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

Helen Jensen, Brent Kreider, and Alex Zhylyevskyy
Iowa State University

NBER Pre-Conference

February 3, 2017

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 SNAP or WIC alone?

(2) How much can we learn by combining self-reported survey data on program participation with auxiliary administrative data in FoodAPS?

- Objective: Derive sharp bounds on average treatment effects (ATE) of multiple program participation
  - Must be logically consistent with the observed data and any imposed statistical and behavioral assumptions

FoodAPS

- Sample of 4,826 households that participated during one week between April 2012 and January 2013:
  - SNAP participants, low-income nonparticipants, higher income nonparticipants
- We focus on impacts of SNAP and WIC on food security
- Two data features of especially high value for our research:
  - FoodAPS contains ***administratively verified*** info on SNAP participation
  - FoodAPS-GC provides local food environment data: can construct Monotone Instrumental Variables (MIVs) related to food expenditures and food retailer availability

Identifying causal effects of program participation presents methodological challenges:

- Endogenous self-selection of households into the programs (not randomly assigned)
  - unobserved characteristics such as expected health, human capital, and financial stability may be related to both participation and food security outcomes
- Systematic underreporting of food assistance
  - propensity to misreport varies with unobserved characteristics

To identify causal impacts, extend recently developed nonparametric treatment effect methods:

- Account for endogenous selection and misclassification in a unifying potential outcomes framework
  - extend binary treatment methods to accommodate the case of a *partially ordered treatment*
- Exploit a unique feature of FoodAPS: provides both self-reported and administratively verified SNAP participation data
  - auxiliary information tightens inferences

To further tighten inferences, exploit information in the geographical component of FoodAPS to construct monotone instrumental variables (MIVs)

- Unlike standard instrumental variables (IVs), MIVs require no a priori exclusion restrictions
  - only require that the latent food insecurity outcomes vary monotonically with the MIV
- Will compare results with those obtained using conventional IV techniques
  - extract state-level IVs for program participation from the USDA ERS's SNAP Policy Database

Motivation for Our Methodology

Compare with a simple parametric approach:

$$\underbrace{Y_i}_{\text{Outcome}} = \gamma \cdot \underbrace{S_i}_{\text{Treatment}} + \underbrace{x_i}_{\text{Covariates}} \cdot \beta + \underbrace{\varepsilon_i}_{\text{Error term}}$$

- Treatment S_i is **binary**. Say, $S_i = 1$ if i is on SNAP, 0 if not
- If same unobservables affect S_i and Y_i, then $cov(S_i, \varepsilon_i) \neq 0$ and OLS is inconsistent due to **endogeneity**
- Measurement error in S_i is **nonclassical**. Thus, standard IV estimation is inconsistent as well
- Our **nonparametric bounding** methodology handles endogeneity, misreporting, and multiple treatments (not just binary S_i). Also, allows for heterogeneous response to treatment across i

Basics of Our Approach: Notation

$\boldsymbol{S^*}$: **true** program participation status is **partially ordered**

$S^* = 0$: neither SNAP nor WIC

$S^* = 1$: SNAP alone

$S^* = 2$: WIC alone

$S^* = 3$: both SNAP and WIC

$\boldsymbol{S}$: **reported** program participation; S need not equal S^*

Potential outcomes framework:

$\boldsymbol{Y(S^*)}$: potential outcome under treatment S^*

$Y = 1$ if household is food secure, $Y = 0$ otherwise

$\boldsymbol{X}$: covariates (some used as instruments)

SNAP Verification Status

- A fraction of households in FoodAPS was matched to administrative records. In such cases, we can verify whether a household received SNAP benefits in past month

Verification Status	Sample Fraction (Weighted)
Matched households:	
Confirmed participation	57.6%
Confirmed nonparticipation	2.6%
Unmatched households:	
Not matched to administrative data	37.5%
Withheld consent to be matched	2.3%

Reported Program Participation

- Weighted sample distribution by reported participation when SNAP participation indicator incorporates administrative data [modified using SNAPNOWHH]:

		WIC	
		No	*Yes*
SNAP	*No*	15.3% [13.0%]	16.6% [13.6%]
	Yes	31.4% [33.6%]	36.7% [39.7%]

- Our sample (*N* = 460) includes FoodAPS households with:
  - income below 130% poverty, and
  - a pregnant woman, or a child aged < 5 years

Food Security by Participation

- Weighted prevalence of food security status by food program participation [modified using SNAPNOWHH]:

Proportion food secure

		WIC	
		No	*Yes*
SNAP	*No*	53.2% [55.1%]	54.5% [50.5%]
	Yes	52.2% [51.6%]	58.5% [59.5%]

- Food security measure is based on USDA's 30-day Adult Food Security Scale

Food Security by Participation

- Weighted prevalence of food security status by food program participation [modified using SNAPNOWHH]:

Proportion not "very low food secure"

		WIC	
		No	*Yes*
SNAP	*No*	74.6% [73.4%]	83.7% [81.3%]
	Yes	78.7% [78.9%]	90.4% [90.7%]

- Food security measure is based on USDA's 30-day Adult Food Security Scale

Basics of Our Approach: ATE

- We focus on **average treatment effects** (**ATEs**):

$$ATE_{jk} = P[Y(S^* = j) = 1 | X] - P[Y(S^* = k) = 1 | X] \text{ for } j \neq k$$

- For example, consider ATE_{31}:

$$ATE_{31} = P[Y(S^* = 3) = 1 | X] - P[Y(S^* = 1) = 1 | X]$$

- ATE_{31} measures by how much likelihood of food security would change if household were to participate in both SNAP and WIC vs. in SNAP alone

- There are no regression orthogonality conditions to satisfy

- Covariates are only used to specify subpopulations

Decomposition Strategy

- ATE cannot be point-identified without assumptions even if $S \equiv S^*$
- We decompose formulas into what is identified and what is not
- Simplify notation: $ATE_{31} = P[Y(3)=1] - P[Y(1)=1]$
- Consider decomposition:

$$P[Y(3)=1] = \underbrace{P[Y(3)=1 \mid S^* = 3]}_{\textbf{identified}}\underbrace{P(S^* = 3)}_{\textbf{identified}} + \underbrace{\overset{\in [0,1]}{P[Y(3)=1 \mid S^* \neq 3]}}_{\textbf{not identified}}\underbrace{P(S^* \neq 3)}_{\textbf{identified}}$$

- Data cannot identify $P[Y(3)=1 \mid S^* \neq 3]$ because it refers to an unobserved **counterfactual**.
- However, extending methods of Manski (1995), we derive worst-case bounds for $P[Y(3)=1]$, $P[Y(1)=1]$, and ATE_{31}

Addressing Misreporting

- When S may deviate from S^*, define: $\theta_i^{j,k} \equiv P(Y=i, S=j, S^*=k)$
- $P[Y(3)=1]$ becomes:

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

- ATE_{31} can be bounded as:

$$-P(Y=0, S \neq 1) - P(Y=1, S \neq 3) + \Theta_{3,1}^{LB}$$
$$\leq ATE_{3,1} \leq$$
$$P(Y=0, S \neq 3) + P(Y=1, S \neq 1) + \Theta_{3,1}^{UB}$$

unobserved

$$\Theta_{3,1}^{LB} \equiv \theta_1^{-3,3} - \theta_1^{3,-3} + \theta_0^{-1,1} - \theta_0^{1,-1}, \quad \Theta_{3,1}^{UB} \equiv -\theta_0^{-3,3} + \theta_0^{3,-3} - \theta_1^{-1,1} + \theta_1^{1,-1}$$

Tightening Bounds

- Without assumptions, bounds on ATEs are wide
- To **tighten** the bounds, we can impose restrictions on
  1) Misreporting process
  2) Selection process
- Consider **restricting the misreporting** process:
  - Exploit logical constraints on probabilities and auxiliary data to restrict $\{\theta\}$:

FoodAPS SNAP verification

Ex. $\theta_0^{-1,1} \leq \min\{P(Y=0, S \neq 1, V_{SNAP} \neq 0), P(S^* = 1)\}$

  - exploits both the self-reported and administrative data in FoodAPS

Tightening Bounds, cont.

To **restrict selection process**, we can employ:

- **Exogenous selection** assumption (often does not hold, though)
- Monotone treatment selection (**MTS**) (Manski & Pepper, 2000)
- Monotone treatment response (**MTR**) (Manski, 1995)
  - We extend MTS and MTR to partially ordered unobserved treatments
- Monotone instrumental variables (**MIVs**, Manski & Pepper, 2000)
- Instrumental variables (**IVs**). E.g., IVs for SNAP (Ratcliffe et al., 2011)

We can **combine assumptions** to further tighten bounds on ATEs

Example of Analytical Results

Proposition 2:

Under "**no-stigma verification**" with **endogenous selection**, bounds on $ATE_{3,1}$ are given as follows:

- **Lower bound**:

$$ATE_{3,1}^{LB} = -P(Y=1, S \neq 3) - P(Y=0, S \neq 1)$$
$$+ \max\{0, \Delta_3 - P_{000}\} + \max\{0, \Delta_1 - P_{100}\}$$

- **Upper bound**:

$$ATE_{3,1}^{UB} = P(Y=0, S \neq 3) + P(Y=1, S \neq 1)$$
$$- \max\{0, \Delta_3 - P_{100}\} - \max\{0, \Delta_1 - P_{000}\}$$

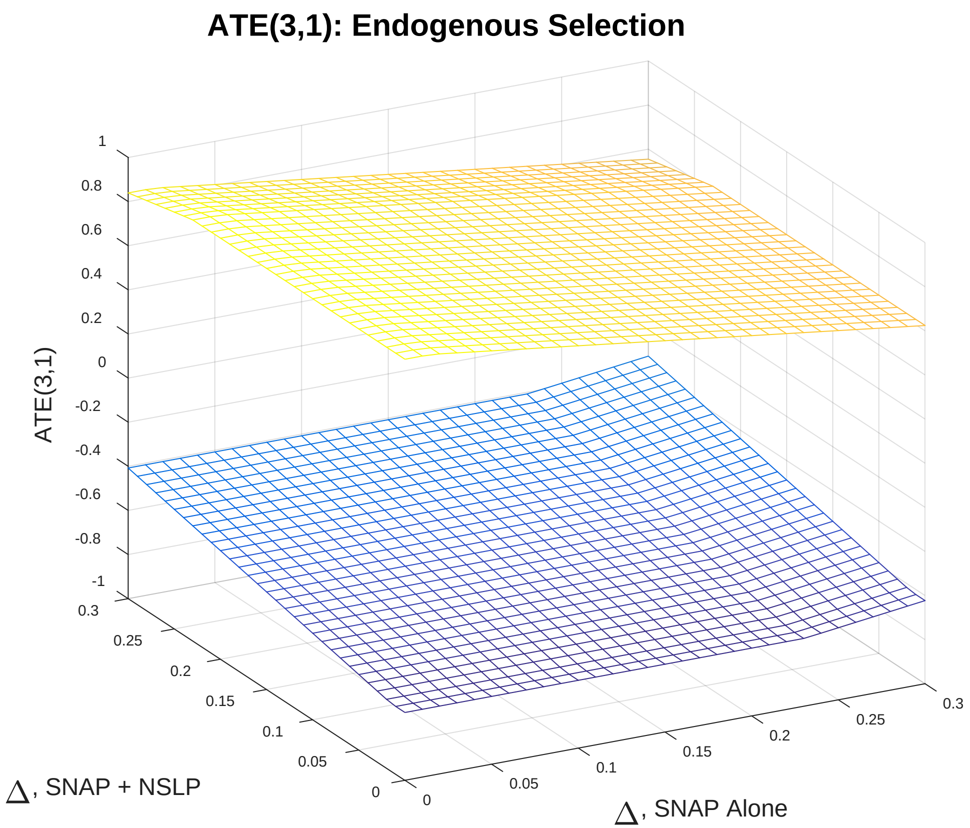
$\Delta_1 \equiv P_1^* - P_1,\ \Delta_3 \equiv P_3^* - P_3,$

FoodAPS verification

$P_{000} \equiv P(Y=0, S=0, V_{SNAP} \neq 1),\ P_{100} \equiv P(Y=1, S=0, V_{SNAP} \neq 1)$

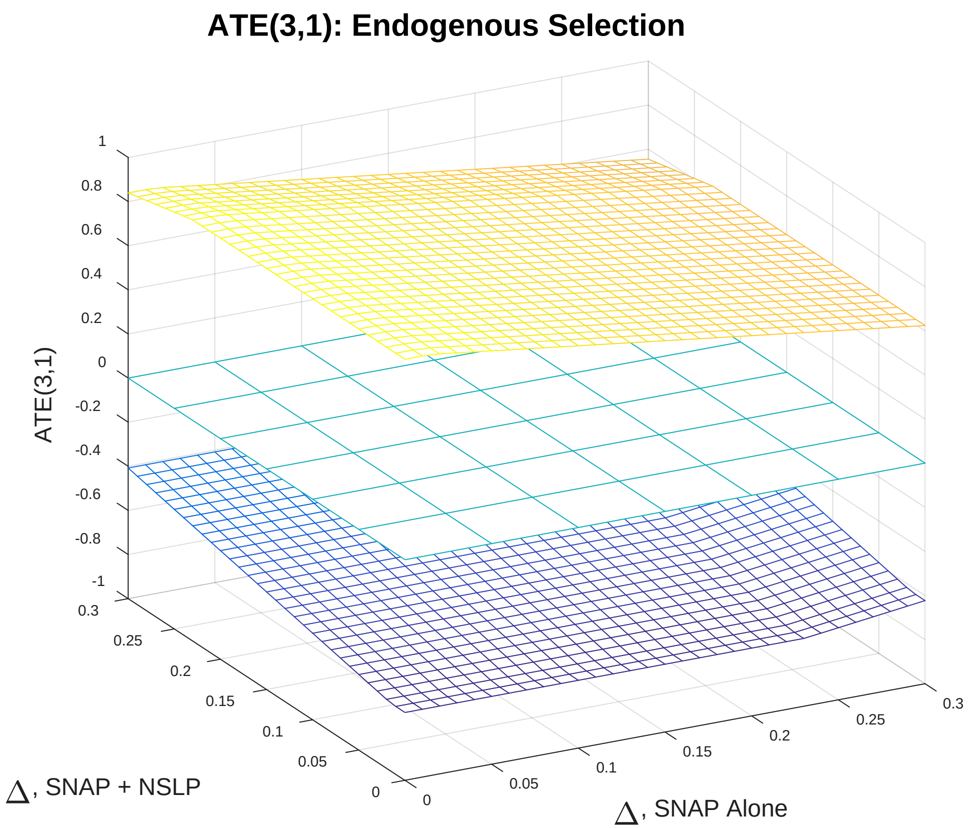
Illustrative Worst-Case Bounds (CPS, not FoodAPS)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



Illustrative Worst-Case Bounds (CPS, not FoodAPS)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



Exogenous Selection: Definition

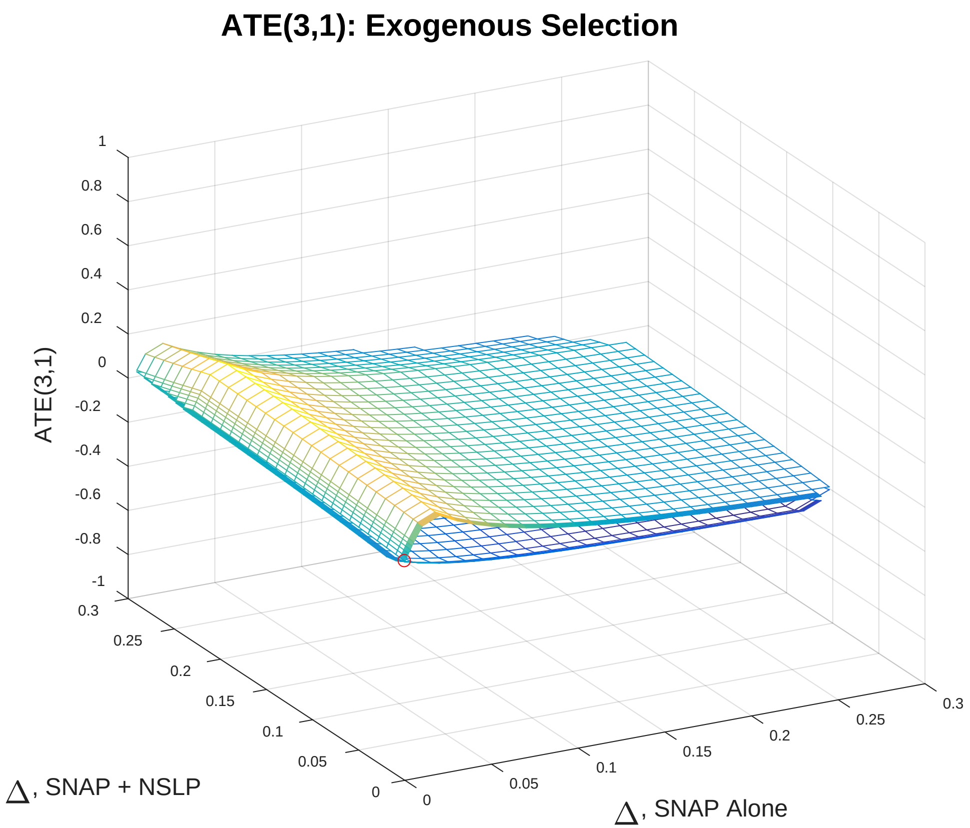
Exogenous selection:

$$P[Y(j)=1] = P[Y(j)=1 \mid S^{*}=k] \quad \forall j,k$$

- Holds if potential outcomes do not depend on realized treatment
- Assumption applies when assignment to programs is truly **random**

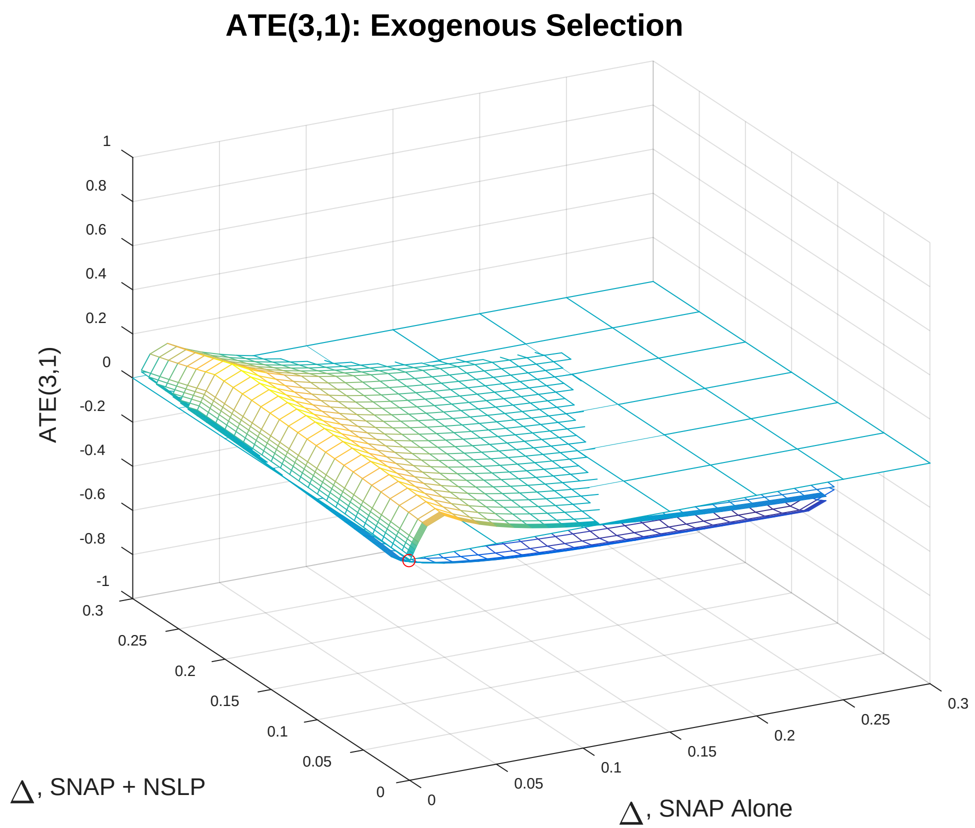
Exogenous Selection (CPS data)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



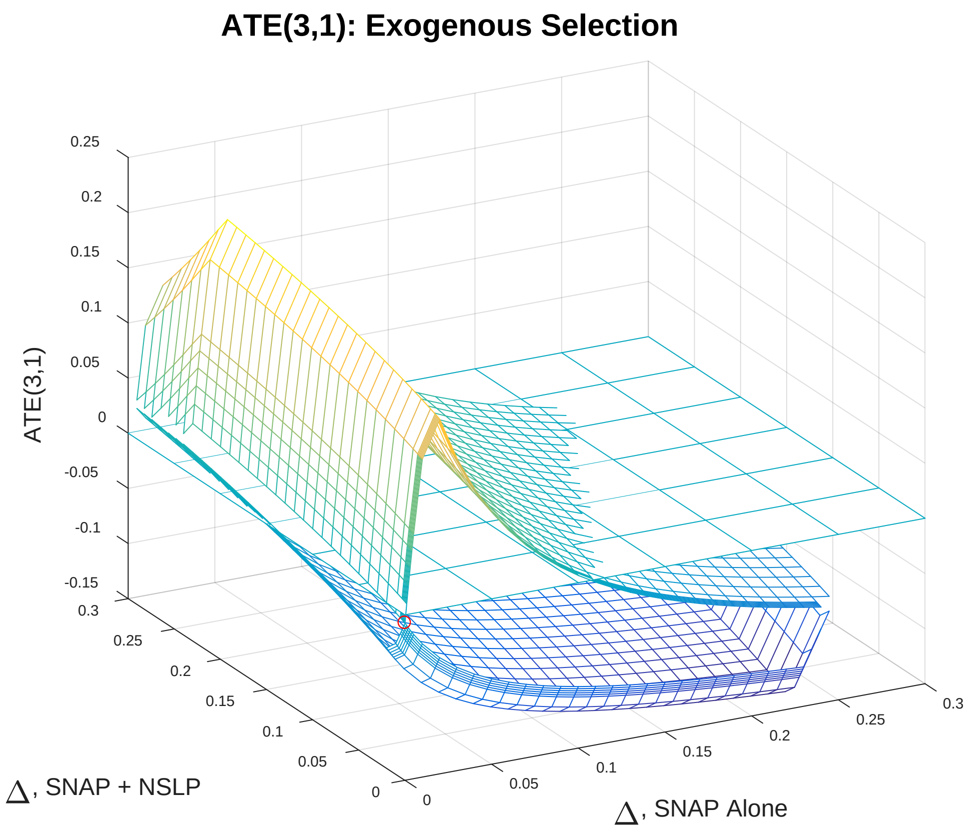
Exogenous Selection (CPS data)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



Exogenous Selection (CPS data): Closer View

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



Exogenous Selection: Identification Decay (CPS)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:

		$\Delta_1 = 0$			$\Delta_1 = 0.01$			$\Delta_1 = 0.10$		
		LB	UB	width	LB	UB	width	LB	UB	width
$\Delta_3 = 0$	p.e.	[-0.007,	-0.007]	0.000	[-0.029,	0.14]	0.167	[-0.094,	0.007]	0.101
	CI	[-0.040,	0.026]		[-0.051,	0.16]		[-0.106,	0.022]	
$\Delta_3 = 0.01$	p.e.	[-0.031,	-0.004]	0.028	[-0.053,	0.14]	0.195	[-0.118,	0.010]	0.129
	CI	[-0.057,	0.022]		[-0.075,	0.17]		[-0.130,	0.025]	
$\Delta_3 = 0.10$	p.e.	[-0.010,	0.023]	0.034	[-0.032,	0.17]	0.201	[-0.097,	0.037]	0.134
	CI	[-0.036,	0.049]		[-0.054,	0.19]		[-0.108,	0.051]	

Identification deteriorates with extent of underreporting of SNAP

MTS: Definition

Monotone treatment selection (**MTS**):

$$P[Y(j)=1 \mid S^* = 3]$$

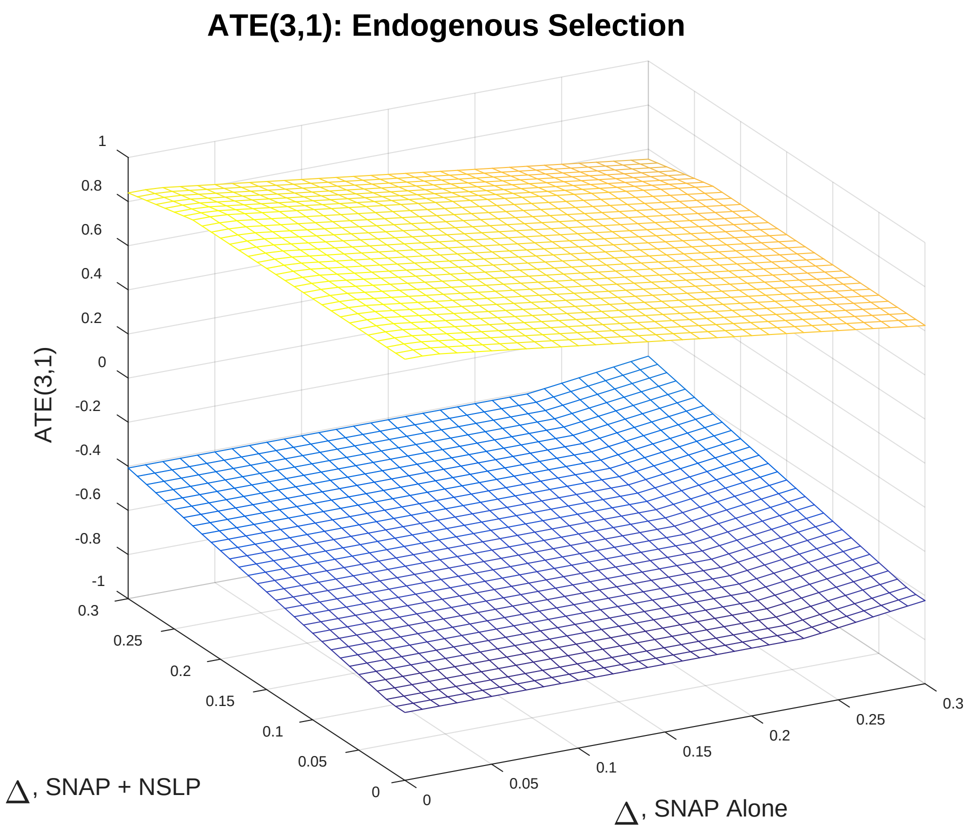
$$\leq P[Y(j)=1 \mid S^* = k] \leq$$

$$P[Y(j)=1 \mid S^* = 0] \quad \forall\ j;\ k = 1,2$$

- Under MTS assumption, decision to participate is monotonically related to food insecurity: on average, households choose to participate in more programs if anticipating worse food security situation

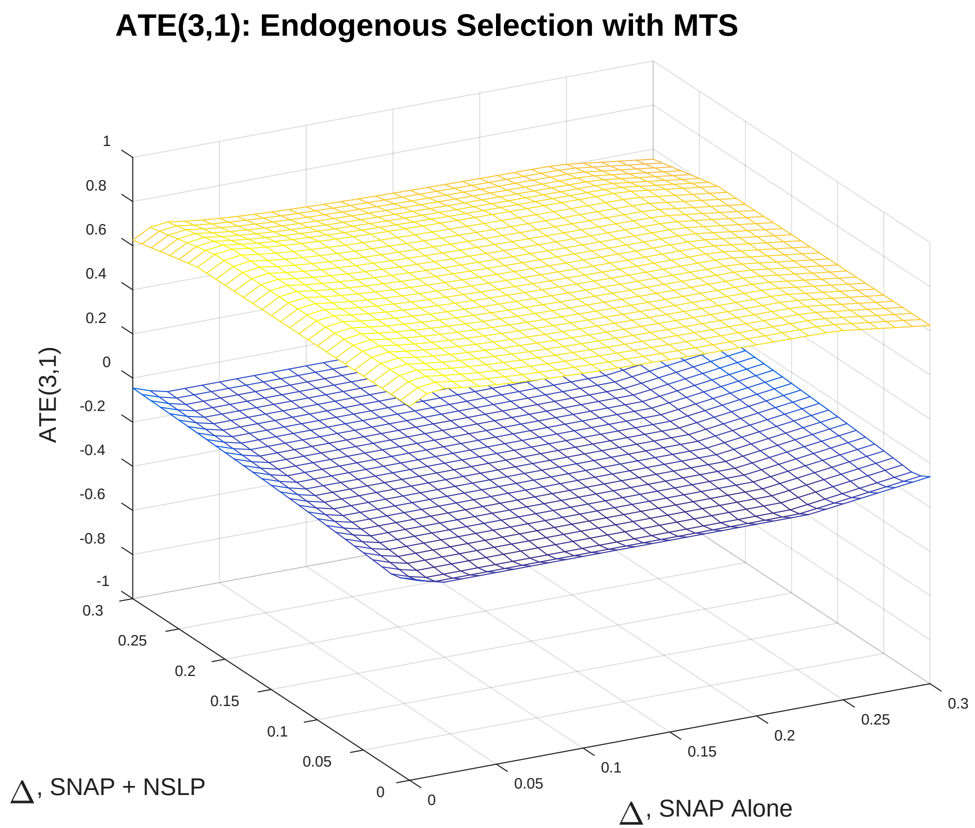
Recall: Worst-Case Bounds (CPS)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



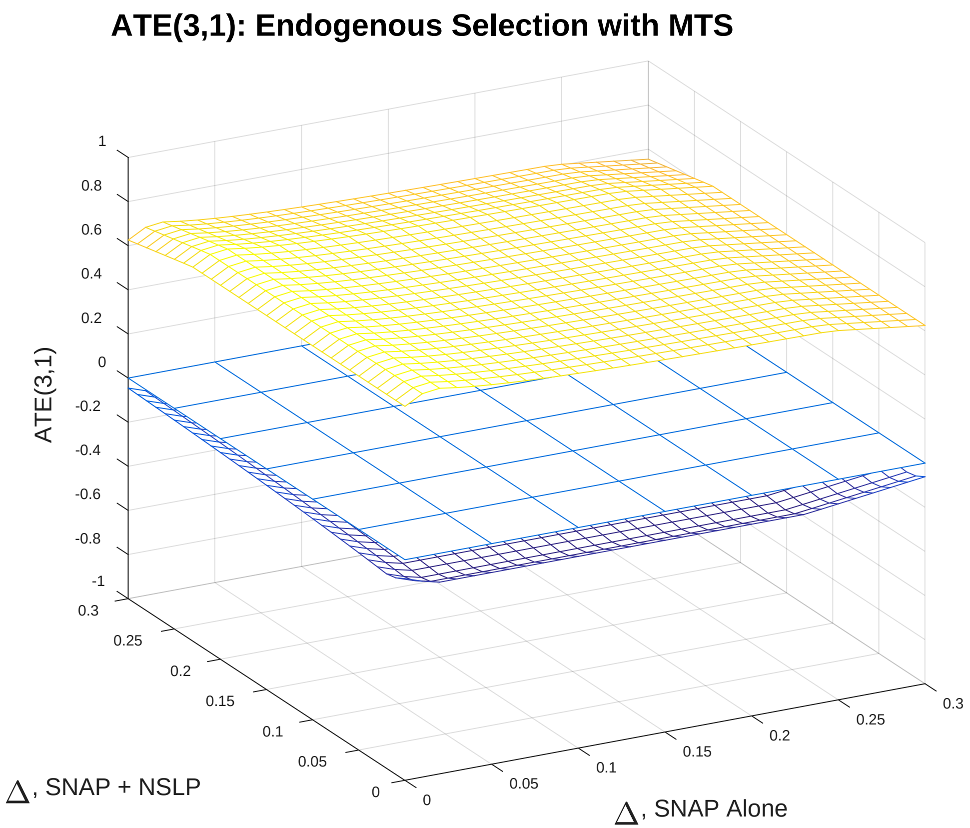
Endogenous Selection with MTS (CPS)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



Endogenous Selection with MTS (CPS)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



MTR: Definition

Monotone treatment response (**MTR**):

$$P[Y(3)=1 \mid S^{*}] \geq P[Y(1)=1 \mid S^{*}] \geq P[Y(0)=1 \mid S^{*}]$$

$$P[Y(3)=1 \mid S^{*}] \geq P[Y(2)=1 \mid S^{*}] \geq P[Y(0)=1 \mid S^{*}]$$

- On average, participation in more food programs would not harm food security (but might not help)

MIV: Definition

Monotone instrumental variable (**MIV**):

$$
\begin{aligned}
&u_1 \le u \le u_2 \Rightarrow \\
&\qquad P[Y(j)=1 \mid v=u_1] \\
&\qquad\qquad \le P[Y(j)=1 \mid v=u] \le \\
&\qquad\qquad\qquad P[Y(j)=1 \mid v=u_2]
\end{aligned}
$$

We construct and use:

$$v = \frac{\text{Actual food-at-home expenditures}}{\text{TFP-based food expenditures}}$$

Assumption: higher v would not harm food security on average

IV: Definition

Instrumental variable (**IV**):

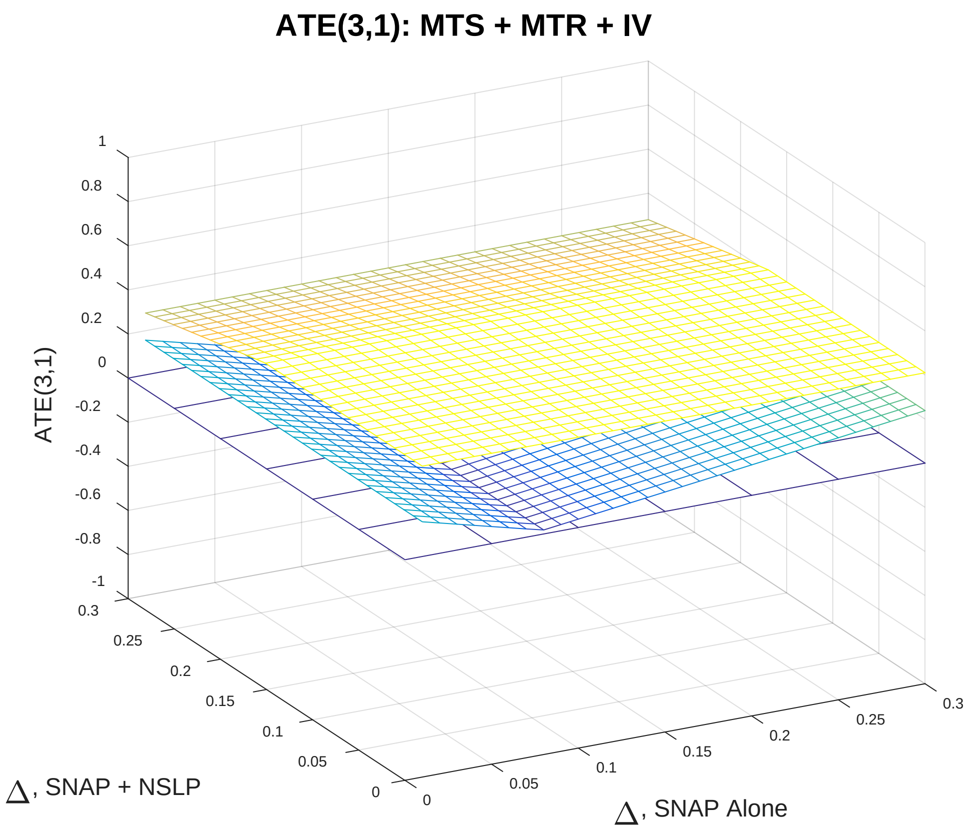
$$\forall u_1, u_2 :$$

$$P[Y(j)=1 \mid v=u_1] = P[Y(j)=1 \mid v=u_2]$$

- IV is a special case of MIV
- We employ **SNAP Policy Database** to construct conventional IVs used in previous literature to instrument for SNAP participation. Many such IVs are binary
- We create a scalar IV with many values by combining seven conventional IVs

Bounds under MTS + MTR + IV (CPS)

Bounds on ATE of participating in SNAP and NSLP vs. SNAP alone:



Summary

Motivating question: How do existing food programs interact in creating a food safety net?

Research objective: Quantify by how much SNAP+WIC improves household food security vs. SNAP alone, or WIC alone, or nonparticipation

FoodAPS data: Partial verification of SNAP participation

Methodology: Nonparametric bounding approach handles endogeneity, misreporting, multiple partially ordered treatments

Appendix: Supplementary Data Sources

SNAP Policy Database provides state-level policies regarding SNAP eligibility, reporting requirements, use of biometric technology, etc.

- Coverage: every state, every month, 1996–
- Allows us to construct **IVs** for SNAP participation used in the literature:
  - Continuous: e.g., SNAP outreach spending per capita
  - Binary: e.g., fingerprinting, phone certification