Causal Effects of Mental Health on Food Security: A Partial Identification Analysis

Helen Jensen
Iowa State University

Brent Kreider
Iowa State University

John Pepper
University of Virginia

Oleksandr Zhylyevskyy
Iowa State University

Kimberly Greder
Iowa State University

MVEA

October 7, 2022

Motivation

Mental health conditions affect many individuals:

- 21% of U.S. adults suffer from mental illness

Studies find **correlation** between mental health and food security:

- Depressive symptoms and food insecurity
- Mother's mental health problems and food insecurity

No known research on **causal** effect of mental health on food security that **simultaneously** accounts for:

- **Misreporting** of true mental health status
- **Endogeneity** of mental health

Preview of Main Findings

We focus on **nonspecific psychological distress** (**NPD**) in adults who are 18-64 years old and below 130% of poverty line

Self-reported rate of moderate to severe NPD is **23.5%** (in 2011–2014)

NPD is **underreported**: we develop an approach to approximate the **true** rate of moderate to severe NPD and find it to be **30%**

Only **68%** of families of these adults are food secure

Under mild assumptions, alleviating NPD would increase food security rate by at least **14 p.p.**, i.e., by at least **24%**

Main Data Source

National Health Interview Survey (**NHIS**):

- Principal source of information on health of U.S. civilian population
- Cross-sectional, nationally representative, 80% response rate
- Annual sample of 35,000 households containing 87,500 individuals

Core components of NHIS questionnaire:

- Household: basic demographics, geocodes (restricted access)
- **Family**: demographics, **food security**, program participation, health status, injuries, healthcare utilization, health insurance
- **Sample adult** (one randomly selected adult per family): **psychological distress**, selected **mental health problems**, other aspects of health status, health care services, health behaviors
- Sample child (one randomly selected child per family): health status, health care services, health behaviors

NHIS also provides imputed income and covers additional topics

Analytical Sample

We pool linked sample adult–family records, NHIS 2011–14:

- Sample adult is aged **18–64** years (working age)
- Every family member is a U.S. citizen
- Family income ≤ **130% of poverty line** (gross income cutoff for SNAP)
- ***N* = 21,520**
- *Note 1:* NHIS started administering food security survey module in 2011
- *Note 2:* In 98% of cases, "household" is identical to "family"

Selected characteristics:

Variable	Mean (weighted)
Adult's **age** (years)	37.1
Adult is **male** (indicator)	0.436
Income-to-poverty ratio (%)	68.9
SNAP participation (indicator)	0.485

Food Security Indicators

NHIS includes a **10-item** food security survey module:

- Referenced to **last 30 days**
- Includes family- and adult-specific questions (no child questions)

We create two indicators of family **food security status**:

1) **Food secure**: 1 if raw score ≤ 2 (i.e., "high" or "marginal" food security)
2) **Not very low food secure**: 1 if raw score ≤ 5 (i.e., absence of "very low" food security)

Descriptive statistics:

Indicator	% (weighted)
Food secure family	67.7
Not very low food secure family	83.1

Indicators of Psychological Distress

NHIS administers six questions underlying **Kessler (K-6) nonspecific psychological distress (NPD) scale**:

- How frequently in past 30 days one felt sad, nervous, restless, hopeless, that everything was an effort, worthless (5-point Likert scale for answers)
- K-6 is a standardized and validated measure of NPD (CDC, 2013)

We follow McMorrow et al. (2016) and create two NPD indicators:

1) Adult is in **moderate to severe distress**: 1 if K-6 scale ≥ 8
2) Adult is in **severe distress**: 1 if K-6 scale ≥ 13 (max is 24)

Descriptive statistics:

Indicator	% (weighted)
Moderate to severe NPD	23.5
Severe NPD	10.0

Additional mental health measures are in the appendix

Methodological Challenge

Identifying **causal effect** of NPD is difficult:

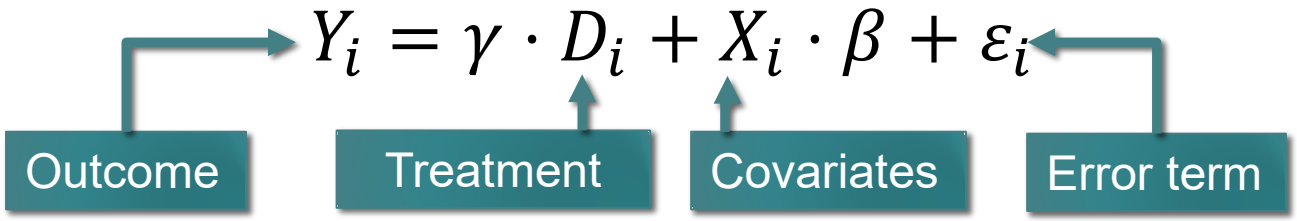
- **Endogeneity**: same unobservables simultaneously affect food security and NPD
  - Simple regression methods produce **inconsistent** estimates of causal effects
- **Measurement error**: stigma leads to misreporting of mental health problems, survey instruments have flaws, etc.
  - Treatment variables (i.e., NPD indicators) are binary → error is **non-classical**
  - Standard IV methods produce **inconsistent** estimates

While psychological distress and food security are **negatively correlated**, quantifying the **causal effect** of NPD on food security presents a significant methodological challenge

We adopt the **partial identification methodology** of Kreider et al. (2012) to bound the causal effect in the presence of indicated endogeneity and measurement error problems

Motivation for Our Methodology

Simple parametric approach:



Treatment D_i is **binary**: $D_i = 1$ if adult i is in distress, 0 if not

If same unobservables affect D_i and Y_i, then $cov(D_i, \varepsilon_i) \neq 0$ and OLS is inconsistent due to **endogeneity**

Measurement error in D_i is **non-classical** → standard IV estimation is inconsistent as well

The **partial identification** (**nonparametric bounding**) methodology handles endogeneity and misreporting in a unified framework and produces **a range of values** to which the causal effect belongs

Basics of Our Approach

Define: D^* = 1 if adult is truly in distress, = 0 otherwise; D is self-reported measure of D^*

We quantify **average treatment effect** (**ATE**) of psychological distress on food security:

$$ATE(1,0 \mid S,X) = P[Y(D^* = 1) = 1 \mid S,X] - P[Y(D^* = 0) = 1 \mid S,X]$$

Y = 1: family is food secure, Y = 0: insecure

$Y(D^* = 1)$ indicates **potential** food security outcome if adult were to be in distress. $Y(D^* = 0)$ denotes potential outcome if adult were not to be in distress

S, X specify subpopulation of interest (e.g., specific SNAP status and income level)

Not a regression framework: covariates are not regressors, no regression error term here, no orthogonality conditions to satisfy

Decomposition Strategy

ATE cannot be point-identified without assumptions **even if** $D^* = D$

We decompose every formula into what is identified and what is not

Simplify notation: $ATE = P[Y(1)=1] - P[Y(0)=1]$

Consider a decomposition under no misreporting of distress:

$$P[Y(1)=1\ ] = P[Y(1)=1 \mid D^*=1]P(D^*=1) + P[Y(1)=1 \mid D^*=0]P(D^*=0).$$

identified **identified** **not identified** **identified**

Data cannot identify $P[Y(1)=1 \mid D^*=0]$ because it refers to unobserved **counterfactual**. We only know it must lie within the interval [0,1]

However, using methods of Manski (1995), we can still find worst-case bounds for $P[Y(1)=1]$, $P[Y(0)=1]$, and ATE

Addressing Misreporting

$$P[Y(1)=1] = P(Y=1, D^*=1) + P[Y(1)=1 \mid D^*=0]P(D^*=0)$$

$$= P(Y=1, D=1) + \theta_1^- - \theta_1^+ + P[Y(1)=1 \mid D^*=0]P(D^*=0)$$

where $\theta_1^- \equiv P(Y=1, D=0, Z^*=0)$ $\quad \in [0,1]$

$\theta_1^+ \equiv P(Y=1, D=1, Z^*=0)$

$Z^* = 0$ indicates that self-reported NPD status is incorrect

→ Sharp **bounds** on ATE:

$$P(Y=1, D=1) - P(Y=1, D=0) - P^* + 2(\theta_1^- - \theta_1^+)$$

$$\leq ATE \leq$$

$$P(Y=1, D=1) - P(Y=1, D=0) + (1-P^*) + 2(\theta_1^- - \theta_1^+)$$

Note: $P^* \equiv P(D^*=1)$

Tightening Bounds

Without assumptions, ATE bounds are **wide** and **contain zero**

To tighten them, we can:

- Use logical constraints on probabilities and auxiliary data to restrict θ's
- Apply "**no false positives**" assumption $\rightarrow \theta_1^+ = \theta_0^+ = 0$
- Impose (mild) restrictions on potential food security outcomes:
  - Monotone treatment selection (**MTS**)
  - Monotone instrumental variable (**MIV**)
  - Monotone treatment response (**MTR**)

More on these monotonicity assumptions in the appendix

Approximating P^* (True Rate of NPD)

True prevalence rate of NPD, P^*, plays an important role in bounding *ATE* (an appendix figure provides an illustration)

If NPD is misreported, $P^* \neq P = 0.235$. Unfortunately, there are no estimates of P^* available in the literature

We develop **an approximation approach**:

$$P^* \approx \frac{\tilde{P}^*}{\tilde{P}} \cdot P = \frac{0.268}{0.207} \cdot 0.235 = 0.304.$$

$\tilde{P}^*$ and $\tilde{P}$ are true and self-reported rates of a related mental health measure, namely, an indicator of **any mental illness**

$\tilde{P}^*$ comes from SAMHSA's official tables

$\tilde{P}$ is calculated by us using NHIS 2012 data

ATE Bounds Under Few Assumptions

		Self-reported prevalence rate: $P^* = P = 0.235$			True prevalence rate: $P^* = 0.304$		
Endogenous selection		LB	UB	width	LB	UB	width
(a) Arbitrary errors	p.e.†	[-0.912,	0.558]	1.469	[-0.981,	0.627]	1.608
	CI‡	[-0.919	0.567]		[-0.989	0.636]	
(b) No false positives	p.e.	[-0.710,	0.290]	1.000	[-0.779,	0.359]	1.138
	CI	[-0.716	0.296]		[-0.786	0.365]	

† Point estimates of the population bounds.

‡ Imbens-Manski 95% confidence intervals around the true ATE.

- These bounds are **too wide** to determine the sign of ATE

ATE Bounds Under MTS+MIV+MTR

		Self-reported prevalence rate: $P^* = P = 0.235$			True prevalence rate: $P^* = 0.304$		
MTS + Food Density MIV + MTR:		LB	UB	width	LB	UB	width
(a) Arbitrary errors	p.e.	**[-0.852,**	**-0.142]**	0.710	**[-0.861,**	**-0.142]**	0.719
	CI	[-0.894	-0.054]		[-0.965	-0.054]	
(b) No false positives	p.e.	[**-0.224,**	**-0.142**]	0.083	[**-0.422,**	**-0.142**]	0.280
	CI	[-0.340	-0.054]		[-0.448	-0.054]	

Strictly negative average treatment effects in **bold.**

- Alleviating NPD would increase food security rate by at least 14 p.p.

Thank you!

Appendix

Indicators of Mental Health Problems

NHIS asks sample adults about degree of **difficulty** with 12 daily activities (e.g., walking) and what health problem caused this

NHIS also asks whether adults are **limited** in performing 7 activities (e.g., personal care) and what health problem caused this

We create indicators for existence of:

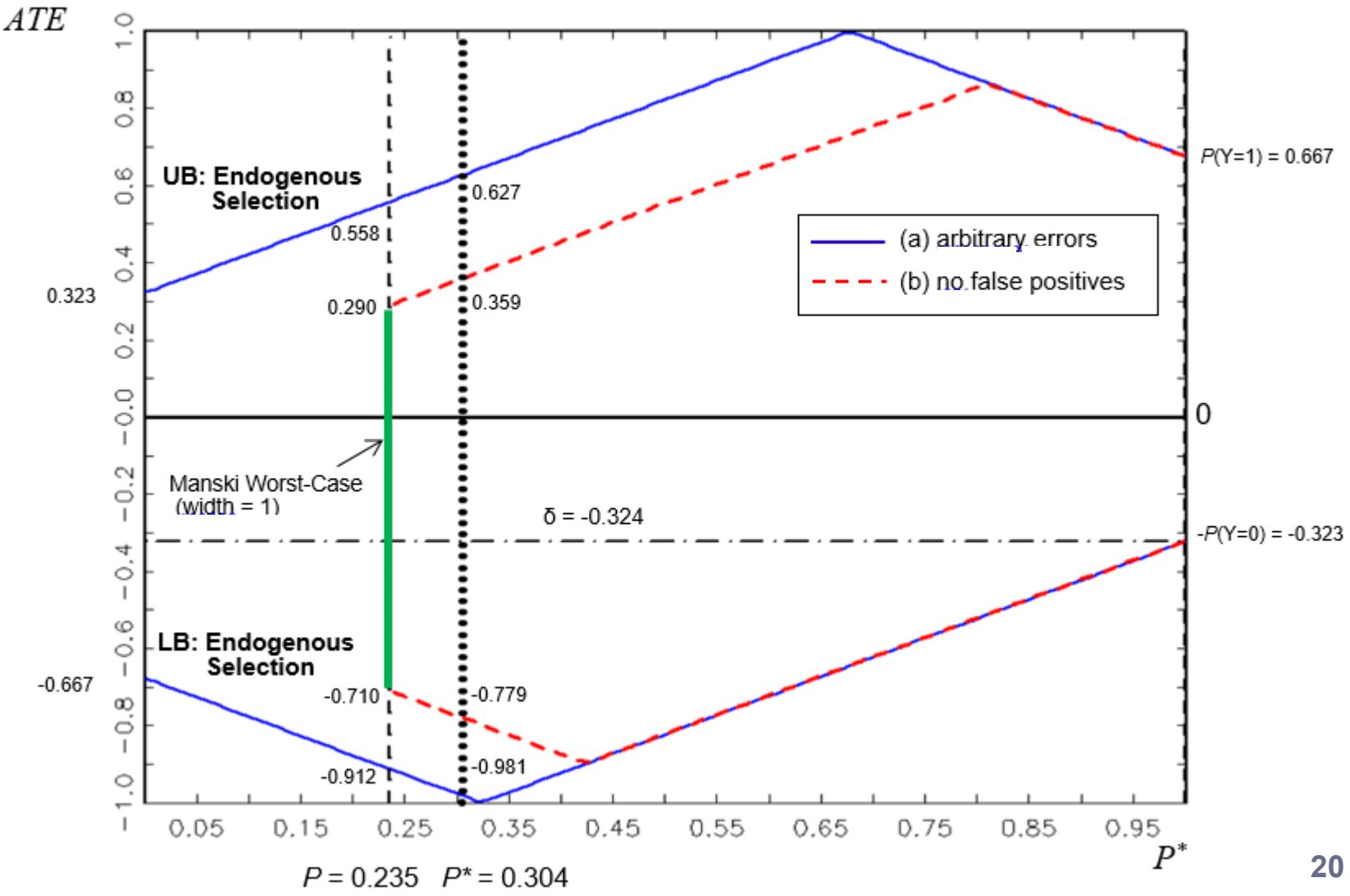
1) **Mental health problem causing difficulty** with activities
2) **Mental health problem causing limitation** in activities

'Problem' includes depression, anxiety, ADD, bipolar, schizophrenia, etc.

Selected descriptive statistics:

Indicator	% (weighted)
Adult has mental health problem causing **difficulty**	6.9
Adult has mental health problem causing **limitation**	8.3

Bounds Under Endogeneity of NPD



Monotonicity Assumptions

Monotone treatment selection (**MTS**):

$$P[Y(j)=1 \mid D^{*}=1] \leq P[Y(j)=1 \mid D^{*}=0],\ j=1,0.$$

Monotone instrumental variable (**MIV**):

Let *v* be food store density. Higher *v* would not harm food security:

$$u_1 \leq u \leq u_2 \Rightarrow$$

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

Monotone treatment response (**MTR**):

Psychological distress would not improve food security on average:

$$P[Y(1)=1 \mid D^{*}] \leq P[Y(0)=1 \mid D^{*}].$$