An Analysis of the Local Food Environment Factors Affecting Produce Consumption of Individuals in FoodAPS Using Directed Acyclic Graphs

Mark C. Senia, Agricultural Economics, Texas A&M University
Senarath Dharmasena, Agricultural Economics, Texas A&M University

Background

The main result of eating a large amount of calorie-dense foods is a rise in obesity. As explained by Finkelstein, Ruhm, and Kosa (2005) there are numerous economic causes and consequences of obesity among adults and children in the United States. Obesity is a major risk factor for diabetes, cardiovascular disease, cancer, sleep apnea, nonalcoholic fatty liver disease, osteoarthritis, and other problems (Ahima and Lazar, 2013). According to Dharmasena and Capps (2012), two-thirds of adults in the United States are either overweight or obese. Ogden *et al*., (2014) shows that childhood obesity has more than doubled in children and quadrupled in adolescents in the past 30 years in the United States. This rate has slowed, as there has been no significant changes in the obesity prevalence in youth or adults between 2003-2004 and 2011-2012.

The food consumption behavior of SNAP households has been studied. Liu, Kasteridis, and Yen (2013) find no evidence that SNAP participation promotes consumption of less healthy food away from home. Wilde and Ranney (2000) conclude that food spending by SNAP households peaks sharply in the first three days after benefits are received. For those who conduct major grocery shopping trips only once per month (42% of all SNAP households), calorie intake drops by the fourth week of the month. Results on the relationship between SNAP participation and food security are subject to problems of selection bias and endogeneity (Gundersen and Oliveira 2001; Jensen 2002). Gundersen and Oliveira (2001) show that once one controls for adverse selection, food stamp recipients have the same probability of food insufficiency as non-recipients.

Taylor and Villas-Boas (2016) use the new National Household Food Acquisition and Purchase Survey (FoodAPS) with a multinomial mixed logit model to estimate food store choices as a function of type and household attributes. They find that households are willing to pay between $12 and $17 per week in distance traveled for Superstores, Supermarkets, and Fast Food, while they are willing to pay significantly less for the remaining outlets. They conclude that policymakers should consider incentivizing the building of the outlets that consumers are willing to pay more. This research will focus on food choices instead of store choices.

Dharmasena, Bessler, and Capps (2016) use directed acyclic graphs to model the food environment in the United States. The authors use state level data to model the food environment in contrast to the individual level data used in this research. The results indicate that food insecurity and SNAP participation are related but have no direct causal link. The authors also find that poverty and SNAP participation are related through back-door paths (food insecurity, unemployment, race and food taxes).

The objectives of this paper are to: (1) use local food environment variables to identify the causal structure associated with an individual's consumption of produce and (2) compare the causal structure of SNAP and non-SNAP households

Proposed Model and Estimation Procedure

Recent literature has begun to describe the inference of causal relationships from observational data in the absence of controlled experiments (Pearl, 2000; Spirtes et al., 2000). These methods rely on algorithms that allow causal inferences to arise without explicitly formed

hypotheses. The causal structures that arise from these algorithms can be represented in graphical form (the directed acyclic graph).

The TETRAD V software developed by Glymour et al. (2016) includes a number of algorithms to create casual structures. The FCI (Fast Causal Inference) algorithm uses a data set that can be represented by a graphical structure and is assumed to be represented by a linear structural equation model with normal error terms (Spirtes, Meek, and Richardson, 1999). The FCI algorithm allows undetermined edge when the algorithm cannot assign a direction to an edge. This algorithm allows for unknown latent variables and sample selection bias. The need to account for selection bias is especially important when evaluating government program such as SNAP. Not accounting for this would challenge the external validity of any results.

Data

The National Household Food Acquisition and Purchase Survey (FoodAPS) is a nationally representative panel of 4,826 U.S households containing information about each household's food purchases and acquisitions (ERS-USDA, 2016). Details were collected about foods purchased or acquired for consumption at home and away from home, including nutrition assistance programs. The survey is unique in that it oversamples SNAP households and low-income households not participating in SNAP in relation to higher income households.

FoodAPS was fielded between April 2012 and January 2013 and collected information on all food acquisitions and purchases by all members of the household over a seven-day period. Households had to scan barcodes, save receipts, and record other information in food journals. Information obtained from the household includes the quantities, prices, and expenditures for all at home and away from home foods and beverages purchased or acquired by all household members, eating occasions by household members, household characteristics (e.g. income, program participation, food security, health status, etc.), and household access to food (e.g. location of purchase and distance to food stores and restaurants) (ERS-USDA, 2016). The USDA added information about nutrient content of purchased food and the local retail environmental based on scanned barcodes and household locations.

Expected results and Potential for discussion

Few studies have combined directed acyclic graphs with individual level consumption and local food environment data. The expected results will shed light on the causal structure linking produce consumption of U.S. individuals to the local food environment. These results will guide policy makers because the complex casual structure of the local food environment means that these factors must be considered jointly and cannot be dealt with individually. For example, we will generate more detail about the casual structure linking obesity and produce consumption. If we find that a causal chain links produce consumption and obesity with some third local food environmental variable in the middle, conditioning on this third variable in a model would block the path from produce consumption and obesity and one would not be able to identify the effect of produce consumption on obesity.

These results will also cause discussion to be generated about the future of using directed acyclic graphs or some other artificial intelligence methodology to model complex local environments. Discussion about the best policy intervention given the causal structure found is an expected result of this research. It is also important to consider how these results could be useful for food retailers as well as government policymakers to identify strategies to improve health in the United States.

References

Ahima, R.S., and M.A. Lazar. 2013. "The Health Risk of Obesity-Better Metrics Imperative." *Science* 23:856-858.

Dharmasena, S., D.A. Bessler., and O. Capps Jr. 2016. "Food Environment in the United States as a Complex Economic System." *Food Policy* 61: 163-175.

Dharmasena, S., and O. Capps, Jr. 2012. "Intended and Unintended Consequences of a Proposed National Tax on Sugar-Sweetened Beverages to Combat the U.S. Obesity Problem." *Health Economics* 21(6):669-694

Economic Research Service, U.S. Department of Agriculture (ERS-USDA). 2016 National Household Food Acquisition and Purchase Survey (FoodAPS). http://www.ers.usda.gov/data-products/foodaps-national-household-food-acquisition-and-purchase-survey.aspx

Finkelstein, E.A., C.J. Ruhm, and K.M. Kosa. 2005. "Economic Causes and Consequences of Obesity." *Annual Rev. of Pub. Health.* 26:239-257.

Glymour, C., R. Scheines, P. Spirtes, and J. Ramsey. 2016. The TETRAD Project: Causal Models and Statistical Data. Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA, http://www.phil.cmu.edu/tetrad/current.html.

Gundersen, C., and V. Oliveira. 2001. "The Food Stamp Program and Food Insufficiency." *Amer. J. Agr. Econ.* 83(4): 875-887.

Jensen, H. H. 2002. "Food Insecurity and the Food Stamp Program." *American Journal of Agricultural Economics* 84 (5): 1215–1228.

Liu, M., P. Kasteridis, and S.T. Yen. 2013. "Breakfast, Lunch, and Dinner Expenditures Away From Home in the United States." *Food Policy* 38:156-164.

Ogden CL, Carroll MD, Kit BK, Flegal KM. 2014. "Prevalence of Childhood and Adult Obesity in the United States, 2011-2012." *Journal of the American Medical Association,* 311(8): 806-814.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference,* New York: Cambridge University Press.

Spirtes, P., C.N. Glymour, and R. Scheines. 2000. *Causation, Prediction, and Search,* Cambridge, MA: MIT Press.

Spirtes, P., Meek, C. and Richardson, T. 1999. "An Algorithm for Causal Inference in the Presence of Latent Variables and Selection Bias." In *Computation, Causation, and Discovery* 211–252. AAAI Press, Menlo Park, CA.

Taylor, R., and S.B. Villas-Boas. 2016. "A Discrete Choice Analysis of the National Household Food Acquisition and Purchase Survey (FoodAps)." *Amer. J. Agr. Econ.* 98(2): 513-532.

Wilde, P., and C.K. Ranney. 2000. "The Monthly Food Stamp Cycle: Shopping Frequency and Food Intake Decisions in an Endogenous Switching Regression Framework." *Amer. J. Agr. Econ.* 82: 200-213.