Food Insecurity, Poverty, Unemployment and Obesity in the United States: Effect of (Not) Considering Back-Door Paths in Policy Modeling

Senarath Dharmasena
Department of Agricultural Economics
Texas A&M University
College Station, TX 77843-2124
sdharmasena@tamu.edu

David A. Bessler
Department of Agricultural Economics
Texas A&M University
College Station, TX 77843-2124
d-bessler@tamu.edu

Selected Paper prepared for presentation at the Southern Agricultural Economics Association's 2018 Annual Meeting, Jacksonville, Florida; February 2-6, 2018

Food Insecurity, Poverty, Unemployment and Obesity in the United States: Effect of (Not) Considering Back-Door Paths in Policy Modeling

Abstract

The causes and consequences of food environment factors such as food insecurity, poverty, unemployment and obesity in the United States are complex. Once causality patterns with regards to these variables are identified, it is important to recognize *front-door* (Pearl, 2000) and *back-door paths* (Pearl, 2000) associated with these variables to make sensible and credible policy decisions. These policy interventions are known as performing *do*-Calculus (Pearl 2000, Spirtes et al., 2000) in causality literature. In this study we use the complex interactions of four food environment variables in the United States (food insecurity, poverty, unemployment and obesity) estimated using artificial intelligence and directed acyclic graphs by Dharmasena, Bessler and Capps (2016) and perform several policy interventions, recognizing *front-door* and *back-door* paths. Such policy simulations are vital for agencies not only to design appropriate policies for food assistance, poverty alleviation, combating food insecurity and obesity, but also to recognize effects of policy prior to the desired intervention. Preliminary analysis shows that there are two *front-door* paths from income to food insecurity, via poverty and via unemployment. Also, there is a *front-door* path from poverty to food insecurity, while there is an important *back- door* path from poverty to food insecurity via unemployment.

Key Words: Front-door paths, back-door paths, do-Calculus, machine learning, directed acyclic graphs

JEL Classification: C45, C52, C82, D85,I38

Food Insecurity, Poverty, Unemployment and Obesity in the United States: Effect of (Not) Considering Back-Door Paths in Policy Modeling

Background Information

The causes and consequences of food environment factors such as food insecurity, poverty, unemployment and obesity in the United States are complex. Research is emerging with regards to our understanding of complex interactions of aforementioned factors of food environment (for example Dharmasena, Bessler and Capps, 2016 and Senia, Dharmasena and Todd, 2017). Once causality patterns with regards to these variables are identified, it is important to recognize *front-door* (Pearl, 2000, page 82) and *back-door paths* (Pearl, 2000, page 79) associated with these variables to make sensible and credible policy decisions. These policy interventions are known as performing *do*-Calculus (Pearl 2000, Spirtes et al., 2000) in causality literature (explained below). Let X, Z and Y represent three observed variables, assuming no unobserved variables within the causality structure (causal chain) as follows, X→Z→Y; X causes Y via Z. This is a *front-door* path from X to Y. If one wants to find the effect of X on Y, simply Y can be regressed on X. Since conditioning on Z makes Y and X orthogonal (Dharmasena, Bessler and Capps, 2016), one should not include Z among the right-hand side variables in this regression. Now assume a common cause, R, causing both X and Y on top of the relation X→Z→Y. This common cause R creates a *back-door* path from X to Y. If one wants to find the effect of X on Y, now, not only has one to condition on X, but also on R (put R in the right-hand side of the regression). Ignoring this *back-door* path or not being able to observe this *back-door* path variable introduces bias to the parameter estimate associated with X, if one regress Y on X (also explained below). In the absence of *back-door* paths, if one wants to find out the effect of policy on X (meaning set X to some *x*; or do(X=*x*); also known as performing *do*-Calculus; Pearl (2000)), the regression Y on X and the associated the parameter estimate with respect to X (or the marginal effect of X on Y) will be used to simulate the effect of this policy intervention. However, if the aforementioned *back-door* path from X to Y via R exists, now ignoring R while performing the do(X=*x*) operation would introduce bias to the parameter estimate associated with X (marginal effect of X on Y), hence would result in an error in the policy simulation of the intervention.

In this study we use the complex interactions of four food environment variables in the United States (food insecurity, poverty, unemployment and obesity) estimated using artificial intelligence and directed acyclic graphs by Dharmasena, Bessler and Capps (2016) and perform several policy interventions, recognizing *front-door* and *back-door* paths associated with these policy variables. Such policy simulations are vital for agencies not only to design appropriate policies for food assistance, poverty alleviation, combating food insecurity and obesity in the United States, but also to recognize effects of policy interventions on those variables on effect variables prior to the desired intervention.

Objectives

The broad objective of this study is to perform policy interventions on several food environment variables in the United States and measure their effects while considering *front-door* and *back-door* paths associated with these variables. More specifically, using Figure 2 of Dharmasena, Bessler and Capps (2016) to (i) estimate effect of income on food insecurity; poverty on food insecurity; unemployment on food insecurity; income on obesity; income on participation in Supplemental Nutrition Assistance Program (SNAP); unemployment on SNAP; all just considering the *front-door* paths; and (ii) estimate all relationships in (i) considering all *back-door* paths associated with these variables; and (iii) to conduct policy interventions (perform *do*-Calculus) with respect to reduced poverty, increased income, reduced unemployment, on food insecurity, obesity and SNAP participation for both types of models, considering only *front-door* paths, and considering both *front*- and *back-door* paths.

Data and Methodology

Figure 2 of Dharmasena, Bessler and Capps (2016) shows complex causality structures associated with four food environment variables in the United States, i.e. food insecurity, poverty, unemployment and obesity. This directed acyclic graph and associated data underlying this graph (as explained in Dharmasena, Bessler Capps, (2016), Table 1 and 2; sixteen food environment variables in the United States from 2008-2010 extracted from various government sources) are used in this study. First, using ordinary least squares, simple linear regression models are estimated to satisfy relationships explained in specific objective (i), just considering *front-door* paths. Second, several regression models are developed to satisfy relationships explained in specific objective (ii), considering not only *front-door*, but also *back-door* paths associated with these variables. Once regression models are developed, the marginal effects of the dependent variable in each case with respect to conditioning variables, along with their statistical significance, are compared across models. These comparisons are offered, considering both *front-door* paths, as well as *front*- and *back-door* paths. Next, several policy interventions are performed, such as reducing poverty, increasing income, and reducing unemployment. Finally marginal effects of dependent variables (food insecurity, obesity and SNAP participation) and their statistical significance resulting from such policy interventions are compared across both types of models, both *front-door* path as well as *front*- and *back-door* path models.

Results and Implications

Preliminary analysis shows that there are two *front-door* paths from income to food insecurity. They are via poverty and via unemployment. That is income→poverty→food insecurity; income→unemployment→food insecurity. Also, there is a *front-door* path from poverty to food insecurity, while there is an important *back- door* path from poverty to food insecurity via common cause variable, unemployment (unemployment is a common cause for both food insecurity and poverty; poverty←unemployment→food insecurity). Ignoring the effect of unemployment on food insecurity, while regressing food insecurity on poverty, produces an overestimated marginal effect of poverty on food insecurity. As shown in Table 1, the OLS estimate of regression of poverty on food insecurity resulted in parameter estimate of 0.636. However, as shown in Table 2, the OLS estimate of poverty on food insecurity when both poverty and unemployment are regressed on food insecurity is estimated to be 0.561. This shows that not including the unemployment among the substantive information (right hand side variables) or ignoring the back-door path from unemployment to food insecurity, overestimated the effect of poverty on food insecurity. If this marginal effect is used to summarize the effect of a poverty reduction policy intervention on food insecurity in the United States, one would overestimate the effect of this poverty reduction policy on food insecurity, thereby potentially implementing a suboptimal policy. Similar analysis is performed to the rest of the *front-door* and *back-door* paths with regards to complex interactions among variables and policy implications are explored.

Table 1: Regression Estimates: Poverty on Food Insecurity

OLS Regression Statistics for Food Insecurity Rate			
F-test	56.046	**Prob(F)**	0.000
R^2	0.549		
$RBar^2$	0.539		
Akaike Information Criterion	1.065		
Schwarz Information Criterion	1.104		

	Intercept	Poverty Rate
Beta	**5.024**	**0.636**
S.E.	1.206	0.085
t-test	4.166	7.486
Prob(t)	0.000	0.000

Regression Analysis produced using Simetar statistical package, 2009

Table 2: Regression Estimates: Poverty and Unemployment on Food Insecurity

OLS Regression Statistics for Food Insecurity Rate			
F-test	34.167	**Prob(F)**	0.000
R^2	0.603		
$RBar^2$	0.585		
Akaike Information Criterion	0.980		
Schwarz Information Criterion	1.058		

	Intercept	Poverty Rate	Unemployment Rate
Beta	**3.400**	**0.561**	**0.354**
S.E.	1.320	0.086	0.144
t-test	2.576	6.510	2.467
Prob(t)	0.013	0.000	0.017

Regression Analysis produced using Simetar statistical package, 2009

References

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

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

Senia, M., S. Dharmasena and J. Todd, 2017, “A Complex Model of Consumer Food Acquisitions: Applying Machine Learning and Directed Acyclic Graphs to the National Household Food Acquisition and Purchase Survey (FoodAPS)” Working Paper, Texas A&M University

Spirtes, P., Glymour, C.N., Scheines, R., 2000. Causation, Prediction, and Search. The MIT Press.