Behavior within a Clinical Trial and Implications for Mammography Guidelines

Amanda E. Kowalski
Gail Wilensky Professor of Applied Economics and Public Policy
University of Michigan

March 2019
U.S. Preventive Services Task Force (USPSTF) 2016 Guidelines for Women in 40’s:

“The decision to start screening mammography in women prior to age 50 years should be an individual one. Women who place a higher value on the potential benefit than the potential harms may choose to begin biennial screening between the ages of 40 and 49 years”
USPSTF 2016 Guidelines Based on RCT’s

### Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation

Heidi D. Nelson, MD, MPH; Rochelle Fu, PhD; Amy Cantor, MD, MPH; Miranda Pappas, MA; Monica Daeges, BA; and Linda Humphrey, MD, MPH

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Trial Name</th>
<th>Mean Follow-up, y</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women aged 39–49 y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST II</td>
<td>11.2</td>
<td>0.64 (0.39–1.06)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Kopperberg</td>
<td>12.5</td>
<td>0.73 (0.37–1.41)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Östergötland</td>
<td>12.5</td>
<td>1.02 (0.52–1.99)</td>
</tr>
<tr>
<td>Moss et al, 2015 (27)</td>
<td>Age</td>
<td>17.5</td>
<td>0.93 (0.80–1.09)</td>
</tr>
<tr>
<td>Bjurstam et al, 2003 (25)</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.69 (0.45–1.05)</td>
</tr>
<tr>
<td>Habbema et al, 1986 (29)</td>
<td>HIP</td>
<td>14.0</td>
<td>0.75 (0.53–1.05)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>Stockholm</td>
<td>14.3</td>
<td>1.52 (0.80–2.88)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST I</td>
<td>18.2</td>
<td>0.74 (0.42–1.29)</td>
</tr>
<tr>
<td>Miller et al, 2014 (15)</td>
<td>CNBSS-1</td>
<td>21.9</td>
<td>1.04 (0.87–1.24)</td>
</tr>
<tr>
<td>Overall (I² = 25%; P = 0.230)</td>
<td></td>
<td></td>
<td>0.92 (0.75–1.02)</td>
</tr>
</tbody>
</table>

*Random-effects model*
CNBSS Consistent with Meta-analysis of RCT’s

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Trial Name</th>
<th>Mean Follow-up, y</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST II</td>
<td>11.2</td>
<td>0.64 (0.39–1.06)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Kopperberg</td>
<td>12.5</td>
<td>0.73 (0.37–1.41)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Östergötland</td>
<td>12.5</td>
<td>1.02 (0.52–1.99)</td>
</tr>
<tr>
<td>Moss et al, 2015 (27)</td>
<td>Age</td>
<td>17.5</td>
<td>0.93 (0.80–1.09)</td>
</tr>
<tr>
<td>Bjurstem et al, 2003 (25)</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.69 (0.45–1.05)</td>
</tr>
<tr>
<td>Habbema et al, 1986 (29)</td>
<td>HIP</td>
<td>14.0</td>
<td>0.75 (0.53–1.05)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>Stockholm</td>
<td>14.3</td>
<td>1.52 (0.80–2.88)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST I</td>
<td>18.2</td>
<td>0.74 (0.42–1.29)</td>
</tr>
<tr>
<td><strong>Miller et al, 2014 (15)</strong></td>
<td>CNBSS-1</td>
<td><strong>21.9</strong></td>
<td><strong>1.04 (0.87–1.24)</strong></td>
</tr>
<tr>
<td>Overall ($I^2 = 25%; P = 0.230$)</td>
<td></td>
<td></td>
<td><strong>0.92 (0.75–1.02)</strong></td>
</tr>
</tbody>
</table>
Meta-analysis for All Cause Mortality

**Evidence Synthesis**

Number 124

Screening for Breast Cancer: A Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study</th>
<th>Mean followup, year</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>39 to 49 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tabár et al., 1989</td>
<td>Kopparberg</td>
<td>7.9</td>
<td>1.33 (1.01 to 1.77)</td>
</tr>
<tr>
<td>Tabár et al., 1989</td>
<td>Östergötland</td>
<td>7.9</td>
<td>0.93 (0.76 to 1.12)</td>
</tr>
<tr>
<td>Bjurstam et al., 1997</td>
<td>Gothenburg</td>
<td>10.0</td>
<td>0.98 (0.86 to 1.12)</td>
</tr>
<tr>
<td>Frisell et al., 1997</td>
<td>Stockholm</td>
<td>11.0</td>
<td>1.12 (0.55 to 2.41)</td>
</tr>
<tr>
<td><strong>Miller et al., 2002</strong></td>
<td><strong>CNBSS-1</strong></td>
<td>13.0</td>
<td>1.00 (0.87 to 1.15)</td>
</tr>
<tr>
<td>Nyström et al., 2002</td>
<td>Malmö II</td>
<td>9.1</td>
<td>1.03 (0.89 to 1.20)</td>
</tr>
<tr>
<td>Moss et al., 2006</td>
<td>Age</td>
<td>10.7</td>
<td>0.97 (0.89 to 1.04)</td>
</tr>
<tr>
<td><strong>Subtotal (I²=0.0%, p=0.478)</strong></td>
<td></td>
<td></td>
<td>0.99 (0.94 to 1.06)</td>
</tr>
</tbody>
</table>

Favors Screening Group

Favors Control Group
I Examine Micro Data from the CNBSS

• Canadian National Breast Screening Study
  – 89,835 patients enrolled
  – Patients received mammograms for 4 to 5 years during active study period
  – Recorded mammogram receipt, even in control group
  – Followed patient outcomes from 1980 to 2005 through cancer registry and death records (no attrition)
  – Collected risk factors and demographic data
I Examine Behavior within a Clinical Trial

- Medical literature examines clinical trials
  - See Nelson (2016)
- Economics literature examines mammogram takeup
  - Strumpf, Chai, and Kadiyala (2010)
  - Kadiyala and Strumpf (2011, 2016)
  - Zanella and Banerjee (2016)
  - Buchmueller and Goldzahl (2018)
  - Kim and Lee (2017)
I Examine Behavior within a Clinical Trial

• I show two relationships between biology and behavior in existing clinical trial data by building on LATE and MTE literatures from economics
  – Bjorklund and Moffitt (1987)
  – Imbens and Angrist (1994)
  – Vytlacil (2002)
  – Brinch, Mogstad, Wiswall (2015)
I Examine Behavior within a Clinical Trial


“Extrapolation Using Selection and Moral Hazard Heterogeneity from Within the Oregon Health Insurance Experiment.” *NBER WP 24647.*

“How to Examine External Validity Within an Experiment” *NBER WP 24834.*

“Behavior within a Clinical Trial and Implications for Mammography Guidelines.” *NBER WP 25049.*
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Mortality

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be harmed by them

• Robustness
• Conclusions
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]

Assumptions:

A.1. (Continuity) \( F(\cdot) \): absolutely continuous with respect to the Lebesgue measure
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

Assumptions:

A.1. (Continuity) \( F(\cdot) \): absolutely continuous with respect to the Lebesgue measure

**Proof:** \( U_D \sim U[0, 1] \)

\[
F_{U_D}(u) = P(U_D \leq u) \\
= P(F(\nu_D) \leq u) \\
= P(\nu_D \leq F^{-1}(u)) \\
= F(F^{-1}(u)) = u
\]

\( (F(\cdot) \) absolutely continuous by A.1)
First Stage:

\[ V = V_U + (V_T - V_U) D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]

\[ U_D = F(\nu_D), \ U_D \sim U[0,1] \]

Assumptions:

A.1. (Continuity) \( F(\cdot) \): absolutely continuous with respect to the Lebesgue measure

A.2. (Independence) \((U_D, \gamma_T)\) and \((U_D, \gamma_U) \perp Z\)
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

Assumptions:

A.1. (Continuity) \( F(\cdot) \): absolutely continuous with respect to the Lebesgue measure

A.2. (Independence) \((U_D, \gamma_T)\) and \((U_D, \gamma_U) \perp Z\)

Proof: \( D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \)

\[ D = 1\{0 \leq V_T - V_U\} \]
\[ = 1\{0 \leq \mu_D(Z) - \nu_D\} \]
\[ = 1\{\nu_D \leq \mu_D(Z)\} \]
\[ = 1\{F(\nu_D) \leq F(\mu_D(Z))\} \quad \text{(definition of } F(\cdot) \text{ from A.1)} \]
\[ = 1\{U_D \leq F(\mu_D(Z))\} \quad \text{(} U_D = F(\nu_D) \text{ by definition)} \]
\[ = 1\{U_D \leq P(D = 1 \mid Z = z)\}, \]

where the last equality follows from

\[ F(\mu_D(Z)) = P(\nu_D \leq \mu_D(Z)) \]
\[ = P(\nu_D \leq \mu_D(z) \mid Z = z) \quad (\nu_D \perp Z \text{ by A.2}) \]
\[ = P(0 \leq \mu_D(Z) - \nu_D \mid Z = z) \]
\[ = P(0 \leq V_T - V_U \mid Z = z) \]
\[ = P(D = 1 \mid Z = z). \]
First Stage:

\[ V = V_U + (V_T - V_U) D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]

\[ U_D = F(\nu_D), \ U_D \sim U[0, 1] \]

Assumptions:

A.1. (Continuity) \( F(\cdot) \): absolutely continuous with respect to the Lebesgue measure

A.2. (Independence) \((U_D, \gamma_T)\) and \((U_D, \gamma_U) \perp Z\)

A.3. (Instrument Relevance) \( \mu_D(Z) \): nondegenerate random variable
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \quad \text{...} \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 | Z = z)\} \]
\[ Z = 0 : \quad D = 1\{U_D \leq p_C\}, \quad p_C = P(D = 1 | Z = 0) \]
\[ Z = 1 : \quad D = 1\{U_D \leq p_I\}, \quad p_I = P(D = 1 | Z = 1) \]

Assumptions:

A.1. (Continuity) \( F(\cdot) \): absolutely continuous with respect to the Lebesgue measure
A.2. (Independence) \((U_D, \gamma_T)\) and \((U_D, \gamma_U) \perp Z\)
A.3. (Instrument Relevance) \(\mu_D(Z)\): nondegenerate random variable
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]

\[ Z = 0:\quad D = 1\{U_D \leq p_C\}, \quad p_C = P(D = 1 \mid Z = 0) \]

\[ Z = 1:\quad D = 1\{U_D \leq p_I\}, \quad p_I = P(D = 1 \mid Z = 1) \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

\[ U_D: \text{ unobserved net cost of treatment} \]
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 | Z = z)\} \]

\[ Z = 0 : \quad D = 1\{U_D \leq p_c\}, \quad p_c = P(D = 1 | Z = 0) \]
\[ Z = 1 : \quad D = 1\{U_D \leq p_I\}, \quad p_I = P(D = 1 | Z = 1) \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

\[ Z = 0 \]
\[ D = 1 \]
\[ 0 \leq p \leq p_c \]

\[ p_c = 0.19 \]

Always Takers

\[ U_D: \ \text{unobserved net cost of treatment} \]
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]

\[ Z = 0 : \quad D = 1\{U_D \leq p_c\}, \quad p_c = P(D = 1 \mid Z = 0) \]
\[ Z = 1 : \quad D = 1\{U_D \leq p_I\}, \quad p_I = P(D = 1 \mid Z = 1) \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

\[ U_D: \text{unobserved net cost of treatment} \]
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1 \{ 0 \leq V_T - V_U \} \]
\[ \Rightarrow D = 1 \{ U_D \leq P(D = 1 \mid Z = z) \} \]

\[ Z = 0 : \quad D = 1 \{ U_D \leq p_c \}, \quad p_c = P(D = 1 \mid Z = 0) \]
\[ Z = 1 : \quad D = 1 \{ U_D \leq p_I \}, \quad p_I = P(D = 1 \mid Z = 1) \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 | Z = z)\} \]

\[ Z = 0: \quad D = 1\{U_D \leq p_C\}, \quad p_C = P(D = 1 | Z = 0) \]
\[ Z = 1: \quad D = 1\{U_D \leq p_I\}, \quad p_I = P(D = 1 | Z = 1) \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

\( D = 1 \) \( 0 \leq p \leq p_I \)
\( p_I < p \leq 1 \)

\( D = 0 \) \( p_C < p \leq 1 \)

\( 0.00 \quad p_C = 0.19 \) \( \text{Always Takers} \)
\( \) \( p_I = 0.95 \) \( 1.00 \) \( \text{Never Takers} \)

\( U_D \): unobserved net cost of treatment
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]

\[ Z = 0 : \quad D = 1\{U_D \leq p_C\}, \quad p_C = P(D = 1 \mid Z = 0) \]
\[ Z = 1 : \quad D = 1\{U_D \leq p_I\}, \quad p_I = P(D = 1 \mid Z = 1) \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

\[ Z = 1 \]
\[ D = 1 \]
\[ 0 \leq p \leq p_I \]
\[ D = 0 \]
\[ p_I < p \leq 1 \]

\[ Z = 0 \]
\[ D = 1 \]
\[ 0 \leq p \leq p_C \]
\[ D = 0 \]
\[ p_C < p \leq 1 \]

Always Takers \hspace{2cm} Compliers \hspace{2cm} Never Takers

\[ p_C = 0.19 \hspace{2cm} p_I = 0.95 \hspace{2cm} 1.00 \]

\[ U_D: \text{unobserved net cost of treatment} \]
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Mortality

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be harmed by them

• Robustness
• Conclusions
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 | Z = z)\} \]
\[ Z = 0: \quad D = 1\{U_D \leq p_C\}, \quad p_C = P(D = 1 | Z = 0) \]
\[ Z = 1: \quad D = 1\{U_D \leq p_I\}, \quad p_I = P(D = 1 | Z = 1) \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

Second Stage:

\[ Y = Y_U + (Y_T - Y_U)D \]
\[ Y_T = g_T(U_D, \gamma_T) \]
\[ Y_U = g_U(U_D, \gamma_U) \]
\[ Z \perp (\gamma_T, \gamma_U) \text{ by A.2.} \]

Assumptions (Second Stage):

A.4. (Treated and Untreated) \( 0 < P(D = 1) < 1 \)
A.5. (Finite Average Outcomes) \( E[Y_T], E[Y_U] \) are finite
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]

\[ \Rightarrow D = 1\{U_D \leq P(D = 1| Z = z)\} \]

\[ Z = 0: \quad D = 1\{U_D \leq p_C\}, \quad p_C = P(D = 1| Z = 0) \]
\[ Z = 1: \quad D = 1\{U_D \leq p_I\}, \quad p_I = P(D = 1| Z = 1) \]

\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

Second Stage:

\[ Y = Y_U + (Y_T - Y_U)D \]
\[ Y_T = g_T(U_D, \gamma_T) \]
\[ Y_U = g_U(U_D, \gamma_U) \]

\[ Z \perp (\gamma_T, \gamma_U) \text{ by A.2.} \]
All Cause Deaths 20 Years After Enrollment (per 10,000)

- Treated
- Untreated

$p_C = 0.19$
$p_I = 0.95$

Always Takers: $Z = 0$
Compliers: $U_D$: unobserved net cost of treatment
Never Takers: $D = 1$

$D = 0$

27 of 60
All Cause Deaths 20 Years After Enrollment (per 10,000)

\[
\frac{(1-0.19) \times 463 - (1-0.95) \times 990}{0.95-0.19} \\
\frac{(0.95 \times 422 - 0.19 \times 451)}{0.95-0.19}
\]

Always Takers | Compliers | Never Takers
---|---|---
\( Z = 0 \) | \( U_D : \) unobserved net cost of treatment | \( Z = 1 \)
All Cause Deaths 20 Years After Enrollment (per 10,000)

- Green dots: treated
- Blue dots: untreated

$p_C = 0.19$
$p_I = 0.95$

$U_D$: unobserved net cost of treatment

Always Takers
Compliers
Never Takers
All Cause Deaths 20 Years After Enrollment (per 10,000)

- treated
- untreated

$LATE = -13$

$p_C = 0.19$

$p_I = 0.95$

$U_D$: unobserved net cost of treatment
All Cause Deaths 20 Years After Enrollment (per 10,000)

- **treated**
- **untreated**

\[ U_D : \text{unobserved net cost of treatment} \]

\( p_C = 0.19 \)

\( p_I = 0.95 \)

\( LATE = -13 \)
Model
– First Stage: Mammography
– Second Stage: Mortality

Results
1. Selection Heterogeneity
   • Women more likely to receive mammograms are healthier

2. Treatment Effect Heterogeneity
   • Women more likely to receive mammograms are more likely to be harmed by them

• Robustness
• Conclusions
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]
\[ U_D = F(\nu_D), \quad U_D \sim U[0, 1] \]

Second Stage:

\[ Y = Y_U + (Y_T - Y_U)D \]
\[ Y_T = g_T(U_D, \gamma_T) \]
\[ Y_U = g_U(U_D, \gamma_U) \]
\[ Z \perp (\gamma_T, \gamma_U) \text{ by A.2.} \]

Untreated Outcome Test:

\[ E[Y_U \mid p_C < U_D \leq p_I] - E[Y_U \mid p_I < U_D \leq 1] = \int_0^1 (\omega(p, p_C, p_I) - \omega(p, p_I, 1)) \text{MUO}(p) \, dp \]

where \( \omega(p, p_L, p_H) = 1\{p_L \leq p < p_H\}/(p_H - p_L) \)

(Bertanha and Imbens (2014); Guo, Cheng, Lorch, and Small (2014); Black, Joo, LaLonde, Smith, and Taylor (2015); Mogstad, Santos, and Torgovitsky (2018).)
untreated outcome test: -562*** ≠ 0

LATE = -13

$U_D$: unobserved net cost of treatment
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Mortality

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be harmed by them

• Robustness
• Conclusions
First Stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T - V_U = \mu_D(Z) - \nu_D \]
\[ D = 1\{0 \leq V_T - V_U\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]

U_D = F(\nu_D), U_D \sim U[0, 1]

Second Stage:

\[ Y = Y_U + (Y_T - Y_U)D \]
\[ Y_T = g_T(U_D, \gamma_T) \]
\[ Y_U = g_U(U_D, \gamma_U) \]

Z \perp (\gamma_T, \gamma_U) by A.2.

Ancillary Assumption - Weak Monotonicity of the MUQ Function

For all \( p_1, p_2 \in [0, 1] \) such that \( p_1 < p_2 \):

\[ E[Y_U \mid U_D = p_1] \leq E[Y_U \mid U_D = p_2] \text{ or } E[Y_U \mid U_D = p_1] \geq E[Y_U \mid U_D = p_2] \]

(Brinch, Mogstad, and Wiswall (2017).)
test rejects treatment effect homogeneity: $1\{23 > -13\} = 1.00$

always take average treatment effect lower bound $= 23$

$U_D$: unobserved net cost of treatment
“I never, though, had a patient whose worry about those side effects came close to her worry about the disease. Being preoccupied with saving one’s life produces a myopia, in which other worries unrelated to one’s possibly imminent death fall away.”

“And so, unable to say whether any particular patient will benefit, we have no choice but to overtreat.”

-Mukherjee, New Yorker, September 11, 2017
## Baseline Covariates Support Ancillary Assumption

<table>
<thead>
<tr>
<th>Baseline Socioeconomic Status</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Always Takers</td>
</tr>
<tr>
<td>University, trade or business school</td>
<td>0.50</td>
</tr>
<tr>
<td>In work force</td>
<td>0.65</td>
</tr>
<tr>
<td>Age at first birth</td>
<td>24.28</td>
</tr>
<tr>
<td>No live birth</td>
<td>0.16</td>
</tr>
<tr>
<td>Married</td>
<td>0.80</td>
</tr>
<tr>
<td>Husband in work force / alive</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Baseline Health Behavior</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Always Takers</td>
</tr>
<tr>
<td>Non-Smoker</td>
<td>0.78</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>23.87</td>
</tr>
<tr>
<td>Used oral contraception</td>
<td>0.74</td>
</tr>
<tr>
<td>Used estrogen</td>
<td>0.13</td>
</tr>
<tr>
<td>Mammograms prior to enrollment</td>
<td>0.23</td>
</tr>
<tr>
<td>Practiced breast self examination</td>
<td>0.47</td>
</tr>
</tbody>
</table>
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Mortality

• Results
  1. Selection Heterogeneity
     • *Women more likely to receive mammograms are healthier*
  2. Treatment Effect Heterogeneity
     • *Women more likely to receive mammograms are more likely to be harmed by them*

• Robustness
• Conclusions
test rejects treatment effect homogeneity: $1\{23 > -13\} = 1.00$

always taker average treatment effect lower bound $= 23$

untreated outcome test: $-562^{***} \neq 0$

LATE $= -13$

$U_D$: unobserved net cost of treatment

$p_C = 0.19$

$p_I = 0.95$  1

Always Takers

Compliers

Never Takers
<table>
<thead>
<tr>
<th>Main Specification</th>
<th>N</th>
<th>Untreated Outcome Test</th>
<th>Always Taker Average Treatment Effect Lower Bound</th>
<th>LATE</th>
<th>Test Rejects Treatment Effect Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Specification</td>
<td>19,505</td>
<td>-562 (147)</td>
<td>23 (59)</td>
<td>-13 (38)</td>
<td>1.00 (0.48)</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• **Sample Restrictions**
  – Excluded participants aged 40-49
  – All participants aged 40-49
  – All participants aged 50-59

• **Definitions of mammography**
  – Narrower
  – Broader

• **Outcomes**
  – Breast cancer mortality
  – Mortality at alternative follow-up lengths
Robust to Alternative Sample Restrictions

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Untreated Outcome Test</th>
<th>Always Taker Average Treatment Effect Lower Bound</th>
<th>LATE</th>
<th>Test Rejects Treatment Effect Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Specification</strong></td>
<td>19,505</td>
<td>-562</td>
<td>23</td>
<td>-13</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(147)</td>
<td>(59)</td>
<td>(38)</td>
<td>(0.48)</td>
</tr>
<tr>
<td><strong>Alternative Sample Restrictions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All excluded participants aged 40-49 at enrollment</td>
<td>30,925</td>
<td>-759</td>
<td>60</td>
<td>27</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(135)</td>
<td>(39)</td>
<td>(40)</td>
<td>(0.47)</td>
</tr>
<tr>
<td>All participants aged 40-49 at enrollment</td>
<td>50,430</td>
<td>-672</td>
<td>53</td>
<td>9</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(103)</td>
<td>(31)</td>
<td>(27)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>All participants aged 50-59 at enrollment</td>
<td>39,405</td>
<td>-1,216</td>
<td>-83</td>
<td>15</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(154)</td>
<td>(51)</td>
<td>(46)</td>
<td>(0.26)</td>
</tr>
</tbody>
</table>
test rejects treatment effect homogeneity: $1\{-82 > 16\} = 0.00$

untreated outcome test: $-1217^{***} \neq 0$

always take average treatment effect lower bound $= -82$

$LATE = 16$

$U_D : \text{unobserved net cost of treatment}$
Results Are Robust Along Many Dimensions

• Sample Restrictions
  – Excluded participants aged 40-49
  – All participants aged 40-49
  – All participants aged 50-59

• Definitions of mammography
  – Narrower
  – Broader

• Outcomes
  – Breast cancer mortality
  – Mortality at alternative follow-up lengths
Robust to Alternative Definitions of Mammography

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Untreated Outcome Test</th>
<th>Always Taker Average Treatment Effect Lower Bound</th>
<th>LATE</th>
<th>Test Rejects Treatment Effect Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammogram in at least one year after enrollment during the active study period, missing in year = no mammogram in year</td>
<td>19,505</td>
<td>-562</td>
<td>23</td>
<td>-13</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(147)</td>
<td>(59)</td>
<td>(38)</td>
<td>(0.48)</td>
</tr>
<tr>
<td><strong>Narrower Definitions of Mammography</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammogram in more than one year after enrollment during the active study period, missing in year = no mammogram in year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least two active study period years</td>
<td>19,505</td>
<td>-465</td>
<td>-27</td>
<td>-12</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(106)</td>
<td>(77)</td>
<td>(35)</td>
<td>(0.49)</td>
</tr>
<tr>
<td>At least three active study period years</td>
<td>19,505</td>
<td>-420</td>
<td>56</td>
<td>-12</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(94)</td>
<td>(145)</td>
<td>(36)</td>
<td>(0.48)</td>
</tr>
<tr>
<td>All active study period years</td>
<td>19,505</td>
<td>-225</td>
<td>-135</td>
<td>-15</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(75)</td>
<td>(138)</td>
<td>(42)</td>
<td>(0.37)</td>
</tr>
<tr>
<td><strong>Broader Definition of Mammography</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammogram in at least one year after enrollment during the active study period</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing in year = mammogram in year</td>
<td>19,505</td>
<td>-776</td>
<td>103</td>
<td>-24</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(835)</td>
<td>(43)</td>
<td>(69)</td>
<td>(0.43)</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• Sample Restrictions
  – Excluded participants aged 40-49
  – All participants aged 40-49
  – All participants aged 50-59

• Definitions of mammography
  – Narrower
  – Broader

• Outcomes
  – Breast cancer mortality
  – Mortality at alternative follow-up lengths
# Robust to Alternative Outcomes

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Untreated Outcome Test</th>
<th>Always Taker Average Treatment Effect Lower Bound</th>
<th>LATE</th>
<th>Test Rejects Treatment Effect Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Specification</td>
<td>19,505</td>
<td>-562</td>
<td>23</td>
<td>-13</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(147)</td>
<td>(59)</td>
<td>(38)</td>
<td>(0.48)</td>
</tr>
<tr>
<td><strong>Alternative Outcome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast Cancer Mortality</td>
<td>19,505</td>
<td>-43</td>
<td>30</td>
<td>-12</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(47)</td>
<td>(25)</td>
<td>(13)</td>
<td>(0.43)</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• Sample Restrictions
  – Excluded participants aged 40-49
  – All participants aged 40-49
  – All participants aged 50-59

• Definitions of mammography
  – Narrower
  – Broader

• Outcomes
  – Breast cancer mortality
  – Mortality at alternative follow-up lengths
Robust to Mortality at Alternative Follow-Up Lengths: 11-20

<table>
<thead>
<tr>
<th>Years Since Enrollment</th>
<th>N</th>
<th>Untreated Outcome Test</th>
<th>Always Taker Average Treatment Effect Lower Bound</th>
<th>LATE</th>
<th>Test Rejects Treatment Effect Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main specification: 20</td>
<td>19,505</td>
<td>-562 (147)</td>
<td>23 (59)</td>
<td>-13 (38)</td>
<td>1.00 (0.48)</td>
</tr>
<tr>
<td>19</td>
<td>19,505</td>
<td>-485 (142)</td>
<td>50 (58)</td>
<td>-13 (37)</td>
<td>1.00 (0.40)</td>
</tr>
<tr>
<td>18</td>
<td>19,505</td>
<td>-492 (139)</td>
<td>54 (56)</td>
<td>-8 (35)</td>
<td>1.00 (0.41)</td>
</tr>
<tr>
<td>17</td>
<td>19,505</td>
<td>-456 (135)</td>
<td>18 (50)</td>
<td>-8 (33)</td>
<td>1.00 (0.48)</td>
</tr>
<tr>
<td>16</td>
<td>19,505</td>
<td>-471 (134)</td>
<td>15 (46)</td>
<td>-16 (31)</td>
<td>1.00 (0.47)</td>
</tr>
<tr>
<td>15</td>
<td>19,505</td>
<td>-480 (131)</td>
<td>-11 (42)</td>
<td>-15 (31)</td>
<td>1.00 (0.50)</td>
</tr>
<tr>
<td>14</td>
<td>19,505</td>
<td>-396 (121)</td>
<td>-38 (38)</td>
<td>-21 (30)</td>
<td>0.00 (0.45)</td>
</tr>
<tr>
<td>13</td>
<td>19,505</td>
<td>-365 (115)</td>
<td>-30 (36)</td>
<td>-24 (28)</td>
<td>0.00 (0.49)</td>
</tr>
<tr>
<td>12</td>
<td>19,505</td>
<td>-334 (106)</td>
<td>-23 (32)</td>
<td>-27 (27)</td>
<td>1.00 (0.50)</td>
</tr>
<tr>
<td>11</td>
<td>19,505</td>
<td>-351 (105)</td>
<td>-30 (28)</td>
<td>-10 (25)</td>
<td>0.00 (0.42)</td>
</tr>
</tbody>
</table>
## Robust to Mortality at Alternative Follow-Up Lengths: 1 - 10

<table>
<thead>
<tr>
<th>Years Since Enrollment</th>
<th>N</th>
<th>Untreated Outcome Test</th>
<th>Always Taker Average Treatment Effect Lower Bound</th>
<th>LATE</th>
<th>Test Rejects Treatment Effect Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>19,505</td>
<td>-306 (97)</td>
<td>-41 (25)</td>
<td>-15 (23)</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>19,505</td>
<td>-314 (97)</td>
<td>-35 (21)</td>
<td>-12 (20)</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>19,505</td>
<td>-340 (97)</td>
<td>-14 (21)</td>
<td>-2 (18)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>19,505</td>
<td>-351 (97)</td>
<td>-15 (18)</td>
<td>-6 (17)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>19,505</td>
<td>-317 (93)</td>
<td>-24 (16)</td>
<td>-5 (15)</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>19,505</td>
<td>-269 (86)</td>
<td>-12 (15)</td>
<td>-5 (13)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>19,505</td>
<td>-218 (77)</td>
<td>-3 (14)</td>
<td>-9 (11)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>19,505</td>
<td>-209 (76)</td>
<td>-3 (11)</td>
<td>-6 (9)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>19,505</td>
<td>-194 (67)</td>
<td>-3 (9)</td>
<td>-3 (9)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>19,505</td>
<td>-55 (40)</td>
<td>-5 (5)</td>
<td>-5 (5)</td>
</tr>
</tbody>
</table>
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Mortality

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be harmed by them

• Robustness
• Conclusions
U.S. Preventive Services Task Force (USPSTF) 2016 Guidelines for Women in 40’s:

“The decision to start screening mammography in women prior to age 50 years should be an individual one. Women who place a higher value on the potential benefit than the potential harms may choose to begin biennial screening between the ages of 40 and 49 years”
U.S. Preventive Services Task Force (USPSTF) “C recommendation”

“The USPSTF recommends selectively offering this service to individual patients based on professional judgment and patient preferences”
CNBSS Protocols Varied by Age

• Patients aged 40-49:
  – Intervention group: mammography + physical examination each year for 4-5 years, then return to usual care
  – Control group: usual care

• Patients aged 50-59:
  – Intervention group: mammography + physical examination each year for 4-5 years, then return to usual care
  – Control group: physical examination each year for 4-5 years, then return to usual care
USPSTF Recommendations Differ for Women in 40’s and 50’s

- The U.S. Preventive Services Task Force (USPSTF) Assigns “grades”
  - “A” and “B” grades fully-covered under ACA

- Different grades for 40’s and 50+ (Siu, 2016)
  - “The decision to start screening mammography in women prior to age 50 years should be an individual one. (Grade C recommendation)”
  - “The USPSTF recommends biennial screening mammography for women aged 50 to 74 years. (Grade B recommendation)”