Causal Effects of Multiple Food Assistance Program Participation on Child Food Insecurity

Helen H. Jensen, *Iowa State University*
Brent Kreider, *Iowa State University*
Oleksandr Zhylyevskyy, *Iowa State University*

Southern Economic Association Annual Meeting

Washington, DC
November 20, 2016

Motivation

Food insecurity harms child's physical, intellectual, social development and health (Gundersen et al., 2011)

Prevalence of food insecurity in low-income population is high. Among households with children and income below 130% poverty (Coleman-Jensen et al., 2016):

- **29%** had low food security
- **12%** had very low food security

} **3.2M** food-insecure households

- Also, **20%** (**1.5M** households) had food-insecure children

Existing programs (e.g., SNAP, NSLP, WIC) aim at reducing food insecurity. Most papers focus on only one program. Few study **multiple** programs (see Schmidt et al., 2016)

Many assistance recipients participate in multiple programs. How do various programs interact in creating a **food safety net**?

Methodological Challenge

Identification of **causal** effect is difficult even for a single program:

- **Nonrandom selection**: unobservables simultaneously affect food security and program participation
  - Simple regression methods produce **inconsistent** estimates of causal effects
- **Nonclassical measurement error**: households systematically underreport benefits, misreporting varies across households with different attributes
  - Standard IV methods produce **inconsistent** estimates as well

Allowing for **multiple** programs adds another layer of complexity:

- Participation can no longer be modeled using a binary variable
- Dimensionality of measurement error problem increases

Our approach and methodological contribution:

- Introduce a **partially-ordered** multiple treatment variable to model participation
- Extend partial identification methods of Kreider & Hill (2009), Kreider et al. (2012), which account for selection and measurement error in a single framework

Research Focus and Relevance

We develop methodology to study **two** programs

In application, we focus on:

- **SNAP**: Supplemental Nutrition Assistance Program (food stamps)
- **NSLP**: National School Lunch Program (school lunches)

Both are large programs. In 2015 (Oliveira, 2016):

- 46M people participated in SNAP on average per month
- 22M children received free/reduced-price school lunches on average per day
- Annual federal expenditures on SNAP: $74B, NSLP: $13B

Receipt of benefits is underreported in surveys (Meyer et al., 2015):

- 40% of SNAP benefits are not reported in CPS; 45% underreporting for NSLP

Our goal is to account for selection and misreporting and quantify:

- To what extent participation in SNAP+NSLP improves food security compared to no program participation
- To what extent participation in both augments effect of either program alone

Data Sources

Main source: ***Food Security Supplement*** of **CPS**

FSS is administered in December; we pool years 2002–2010

FSS/CPS provides info on food security, food program participation, food expenditures, socioeconomic characteristics

Analytical sample: households with **school-age** children and income below **130%** of federal poverty line, ***N*** **= 10,390**

Additional sources (data on IVs and MIVs):

- Quarterly Food-at-Home Price Database (**QFAHPD**) provides **prices** for 50+ food groups across 35 geographic areas
- **SNAP Policy Database** provides state-level info on **policies** regarding eligibility, reporting requirements, use of biometrics

Child Food Security Measurement

FSS has 18-item **Household Food Security Survey Module**

8 items are child specific; responses by adult proxy

Examples of questions (referenced to past month):

- *Did any of the children ever skip a meal because there wasn't enough money for food?* Yes/No
- *Did any of the children ever not eat for a whole day because there wasn't enough money for food?* Yes/No

Responses are scored (0-1) and summed. Summary score is used to construct categories of child food security:

food secure (88.3%)
- (1) high: score = 0
- (2) marginal: score = 1

food insecure (11.7%)
- (3) low: score = 2, 3, or 4
- (4) very low: score = 5, 6, 7, or 8 (1.5%)

Reported Program Participation

Weighted sample distribution by program participation, *N* = 10,390:

		SNAP	
		yes	*no*
NSLP	*yes*	34.9%	35.6%
	no	5.0%	24.6%

- Reference period for food assistance program participation: **past month**
- Sample: households with 1+ school-age child, income below 130% poverty

Motivation for Our Methodology

A simple parametric approach:

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

Outcome | Treatment | Covariates | 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, it allows for heterogeneous response to treatment across i

Basics of Our Approach

$\boldsymbol{S^*}$: **true** program participation; $S^* = 0$: none, $S^* = 1$: SNAP alone, $S^* = 2$: NSLP alone, $S^* = 3$: SNAP & NSLP; S^* is *partially ordered*

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

Potential outcomes framework:

$\boldsymbol{Y(S^*)}$: potential outcome under treatment S^*; $Y = 1$ if FS, 0 otherwise

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

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

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

Say, ATE_{31} measures how likelihood of FS would change if household were to participate in SNAP & NSLP vs. 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 every formula into what is identified and what isn't

Let's simplify notation: $ATE_{31} = P[Y(3)=1] - P[Y(1)=1]$

Consider decomposition:

$$P[Y(3)=1] = \underbrace{P[Y(3)=1 \mid S^*=3]}_{\text{identified}}\underbrace{P(S^*=3)}_{\text{identified}} + \underbrace{P[Y(3)=1 \mid S^* \neq 3]}_{\text{not identified}}\underbrace{P(S^* \neq 3)}_{\text{identified}}$$

Data cannot identify $P[Y(3)=1 \mid S^* \neq 3]$ because it refers to unobserved **counterfactual**. We only know $P[Y(3)=1 \mid S^* \neq 3] \in [0,1]$

However, using methods of Manski (1995), we can still find 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 on ATEs

Without assumptions, bounds on ATEs are wide and **contain zero**

To **tighten** them, we can:

- Exploit logical constraints on probabilities and auxiliary data to restrict θ's. Say:

$$\theta_1^{-1,1} \leq \min\{P(Y=1, S\neq 1), P(S^*=1)\}$$

- Restrict prevalence of misreporting. Say, constrain value of $\Delta_j \equiv P_j^* - P_j$

- Restrict pattern of misreporting. Say, impose "**No-stigma verification**" assumption: Household with $S\neq 0$ is presumed to provide accurate participation response for both SNAP and NSLP. This assumption zeroes out several θ's

- Restrict selection process by imposing **exogenous selection**, monotone treatment selection (**MTS**), monotone treatment response (**MTR**), monotone instrumental variables (**MIVs**), instrumental variables (**IVs**)

By layering progressively stronger assumptions we demonstrate how they shape inference

Example of Analytical Results

Proposition 2(ii)(B):

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

- **Lower bound**:

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

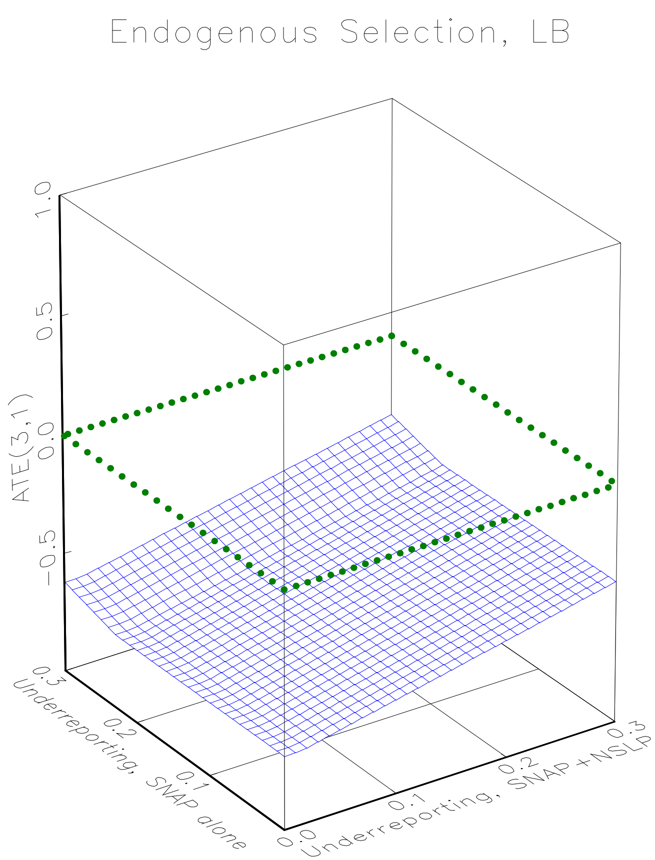
- **Upper bound**:

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

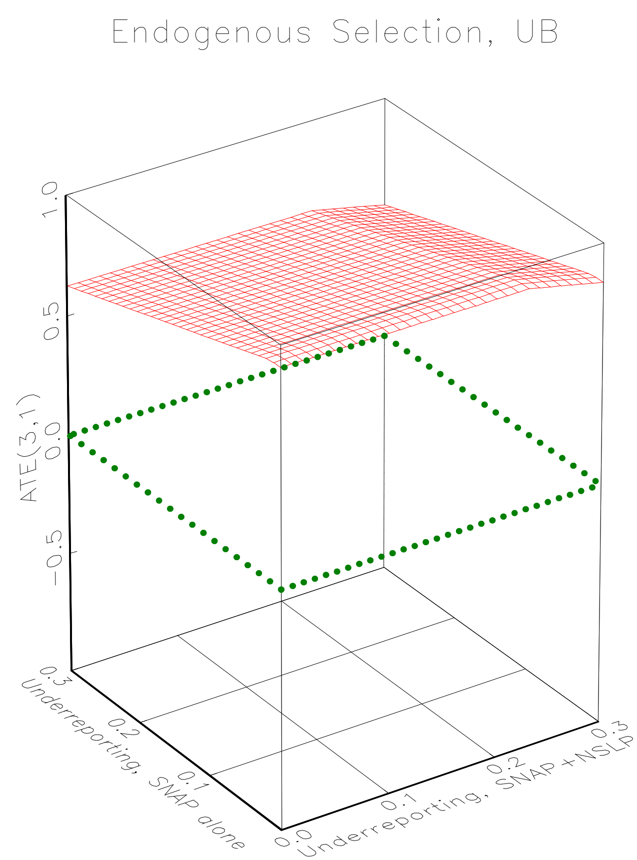
$$\Delta_1 \equiv P_1^* - P_1,\ \Delta_3 \equiv P_3^* - P_3,\ P_{000} \equiv P(Y=0, S=0, V=0),$$
$$P_{100} \equiv P(Y=1, S=0, V=0)$$

Results: Worst-Case Bounds

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



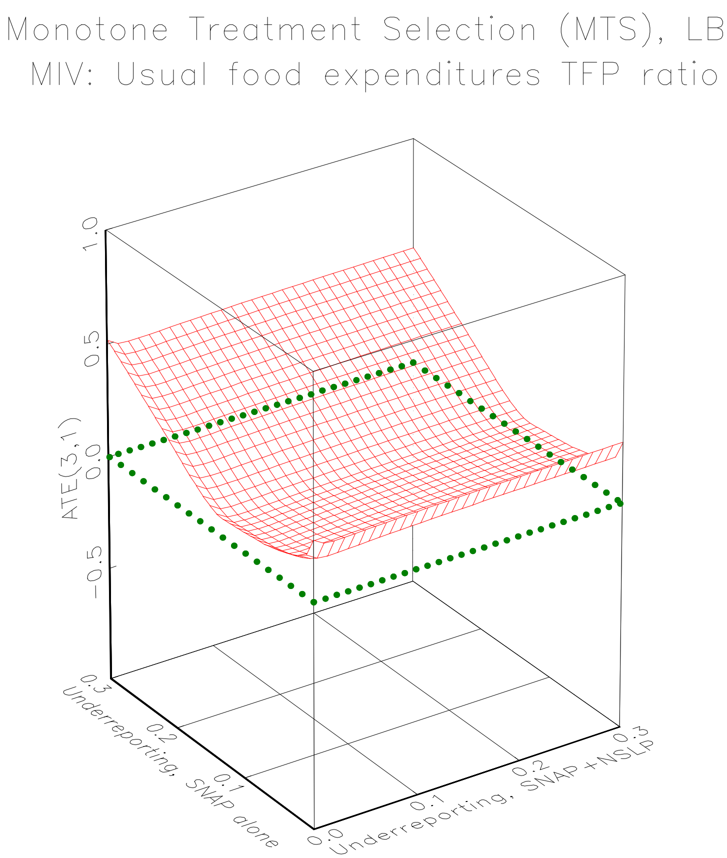
Lower bound



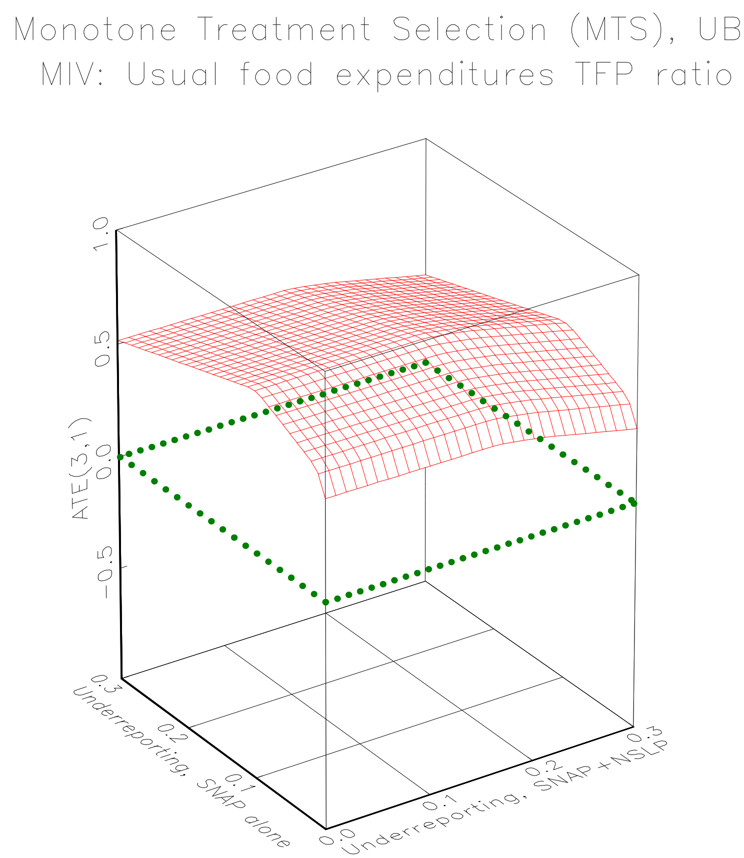
Upper bound

Results: Bounds under MTS + MIV

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



Lower bound



Upper bound

Summary

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

Research objective: Quantify by how much SNAP+NSLP improves child food security relative to SNAP alone or NSLP alone and relative to nonparticipation

Data: Large national sample drawn from FSS/CPS

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

Selected result: Bounding under MTS and MIV indicates SNAP+NSLP improves child food security on top of effect of SNAP alone $(ATE_{3,1} > 0)$

Thank you!

Appendix

More on 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 status is assigned based on a survey module with questions on food-related behaviors under lack of resources:

- Example: "*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, other surveys may 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

Prevalence of Child Food Security

Unweighted prevalence by food program participation:

NSLP \ SNAP	*yes*	*no*
yes	0.8634	0.8777
no	0.8691	0.9374

Weighted prevalence by food program participation:

NSLP \ SNAP	*yes*	*no*
yes	0.8626	0.8661
no	0.8700	0.9386

- Reference period: **past month**. All variables are as reported
- Each cell shows fraction of households with given condition in subsample

Prevalence of No Very Low Child FS

Unweighted prevalence by food program participation:

NSLP \ SNAP	yes	no
yes	0.9783	0.9881
no	0.9863	0.9919

Weighted prevalence by food program participation:

NSLP \ SNAP	yes	no
yes	0.9782	0.9877
no	0.9870	0.9921

- Reference period: **past month**. All variables are as reported
- Each cell shows fraction of households with given condition in subsample

Reported Program Participation (II)

Unweighted sample distribution by program participation, N = 10,390:

		SNAP	
		yes	*no*
NSLP	*yes*	34.5%	35.5%
	no	4.9%	25.0%

- Reference period for food assistance program participation: **past month**
- Sample: households with 1+ school-age child, income below 130% poverty line

QFAHPD, SNAP Policy Database: Details

QFAHPD is based on Nielsen Homescan: food purchase transactions by a large panel of households. ERS aggregated data within/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 regarding program eligibility, reporting requirements, use of biometric technology, etc.

- Coverage: every state and DC, every month between 1996 and 2014
- 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

Restricting Selection Process

Exogenous selection: expected potential outcomes do not depend on realized treatment:

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

Monotone instrumental variable (**MIV**):

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

Monotone treatment selection (**MTS**) is a special case of MIV: decision to participate is monotonically related to food insecurity:

$$P[Y(j)=1 \mid S^*=3] \le P[Y(j)=1 \mid S^*=k] \le P[Y(j)=1 \mid S^*=0] \quad \forall\, j;\, k=1,2$$

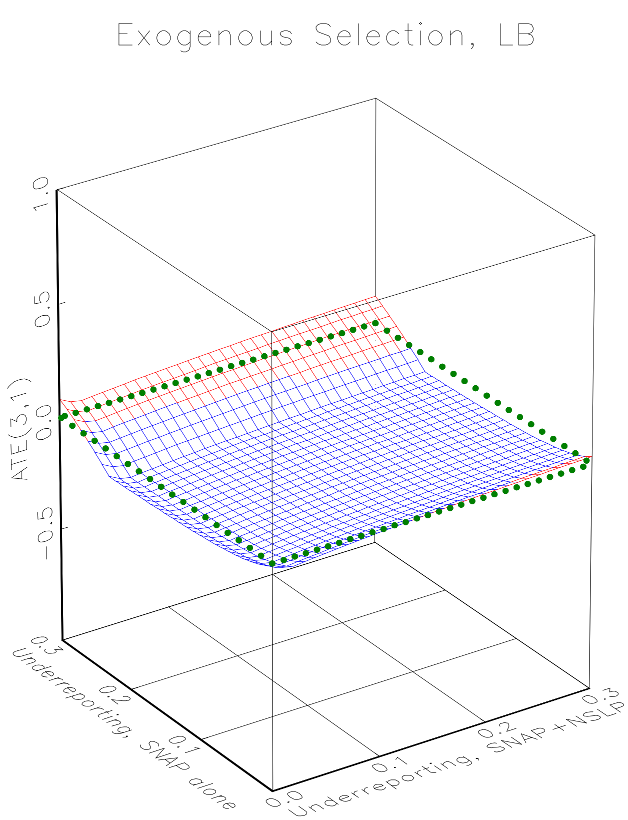
Monotone treatment response (**MTR**): potential participation in food programs would not harm food security:

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

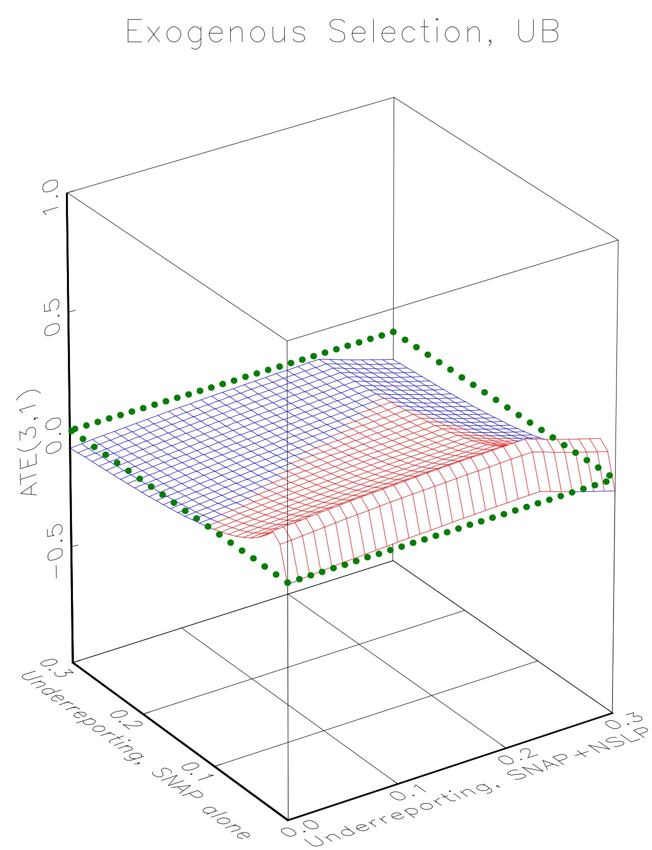
$$P[Y(3)=1 \mid S^*] \ge P[Y(2)=1 \mid S^*] \ge P[Y(0)=1 \mid S^*]$$

Results: Exogenous Selection (I)

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



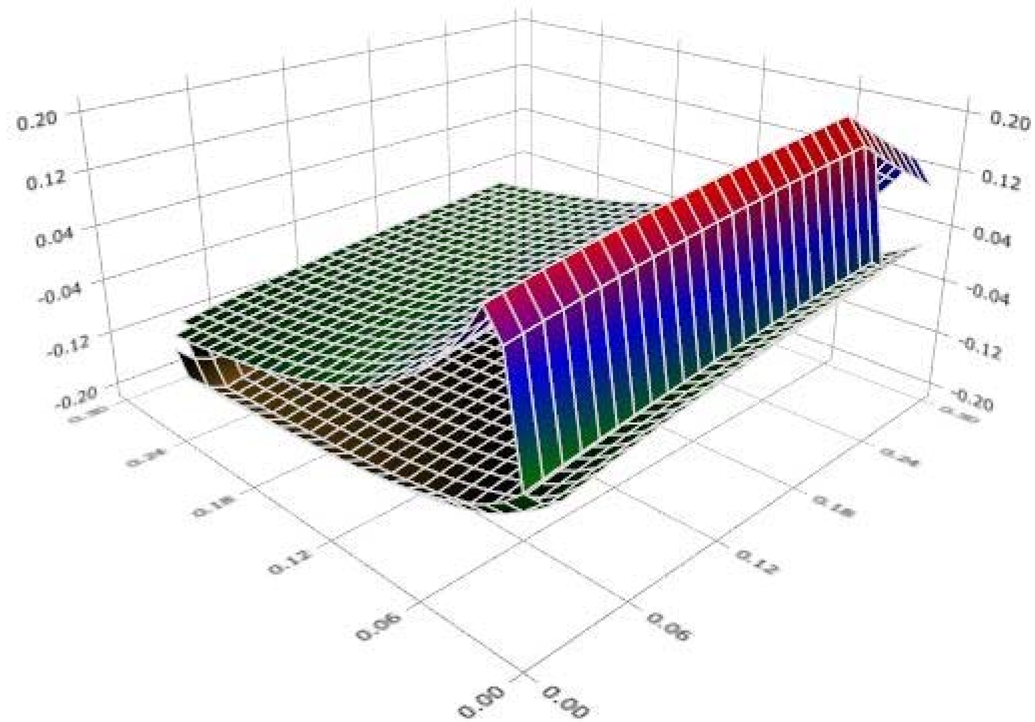
Lower bound



Upper bound

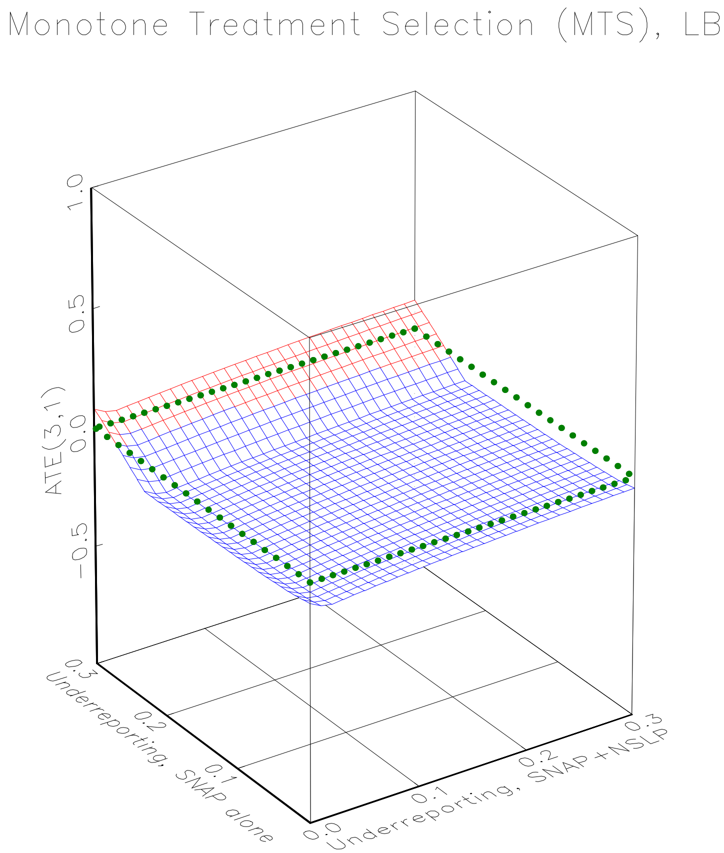
Results: Exogenous Selection (II)

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

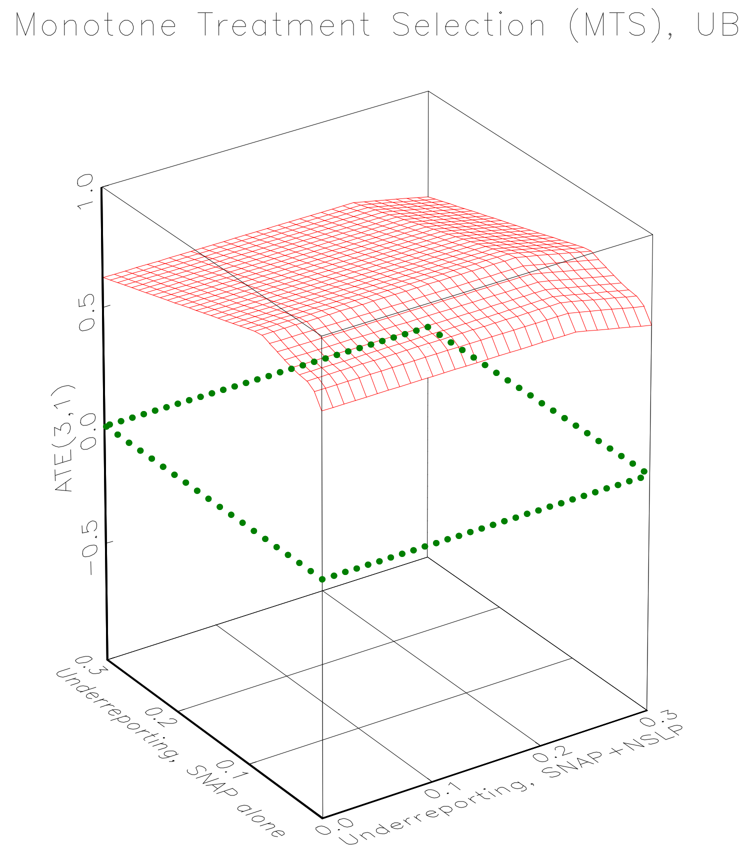


Results: MTS

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



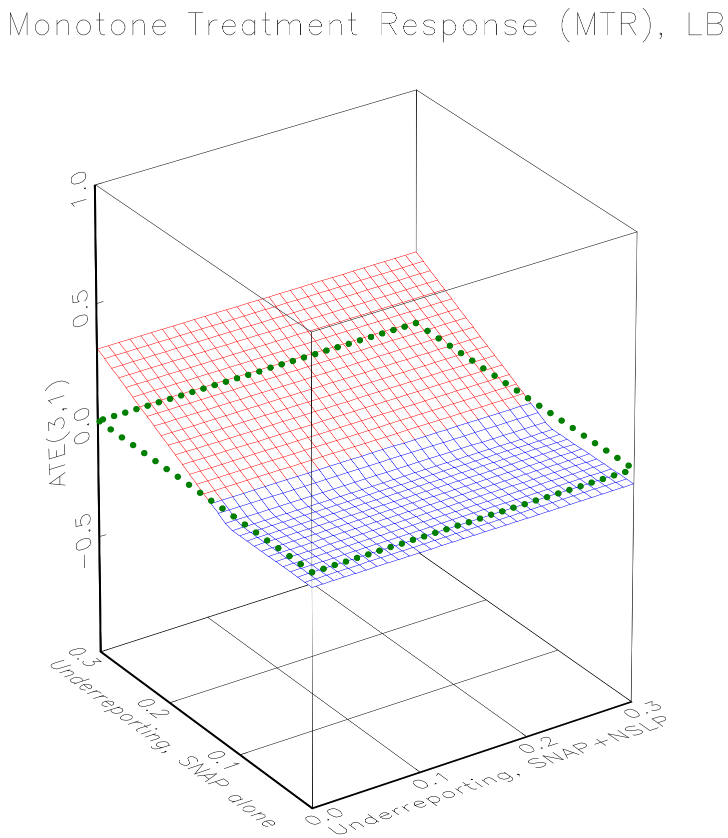
Lower bound



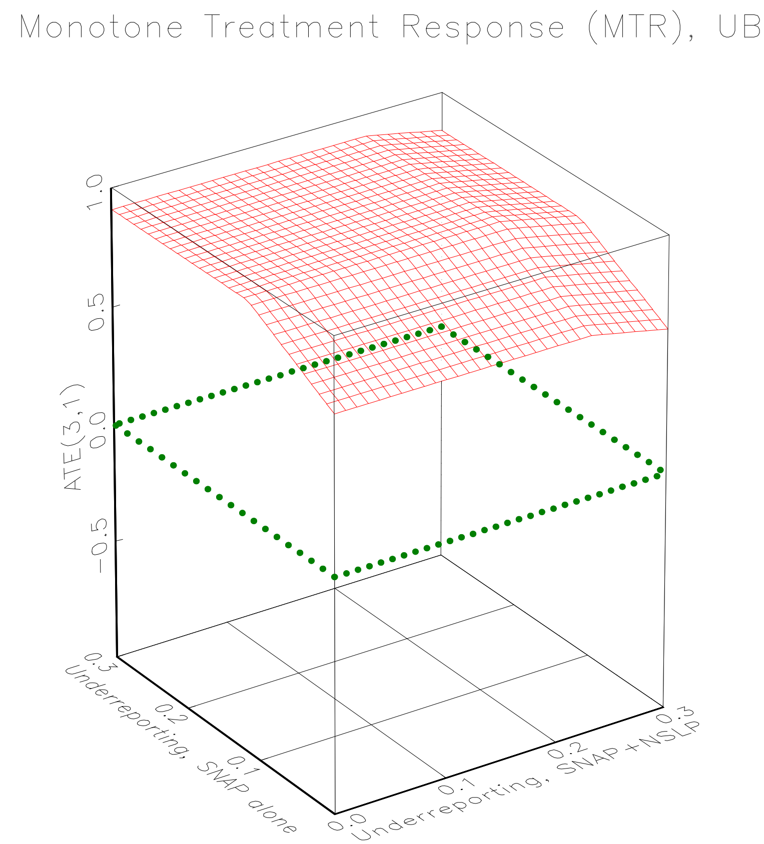
Upper bound

Results: MTR

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



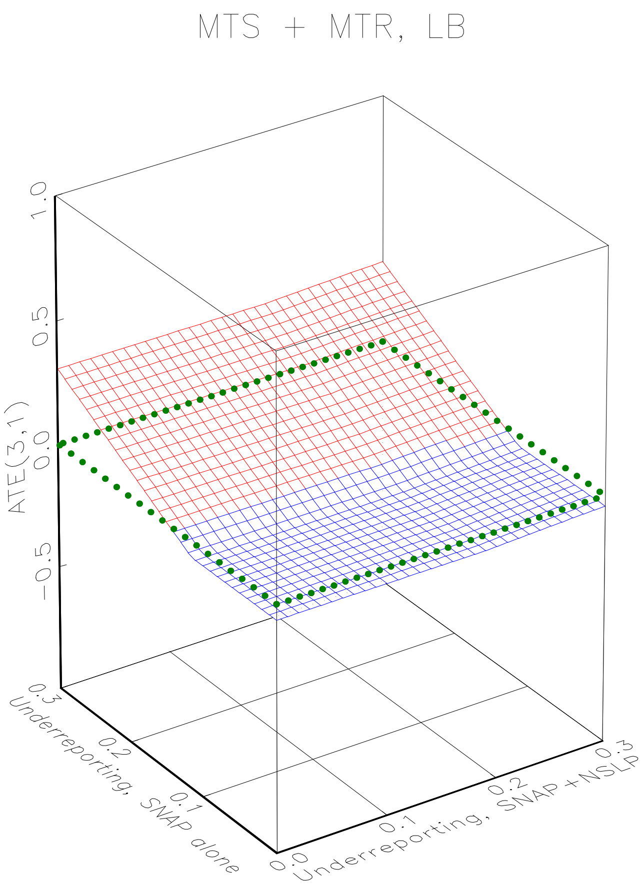
Lower bound



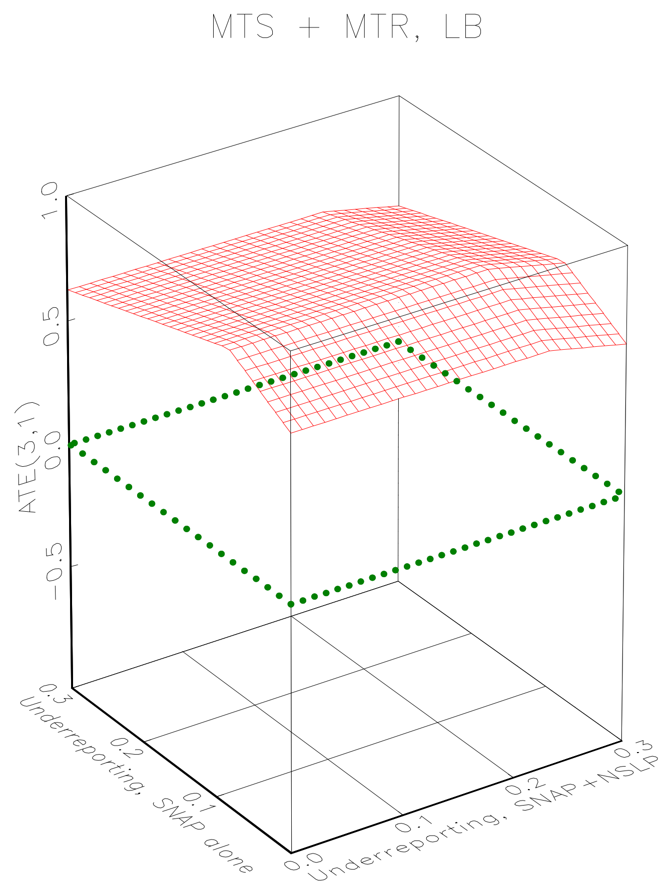
Upper bound

Results: MTS + MTR

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



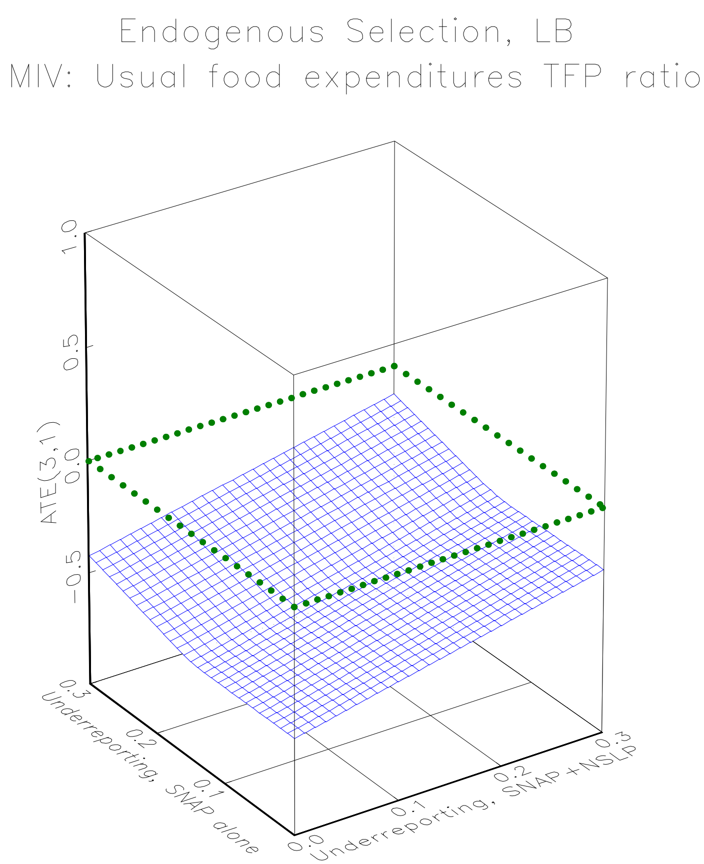
Lower bound



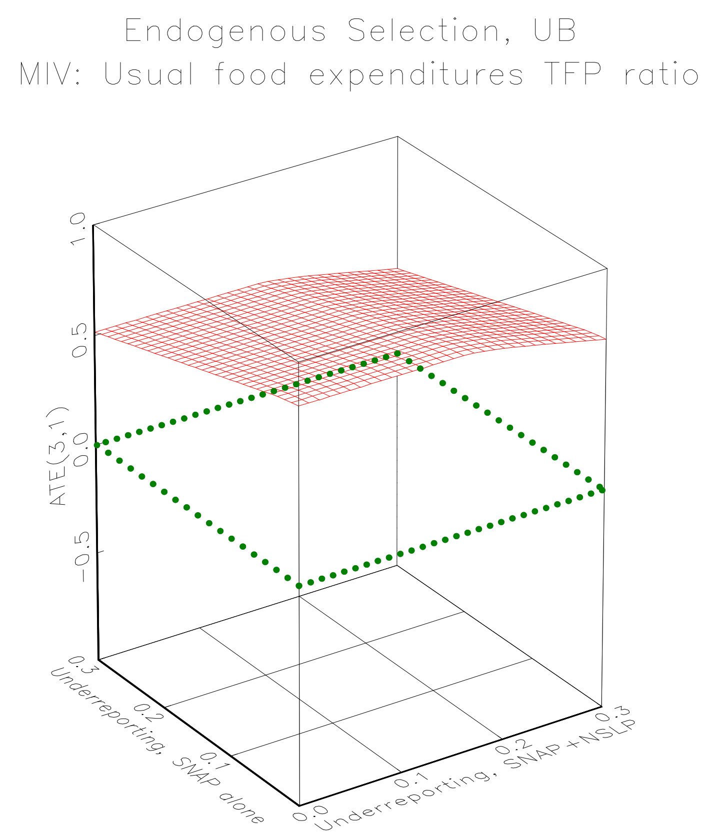
Upper bound

Results: MIV

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



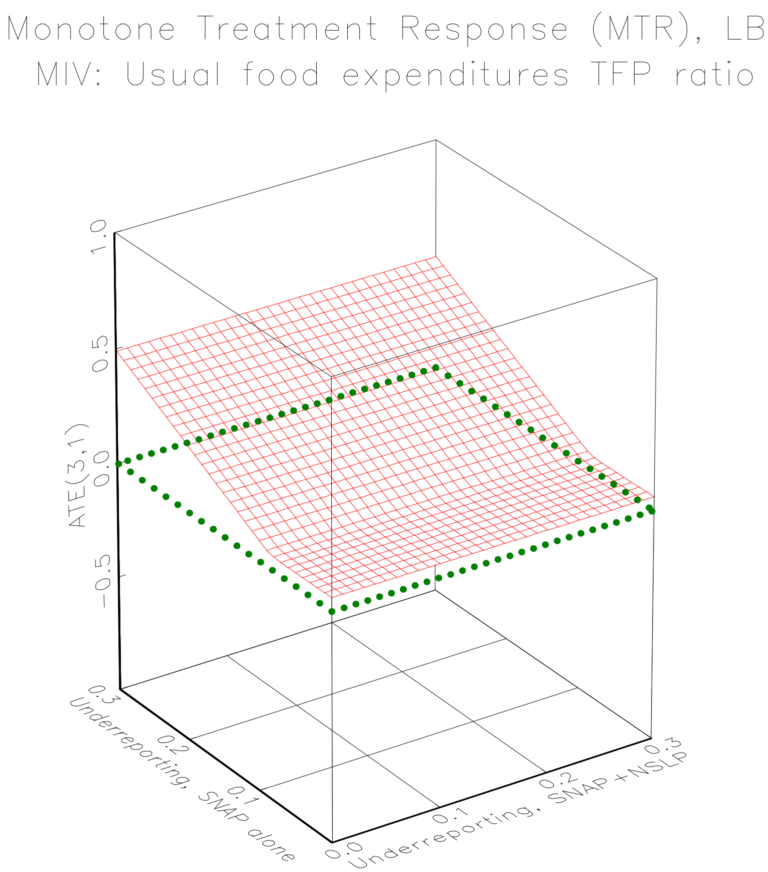
Lower bound



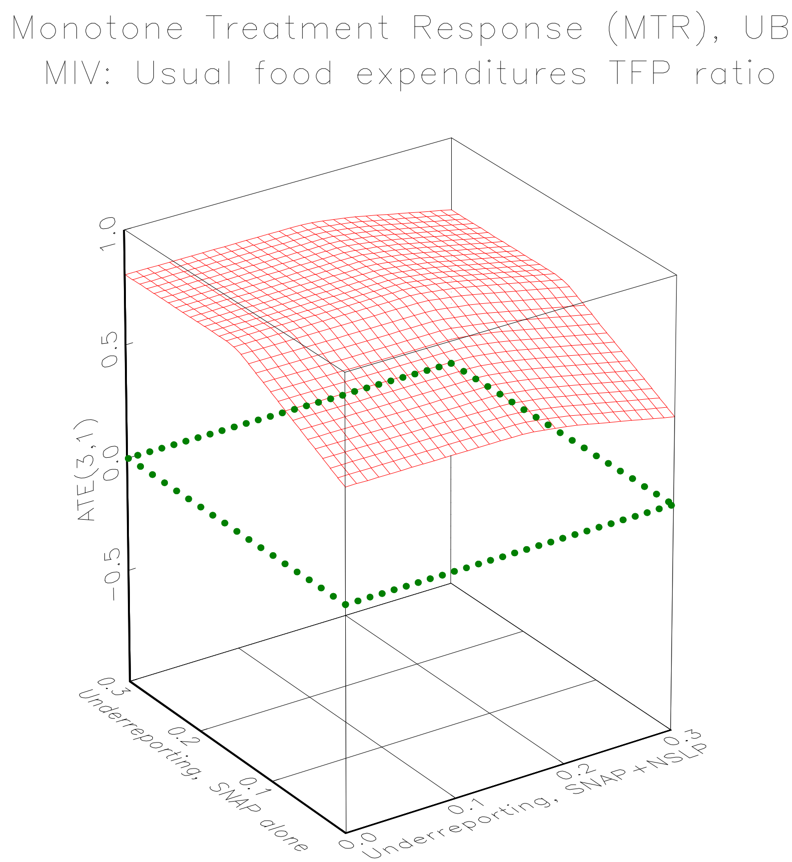
Upper bound

Results: MTR + MIV

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



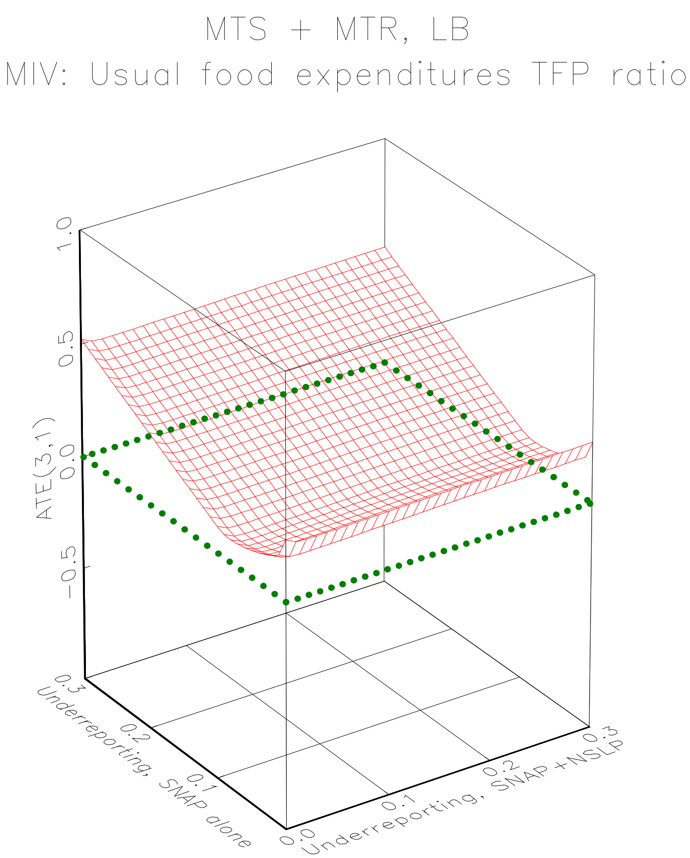
Lower bound



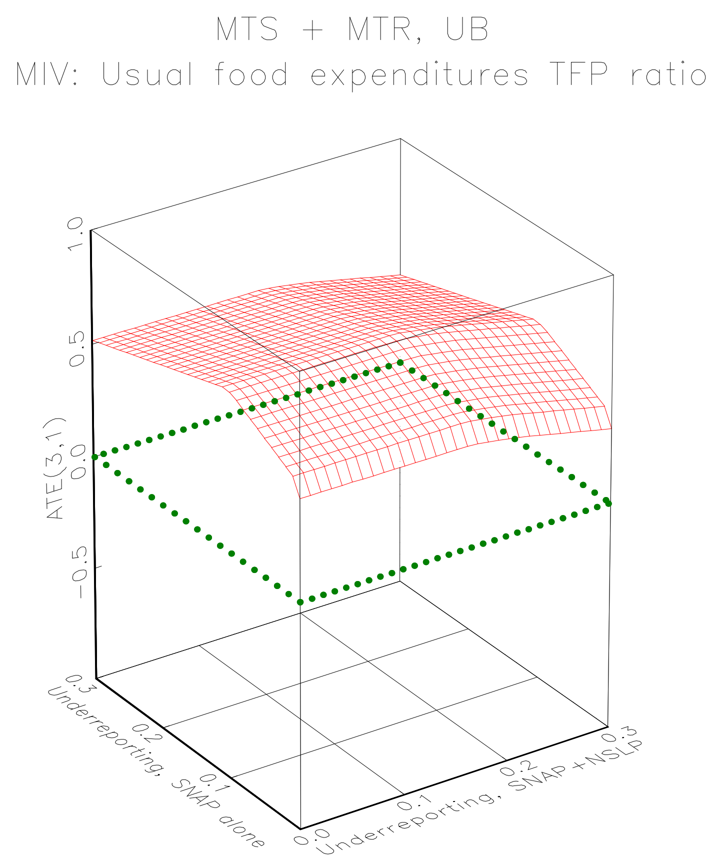
Upper bound

Results: MTS + MTR + MIV

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



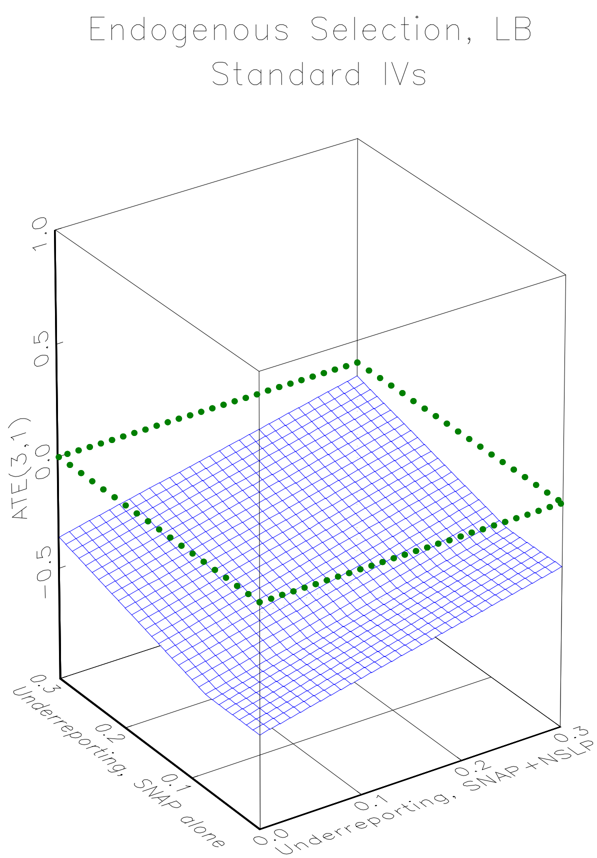
Lower bound



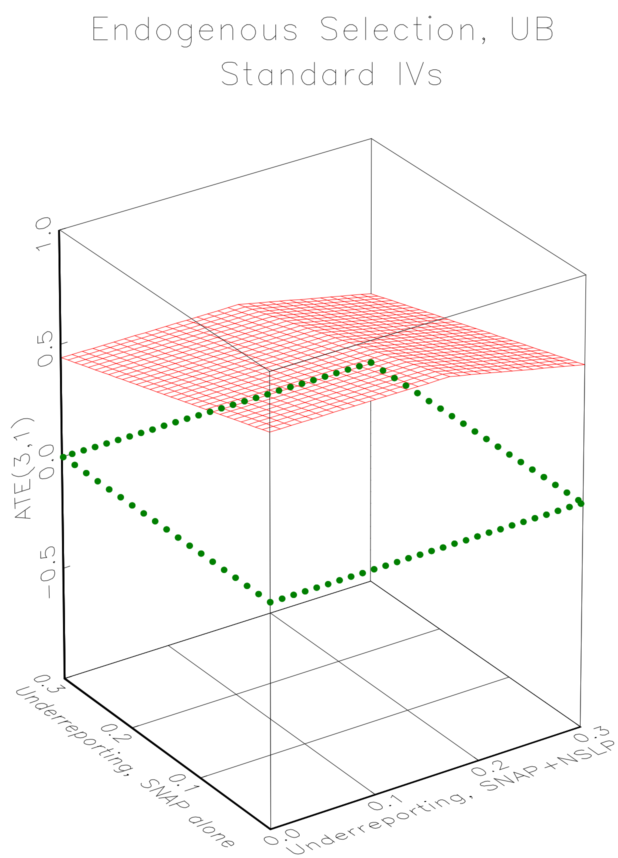
Upper bound

Results: Instrumental Variable (IV)

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



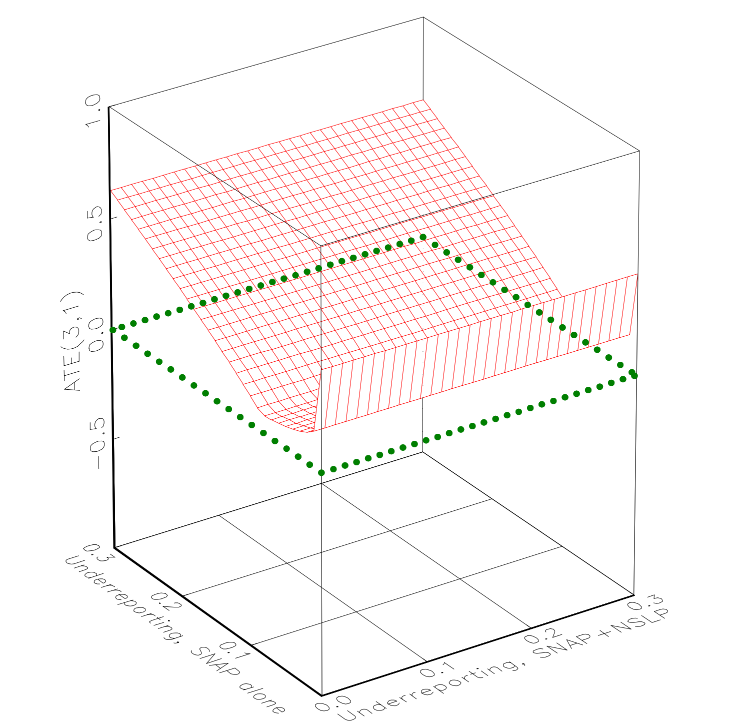
Lower bound



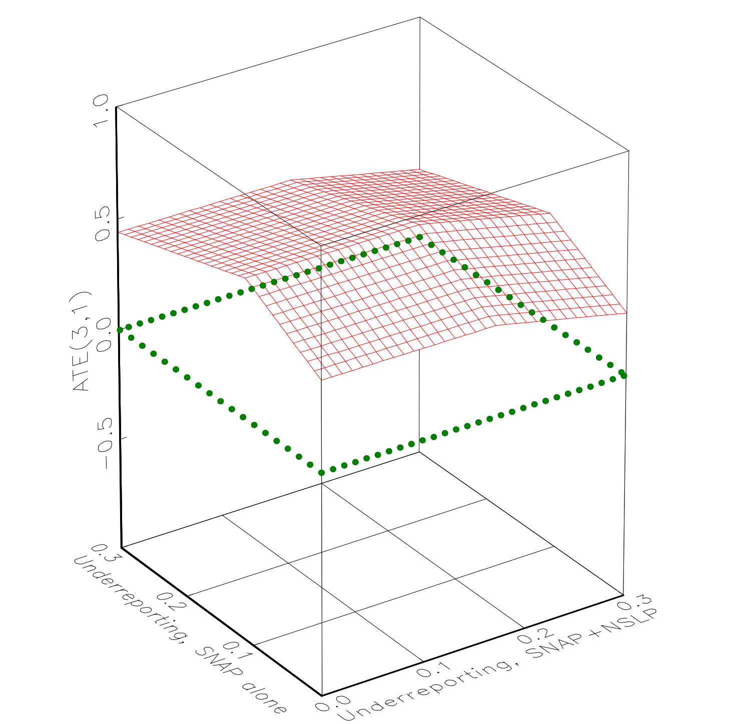
Upper bound

Results: MTS + IV

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



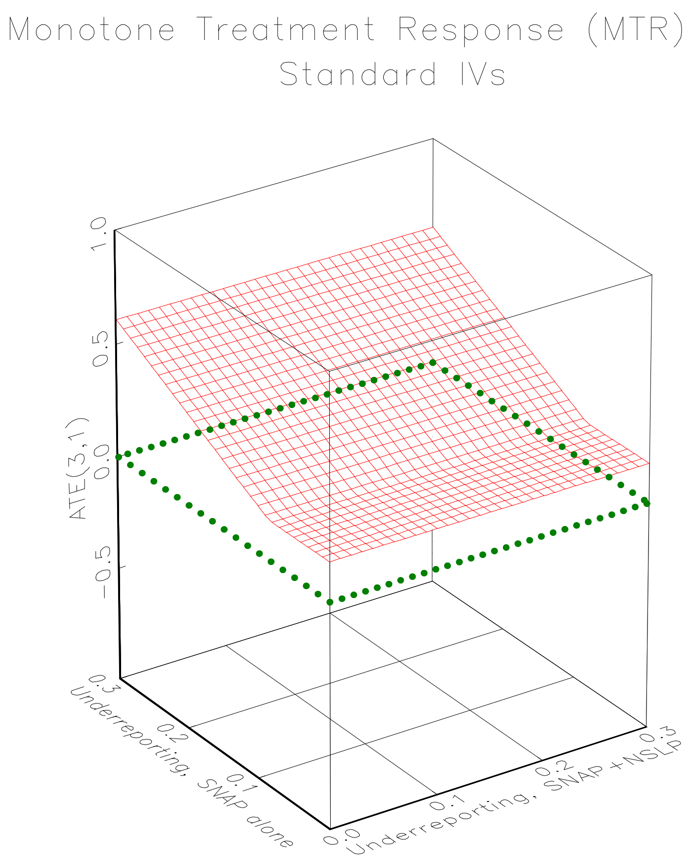
Lower bound



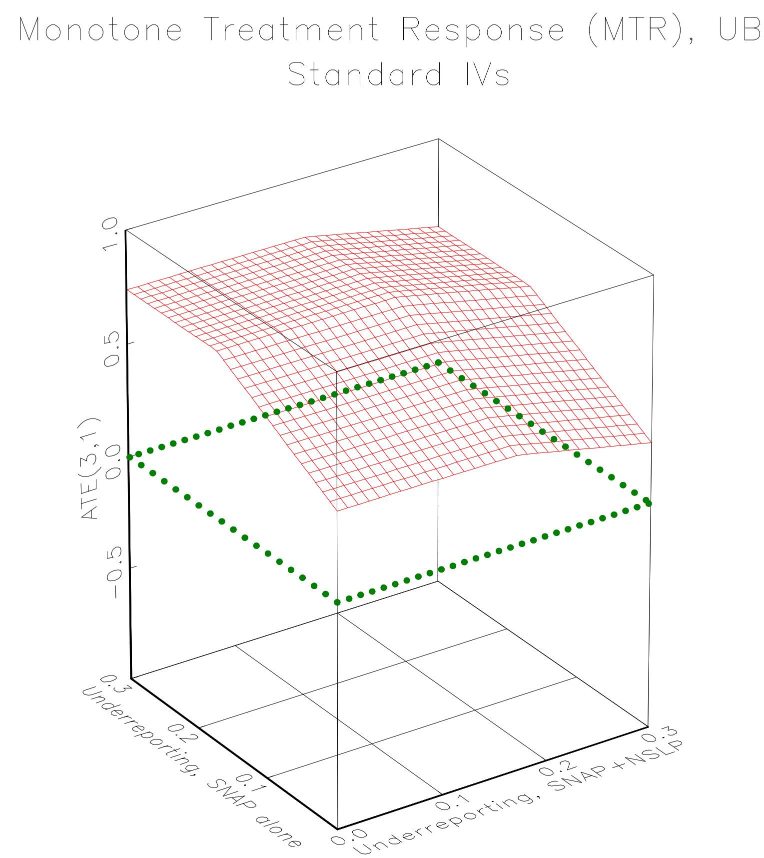
Upper bound

Results: MTR + IV

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



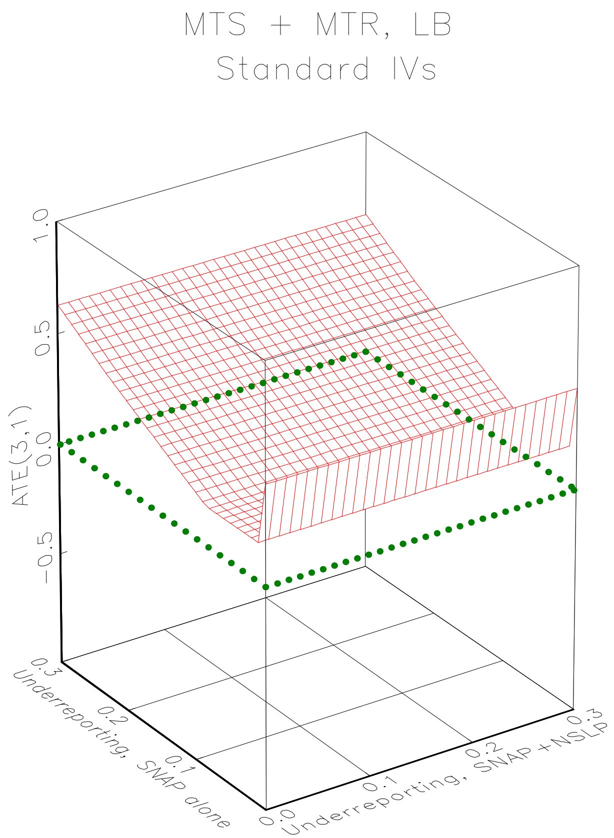
Lower bound



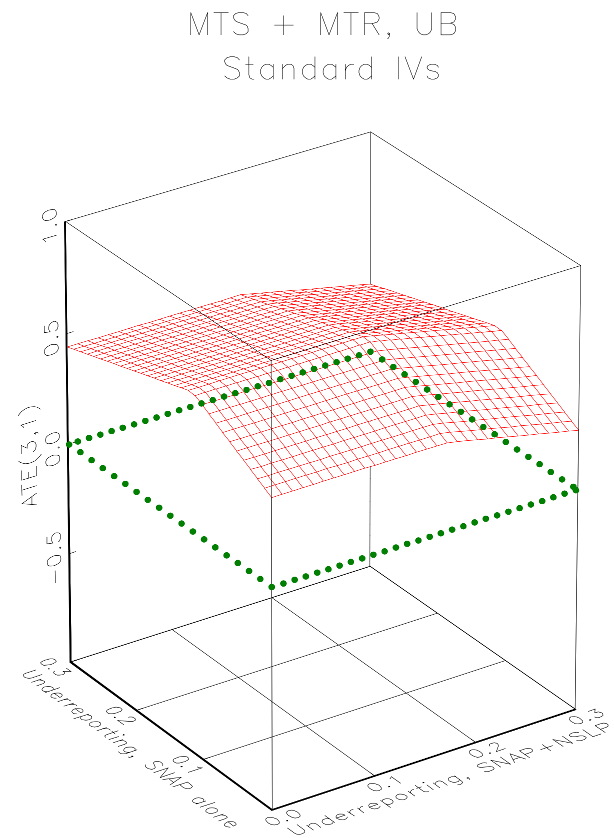
Upper bound

Results: MTS + MTR + IV

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



Lower bound



Upper bound

Abbreviations

ATE: average treatment effect

CPS: Current Population Survey

ERS: Economic Research Service of USDA

FSM: food security module

FSS: Food Security Supplement of CPS

MIV: monotone instrumental variable

MTR: monotone treatment response

MTS: monotone treatment selection

NSLP: National School Lunch Program

QFAHPD: Quarterly Food-at-Home Price Database

SNAP: Supplemental Nutrition Assistance Program

TFP: Thrifty Food Plan