Investigating Causal Effects of SNAP and WIC on Food Insecurity Using FoodAPS

Helen Jensen, Brent Kreider, and Oleksandr Zhylyevskyy (Iowa State University)

1. Research Question

In 2014, 37.4% of households with income below 130% of the poverty line experienced low or very low food security (Coleman-Jensen et al., 2015). Substantial prevalence of food insecurity among the low-income population is a matter of intense public concern since food insecurity can be detrimental to health and well-being (for a literature review, see Gundersen et al., 2011).

Many food assistance recipients participate in multiple food programs, but little is known about how such programs interact in creating a food safety net. Gaining this knowledge can help improve the design of food programs and make them more effective. Our specific interest is in assessing the degree to which a household's participation in *multiple* food assistance programs augments the benefits of participating in a single program in terms of reducing food insecurity.

We leverage unique features of FoodAPS to focus on two specific programs—the Supplemental Nutrition Assistance Program (SNAP) and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)—and address the following research questions:

(1) To what extent does participation in both SNAP and WIC lead to a reduction in household food insecurity compared with participation in only SNAP or WIC?

(2) To what extent does combining self-reported survey data on program participation with auxiliary administrative data in FoodAPS tighten inferences on causal treatment effects?

Since the 1930s, the United States has had several food programs in place designed to alleviate food insecurity. While target populations and forms of benefits differ, most programs aim to provide a nutritionally adequate diet and resources to increase the food supply for low-income households. Yet, 53.7% of households on SNAP and 41.1% on WIC were food insecure in 2014. Moreover, the rate of food insecurity among SNAP recipients was more than twice that among low-income nonrecipients. Improving methods to study the impacts of participation in multiple programs will help researchers and policymakers better understand the overall efficacy of the U.S. food safety net.

Identifying *causal*, rather than associative, effects of program participation presents a methodological challenge because of endogenous self-selection of households into the programs (households are not randomly assigned) and underreporting of food assistance in national surveys. For example, unobserved characteristics such as expected future health status, human capital, and financial stability can be related both to participation in food programs and to food security outcomes. This simultaneity precludes the use of standard regression techniques to estimate causal effects (e.g., see Gundersen and Oliveira, 2001). Moreover, households are known to systematically underreport the receipt of food assistance in national surveys, and the propensity to misreport varies with unobserved characteristics. Meyer et al. (2009) find that only about two-thirds of benefits are recorded for SNAP. Bitler et al. (2003) find evidence of a substantial underreporting for WIC.

To estimate causal effects under these conditions, we apply nonparametric treatment effect methods of Kreider, Pepper, Gundersen, and Jolliffe (2012, hereafter KPGJ) that account for endogenous selection and misclassification in a unifying framework. We extend their potential outcomes framework, which is designed to handle a binary treatment variable, to accommodate the case of *partially ordered multiple treatments*. Also, we exploit a unique feature of FoodAPS in that it

provides both self-reported and administratively verified SNAP participation data. We incorporate this auxiliary administrative information to both tighten inferences on the causal effects of program participation and assess the broader impacts of misreporting when such information is not available.

To further tighten inferences, we will exploit information in the geographical component of FoodAPS to construct monotone instrumental variables (MIVs) (Manski and Pepper, 2000). Unlike standard instrumental variables (IVs), less restrictive MIVs require no *a priori* exclusion restrictions; instead, we only require that the latent food insecurity outcomes vary (weakly) monotonically with the MIV, on average, conditional on the treatment. Additionally, to compare our results with those obtainable using conventional IV techniques, we will extract state-level IVs for program participation from the USDA ERS's SNAP Policy Database (see Ratcliffe et al., 2011; Gregory and Deb, 2015).

Deliverables for this project include a manuscript describing the methodological approach, data, estimation results, and implications for public policy, as well as presentations at the pre-conference and final research conference. We also intend to disseminate our findings at other universities and national conferences.

2. Methodology

To estimate causal treatment effects of multiple program participation, we apply nonparametric bounding methods developed in KPGJ that account for the dual identification problems of endogenous selection and classification error. We extend their framework to introduce partially ordered multiple treatments, which is necessary to model multiple program participation. We will derive sharp bounds on average treatment effects that are logically consistent with the observed data and imposed statistical and behavioral assumptions. As described below, we will exploit the unique feature of FoodAPS that it contains both self-reported and administratively verified data on SNAP participation. This verification information is valuable in part because it will allow us to reduce the dimensionality of the classification error problem.

For the outcome, let $Y = 1$ indicate that a household is food secure, with $Y = 0$ otherwise. We will also consider related outcomes such as very low food security. Let S^* be an unobserved indicator of true program participation where $S^* = 0$ denotes no participation in SNAP or WIC, $S^* = 1$ denotes participation in SNAP alone, $S^* = 2$ denotes participation in WIC alone, and $S^* = 3$ denotes participation in both SNAP and WIC. This treatment variable is partially ordered: $S^* = 1$ or 2 denotes some participation while $S^* = 0$ does not, and $S^* = 3$ involves more participation. (Since $S^* = 1$ and $S^* = 2$ represent different programs, these two treatments are not ordered.)

Instead of observing S^* in the data, we observe a self-reported counterpart, *S*. We also observe FoodAPS administrative information on SNAP. Let $V_{SNAP} = 1$ denote verification that a household truly received SNAP (using the FoodAPS variable "*snapnowadmin*"), implying that $S^* = 1$ or 3, with $V_{SNAP} = 0$ conversely implying that $S^* = 0$ or 2.

We focus on average treatment effects (ATEs) associated with participating in both food assistance programs versus a single program, or compared with no participation:

$$ATE_{jk} = P[Y(S^* = j) = 1 \mid X] - P[Y(S^* = k) = 1 \mid X] \text{ for } j,k \in \{0,1,2,3\},\ j \neq k, \qquad (1)$$

where $Y(S^*)$ indicates the (latent) potential food security outcome under treatment S^*, X denotes

any covariates of interest, and *P* denotes the probability of an outcome.[1,2] To simplify notation, we suppress the conditioning on *X* and write $P[Y(S^* = j) = 1]$ more compactly as $P[Y(j)]$.

In what follows, we illustrate the case of $j = 3$ vs. $k = 1$. Here, $ATE_{31} = P[Y(3)] - P[Y(1)]$ measures how the prevalence of food security would change if all eligible households participated in both SNAP and WIC rather than in SNAP alone.[3] One cannot identify ATE_{31} without additional assumptions, even if *S* is accurately reported, because the potential outcome $Y(S^* = 3)$ is observed only for households that chose to participate in both SNAP and WIC, while $Y(S^* = 1)$ is observed only for households that chose to participate in SNAP alone. The decomposition $P[Y(3)] = P[Y(3) \mid S^* = 3]P(S^* = 3) + P[Y(3) \mid S^* \neq 3]P(S^* \neq 3)$ highlights the selection problem: the term $P[Y(3) \mid S^* \neq 3]$ represents an unobserved counterfactual outcome.

As a further identification problem, households are thought to systematically underreport program participation in national surveys, and such misreporting may be related to personal characteristics (including the food security outcome). Allowing *S* to deviate from S^*, let $\theta_i^{j,k} \equiv P(Y = i, S = j, S^* = k)$ for $j,k = \{0,1,2,3\}$ denote the fraction of households with food security status *i* reporting participation status *j* when true participation status is *k*. Using the law of total probability, the first term in ATE_{31} becomes

$$P[Y(3)] = P(Y = 1, S = 3) + \theta_1^{-3,3} - \theta_1^{3,-3} + P[Y(3) \mid S^* \neq 3]\left\{ P(S \neq 3) + \sum_{j \neq 3} (\theta_1^{-j,j} + \theta_0^{-j,j} - \theta_1^{j,-j} - \theta_0^{j,-j}) \right\}$$

where $\theta_i^{-j,k} \equiv P(Y = i, S \neq j, S^* = k)$ and $\theta_i^{j,-k} \equiv P(Y = i, S = j, S^* \neq k)$. An analogous expression can be derived for $P[Y(1)]$.

Without further assumptions, we can show that the marginal impact on food security associated with participating in both SNAP and WIC, compared with participating in SNAP alone, is bounded as follows:

$$-P(Y = 1, S \neq 3) - P(Y = 0, S \neq 1) + \theta_{31} \leq ATE_{31} \leq P(Y = 1, S \neq 1) - P(Y = 0, S \neq 3) + \theta_{31} \qquad (2)$$

where $\theta_{31} \equiv \theta_1^{-3,3} + \theta_0^{-1,1} - \theta_1^{3,-3} - \theta_0^{1,-1}$. Terms like $P(Y = 1, S \neq 3)$ are observed from the data, while the unobserved components $\{\theta\}$ are logically bounded. For example, we can show that $0 \leq \theta_0^{-1,1} \leq \min\{P(Y = 0, S \neq 1, V_{SNAP} = 1), P(S^* = 1)\}$ which exploits both the self-reported and administrative data in FoodAPS. Specifically, the first term constrains $\theta_0^{-1,1}$ to be no larger than the fraction of households that (i) are classified as being food insecure, (ii) report not solely participating in SNAP (i.e., report no participation, participation in WIC alone, or participation in both SNAP and WIC), and (iii) are documented by the FoodAPS administrative variable "*snapnowadmin*" to be participating at least in SNAP.

To narrow these bounds, we will consider the identifying power of relatively weak monotonicity

[1] There are no regression orthogonality conditions to be satisfied in this framework. Thus, there is no need to include covariates as a means of avoiding omitted variable bias.

[2] Our framework can be extended to handle continuous outcomes or number of positive responses to the food insecurity questions in the survey.

[3] Note that we are not restricting a treatment effect to be the same across different households. As emphasized by Moffitt (2005), the classical linear response model assumption, for example, is difficult to justify when considering government assistance programs that are thought to have heterogeneous effects.

restrictions such as Monotone Treatment Selection (Manski and Pepper, 2000; KPGJ) and Monotone Treatment Response (Manski, 1997; KPGJ). The MTS assumption formalizes the notion that unobserved factors associated with food insecurity are likely to be positively associated with the decision to take up food assistance programs. Under the MTR assumption, participating in SNAP and WIC would not harm food security, at least on average, conditional on treatment selection.

We will further narrow the bounds using MIVs. Monotone instruments are often easier to motivate than standard IVs because they do not require any orthogonality/exclusion restrictions. In our application, we merely require that the instrument leads to weakly improved latent food security outcomes, on average, conditional on the treatment. As MIVs, we plan to use variables reflective of important aspects of local food environment, as recorded in the geographical component of FoodAPS. In particular, we will use the number of SNAP-authorized retailers within a mile (or other distance) of the household and will investigate using the ratio of actual food expenditures to expenditures consistent with the Thrifty Food Plan's (TFP) recommendations and local food prices. We will also use a variety of other MIVs, including ones based on household income and, similar to a regression discontinuity design, based on program eligibility criteria.[4]

By layering successively stronger combinations of assumptions, we will provide successively tighter bounds on the causal impacts of participating in multiple food assistance programs. We will also compare our ATE results with those obtained using conventional parametric IV techniques.

3. Data

The National Household Food Acquisition and Purchase Survey (FoodAPS) is the primary dataset for this analysis. The survey provides data that were obtained from a nationally representative, stratified sample of 4,826 households during the period between April 2012 and January 2013 and includes household demographics, information on household food purchases and acquisitions during the survey week, and local food environmental information. The survey also collects information on income, receipt of SNAP benefits (current, last 12 months, date of last SNAP benefit receipt), confirmation of SNAP receipt through administrative match, and reported WIC receipt along with information on demographics and categorical eligibility to determine WIC eligibility. Data are also available on the amount and source of payment for expenditures for foods consumed at home.

In addition, households in FoodAPS responded to 10 food insecurity-related questions, which are used to create the raw score for the household's Food Security Scale. Based on USDA's 30-day Adult Food Security Scale, households are assigned to categories: "food secure" (raw score 0-2), "low food security" (raw score 3-5), and "very low food security" (raw score 6-10). Food insecure households are those with the raw score of 3 or more. We plan to employ rich SNAP participation/ receipt data collected by FoodAPS to study how varying the definition of "program participation" can affect the robustness of our inferences on program effects.

We will also request access to the geographic component of FoodAPS to include data collected on the location of different types of retailers, measures of access, and food prices. Available reported data from FoodAPS indicate that it includes 461 WIC households and 1,581 SNAP households (Ver Ploeg et al., 2015). About 48% of WIC households also participated in SNAP (IOM, 2015). Also, we will use data from the USDA ERS's SNAP Policy Database to construct IVs for program participation and merge them with the FoodAPS records using location (i.e., state of household residence) and time identifiers.

[4] Gundersen et al. (2012) show how health outcomes among income-ineligible households can be used to bound counterfactual health outcomes among eligible households.

4. Budget, Time Schedule, and Conflict of Interest

The overall budget for the project is $40,000 which includes $25,000 salary support for the three project investigators and $15,000 for research assistant/post doc support.

The project will begin on notification of selection (spring 2016). The project team will participate in a conference call in May 2016, present initial results at a pre-conference in early 2017, and provide a completed manuscript on the project results and executive summary for presentation at a final research conference in late 2017.

None of the three investigators (H. Jensen, B. Kreider, and O. Zhylyevskyy) has any potential conflicts of interest (financial or other) that might arise from this project.

5. References

Bitler, M., Currie, J., Scholz, J. 2003. WIC Eligibility and Participation. *Journal of Human Resources* 38 (5), 1139–1179.

Coleman-Jensen, A., Rabbitt, M.P., Gregory, C., Singh, A. 2015. Statistical Supplement to Household Food Security in the United States in 2014, AP-069, USDA, ERS, September.

Gregory, C.A., Deb P. 2015. Does SNAP Improve Your Health? *Food Policy* 50, 11–19.

Gundersen, C., Kreider, B., Pepper, J. 2011. The Economics of Food Insecurity in the United States. *Applied Economic Perspectives and Policy* 33 (3), 281–303.

Gundersen, C., Kreider, B., Pepper, J. 2012. The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis. *Journal of Econometrics* 166 (1), 79–91.

Gundersen, C., Oliveira, V. 2001. The Food Stamp Program and Food Insufficiency. *American Journal of Agricultural Economics* 84 (3), 875–887.

Institute of Medicine (IOM). 2015. Review of WIC Food Packages: Proposed Framework for Revisions: Interim Report. K.M. Rasmussen, M.E. Latulippe, A.L. Yaktine (eds.).

Kreider, B., Pepper, J., Gundersen, C., Jolliffe, D. 2012. Identifying the effects of SNAP (Food Stamps) on Children's Health Outcomes When Participation is Endogenous and Misreported. *Journal of the American Statistical Association* 107 (499), 958–975.

Manski, C.F. 1997. Monotone Treatment Response. *Econometrica* 65 (6), 1311–1334.

Manski, C.F., Pepper, J.V. 2000. Monotone Instrumental Variables: With an Application to the Returns to Schooling. *Econometrica* 68 (4), 997–1010.

Meyer, B.D., Mok, W.K.C., Sullivan, J.X. 2009. *The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences*. NBER Working Paper 15181.

Moffitt, R. 2005. Remarks on the Analysis of Causal Relationships in Population Research. *Demography* 42 (1), 91–108.

Ratcliffe, C., McKernan, S.-M., Zhang, S. 2011. How Much Does the Supplemental Nutrition Assistance Program Reduce Food Insecurity? *American Journal of Agricultural Economics* 93 (4), 1082–1098.

Ver Ploeg, M., Mancino, L., Todd, J.E., Clay, D.M., Scharadin, B. 2015. Where Do Americans Usually Shop for Food and How Do They Travel To Get There? Initial Findings from the National Household Food Acquisition and Purchase Survey, EIB-138, USDA, ERS, March.