Towards Improving Local Policy Responses to Food Insecurity: Exploratory and Predictive Analytics

Henning Tovar
Lakshmi Iyer
Charlie C. Chen

Follow this and additional works at: https://aisel.aisnet.org/sigdsa2019
Towards Improving Local Policy Responses to Food Insecurity: Exploratory and Predictive Analytics

Henning Tovar
Appalachian State University
tovarh@appstate.edu

Lakshmi Iyer, PhD
Appalachian State University
iyerls@appstate.edu

Charlie Chen, PhD
Appalachian State University
chench@appstate.edu

Abstract

Academic research and policy proposals that aim to alleviate food insecurity in the United States can benefit tremendously from new developments in analytics. Much academic literature has singled out predictors for food insecurity. However, the research so far often focuses on either the individual or the aggregated state or country level. Both levels of analysis are either too granular or too broad to inform public policy effectively. The research that does focus on medium-size units of analysis, on the other hand, lacks incorporation of non-linear effects at the community level. Our research aims at incorporating the literature of social capital to account for potential non-linearity. We propose a case study at the county level in North Carolina that can be extended to a national scale.

Keywords

Food Insecurity, Predictive Model, Non-Linearity, Social Capital

Introduction

The reference after the next sentence by Coleman-Jensen et al is an official report by USDA. The definition of food insecurity is in the report.

Food insecurity defined by the United States Department of Agriculture (USDA) is insufficient access to food for an active and healthy lifestyle is a widespread problem. Currently, about 11.1 percent of households in the United States experience some level of food insecurity (Coleman-Jensen, Rabbitt, Gregory, & Singh, 2019). Among the various states, the population that lacks consistent food security differs considerably. In New Mexico, almost 17 percent of the households experience some degree of food insecurity, while less than 8 percent of the households residing in New Hampshire lack consistent access to food. Furthermore, research has pointed towards geographic determinants of food insecurity, so-called food deserts (Beaulac, Kristjansson, & Cummins, 2009); Walker, Keane, and Burke (2010). Within mostly rural areas, access to fresh and healthy food is especially limited, leaving these areas particularly vulnerable.

Inconsistent access to sufficient nutrition has many negative consequences. Among children low food security has detrimental health outcomes and threatens overall child development on a physical and emotional level (Howard, 2011; Jyoti, Frongillo, & Jones, 2005; Wight, Kaushal, Waldfogel, & Garfinkel, 2014). A lack of food security can negatively impact children’s educational attainment (McIntyre, Kwok, & Patten, 2017) particularly. Furthermore, minorities and households with low socioeconomic status are especially at risk of experiencing food insecurity. The lack of food forces these groups to make a tough decision between food and for example health care. Overall, the limited access to nutrition creates a vicious circle that reinforces poverty among already vulnerable groups (Dean & Sharkey, 2011; Sharkey, 2005; Wight et al., 2014). The need to formulate a policy response to food insecurity is thus very clear. Ending hunger, achieving food security and improved nutrition, and promoting sustainable agriculture also fall under Goal #2 of the United Nations 17 Sustainable Goals (sustainabledevelopment.un.org/sdg2). In the
U.S., apart from federal programs like the Supplemental Nutrition Assistance Program (SNAP), most state-level policies rely on charity efforts to improve food security. Charities, however, face considerable uncertainty when relying on in-kind donations to alleviate hunger (Davis, Jiang, Morgan, Nuamah, & Terry, 2016). Our research, hence, strives to inform policymakers that aim at tackling food insecurity below the federal level. We firmly believe that analytics can be used to improve policy proposals based on carefully selected data.

The variety of accessible data makes it possible for analysts to tap into unusual data sources (Du, Gebremedhin, & Taylor, 2019; Kolak et al., 2018). Furthermore, researchers can gain insight from blending different kinds of data, including classical survey data as well as geospatial data (Pokhriyal & Jacques, 2017). The existing literature on food insecurity already makes use of newly developed data science techniques and machine learning algorithms in some cases. Often these machine learning applications are implemented to improve decision on the federal or global level (Christensen, Srinivasan, Hart, & Marshall-Colon, 2018; Pokhriyal & Jacques, 2017; Vega, Hinojosa, & Nguyen, 2017) or at the most granular, individual level (Hossain, Mullally, & Asadullah, 2019). However, while informative in an academic setting, these levels of analysis are often too broad or granular for policy makers. A burgeoning strain of research has focused on food insecurity on the census tract or county level which seems to be a size optimal for efficiently tailored policy solutions (Borders, Ferris, Beeby, & McCahill, 2018). Some more advanced models with a nationwide scope, even account for differences among states and counties with multilevel modeling (Gundersen, Engelhard, & Waxman, 2014; Gundersen & Ziliak, 2014). However, even these advanced models do not allow for non-linear effects among indicators of food insecurity. Building on the theory of social capital (Bourdieu, 1986), we suspect that tight-knit neighborhoods will be more resilient when facing food security threats compared to less tight-knit communities. Our research thus focuses on the county level modeling of food-insecurity. At the current point in our project, we are developing a case study in North Carolina and will then continue to broaden the scope of our project to the national level.

**Literature Review**

The literature on food insecurity has established a selection of variables that indicate low food security (Borders et al., 2018; Gundersen et al., 2014; Hunt, Benjamins, Khan, & Hirschtick, 2019). Among the most influential variables is income, implying that groups with lower socioeconomic status (SES) are at higher risk of food insecurity (Wight et al., 2014). Further, financial management skills contribute to this variable. Households with poor financial management skills (i.e., budgeting, management of bills, etc.) are at a relatively higher risk of food insecurity (Gundersen & Garasky, 2012). As pointed out by Leung et al. (2014), SNAP participation rates can be used as a proxy for estimating poverty. As a federal social welfare program SNAP is designed for income groups below a certain threshold.

Furthermore, the income of local peers seems to influence the risk of food insecurity. As a study conducted by Morrissey, Oellerich, Meade, Simms, and Stock (2016) has highlighted, poor neighborhoods impose somewhat of a neighborhood effect. Individuals living in those neighborhoods and put them at an elevated risk of food insecurity.

Racial minorities are, furthermore, generally at an increased risk of lacking stable access to nutrition (Sano, Garasky, Greder, Cook, & Browder, 2011; Vedovato et al., 2016). This increased risk can be found across minority groups. Some specific marginalized groups, however, are at high risk of lacking food. As Quandt, Arcury, Early, Tapia, and Davis (2004) have pointed out, migrant seasonal and farmworkers are very likely to experience food insecurity since they are in low-paying jobs and, furthermore, often send money back to their home country. It should be noted that using race as an indicator requires other variables as control. In general, racial minorities receive lower salaries and are thus already at an elevated risk of food insecurity (Mitra & Brucker, 2017).

The detrimental effect of food insecurity does not only affect children and youth harder than adults, but youth are also at higher risk of food insecurity (Baer, Scherer, Fleegler, & Hassan, 2015). Since children lack independent income, they are dependent on their parents and thus somewhat inherit food insecurity. This risk is demonstrated and alleviated among children of single mothers (Wight et al., 2014). Furthermore, housing circumstances have a substantial influence on the risk of low food security. Homeownership is generally associated with a lower probability of low food security (Guo, 2011). In particular, households that
spend more than 30 percent of their income on rent are at a high risk of experiencing some level of food insecurity (Kirkpatrick & Tarasuk, 2011).

Social capital is a variable not commonly included in models of food insecurity. A concept often applied in sociology, social capital indicates the resources available to an individual based on their social setting and connections (Bourdieu, 1986). Similar to financial capital, an individual with many social connections can utilize this form of capital to advance in life. In a professional context, a more colloquial phrase for social capital would be network. In the context of food insecurity however, social capital is the sum of relationships an individual can rely on to alleviate limited access to food (Dean & Sharkey, 2011). Family and community networks can supply food-insecure households with the necessary resources to overcome hardship (Martin, Rogers, Cook, & Joseph, 2004). Research has found consistent evidence for the mitigating effect of social capital on food insecurity (Johnson, Sharkey, & Dean, 2010; Locher et al., 2005; Morrissey et al., 2016). These evidence point towards our assumption of non-linear effects among indicators of food insecurity. Black neighborhoods, for example, above a certain threshold, might have a mitigating impact on food insecurity due to close-knit neighborhood effects. Garasky, Morton, and Greder (2006) have found that social capital networks among rural communities have an alleviating impact on food insecurity.

Methods & Data

As our project is research in progress, we have not yet collected our final datasets. Our preliminary analysis is limited to counties in North Carolina (N=100) for the year 2017. We collected county-level data on gender, median income, urban vs. rural, and race from the American Community Survey (ACS) administered yearly by the US Census Bureau. ACS data is collected on questions that are also part of the decennial census. The ACS is not a full census but rather pools data over 12-, 36-, 60-month intervals (U.S. Census Bureau, 2013).

Further, we collected data on unemployment rates from the North Carolina Chamber of Commerce. Another variable we collected from the NC Chamber of Commerce was information on the economic tier a county has. Economic tiers are a compound measure issued by the NC Chamber of Commerce to compare the economic performance of counties amongst each other. The tiers range from 1 for the worst-performing counties through 3 for the best performing counties. NC Chamber of Commerce uses this metric to offer tax incentives to companies that are interested in investing in an economically suppressed county (North Carolina Chamber of Commerce, 2019). Finally, we retrieved data on county-level food insecurity from Feeding America’s Map the Meal project. The same data sources could also be used when widening the scope of our project to a national level. We used R for data processing, analysis, modeling and Tableau for interactive visualizations.

While we use social capital as a foundation, we have not yet included a measurement of social capital. Measuring social capital on the county level is particularly challenging because of the inherently contextual nature of social capital. Firstly, measuring the phenomenon of social capital might require a somewhat mixed-methods approach to incorporate the realities of those bearing social capital. Furthermore, switching between aggregated county-level data and individual-level social capital data introduces the risk of running into an ecological fallacy (Robinson, 1950). Ecological fallacies are a common risk in the social sciences that occur when using aggregated level data to explain individual-level behavior. An example of this would be utilizing the average SAT score of a college to conclude the SAT of a single student attending the college (Singleton & Straits, 2010). However, recent academic research has found evidence that when using multilevel regression models, ecological fallacies can be avoided while still utilizing data from different levels of analysis (Subramanian, Jones, Kaddour, & Krieger, 2009).

Building on the existing literature and indicators found significant for predicting food insecurity; we are aiming at eventually building a multilevel regression model. The final model will allow for variation between states and counties, and furthermore, allows for non-linear effects for input variables. Furthermore, we can include the effects of social capital. Up to this point, we have designed a case study for North Carolina to explore our assumption of non-linearity. We explored the data using Tableau built-in functions and particularly mapped population size against food insecurity, and the compound measure economic tier. We then went on to construct a simple linear regression to explore the signal in our data.
Preliminary Results

Our initial descriptive analysis found a wide spread of food insecurity ranging from 9.5 percent to 24.1 percent. The average food insecurity for North Carolina is 15.4 percent (standard deviation: 3.46 percent). Furthermore, we found a statistically significantly different average food insecurity rates for counties in different economic tiers. Counties that were classified as part of the worst-performing economic tier (Tier 1) had an average food insecurity rate of 18.4 percent (standard deviation: 3.1 percent) with a median value of 18.5 percent. The difference between the economic counties is so stark that on average, counties that are part of the worst-performing economic tier would be considered statistical outliers if they were part of any of the other economic tiers. For full descriptive results for counties and tiers, see Table 1.

<table>
<thead>
<tr>
<th>Counties in Economic Tier</th>
<th>N</th>
<th>Average Food Insecurity</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Interquartile Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic Tier 1</td>
<td>40</td>
<td>18.4%</td>
<td>18.5%</td>
<td>3.1%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Economic Tier 2</td>
<td>40</td>
<td>13.8%</td>
<td>13.5%</td>
<td>1.9%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Economic Tier 3</td>
<td>20</td>
<td>12.2%</td>
<td>12.2%</td>
<td>1.9%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Statewide</td>
<td>100</td>
<td>15.4%</td>
<td>14.6%</td>
<td>3.5%</td>
<td>5.3%</td>
</tr>
</tbody>
</table>

Table 1. Descriptive Statistics on Food Insecurity in North Carolina

We, furthermore, used Tableau’s advanced data visualization applications to gain insight on the distribution of food insecurity rates about population size. Figure 1 shows a county level map of North Carolina. The different counties are colored depending on their level of food insecurity with shades of red for counties above the average rate of food insecurity. Counties colored in shades of blue are below the average rate of food insecurity. The map clearly indicates a concentration of highly food insecure counties in the North-East and South of the state. All other interactive visualizations that we utilized for our exploratory analysis can be found on Tableau Public.
We found that while counties with the smallest population often suffered from the highest food insecurity rates, some of the more heavily populated counties experienced substantially higher food insecurity rates than other counties of a comparable size. To analyze this relationship, we constructed a multiple linear regression model that included indicators discussed above. The full specifications and regression coefficients can be found in the appendix. However, at this stage of the research process, our attempts to allow non-linearity in the model did not yield statistically significant results. We suspect that including data on social capital into the model would yield statistically significant non-linear effects. However, these data are not readily available and will be collected as part of future research.

Conclusion

The results of our research-in-progress demonstrate that there much more ground for us to cover. While we are currently unable to include significant non-linear effects in our regression model, we expect this to change as soon as we can add data on social capital.

Furthermore, we hope to widen the scope of our analysis to the national level, to test our hypothesis against a bigger dataset. From the results above, however, we hope to have already demonstrated that food insecurity is a problem prevalent in North Carolina to which policymakers should pay attention. Our final aim is to provide policy and decision-makers with data-based results that are compelling enough to turn into actionable policy proposals. However, the preliminary results using visualization techniques already imply that breaking food insecurity down to the county level will yield results that allow policy makers to produce more efficient policies.

References


Dean, W. R., & Sharkey, J. R. (2011). Food insecurity, social capital and perceived personal disparity in a predominantly rural region of Texas: An individual-level analysis. *Social Science & Medicine, 72*(9), 1454-1462. doi:https://doi.org/10.1016/j.socscimed.2011.03.015


Appendix

The regression table below shows the full model built for our predictive linear regression. When interpreting the coefficients it is important to keep in mind that variables are unstandardized. Most variables are measured in full units (any variable measuring population or race) and only few are measured as rates (unemployment rate or median age). This means that most coefficients are very small in their influence at first but will have a substantial influence overall. Urban Population in Urban Areas has a coefficient of 0.00001 but with a population of 107971, this is a full percentage increase in predicted food insecurity.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.279†</td>
<td>0.0289</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Urban Population (Urban Areas)</td>
<td>0.000001†</td>
<td>0.000003</td>
<td>0.0445</td>
</tr>
<tr>
<td>Urban Population (Urban Cluster)</td>
<td>0.000001†</td>
<td>0.000003</td>
<td>0.0004</td>
</tr>
<tr>
<td>Rural Population</td>
<td>0.000001†</td>
<td>0.000003</td>
<td>0.0004</td>
</tr>
<tr>
<td>Median Age</td>
<td>-0.00228†</td>
<td>0.0006</td>
<td>0.0005</td>
</tr>
<tr>
<td>White Population</td>
<td>-0.000001†</td>
<td>0.000002</td>
<td>0.0001</td>
</tr>
<tr>
<td>Black Population</td>
<td>-0.000009†</td>
<td>0.000002</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Native American Population</td>
<td>-0.000001†</td>
<td>0.000002</td>
<td>0.0004</td>
</tr>
<tr>
<td>Asian Population</td>
<td>-0.000001†</td>
<td>0.000002</td>
<td>0.0001</td>
</tr>
<tr>
<td>Population Others</td>
<td>-0.000001†</td>
<td>0.000002</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Economic Tier 2</td>
<td>-0.03206†</td>
<td>0.0052</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Economic Tier 3</td>
<td>-0.04248†</td>
<td>0.0078</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>0.0018</td>
<td>(0.0014)</td>
<td>0.1269</td>
</tr>
</tbody>
</table>

Dependent Variable = Food Insecurity Rate (0.00-1.00).

Table 2. Predictive Model, Regression Coefficients and Goodness-of-Fit Statistics

N: 100

R²: 0.73

Adj. R²: 0.69

† significant at p<.05