Bounding Multiple Program Treatment Effects with Unobserved Partially-Ordered Treatments

Helen Jensen

Brent Kreider

Oleksandr Zhylyevskyy

Our Research Projects

(1) Effects of SNAP and NSLP on food security

- Outcomes: low and very low food security among children
- Programs: SNAP (*aka* 'food stamps') and NSLP
- Data sources: FSS/CPS, QFAHPD, SNAP Policy Database

(2) Effects of SNAP and WIC on food security and mental health

- Outcomes: low household food security; anxiety, depression, emotional problems affecting ADLs, mental illnesses among adults
- Programs: SNAP and WIC
- Data sources: NHIS (restricted-access data), SNAP Policy Database

(3) Effects of SNAP and WIC on food security using FoodAPS

- Outcome: low household food security
- Programs: SNAP and WIC
- Data sources: FoodAPS, FoodAPS-GC (restricted-access data)

Empirical Relevance

We study large food assistance programs

- In FY 2015, 46M people participated in SNAP on average per month
- 22M children got free/reduced price lunch via NSLP on ave per day
- 8M people participated in WIC on average per month
- Federal spending: SNAP: $73.9B, NSLP: $13B, WIC: $6.2B
- Nearly half of all children will be on SNAP at some point in childhood

We focus on health-related outcomes with substantial prevalence

- In 2014, among all 124M U.S. households:
  - 86% were food secure
  - 8.4% had low food security (10.5M)
  - 5.6% had very low food security (6.9M)
- Among all 39M HHs with children, 3.7M had food-insecure children
- 18% of U.S. adults (43.6M) had mental illness
- 4.1% (9.8M) had serious mental illness
- 6.6% (15.7M) had at least one major depressive episode (2+ weeks)

Food Security

Conceptually, food security means access to enough food for active, healthy life. It implies:

- Ready availability of nutritionally adequate and safe foods, and
- Assured ability to acquire such foods in socially acceptable ways

In practice, food security is assigned based on a survey module with questions on food-related behaviors under lack of resources

- Ex. "*Did you ever cut the size of your meals or skip meals because there wasn't enough money for food?" (Yes/No)*
- FSS/CPS uses 18 questions, others (e.g., FoodAPS) use ≤ 10 questions
- Questions can focus on household, adults, or children

Answers are converted into # of food-insecure conditions. A threshold separates food secure from food insecure

Food Security Supplement (FSS) of CPS

FSS is added to “core” CPS in December, covers 45K households

- Food purchase behavior, actual and usual weekly food expenditures
- Participation in SNAP, NSLP, SBP, WIC, day-care/Head Start with free food
- 18-item FSM, referenced to last 12 months and last 30 days
- Coping strategies: using food pantries, soup kitchens, etc.

We use FSS to study how SNAP and NSLP affect food insecurity of children among low-income households with school-age children

- Pool FSS data from 2002 through 2010
- Drop irrelevant households and perform data cleaning
- Sample size before imposing income/eligibility constraints: 55,738

We combine FSS with QFAHPD and SNAP Policy Database to construct MIVs and IVs

NHIS

CDC's major data collection program since 1957. Source of info on health, illness/disability trends, federal health programs, etc.

NHIS is cross-sectional; years can be pooled. Structure is complex

- Core components: household, family, sample adult, sample child
- Supplements: in-depth study of core subjects, other topics of interest
- Annual sample: 35K households, 87.5K persons, 34K sample adults, 13K sample children

We focus on effects of SNAP and WIC on household food security and adult mental health conditions. Requisite data are available starting in 2011. Also, 2012 contains a mental health supplement

- Mental health indicators: depression, psychological distress (Kessler's "K6" screen), other disorders, mental problems causing difficulty with ADLs

Parametric approach

Ex. Linear model for a binary treatment:

$$Y_i = \gamma S_i + x_i \beta + \varepsilon_i$$

- Estimation using OLS is subject to omitted variables bias (endogeneity)
- Neither linearity nor homogenous response is grounded in economic theory
- If *S* is mismeasured, then estimates of γ using either OLS or IV will be inconsistent
- In our application, all the classical measurement error assumptions are violated

Average treatment effects (ATE) for multiple treatments:

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

- $Y(S^*)$ is the potential outcome under treatment S^*
- S is the self-reported counterpart in the data
- X denotes any covariates of interest

Ex. Treatment 0 = no participation, 1 = participation in SNAP alone, 2 = participation in NSLP alone, and 3 = both

- $ATE_{31} = P[Y(3)] - P[Y(1)]$ measures how the prevalence of a favorable outcome would change if all eligible households participated in both SNAP and NSLP vs. SNAP alone.

Usual problem: we don’t observe counterfactuals

- What if we don’t observe the treatments either?

For binary outcomes:

$$ATE_{jk} = E[Y(S^* = j) \mid X] - E[Y(S^* = k) \mid X] \tag{1}$$

$$\text{for } j,k \in \{0,1,2,\ldots\},\ j \neq k$$

$$\Rightarrow ATE_{jk} = P[Y(S^* = j) = 1 \mid X] - P[Y(S^* = k) = 1 \mid X] \tag{1'}$$

Definition of errors: $\theta_i^{m,m'} \equiv P(\overset{\text{outcome}}{Y} = i, \overset{\text{reported treatment}}{S} = m, \overset{\text{actual treatment}}{S^*} = m')$

(fraction of households with outcome i reporting treatment m with actual treatment m')

outcome

reported treatment

actual treatment

Definition of errors: $\theta_i^{m,m'} \equiv P(Y = i, S = m, S^* = m')$

(fraction of households with outcome *i* reporting treatment *m* with actual treatment m')

Delta constraints:

true participation rate in program *j*

reported participation rate in program *j*

(D) $\Delta_j \equiv P(S^* = j) - P(S = j) \equiv P_j^* - P_j = \theta^{-j,j} - \theta^{j,-j}$

Definition of errors: $\theta_i^{m,m'} \equiv P(Y=i, S=m, S^*=m')$

outcome, reported treatment, actual treatment

(fraction of households with outcome *i* reporting treatment *m* with actual treatment *m′*)

Delta constraints:

true participation rate in program *j*

reported participation rate in program *j*

(D) $\Delta_j \equiv P(S^*=j) - P(S=j) \equiv P_j^* - P_j = \theta^{-j,j} - \theta^{j,-j}$

Ex. $\Delta_1 \equiv P_1^* - P_1 = \theta^{-1,1} - \theta^{1,-1}$

$$= (\theta_1^{0,1} + \theta_0^{0,1}) + (\theta_1^{2,1} + \theta_0^{2,1}) + (\theta_1^{3,1} + \theta_0^{3,1})$$

$$- (\theta_1^{1,0} + \theta_0^{1,0}) - (\theta_1^{1,2} + \theta_0^{1,2}) - (\theta_1^{1,3} + \theta_0^{1,3})$$

outcome reported treatment actual treatment

Definition of errors: $\theta_i^{m,m'} \equiv P(Y = i, S = m, S^* = m')$

(fraction of households with outcome *i* reporting treatment m with actual treatment m')

Delta constraints:

true participation rate in program *j* reported participation rate in program *j*

$$\text{(D)} \qquad \Delta_j \equiv P(S^* = j) - P(S = j) \equiv P_j^* - P_j = \theta^{-j,j} - \theta^{j,-j}$$

$$\text{Ex.} \qquad \Delta_1 \equiv P_1^* - P_1 = \theta^{-1,1} - \theta^{1,-1}$$

$$= (\theta_1^{0,1} + \theta_0^{0,1}) + (\theta_1^{2,1} + \theta_0^{2,1}) + (\theta_1^{3,1} + \theta_0^{3,1})$$

$$-(\theta_1^{1,0} + \theta_0^{1,0}) - (\theta_1^{1,2} + \theta_0^{1,2}) - (\theta_1^{1,3} + \theta_0^{1,3})$$

from USDA caseloads

- If we don't know P_1^*, can use $P(S^* = 1) \le \min\{(P_1^* + P_3^*), 1 - (P_2^* + P_3^*)\}$

Restrictions on errors based on P_m^* and observed joint probabilities $P(Y=i, S=j, V=0)$:

(R1) $\theta_0^{-m,m} + \theta_1^{-m,m} \le P_m^*,\ m = 0,\ldots,3$

(R1_0) $(\theta_0^{1,0} + \theta_0^{2,0} + \theta_0^{3,0}) + (\theta_1^{1,0} + \theta_1^{2,0} + \theta_1^{3,0}) \le P_0^*$

(R1_1) $(\theta_0^{0,1} + \theta_0^{2,1} + \theta_0^{3,1}) + (\theta_1^{0,1} + \theta_1^{2,1} + \theta_1^{3,1}) \le P_1^*$

(R1_2) $(\theta_0^{0,2} + \theta_0^{1,2} + \theta_0^{3,2}) + (\theta_1^{0,2} + \theta_1^{1,2} + \theta_1^{3,2}) \le P_2^*$

(R1_3) $(\theta_0^{0,3} + \theta_0^{1,3} + \theta_0^{2,3}) + (\theta_1^{0,3} + \theta_1^{1,3} + \theta_1^{2,3}) \le P_3^*$

Restrictions on errors based on P_m^* and observed joint probabilities $P(Y=i,S=j,V=0)$:

$$\text{(R1)}\quad \theta_0^{-m,m}+\theta_1^{-m,m}\le P_m^*,\; m=0,\ldots,3$$

$$\text{(R1\_0)}\quad (\theta_0^{1,0}+\theta_0^{2,0}+\theta_0^{3,0})+(\theta_1^{1,0}+\theta_1^{2,0}+\theta_1^{3,0})\le P_0^*$$
$$\text{(R1\_1)}\quad (\theta_0^{0,1}+\theta_0^{2,1}+\theta_0^{3,1})+(\theta_1^{0,1}+\theta_1^{2,1}+\theta_1^{3,1})\le P_1^*$$
$$\text{(R1\_2)}\quad (\theta_0^{0,2}+\theta_0^{1,2}+\theta_0^{3,2})+(\theta_1^{0,2}+\theta_1^{1,2}+\theta_1^{3,2})\le P_2^*$$
$$\text{(R1\_3)}\quad (\theta_0^{0,3}+\theta_0^{1,3}+\theta_0^{2,3})+(\theta_1^{0,3}+\theta_1^{1,3}+\theta_1^{2,3})\le P_3^*$$

$$\text{(R2)}\quad \theta_i^{m,-m}\le P(Y=i,S=m,V=0),\; m=0,\ldots,3 \quad i=0,1$$

$$\text{(R2\_0)}\quad \theta_i^{0,1}+\theta_i^{0,2}+\theta_i^{0,3}\le P(Y=i,S=0,V=0)\equiv P_{i00}$$
$$\text{(R2\_1)}\quad \theta_i^{1,0}+\theta_i^{1,2}+\theta_i^{1,3}\le P(Y=i,S=1,V=0)\equiv P_{i10}$$
$$\text{(R2\_2)}\quad \theta_i^{2,0}+\theta_i^{2,1}+\theta_i^{2,3}\le P(Y=i,S=2,V=0)\equiv P_{i20}$$
$$\text{(R2\_3)}\quad \theta_i^{3,0}+\theta_i^{3,1}+\theta_i^{3,2}\le P(Y=i,S=3,V=0)\equiv P_{i30}$$

Restrictions based on Q_j^{UB} and Δ_j:

- Can impose an upper bound on the degree of data corruption Q_j^{UB} for each treatment *j*:

$$\theta_1^{-j,j} + \theta_1^{j,-j} + \theta_0^{-j,j} + \theta_0^{j,-j} \leq Q_j^{UB} \in \left[\left| \Delta_j \right|, 1 \right]$$

Restrictions based on Q_j^{UB} and Δ_j:

- Can impose an upper bound on the degree of data corruption Q_j^{UB} for each treatment *j*:

$$\theta_1^{-j,j} + \theta_1^{j,-j} + \theta_0^{-j,j} + \theta_0^{j,-j} \le Q_j^{UB} \in \left[\left| \Delta_j \right|, 1 \right]$$

- logically Q_j^{UB} must be chosen to be at least as large as the magnitude of the gap between the true and reported participation rates, $\left| \Delta_j \right|$.

Restrictions based on Q_j^{UB} and Δ_j:

- Can impose an upper bound on the degree of data corruption Q_j^{UB} for each treatment *j*:

$$\theta_1^{-j,j} + \theta_1^{j,-j} + \theta_0^{-j,j} + \theta_0^{j,-j} \leq Q_j^{UB} \in \left[\left|\Delta_j\right|, 1\right]$$

- logically Q_j^{UB} must be chosen to be at least as large as the magnitude of the gap between the true and reported participation rates, $\left|\Delta_j\right|$.

Lemma 1. *Knowledge of P_j^* places the following restrictions on error rates:*

(L1) $\theta_0^{j,-j} + \theta_1^{j,-j} \leq \frac{1}{2}\left(Q_j^{UB} - \Delta_j\right)$ *and*

(L2) $\theta_0^{-j,j} + \theta_1^{-j,j} \leq \frac{1}{2}\left(Q_j^{UB} + \Delta_j\right)$ *for each treatment j.*

Result. *Under arbitrary errors and arbitrary selection, sharp worst-case bounds on multiple-program ATEs for* $j \neq k$ *are given as follows (adds data errors to Manski's 1995 multiple treatments):*

$$ATE_{j,k}^{LB} = -P(Y=1, S \neq j) - P(Y=0, S \neq k) + \underset{\theta}{\arg\min}(\theta_1^{-j,j} - \theta_1^{j,-j} + \theta_0^{-k,k} - \theta_0^{k,-k})$$

$$ATE_{j,k}^{UB} = \ P(Y=0, S \neq j) + P(Y=1, S \neq k) + \underset{\theta}{\arg\max}(\theta_0^{j,-j} - \theta_0^{-j,j} + \theta_1^{k,-k} - \theta_1^{-k,k})$$

Result. *Under arbitrary errors and arbitrary selection, sharp worst-case bounds on multiple-program ATEs for $j \neq k$ are given as follows (adds data errors to Manski's 1995 multiple treatments):*

$$ATE_{j,k}^{LB} = -P(Y=1, S \neq j) - P(Y=0, S \neq k) + \underset{\theta}{\arg\min}(\theta_1^{-j,j} - \theta_1^{j,-j} + \theta_0^{-k,k} - \theta_0^{k,-k})$$

$$ATE_{j,k}^{UB} = \ P(Y=0, S \neq j) + P(Y=1, S \neq k) + \underset{\theta}{\arg\max}(\theta_0^{j,-j} - \theta_0^{-j,j} + \theta_1^{k,-k} - \theta_1^{-k,k})$$

s.t. $\theta_i^{m,m'} \geq 0, i = 0,1 \ \ \forall m \neq m'$

(D) $\Delta_m = \theta^{-m,m} - \theta^{m,-m}, \ m = 1,\ldots,M$

(R1) $\theta_0^{-m,m} + \theta_1^{-m,m} \leq P_m^*, \ m = 0,\ldots,M$

(R2) $\theta_i^{m,-m} \leq P(Y=i, S=m, V=0), \ m = 0,\ldots,M \quad i = 0,1$

(L1) $\theta_0^{m,-m} + \theta_1^{m,-m} \leq \frac{1}{2}\left(Q_m^{UB} - \Delta_m\right), \ m = 0,\ldots,M$

(L2) $\theta_0^{-m,m} + \theta_1^{-m,m} \leq \frac{1}{2}\left(Q_m^{UB} + \Delta_m\right), \ m = 0,\ldots,M$

Verification assumptions:

(1) Administrative data (FoodAPS, USDA caseload)

Verification assumptions:

(1) Administrative data (FoodAPS, USDA caseload)

(2) Assumptions about reliability of responses

(V1) *No false+*: $S = 1 \Rightarrow S^* \in \{1, 3\}, \quad S = 2 \Rightarrow S^* \in \{2, 3\},$
$S = 3 \Rightarrow S^* = 3$

– A positive program participation response from a household indicates accurate reporting for at least that particular program

Verification assumptions:

(1) Administrative data (FoodAPS, USDA caseload)

(2) Assumptions about reliability of responses

(V1) *No false+*: $S = 1 \Rightarrow S^* \in \{1, 3\}, \quad S = 2 \Rightarrow S^* \in \{2, 3\},$

$S = 3 \Rightarrow S^* = 3$

– A positive program participation response from a household indicates accurate reporting for at least that particular program

(V2) *Strong verification* (lack of stigma): $S > 0 \Rightarrow S^* = S$

– Any positive program participation response from a household indicates accurate reporting for all programs

FoodAPS

- First nationally representative survey to collect comprehensive data on household food purchases and acquisitions; covers food-at-home and away from home. Data have exceptional depth
- A household participated during one week between April 2012 and January 2013. Sample of 4,826 households: SNAP participants, low-income nonparticipants, higher income nonparticipants
- We focus on effects of SNAP and WIC on food security

Two data features of high value:

- FoodAPS contains *administratively verified* info on SNAP participation: reduces dimensionality of classification error problem
- FoodAPS-GC provides local food environment data: can construct MIVs related to food expenditure and food retailer availability

FoodAPS data "partially verifies" SNAP participation

- Matched administrative data establishes whether on SNAP or not, but not a particular value of S^*

SNAP alone

SNAP + WIC

– differentiates between $S^* \in (1,3)$ or $S^* \in (0,2)$

WIC alone

neither

FoodAPS data "partially verifies" SNAP participation

- Matched administrative data establishes whether on SNAP or not, but not a particular value of S^*

SNAP alone

SNAP + WIC

– differentiates between $S^* \in (1,3)$ or $S^* \in (0,2)$

WIC alone

neither

fraction with good outcome that falsely deny being on SNAP alone

$$\text{Ex. } 0 \le \theta_0^{-1,1} \le \min\{P[Y=0, S \neq 1, S^* \in (1,3)], P(S^*=1)\}$$

$$= \theta_0^{0,1} + \theta_0^{2,1} + \theta_0^{3,1}$$

– exploits both self-reported and administrative information in FoodAPS

Counterintuitive result?

Comparing the impacts of any two programs, or combinations, the presence of classification error in the treatment variables under strong verification can only narrow Manski's worst-case ATE bounds:

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

where $\Theta_{j,k}^{LB}\equiv\ \theta_1^{-j,j}+\theta_0^{-k,k}\geq 0$

and $\Theta_{j,k}^{UB}\equiv-\theta_0^{-j,j}-\theta_1^{-k,k}\leq 0$ for $j,k>0$

Consider the potential outcome if participating in program *j*:

$$P[Y(j)=1]=P[Y(S^*=j)=1\,|\,S^*=j]P(S^*=j)$$

$$+P[Y(S^*=j)=1\,|\,S^*\neq j]P(S^*\neq j)$$

Consider the potential outcome if participating in program *j*:

$$P[Y(j)=1]=P[Y(S^*=j)=1\,|\,S^*=j]P(S^*=j)$$

$$+P[Y(S^*=j)=1\,|\,S^*\neq j]P(S^*\neq j)$$

these are actual outcomes, not just potential

$$\Rightarrow P[Y(j)=1]=P(Y=1\,|\,S^*=j)P(S^*=j)+[0,1]_j P(S^*\neq j)$$

Consider the potential outcome if participating in program *j*:

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

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

so far, no restrictions on counterfactuals

$$\Rightarrow P[Y(j)=1]=P(Y=1 \mid S^*=j)P(S^*=j)+[0,1]_j P(S^* \neq j)$$

$$=P(Y=1,S^*=j)+[0,1]_j P(S^* \neq j)$$

Consider the potential outcome if participating in program *j*:

$$P[Y(j)=1]=P[Y(S^*=j)=1\,|\,S^*=j]P(S^*=j)$$

$$+P[Y(S^*=j)=1\,|\,S^*\neq j]P(S^*\neq j)$$

so far, no restrictions on counterfactuals

$$\Rightarrow P[Y(j)=1]=P(Y=1\,|\,S^*=j)P(S^*=j)+[0,1]_j P(S^*\neq j)$$

$$=P(Y=1,S^*=j)+[0,1]_j P(S^*\neq j)$$

$$=P(Y=1,S=j)+\theta_1^{-j,j}-\theta_1^{j,-j}$$

$$+[0,1]_j[P(S\neq j)+(\theta_1^{j,-j}+\theta_0^{j,-j})-(\theta_1^{-j,j}+\theta_0^{-j,j})]$$

Consider the potential outcome if participating in program *j*:

$$P[Y(j)=1]=P[Y(S^*=j)=1\,|\,S^*=j]P(S^*=j)$$
$$+P[Y(S^*=j)=1\,|\,S^*\neq j]P(S^*\neq j)$$

$$\Rightarrow P[Y(j)=1]=P(Y=1\,|\,S^*=j)P(S^*=j)+[0,1]_j P(S^*\neq j)$$
$$=P(Y=1,S^*=j)+[0,1]_j P(S^*\neq j)$$

0 for $j>0$: if say enrolled, then we trust the data

$$=P(Y=1,S=j)+\theta_1^{-j,j}-\cancel{\theta_1^{j,-j}}$$
$$+[0,1]_j[P(S\neq j)+(\cancel{\theta_1^{j,-j}}+\cancel{\theta_0^{j,-j}})-(\theta_1^{-j,j}+\theta_0^{-j,j})]$$

$\Rightarrow$ Worst-case bounds on potential outcomes narrow, not expand, with the degree of data errors:

$$P(Y=1, S=j) + \theta_1^{-j,j}$$

$$\leq P[Y(j)=1] \leq$$

$$P(Y=1, S=j) + P(S \neq j) - \theta_0^{-j,j}$$

- Intuition: what worsens the LB (UB) is the possibility that households with good (bad) outcomes in a program deny participating in the program – but ruling out here
- The basic result holds as long as false negatives don't dominate false positives

Proposition 1. *Under strong verification and no restrictions on the selection process, sharp worst-case bounds on multiple-program ATEs are given as follows:*

(A) Participating in both programs vs. no participation: ($ATE_{3,0}$)

$$-P(Y=1,S\neq 3)-P(Y=0,S\neq 0)+\max\left\{0,\Delta_3-P_{000}\right\}-\min\left\{\Delta_1+\Delta_2+\Delta_3,P_{000}\right\}$$

$$\leq ATE_{3,0}\leq$$

$$P(Y=0,S\neq 3)+P(Y=1,S\neq 0)-\max\left\{0,\Delta_3-P_{100}\right\}+\min\left\{\Delta_1+\Delta_2+\Delta_3,P_{100}\right\}$$

Proposition 1. *Under strong verification and no restrictions on the selection process, sharp worst-case bounds on multiple-program ATEs are given as follows:*

(A) Participating in both programs vs. no participation: ($ATE_{3,0}$)

$$-P(Y=1,S\neq 3)-P(Y=0,S\neq 0)+\max\left\{0,\Delta_3-P_{000}\right\}-\min\left\{\Delta_1+\Delta_2+\Delta_3,P_{000}\right\}$$

$$\leq ATE_{3,0}\leq$$

$$P(Y=0,S\neq 3)+P(Y=1,S\neq 0)-\max\left\{0,\Delta_3-P_{100}\right\}+\min\left\{\Delta_1+\Delta_2+\Delta_3,P_{100}\right\}$$

(B) Participating in both programs vs. one program: (e.g., $ATE_{3,1}$)

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

$$\leq ATE_{3,1}\leq$$

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

Exogenous Selection

- Suppose selection is exogenous such that $E[Y(j)] = E[Y(j) \mid S^*]$:

$$\Rightarrow P[Y(j)=1] = P(Y=1 \mid S^* = j) = \frac{P(Y=1, S^* = j)}{P(S^* = j)}$$

$$= \frac{P(Y=1, S=j) + \theta_1^{-j,j} - \theta_1^{j,-j}}{P_j^*} \in [0,1]$$

Exogenous Selection

- Suppose selection is exogenous such that $E[Y(j)] = E[Y(j) \mid S^*]$:

$$\Rightarrow P[Y(j) = 1] = P(Y = 1 \mid S^* = j) = \frac{P(Y = 1, S^* = j)}{P(S^* = j)}$$

$$= \frac{P(Y = 1, S = j) + \theta_1^{-j,j} - \theta_1^{j,-j}}{P_j^*} \in [0,1]$$

Ex. Participation in both programs compared with no program:

$$ATE_{3,0} = \frac{P(Y = 1, S = 3) + \theta_1^{-3,3} - \theta_1^{3,-3}}{P_3^*} - \frac{P(Y = 1, S = 0) + \theta_1^{-0,0} - \theta_1^{0,-0}}{P_0^*} \quad \text{(E)}$$

Derive sharp **lower bound** on $ATE_{3,0}$ under strong verification:

$$ATE_{3,0} = \frac{P(Y=1,S=3)+\theta_1^{0,3}}{P_3^*} - \frac{P(Y=1,S=0)-(\theta_1^{0,1}+\theta_1^{0,2}+\theta_1^{0,3})}{P_0^*}$$

$$\text{(D)} \quad \Delta_j = \theta_0^{0,j} + \theta_1^{0,j} \quad \forall j$$

$$\text{(R1)} \quad \theta_0^{-m,m} + \theta_1^{-m,m} \le P_m^*, \quad \forall m$$

$$\text{(R2)} \quad \theta_i^{m,-m} \le P(Y=i,S=m,V=0), \quad \forall i,m$$

(a) Minimize $\theta_1^{0,3}$ in the 1st ratio and $(\theta_1^{0,1}+\theta_1^{0,2}+\theta_1^{0,3})$ in the second.

(b) Begin by minimizing $\theta_1^{0,3}$: make $\theta_0^{0,3}$ as close as possible to Δ_3, with $\theta_1^{0,3}$ making up any remaining gap: $\theta_0^{0,3*} = \min\{\Delta_3, P_{000}, \cancel{P_3^*}\}$

$$\Rightarrow \theta_1^{0,3*} = \Delta_3 - \theta_0^{0,3*} = \max\{0, \Delta_3 - P_{000}\}.$$

$$\Rightarrow \quad ATE_{3,0} \geq \frac{P(Y=1,S=3)+\max\{0,\Delta_3 - P_{000}\}}{P_3^*}$$

$$- \frac{P(Y=1,S=0)-\max\{0,\Delta_3 - P_{000}\}-(\theta_1^{0,1}+\theta_1^{0,2})}{P_0^*}$$

Note: LHS ratio ≤ 1 as required:

$$\max\{0,\Delta_3 - P_{000}\} \leq P_3^* - P(Y=1,S=3) \quad \text{(to show)}$$

$$\Leftrightarrow \max\{0,\Delta_3 - P_{000}\} \leq \Delta_3 + P(Y=0,S=3) \text{ using } \Delta_3 \equiv P_3^* - P_3$$

and clearly $\max\{0,\Delta_3 - P_{000}\} \leq \Delta_3$.

$$\Rightarrow \quad ATE_{3,0} \geq \frac{P(Y=1,S=3)+\max\{0,\Delta_3 - P_{000}\}}{P_3^*}$$

$$- \frac{P(Y=1,S=0)-\max\{0,\Delta_3 - P_{000}\}-(\theta_1^{0,1}+\theta_1^{0,2})}{P_0^*}$$

(c) Next, minimize $(\theta_1^{0,1}+\theta_1^{0,2})$ given $\theta_1^{0,3*}$ by maximizing $(\theta_0^{0,1}+\theta_0^{0,2})$

subject to it not occupying the same space as $\theta_0^{0,3*}$.

(d) Can show $\theta_0^{0,1*} = \min\{\Delta_1, \max\{0, P_{000} - \Delta_3\}\}$ and

$\theta_0^{0,2*} = \min\{\Delta_2, \max\{0, \max\{0, P_{000} - \Delta_3\} - \Delta_1\}\}$ which

ultimately implies $\theta_0^{0,1*} + \theta_0^{0,2*} + \theta_0^{0,3*} = \min\{\Delta_1 + \Delta_2 + \Delta_3, P_{000}\}$.

(e) Can verify that the RHS ratio $\in [0,1]$ as required.

Proposition 2. *Under multiple program participation, strong verification, and exogenous selection, sharp bounds on average treatment effects are given as follows:*

(A) Participating in both programs vs. no participation: ($ATE_{3,0}$)

$$\frac{P(Y=1,S=3)+\max\left\{0,\Delta_3-P_{000}\right\}}{P_3^*}-\frac{P(Y=1,S=0)-\max\left\{0,\Delta_1+\Delta_2+\Delta_3-P_{000}\right\}}{P_0^*}$$

$$\leq ATE_{3,0} \leq$$

$$\frac{P(Y=1,S=3)+\min\left\{\Delta_3,P_{100}\right\}}{P_3^*}-\frac{P(Y=1,S=0)-\min\left\{\Delta_1+\Delta_2+\Delta_3,P_{100}\right\}}{P_0^*}$$

(B) Participating in both programs vs. one program (e.g., $ATE_{3,1}$)

$$\frac{P(Y=1,S=3)+\max\left\{0,\Delta_3-P_{000}\right\}}{P_3^*}-\frac{P(Y=1,S=1)+\min\left\{\Delta_1,P_{100}\right\}}{P_1^*}$$
$$\leq ATE_{3,1}\leq$$
$$\frac{P(Y=1,S=3)+\min\left\{\Delta_3,P_{100}\right\}}{P_3^*}-\frac{P(Y=1,S=1)+\max\left\{0,\Delta_1-P_{000}\right\}}{P_1^*}$$

Monotone Instrumental Variables (MIV)

- Formalizes the notion that the expected outcome $E[Y(j)]$ is known to vary monotonically with certain covariates (Manski and Pepper, 2000)

- Traditional IV assumption: mean response is constant across households with different values of a covariate

- MIV instead allows mean response to vary monotonically

$P[Y(j)=1 | v=u_1] \leq P[Y(j)=1 | v=u_2] \ \forall j$ and $u_1 \leq u_2$.

Monotone Treatment Selection (MTS)

- Suppose unobserved factors associated with bad outcomes are positively associated with take-up rates

- For semi-ordered treatments:

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

$\forall\, j$ and $k=1,2$

- E.g., tighten bounds on $P[Y(1)=1]$:

$$LB\text{\_P[1]\_exog}*(P_0^{*}+P_1^{*}) \leq P[Y(1)=1] \leq UB\text{\_P[1]\_exog}*(P_1^{*}+P_3^{*})+(P_0^{*}+P_2^{*})$$

Monotone Treatment Response (MTR)

- Formalize a researcher's belief that an assistance program does no harm on average:

$$P[Y(0)=1] \le P[Y(1)=1] \le [Y(3)=1]$$

$$P[Y(0)=1] \le P[Y(2)=1] \le [Y(3)=1]$$

- E.g., tighten bounds on $P[Y(1)=1]$:

$$P(Y=1,S=1)+P(Y=1,S=0)-\min\{\Delta_2+\Delta_3,P_{100}\}$$
$$\le P[Y(1)=1] \le$$
$$P(Y=1,S=1)+P(Y=1,S=3)+P_0^*+P_2^*+\min\{\Delta_1+\Delta_3,P_{100}\}$$

QFAHPD and SNAP Policy Database

QFAHPD is based on Nielsen Homescan: food purchase transactions by a large panel of households. ERS aggregated data within and across households by food group, area, time period

- Time coverage: every quarter between 1999 and 2010
- 54 food groups: e.g., fresh orange vegetables, low fat cheese
- 35 areas partitioning U.S. = 26 metro areas + 9 non-metro areas
- Food prices are expressed in $ per 100 grams as purchased

SNAP Policy Database is compiled by ERS to provide state-level information on policies relating to program eligibility, reporting requirements, use of biometric technology, etc.

- Coverage: every state and DC, every month between 1996 and 2011
- Allows us to construct nearly all IVs used in previous literature
    - Continuous: e.g., per capita SNAP outreach spending
    - Binary: e.g., fingerprinting, noncitizen eligibility

Food Expenditure-Based MIVs

Recall: v is MIV if for each treatment j and any two values of v, u_1 and u_2, such that $u_1 \leq u_2 \Rightarrow E[Y(j) | v = u_1] \leq E[Y(j) | v = u_2]$

Ex. Household income-to-poverty ratio in the context of SNAP effect on food security

We construct novel food expenditure-based MIVs of the form:

$$\frac{\text{Reported food expenditures}}{\text{Minimum necessary expenditures}}$$

Numerator comes from FSS: actual, usual weekly food spending

Denominator: we calculate weekly cost of TFP given household age-gender composition and local food prices in QFAHPD

Remark: TFP is USDA's standard for nutritious diet at minimal cost

{SNAP + NSLP} vs. NSLP alone

ATE(3,2), **exog**, no MIV

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[**-0.0034, -0.0034**]	0.000	[-0.0071, 0.0202]	0.027	[-0.0199, 0.0172]	0.037
	CI	[-0.0190, 0.0122]		[-0.0191, 0.0322]		[-0.0314, 0.0295]	
0.01	p.e.	[-0.0275, 0.0004]	0.028	[-0.0311, 0.0241]	0.055	[-0.0440, 0.0211]	0.065
	CI	[-0.0394, 0.0124]		[-0.0429, 0.0359]		[-0.0552, 0.0332]	
0.05	p.e.	[-0.0240, 0.0138]	0.038	[-0.0277, 0.0375]	0.065	[-0.0405, 0.0345]	0.075
	CI	[-0.0359, 0.0252]		[-0.0395, 0.0487]		[-0.0518, 0.0460]	

ATE(3,2), **endog**, no MIV

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[-0.651, 0.644]	**1.295**	[-0.651, 0.644]	**1.295**	[-0.651, 0.609]	**1.260**
	CI	[-0.658, 0.651]		[-0.658, 0.651]		[-0.658, 0.616]	
0.01	p.e.	[-0.651, 0.644]	**1.295**	[-0.651, 0.644]	1.295	[-0.651, 0.609]	1.260
	CI	[-0.658, 0.651]		[-0.658, 0.651]		[-0.658, 0.616]	
0.05	p.e.	[-0.616, 0.644]	**1.260**	[-0.616, 0.644]	1.260	[-0.616, 0.609]	1.226
	CI	[-0.623, 0.651]		[-0.623, 0.651]		[-0.623, 0.616]	

ATE(3,2), **endog + MTS**, no MIV

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[-0.0430, 0.431]	0.474	[-0.0430, 0.462]	0.505	[-0.0429, 0.528]	0.571
	CI	[-0.0536, 0.439]		[-0.0536, 0.470]		[-0.0535, 0.535]	
0.01	p.e.	[-0.0670, 0.431]	0.498	[-0.0670, 0.462]	0.529	[-0.0669, 0.528]	0.595
	CI	[-0.0774, 0.439]		[-0.0774, 0.470]		[-0.0773, 0.535]	
0.05	p.e.	[-0.0636, 0.431]	0.495	[-0.0636, 0.462]	0.526	[-0.0635, 0.528]	0.592
	CI	[-0.0740, 0.439]		[-0.0740, 0.470]		[-0.0739, 0.535]	

ATE(3,2), **MTS + MTR**, no MIV

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[-0.0215, 0.413]	0.435	[-0.0215, 0.433]	0.455	[-0.0215, 0.513]	0.535
	CI	[-0.0242, 0.420]		[-0.0242, 0.440]		[-0.0243, 0.520]	
0.01	p.e.	[-0.0215, 0.413]	0.435	[-0.0215, 0.433]	0.455	[-0.0215, 0.513]	0.535
	CI	[-0.0242, 0.420]		[-0.0242, 0.440]		[-0.0243, 0.520]	
0.05	p.e.	[-0.0215, 0.413]	0.435	[-0.0215, 0.433]	0.455	[-0.0215, 0.513]	0.535
	CI	[-0.0242, 0.420]		[-0.0242, 0.440]		[-0.0243, 0.520]	

ATE(3,2), **endog + MTS, MIV**

$Q_{J,K}$		0.00			0.01			0.05		
		LB	UB	width	LB	UB	width	LB	UB	width
0.00	p.e.	[**0.0598,**	**0.404**]	0.344	[**0.0589,**	**0.432**]	0.373	[**0.0561,**	**0.501**]	0.445
	CI	[**0.0196,**	**0.433**]		[**0.0188,**	**0.461**]		[**0.0161,**	**0.529**]	
	bias	+0.030	-0.025		+0.010	-0.013		+0.023	0.003	
0.01	p.e.	[**0.0489,**	**0.404**]	0.355	[**0.0481,**	**0.432**]	0.384	[**0.0453,**	**0.501**]	0.456
	CI	[**0.0034,**	**0.433**]		[**0.0026,**	**0.461**]		[-0.0001,	0.529]	
	bias	+0.032	-0.026		+0.002	-0.012		+0.019	-0.001	
0.05	p.e.	[**0.0525,**	**0.404**]	0.351	[**0.0516,**	**0.432**]	0.380	[**0.0488,**	**0.501**]	0.452
	CI	[**0.0080,**	**0.433**]		[**0.0072,**	**0.461**]		[**0.0045,**	**0.529**]	
	bias	+0.032	-0.025		+0.002	-0.007		+0.015	-0.006	

ATE(3,2), **MTS + MTR, MIV**

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[**0.0704, 0.390**]	0.320	[**0.0695, 0.411**]	0.341	[**0.0663, 0.490**]	0.424
	CI	[**0.0303, 0.418**]		[**0.0295, 0.438**]		[**0.0266, 0.517**]	
	bias	+0.028 -0.025		+0.008 -0.010		+0.028 0.002	
0.01	p.e.	[**0.0590, 0.390**]	0.331	[**0.0580, 0.411**]	0.353	[**0.0548, 0.490**]	0.435
	CI	[**0.0150, 0.418**]		[**0.0142, 0.438**]		[**0.0113, 0.517**]	
	bias	+0.030 -0.024		+0.002 -0.011		+0.023 0.001	
0.05	p.e.	[**0.0623, 0.390**]	0.328	[**0.0614, 0.411**]	0.349	[**0.0582, 0.490**]	0.432
	CI	[**0.0189, 0.418**]		[**0.0181, 0.438**]		[**0.0152, 0.517**]	
	bias	+0.030 -0.025		+0.001 -0.007		+0.019 -0.004	

{SNAP + NSLP} vs. SNAP alone

ATE(3,1), **exog**, no MIV

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[**-0.0074, -0.0074**]	**0.000**	[-0.0292, 0.138]	0.167	[-0.0726, 0.0787]	0.151
	CI	[-0.0383, 0.0236]		[-0.0499, 0.160]		[-0.0871, 0.101]	
0.01	p.e.	[-0.0314, -0.0036]	0.028	[-0.0532, 0.142]	0.195	[-0.0966, 0.0825]	0.179
	CI	[-0.0557, 0.0207]		[-0.0738, 0.164]		[-0.111, 0.104]	
0.05	p.e.	[-0.0280, 0.0098]	0.038	[-0.0498, 0.156]	0.205	[-0.0932, 0.0959]	0.189
	CI	[-0.0521, 0.0336]		[-0.0704, 0.177]		[-0.108, 0.117]	

ATE(3,1), **endog**, no MIV

$Q_{J,K}$		0.00			0.01			0.05		
		LB	UB	width	LB	UB	width	LB	UB	width
0.00	p.e.	[-0.693,	0.909]	**1.601**	[-0.693,	0.909]	**1.601**	[-0.693,	0.874]	**1.567**
	CI	[-0.699,	0.913]		[-0.699,	0.913]		[-0.699,	0.879]	
0.01	p.e.	[-0.693,	0.909]	**1.601**	[-0.693,	0.909]	1.601	[-0.693,	0.874]	1.567
	CI	[-0.699,	0.913]		[-0.699,	0.913]		[-0.699,	0.879]	
0.05	p.e.	[-0.658,	0.909]	**1.567**	[-0.658,	0.909]	1.567	[-0.658,	0.874]	1.532
	CI	[-0.664,	0.913]		[-0.664,	0.913]		[-0.664,	0.879]	

ATE(3,1), **endog + MTS**, no MIV

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[-0.0856, 0.695]	0.781	[-0.0921, 0.753]	0.845	[-0.105, 0.799]	0.904
	CI	[-0.0983, 0.704]		[-0.104, 0.761]		[-0.116, 0.806]	
0.01	p.e.	[-0.110, 0.695]	0.805	[-0.116, 0.753]	0.869	[-0.129, 0.799]	0.928
	CI	[-0.122, 0.704]		[-0.128, 0.761]		[-0.139, 0.806]	
0.05	p.e.	[-0.106, 0.695]	0.801	[-0.113, 0.753]	0.865	[-0.126, 0.799]	0.925
	CI	[-0.119, 0.704]		[-0.124, 0.761]		[-0.136, 0.806]	

ATE(3,1), **MTS + MTR**, no MIV

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[-0.0627, 0.678]	0.741	[-0.0627, 0.698]	0.761	[-0.0627, 0.778]	0.841
	CI	[-0.0681, 0.685]		[-0.0664, 0.705]		[-0.0662, 0.785]	
0.01	p.e.	[-0.0627, 0.678]	0.741	[-0.0627, 0.698]	0.761	[-0.0627, 0.778]	0.841
	CI	[-0.0681, 0.685]		[-0.0664, 0.705]		[-0.0662, 0.785]	
0.05	p.e.	[-0.0627, 0.678]	0.741	[-0.0627, 0.698]	0.761	[-0.0627, 0.778]	0.841
	CI	[-0.0681, 0.685]		[-0.0664, 0.705]		[-0.0662, 0.785]	

ATE(3,1), **endog + MTS, MIV**

$Q_{J,K}$		0.00			0.01			0.05		
		LB	UB	width	LB	UB	width	LB	UB	width
0.00	p.e.	[**0.0044,**	**0.663**]	0.659	[**0.0031,**	**0.727**]	0.724	[-0.0062,	0.774]	0.780
	CI	[-0.0636,	0.696]		[-0.0523,	0.763]		[-0.0531,	0.805]	
	bias	+0.049	-0.029		+0.035	-0.001		+0.011	-0.010	
0.01	p.e.	[-0.0065,	0.663]	0.669	[-0.0077,	0.727]	0.735	[-0.0171,	0.774]	0.791
	CI	[-0.0782,	0.696]		[-0.0675,	0.763]		[-0.0693,	0.805]	
	bias	+0.039	-0.037		+0.037	0.002		+0.001	-0.014	
0.05	p.e.	[-0.0029,	0.663]	0.666	[-0.0042,	0.727]	0.731	[-0.0136,	0.774]	0.787
	CI	[-0.0739,	0.696]		[-0.0631,	0.763]		[-0.0647,	0.805]	
	bias	+0.029	-0.028		+0.030	0.001		+0.006	-0.010	

ATE(3,1), **MTS + MTR, MIV**

$Q_{J,K}$		0.00		0.01		0.05	
		LB UB	width	LB UB	width	LB UB	width
0.00	p.e.	[**0.0072, 0.657**]	0.650	[**0.0072, 0.675**]	0.668	[**0.0072, 0.756**]	0.749
	CI	[-0.0585, 0.688]		[-0.0453, 0.706]		[-0.0337, 0.786]	
	bias	+0.050 -0.030		+0.033 -0.005		+0.013 -0.013	
0.01	p.e.	[-0.0043, 0.657]	0.661	[-0.0043, 0.675]	0.679	[-0.0043, 0.756]	0.760
	CI	[-0.0727, 0.688]		[-0.0599, 0.706]		[-0.0494, 0.786]	
	bias	+0.042 -0.027		+0.033 -0.005		+0.000 -0.017	
0.05	p.e.	[-0.0010, 0.657]	0.658	[-0.0010, 0.675]	0.676	[-0.0010, 0.756]	0.757
	CI	[-0.0689, 0.688]		[-0.0560, 0.706]		[-0.0455, 0.786]	
	bias	+0.029 -0.027		+0.022 -0.006		+0.002 -0.017	