
Examining the Impact of Food Deserts on Diet-related Deaths in Florida

Commissioned By

Florida Department of Agriculture and Consumer Services

September 2014

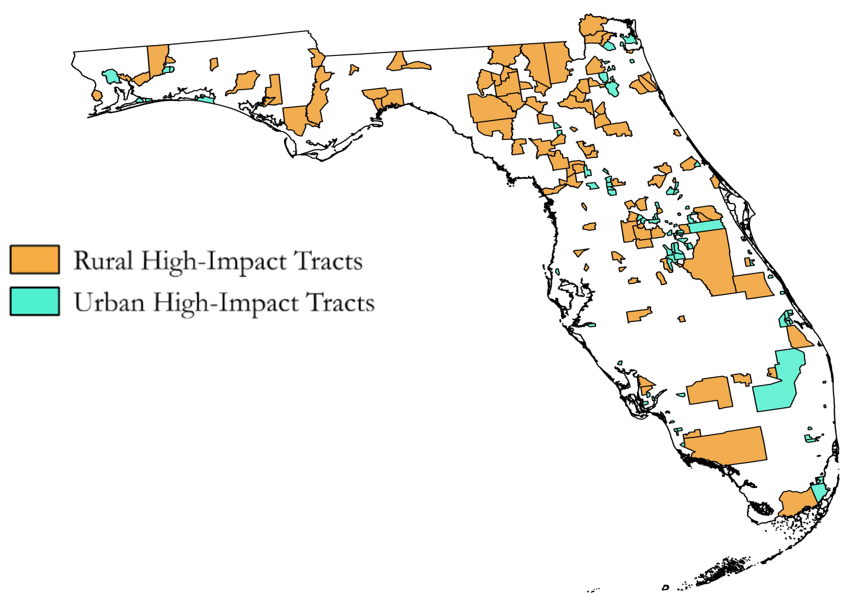
In August of 2014, Mari Gallagher Research & Consulting Group (MG) conducted a statewide analysis of the relationship between food access and diet-related deaths across Florida. The objective was to develop evidence-based recommendations concerning priority areas where immediate support of healthy food access would most likely have the greatest impact on quality and length of life. Florida is a very diverse state in terms of population patterns, land use, and geography. Therefore, the analysis accounts for differences in urban and rural areas and across counties and regions.

KEY FINDINGS

Access to quality retail grocers in Florida is strongly linked to a variety of diet-related health outcomes. Individuals living in places, both urban and rural, where many households reside more than a half-mile from the nearest full-service grocer and lack access to a vehicle are more likely to die prematurely from diabetes, diet-related cancers, stroke, and liver disease than individuals living where grocers are closer and vehicles are more available, after controlling for other key factors. We have identified 100 rural and 100 urban census tracts (Fig. 1a and 1b) where our evidence suggests that policies, programs, and market action that improve healthy food access can substantially and measurably improve public health. We estimate that even modest improvements can make a huge difference. For example, reducing the fraction of the population with insufficient healthy food access in these 200 places by just one single percentage point could prevent nearly 650 premature deaths over a seven-year period and improve the overall health of the 1.2 million urban Floridians and the 780,000 rural Floridians who live in these areas. This is not to suggest that access alone is the silver bullet; programs, such as nutritional education and transportation support, are also needed. But the Floridians in these tracts cannot choose healthy food if they do not have access to it. Access is foundational, and these strategic investments, based on access alone, could yield a substantial return in terms of quality and length of life.

Figure 1a: Florida Census Tracts Where Improved Food Access Can Substantially Improve Diet-Related Health Outcomes

Scroll down to next page for map with county identifiers.



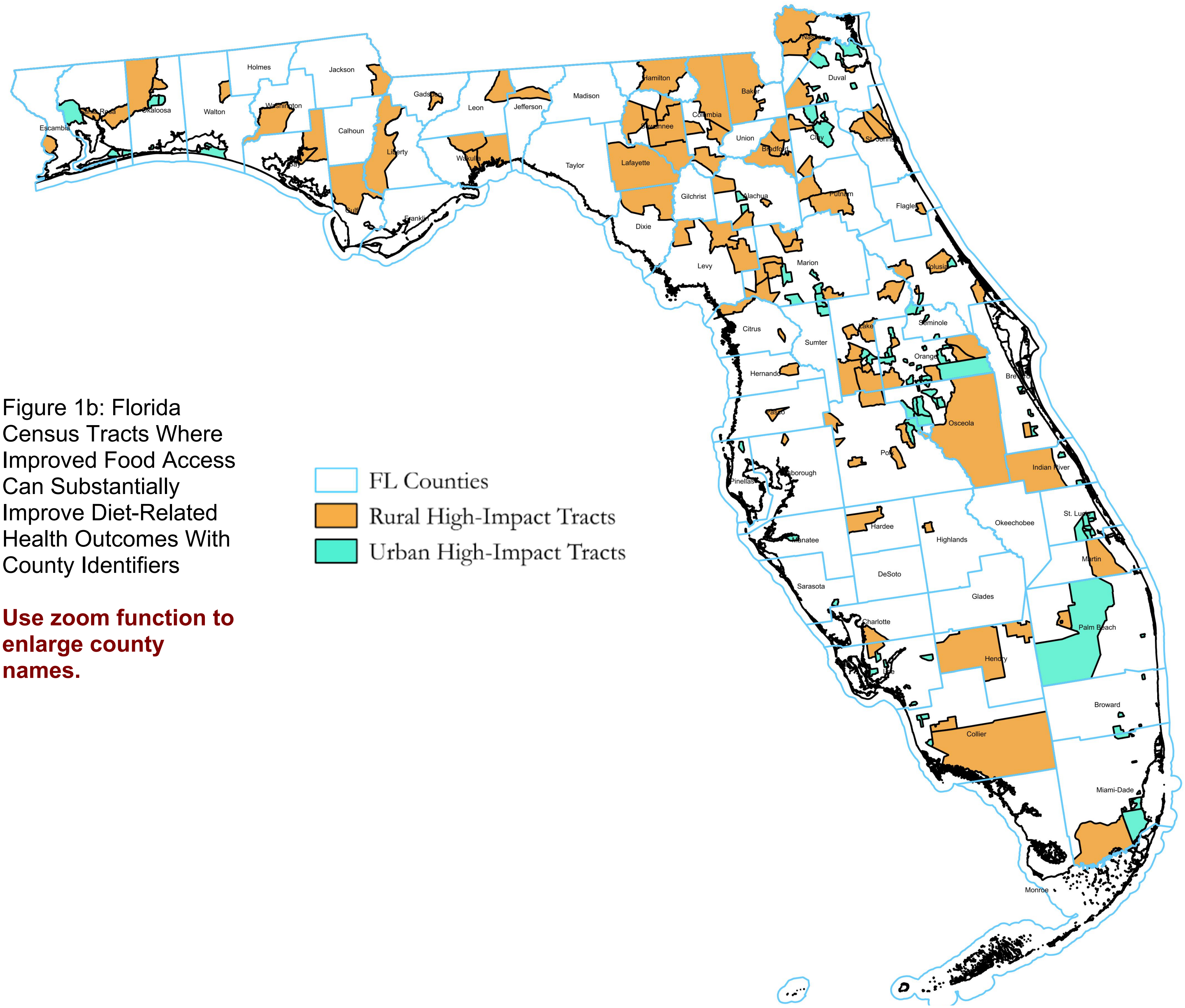


Figure 1b: Florida Census Tracts Where Improved Food Access Can Substantially Improve Diet-Related Health Outcomes With County Identifiers

Use zoom function to enlarge county names.

DETAILS

Data and Data Limitations

Our analysis used three sources of data: (1) deaths by cause for each 2010 Census tract for the years 2007-2013 from the Florida Department of Health; (2) USDA measures of food access by 2010 Census tract; and (3) data and boundary files for 2010 Census tracts based on the 2008-2012 American Community Survey (ACS), available from the National Historical Geographic Information System (NHGIS) at the Minnesota Population Center. Each file contained a unique identifier for each tract, so that they could be easily merged.

The death data were summed across the seven available years for each tract. A total crude death rate (i.e. not adjusted for the age structure of the tract's population or its deaths, as the latter was not available in the death data) per thousand residents was calculated for all diet-related causes as well as for accidents (the latter for validation purposes) using the tract's total population reported in the ACS. The death data have some shortcomings. First, some tracts have greater proportions of nursing and other residential group homes where deaths occur. However, this would only affect our analysis if there was a systematic relationship between food access and where these facilities were located, for which we found no evidence.

The second limitation of the death data is that not all deaths were linked to a Census tract. Errors can occur when hospital staff enters the decedent's home address into the death record. Because of this mistake, when state epidemiologists attempt to geocode all statewide deaths to a Census tract, some attempts fail. Those failed attempts are often listed as "unknowns" and are not assigned to any Census tract. However, an original transcription error might result in a second problem: it might successfully geocode, but to the *wrong* Census tract. For example, in 2007, there were 167,708 deaths recorded in the state of Florida, but only 158,303 of them were linked to a Census tract, leaving 9,405 deaths with the Census tract unknown. Some type of original transcription error likely caused the "unknown" result. However, even among those assigned to a tract, we suspect that a small number were assigned to the *wrong* tract. No dataset is perfect; all have some margin of error. The State of Florida errors were not substantial and did not hinder our ability to analyze the data or report findings with confidence.

The USDA food access measure also has some shortcomings. It takes each tract and calculates the fraction of the population that is more than some specific distance (such as a half-mile, 1 mile, 10 miles, or 20 miles) from the nearest large grocer as the USDA defines it in this measure. This is another shortcoming; first, other types of smaller grocers and other retailers supply healthful food. They need not be large. Second, there is very little data cleaning and no field checks. So some "large grocers" might not actually be large grocers and other true "large grocers" might not be in the data. That is a problem we can correct in any subsequent "drilldown" analysis of smaller, priority geographies as a next step, but not immediately in this analysis.

For each tract, the USDA reports four distance measures. None of these is ideal. The best method, usually only practical for lower geographies such as a county or city, is to calculate all distances from each block to the closest grocer. Given the USDA's objective in providing data across the US, a preferable metric to what they currently employ might be the mean or median distance that must be traveled from each tract to the nearest grocer. That would be doable within the same amount of time and resources.

Despite these shortcomings, the data are still useful and meaningful.

We experiment with each of the four USDA measures in the following analysis in order to determine if any of them is associated with diet-related health. The USDA augments these measures by creating other measures that combine distance to the nearest grocer with other tract-level characteristics. For example, it reports for each tract the fraction of the population that was *both* a half-mile from the nearest grocer *and* without access to a vehicle.

Urban Statewide Analysis

Our first step was to analyze urban and rural tracts separately because the relationship between food access and diet-related health outcomes will vary across densely and sparsely populated areas due to many interrelated factors. One way to adjust for distinct urban and rural characteristics is to test different distance thresholds to grocers within each urban and rural grouping. The USDA uses the U.S. Census definition for urban and rural tracts. It states: "A Census tract is urban if the geographic centroid of the tract is in an area with more than 2,500 people; all other tracts are rural." The USDA Food Access Research Atlas provides two urban food access measures (the fraction of a tract's population that resides more than 0.5 miles and the fraction of a tract's population that resides more than 1.0 miles from a large food retailer) and two rural measures (10 and 20 miles). An analysis that pools urban and rural tracts, even when allowing for different urban and rural distance-to-large-grocer thresholds, assumes that the relationships between diet-related health and all other tract-level characteristics (such as median household income and other variables) are the same in urban and rural places. Although there are additional shortcomings with the USDA's methodology (see the "Data and Data Limitations" section), this particular problem can be overcome by analyzing urban and rural locations separately.

The obvious first step in examining the link between diet-related health and food access is to look at the health metric most closely related to nutrition: diabetes. We analyzed this link for each of Florida's 3,633 urban Census tracts that have enough population to be methodologically considered in the analysis. For each tract, we used the USDA's measures of food access, tract-level characteristics from the 2008-12 American Community Survey (median household income, percent African American population, percent population age 65 and older, and percent population of persons age 25 and older with at least a high school degree), and tract-level death data by cause from years 2007-13. Deaths were aggregated across the seven years and converted to crude (i.e. not age-adjusted) death rates by dividing deaths by the total tract population.

When we used the USDA's calculation of the percent of each tract's population more than a mile from a large grocer, we found no statistically significant relationship between this food access measure and diabetes. The regression coefficient was actually negative, mistakenly suggesting that worse access was associated with lower diabetes death rates. We say “mistakenly” because the coefficient was very small in magnitude: a change of one standard deviation or 0.33 in the percentage a mile or more from a grocer corresponded to a drop in diabetes death rates of 0.003 per thousand, compared to the average diabetes death rate in the sample of 1.92 per thousand. Therefore, this effect was not only small in a substantive sense (i.e. even if taken at face value, its magnitude was negligible) but also not statistically significant. The p-value for the regression coefficient was 0.88, meaning that we can have little confidence in the strength of the finding. In other words, the result was probably a result of noise in the data: as we would suspect, worse healthy food access is not meaningfully associated with a drop in diabetes deaths.

Our next step was to consider whether the lack of an obvious association between food access and diabetes (as described above) was the result of an inadequate food access measure. Therefore, to test this further, we repeated the analysis, except this time we replaced the USDA's one-mile metric with its half-mile metric (i.e. the percentage of each tract's population more than a half-mile from a large grocer). This increased the substantive strength of the relationship between food access and diabetes deaths by a factor of 10 (a one standard deviation or 0.29 improvement in access was associated with a drop in the diabetes death rate of 0.03 per thousand). However, the associated p-value was only 0.13, so we again could not yet be very confident that this was not the result of a pattern in the data unrelated to food access.

However, the physical proximity to the nearest grocer is not the only consideration in the link between food access and diet-related outcomes. For example, a half-mile distance could have a very different association with health when households have either adequate or inadequate means of transportation. We are able to explore this possibility further because the USDA provides yet another food access measure: the percentage of households in each tract who both reside more than a half-mile mile from their large grocer and have no access to a vehicle. This measure had an association with the diabetes death rate that was large in magnitude (a one standard deviation or 0.05 improvement in this access measure was associated with a drop in the diabetes death rate of 0.22 per thousand, or eleven percent of the diabetes death rate). The association was highly statistically significant as well. The p-value was less than 0.0001 so we can be confident that this is not simply the result of noise in the data. The change in the association between access and diabetes deaths highlights the importance of considering not just distance itself, but, instead, distance together with the ability of households to travel that distance more quickly or easily. For example, if you live in an urban area of Florida and your household does not have a car, it might not be easy to take a bus or taxi to the healthy grocery store or to get a ride there from a friend. A return trip carrying bags of groceries may prove even more difficult.

Though the observed link between grocery store access (this time properly measured to account for vehicle access) and death by diabetes is itself an important finding, it by no means captures the full cost to society of inadequate food access. This is because we have until now been considering only deaths in which diabetes was listed as the primary cause. But diabetes also contributes to a great many other deaths, and a great many other causes of death are substantially diet-related. For example, nationally in the year 2010, roughly 70,000 deaths had diabetes listed as the primary cause, but more than three times as many deaths (234,000) included diabetes as a *contributing* cause. This leads us to an additional analysis in which we assess the link between food access and all causes of death that have a known relationship to diet: diabetes, certain cancers, heart disease, stroke, and liver disease.

Moving to this more comprehensive measure of diet-related deaths reveals another strong association between large grocery access and health. Using the half-mile access measure that also accounts for vehicle access and all diet-related deaths, we find the relationship is again highly significant in a statistical sense ($p<0.0001$) and again large in a substantive sense: a one standard deviation or 0.05 improvement in this access measure was associated with a drop in the diet-related death rate of 1.43 per thousand, or ten percent of the diet-related death rate.

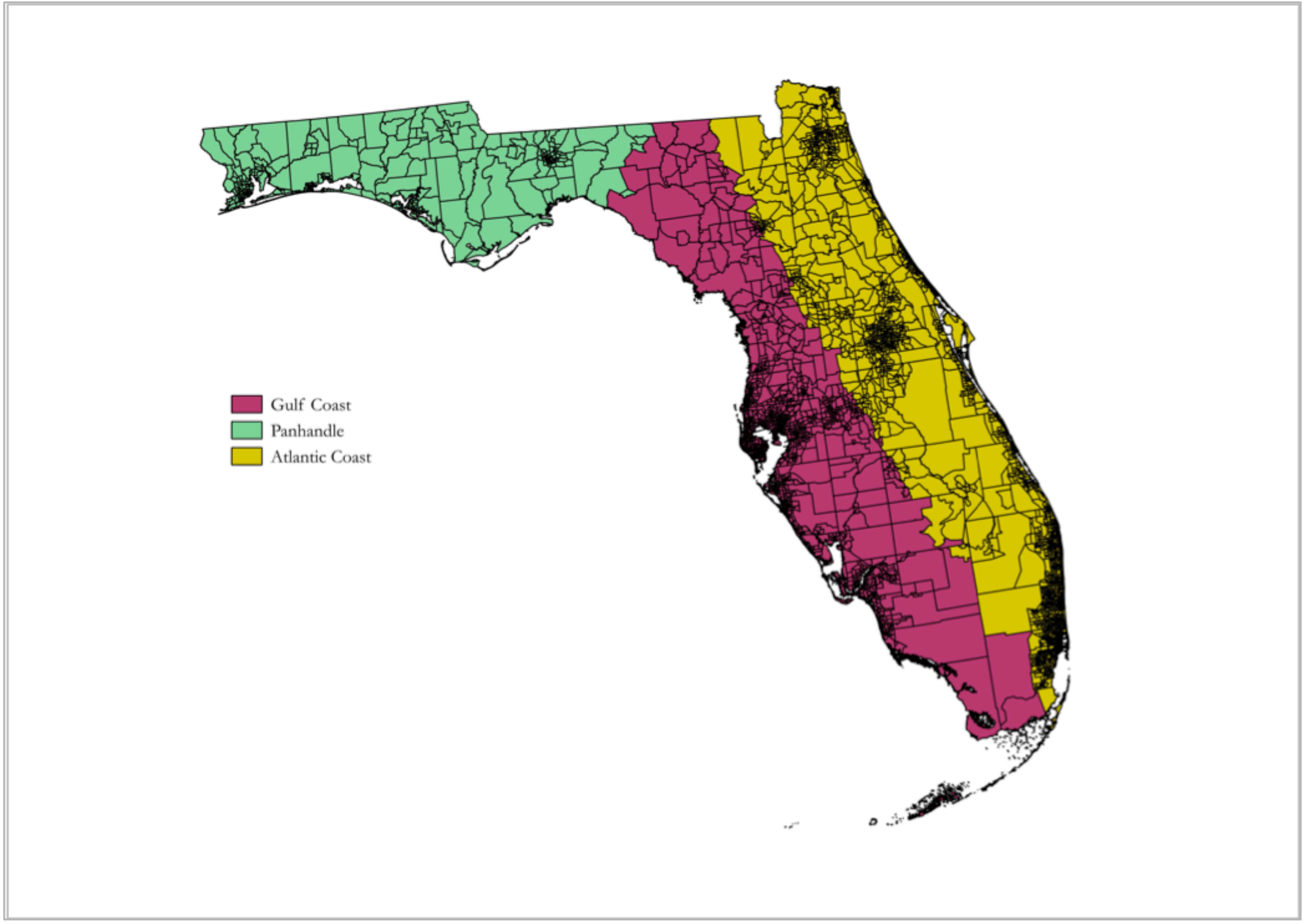
Next, we explore the possibility that food access is simply a proxy for some other, unobserved, tract-level characteristic that is in fact driving tract-level death rates. In this case, the observed link from food access to health may simply reflect the fact that some tracts, perhaps for a variety of reasons, have both higher death rates and worse food access than others. To reduce this concern, we conducted a final urban analysis in which the cause of death examined was motor vehicle accidents. This cause of death should be unrelated to food access, or at least exhibit a much weaker relationship to food access compared to diabetes or all diet-related causes of death. This was in fact the case: the observed relationship was actually negative and very small in magnitude. For example, a one standard deviation or 0.05 improvement in the half-mile food access measure without vehicle access was associated with a drop in the diabetes death rate of 0.003 per thousand, or 0.3 percent of the motor vehicle accident death rate, $p=0.797$.

Finally, in order to account for spatial autocorrelation (the tendency of geographic units that are nearby in physical space to have unobserved characteristics that lead to spurious correlations among observed characteristics), we also re-analyzed the relationship described above, using the correction proposed by Conley (1999) and implemented in Stata by H. Siang's "ols_spatial_HAC" procedure. Tracts within 15 miles of each other were assumed to be spatially correlated. Though this correction increased standard errors and slightly reduced the statistical significance (i.e. led to higher p-values) in some cases, none of the substantive findings reported above were altered. For example, in the analysis of deaths in urban places from all diet-related causes, the food access measure at a half-mile with adjustment for vehicle access had a coefficient of 26.699 and a standard error of 2.011 ($p<0.0001$) in our original analysis; after adjusting for spatial autocorrelation, the coefficient remained unchanged and the standard error rose to 5.671 ($p<0.0001$). References: See T.G. Conley "GMM Estimation with Cross Sectional Dependence," Journal of Econometrics, Vol. 92 Issue 1 (September, 1999), pp. 1-45.

Urban Analysis By Region

Just as we anticipate that the relationship between food access and health would differ between urban and rural places, and that simply using different food access mileage cut-offs would not fully address this concern, we would anticipate these relationships to differ across regions. However, to know for certain, this must be explored and tested. Therefore, we analyzed the urban tract data separately after placing each tract into one of three regions (see Fig. 2): the Panhandle, the Gulf Coast, and the Atlantic Coast.

Figure 2: Key Regions in Florida

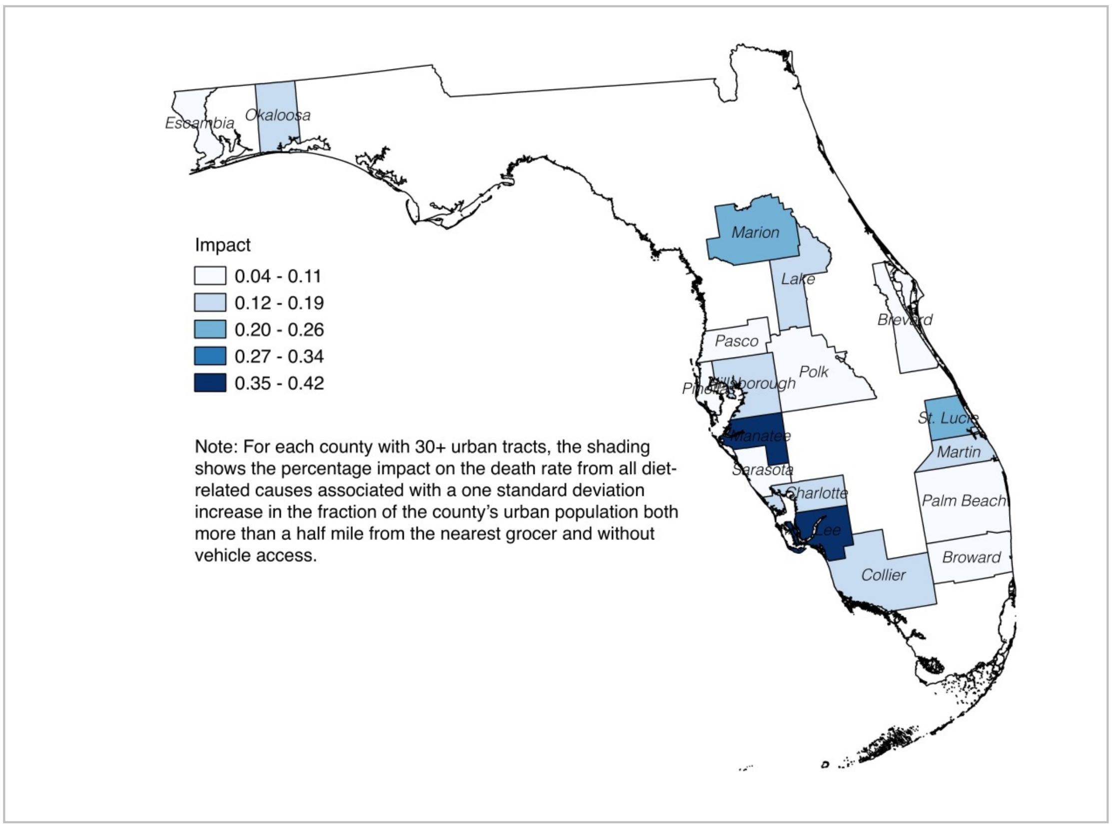


The relationship between the half-mile food access measure (adjusted for vehicle access) and the death rate for all diet-related causes was statistically significant in all three regions ($p<0.01$ for each region). The magnitude of the death rate change associated with a one standard deviation change in food access did differ somewhat across regions: it was seven percent in the Panhandle and in the Atlantic Coast, but fourteen percent in the Gulf Coast.

Urban County-Level Analysis

Twenty-eight of Florida's counties had sufficient urban Census tracts to allow separate urban analyses by county (each of the counties required 30 or more tracts to be methodologically included in the analysis). In each case, the control variables were the same as those used in the statewide and regional analyses, the health measure was the death rate from all diet-related causes, and food access was measured at a half-mile with adjustment for vehicle access. The association between health and food access was statistically significant in 20 of the 28 counties. The percentage change in the death rate associated with a one standard deviation change in food access ranged from 4-5% in Pinellas and Miami-Dade counties to more than 20% in Marion, Lee, and Manatee Counties (Fig. 3).

Figure 3: Percentage Change in the Death Rate for All Diet-Related Causes Associated With a One Standard Deviation Improvement in Food Access



Rural Statewide Analysis

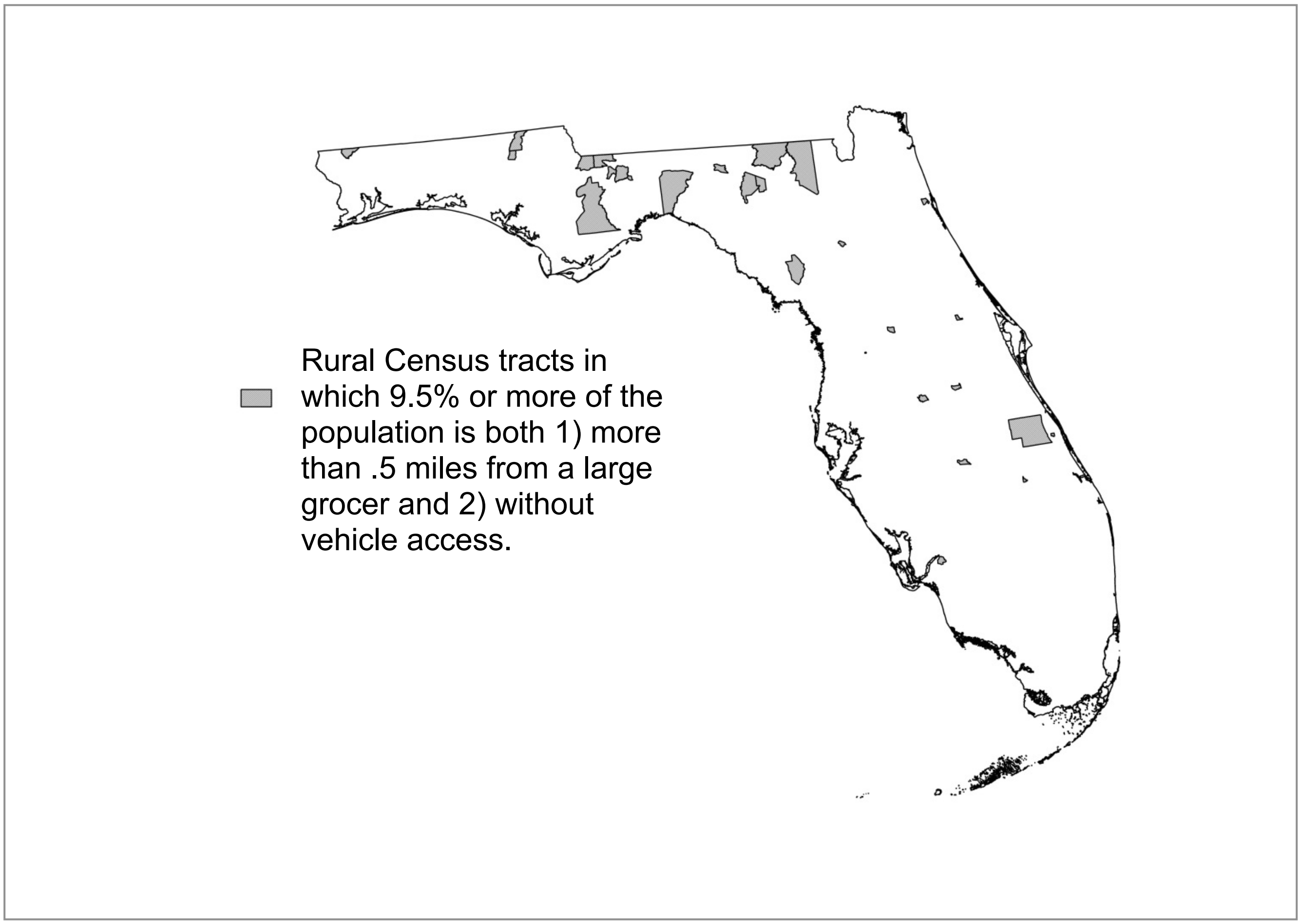
The USDA provides two measures of food access designed specifically for rural places (tracts in which the largest place is less than 2,500 in total population): the fraction more than 10 miles and the fraction more than 20 miles from a large grocer. Neither of these was associated with diabetes deaths nor deaths from all diet-related causes. The USDA also provides the same measures as were used in the urban tract analysis (the fractions more than a half-mile and more than a mile from a large grocer). These latter two measures, after adjustment for vehicle access, were strongly related to diet-related death rates. That is, the fraction of a rural tract's population that was both without vehicle access and more than a half-mile or a mile from a grocer was positively associated with diet-related death rates ($p<0.002$, and a one standard deviation rise in the fraction with poor access and without vehicles was associated with a rise in diet-related deaths of 5.6%). The effect is stronger for tracts where the fraction of the population with poor access by the half-mile distance measure with no vehicle access is above 9.5% than for tracts below 9.5%. In fact, an additional percentage point of poor food and vehicle access was linked to a rise in diet-related deaths 2.5 times greater when the fraction with poor access was greater than 9.5% than when the fraction was below 9.5%.

The absence of a substantial association between access and health at the higher distance measures created by the USDA for rural places and the presence of such a strong association instead at the lower cutoffs created for urban places is somewhat surprising. Recall, however, that effective access to a grocer experienced by households is a function not just of the straight-line distance to the nearest grocer but of the time it takes to travel that distance. Time is a function of distance, but also a function of available transportation and the presence of natural and man-made barriers within the local geography. Real distance is not “as the crow flies” (e.g. the immediate and shorter straight-line path through a forest or a farmer’s field) but the time it takes to reach the bridge that finally crosses the river or the more distant road that finally reaches the grocery store.

For rural households with access to a vehicle, a distance of 10 or 20 miles can be traveled in a matter of minutes; for rural households without access to a vehicle, as for urban households without a vehicle, distance takes on different meanings. For example, for a rural household without a car, 10 miles to the grocery store is a substantial distance. But so is 5 miles. Even a mile or a half-mile for rural dwellers can be a long way to walk and make the return trip carrying groceries. A grocer that is a half-mile or a mile away is likely the same as one that is 20 miles away for some rural households without a car; it is unreachable in terms of access to healthy groceries. In addition, a walk of a half-mile or a mile might actually be more difficult in rural places than in urban places, as rural places seldom have sidewalks and may often be crossed by highways that can be reached only after a significant deviation from the straight-line distance to the grocer. Perhaps some residents, at times, might receive a ride from a friend or relative, but this option might not be consistent, convenient, or reliable. This finding emphasizes the need to identify food access solutions in rural places that are targeted specifically at areas where grocers are more than a mile away and where transportation is not readily available.

Figure 4 shows the 29 rural tracts where the link between access and health is the strongest: those places in which 9.5% or more of the population is both (1) more than a half-mile from a grocer and (2) without vehicle access.

Figure 4: High Impact Rural Tracts in Florida



<u>Sites with a Large Impact of Improved Access</u>

Several of the previous analyses have highlighted particular areas where interventions to improve food access have the potential to generate substantial health improvements. Across the state, tracts with high numbers of households as little as a half-mile from the nearest grocer can benefit from improved grocer access in both urban and rural places. At the regional level, improved grocer access is likely to generate improved diet-related health in some areas more than others. And counties such as Marion, Manatee, and Lee exhibited a particular sensitivity of diet-related health to food access.

We now dig somewhat deeper in identifying tracts across the state likely to yield the highest public health improvement payoff to improving food access. In order to do so, we use the results of the statewide regression analyses we conducted previously for urban and rural tracts. That analysis generated estimates of the relationship between tract-level characteristics (food access accounting for vehicle access, median household income, racial composition, high school graduation rate, and age composition) and an important tract-level health indicator: the death rate from all diet-related causes. The characteristics explained 60% of the variation in the death urban places and 45% in rural places. We can then take the estimated relationships between characteristics and diet-related health together with each tract's actual characteristics and predict the death rate for each tract. For example, the regression analysis reveals the following relationship between diet-related health and tract-level characteristics in urban places:

All Diet-Related Cause Death Rate = 6.65 + (26.70 x Food Access)
+ (-0.09 x Median Income in Thousands) + ...

The "+..." indicates other factors that we will ignore for this example. For a tract with a Food Access score of 0.05 (i.e. 5% of the population is both more than a half-mile from a grocer and without a vehicle) and a Median Income of $25,000, the predicted All Diet-Related Cause Death Rate would then be as follows:

All Diet-Related Cause Death Rate =
6.65 + (26.70 x 0.05) + (-0.09 x 25) + ... = 5.74

We can now predict the impact of improved food access by choosing a different value for Food Access, recalculating the predicted death rate, and comparing this to the original predicted death rate to assess the improvement in diet-related health. For this example, if we improve food access by one percentage point, the new death rate would be:

All Diet-Related Cause Death Rate =
6.65 + (26.70 x 0.04) + (-0.09 x 25) + ... = 5.47

Therefore, the death rate improvement for this tract is 5.74 - 5.47 = 0.27 deaths per thousand population. If the tract has 5,000 residents, the total health improvement for the tract is 0.27 x 5 = 1.35 deaths over 7 years.

We perform this exercise (using the one percentage point improvement) and take the 100 tracts that show the greatest improvement in urban places and the 100 tracts that show the greatest improvement in rural places. These are shown in Fig. 1a and 1b in the first section of this report.