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

November 2019
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.”
Questions Raised by Guidelines

1. Do current guidelines target mammograms to women most likely to benefit from them?
2. Can behavior within a clinical trial inform targeting within guidelines?
Benefits and Harms of Mammography

- **Benefits:** Early detection and treatment of breast cancer that would grow to be life-threatening.

- **Harms:** Overdiagnosis and overtreatment

  “The most important harm is the diagnosis and treatment of noninvasive breast cancer that would otherwise not have become a threat to a woman's health, or even apparent, during her lifetime (that is, overdiagnosis and overtreatment)” (Siu, 2016).
Overdiagnosis

• In the short term, breast cancer incidence is larger in intervention group because mammograms diagnose breast cancers.

• In the long term, breast cancer incidence in the control group should “catch up.”

• Persistent difference is an indication of overdiagnosis.
I Examine Behavior within a Clinical Trial

• Clinical trial literature examines mortality
  – See Nelson (2016)
• Outside clinical trials, literature examines mammography behavior in response to policy interventions, says little about overdiagnosis, perhaps because of data constraints
  Kelaher and Stellman (2000); Habermann et al. (2007); Kadiyala and Strumpf (2011, 2016); Finkelstein et al. (2012); Kolstad and Kowalski (2012); Bitler and Carpenter (2016, 2019); Fedewa et al. (2015); Mehta et al. (2015); Ong and Mandl (2015); Lu and Slusky (2016); Zanella and Banerjee (2016); Cooper et al. (2017); Jacobson and Kadiyala (2017); Kim and Lee (2017); Buchmueller and Goldzahl (2018); Einav et al. (2019); Myerson et al. (2019)
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
  - Followed patient outcomes from 1980 to 2005 (at least 20 years for all participants) through cancer registry and death records (no attrition)
  - Collected risk factors and demographic data
  - Recorded mammogram receipt, even in control group
I Examine Micro Data from the CNBSS

• Heterogeneous selection: are women who are more likely to receive mammograms different from other women?

• Treatment effect heterogeneity: are women who are more likely to receive mammograms more likely to experience better or worse health outcomes because of them?
I Examine Behavior within a Clinical Trial

- I build 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.*

“A Model of a Randomized Experiment with an Application to the PROWESS Clinical Trial.” *NBER WP 25670.*

“Counting Defiers.” *NBER WP 25671.*
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• 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 experience a higher level of overdiagnosis

• Robustness
• Conclusions
$U_D$: unobserved net cost of treatment
Z = 0

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

Always Takers

\[ p_c = 0.19 \]

\[ U_D : \text{unobserved net cost of treatment} \]
$Z = 0$

$D = 1$
$0 \leq p \leq p_c$

$D = 0$
$p_c < p \leq 1$

$p_c = 0.19$
Always Takers

$U_D$: unobserved net cost of treatment
$Z = 1$

$Z = 0$

- $D = 1$, $0 \leq p \leq p_c$
- $D = 0$, $p_c < p \leq 1$

$U_D$: unobserved net cost of treatment

$0.00$ \hspace{1cm} $p_c = 0.19$ \hspace{1cm} $1.00$

Always Takers

Never Takers

$p_l < p \leq 1$

$p_l = 0.95$
Z = 1

D = 1
0 \leq p \leq p_l

D = 0
p_l < p \leq 1

Z = 0

D = 1
0 \leq p \leq p_c

D = 0
p_c < p \leq 1

0.00 \quad p_c = 0.19
Always Takers

\quad p_l = 0.95 \quad 1.00
Never Takers

U_D: unobserved net cost of treatment
$D = 0$
$0 \leq p \leq p_I$

$D = 1$
$p_I < p \leq 1$

$Z = 0$

$D = 1$
$0 \leq p \leq p_c$

$D = 0$
$p_c < p \leq 1$

$Z = 1$

$U_D$: unobserved net cost of treatment

$0.00 \quad p_c = 0.19$

Always Takers

Compliers

Never Takers

$p_I = 0.95 \quad 1.00$
Behavior within a Clinical Trial
and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• 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 experience a higher level of overdiagnosis

• Robustness
• Conclusions
Breast Cancer Incidence 20 Years After Enrollment (per 10,000)

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

\[
\begin{align*}
&\frac{(0.95 \times 453 - 0.19 \times 571)}{(0.95 - 0.19)} \\
&\frac{(1 - 0.19) \times 385 - (1 - 0.95) \times 667}{(0.95 - 0.19)}
\end{align*}
\]

\(U_D\) : unobserved net cost of treatment

- Always Takers
  - \(p_C = 0.19\)
- Compliers
- Never Takers
  - \(p_I = 0.95\)

D = 1

Z = 0

D = 0

Z = 1
Breast Cancer Incidence 20 Years After Enrollment (per 10,000)

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

$U_D$: unobserved net cost of treatment

$p_C = 0.19$

Always Takers

$p_I = 0.95$

Never Takers

Compliers

330 366 424

0 1
Breast Cancer Incidence 20 Years After Enrollment (per 10,000)

- treated
- untreated

$LATE = 58$ (34)

$U_D$: unobserved net cost of treatment

$p_C = 0.19$

Always Takers

$p_I = 0.95$

Never Takers
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• 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 experience a higher level of overdiagnosis

• Robustness
• Conclusions
untreated outcome test: \(366 - 667 = -301\) (119)

LATE = 58 (34)

\(U_D\) : unobserved net cost of treatment
# Baseline Covariates Corroborate Selection Heterogeneity

<table>
<thead>
<tr>
<th>Baseline Socioeconomic Status</th>
<th>Always Takers</th>
<th>Compliers</th>
<th>Never Takers</th>
</tr>
</thead>
<tbody>
<tr>
<td>University, trade or business school</td>
<td>0.50</td>
<td>0.46</td>
<td>0.39</td>
</tr>
<tr>
<td>In work force</td>
<td>0.65</td>
<td>0.64</td>
<td>0.65</td>
</tr>
<tr>
<td>Age at first birth</td>
<td>24.28</td>
<td>23.98</td>
<td>23.57</td>
</tr>
<tr>
<td>No live birth</td>
<td>0.16</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>Married</td>
<td>0.80</td>
<td>0.81</td>
<td>0.75</td>
</tr>
<tr>
<td>Husband in work force and alive</td>
<td>0.81</td>
<td>0.81</td>
<td>0.76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Baseline Health Behavior</th>
<th>Always Takers</th>
<th>Compliers</th>
<th>Never Takers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Smoker</td>
<td>0.78</td>
<td>0.75</td>
<td>0.63</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>23.87</td>
<td>24.42</td>
<td>24.48</td>
</tr>
<tr>
<td>Used oral contraception</td>
<td>0.74</td>
<td>0.71</td>
<td>0.67</td>
</tr>
<tr>
<td>Used estrogen</td>
<td>0.13</td>
<td>0.13</td>
<td>0.15</td>
</tr>
<tr>
<td>Mammograms prior to enrollment</td>
<td>0.23</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>Practiced breast self examination</td>
<td>0.47</td>
<td>0.44</td>
<td>0.38</td>
</tr>
</tbody>
</table>
Natural Experiments Corroborate Selection Heterogeneity

- Einav et al. (2019)
- Kim and Lee (2017)
- Oster (2018)
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• 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 experience a higher level of overdiagnosis

• Robustness

• Conclusions
Breast Cancer Incidence 20 Years After Enrollment (per 10,000)

untreated outcome test: 366 - 667 = -301

LATE = 58

$U_D$: unobserved net cost of treatment
untreated outcome test: $366 - 667 = -301$

$LATE = 58$

$U_D$: unobserved net cost of treatment
test rejects treatment effect homogeneity: \[ 1 \{206 > 58\} = 1.00 \] \[ [0.03] \]

untreated outcome test: \[ 366 - 667 = -301 \] \[ (119) \]

always taker
average
treatment
effect
lower bound = 206 (65)

upper bound

LATE = 58 (34)

\[ U_D \]: unobserved net cost of treatment
Women of higher socioeconomic status are exposed to increased “observational intensity” such that “they are likely to be screened more often and by means of such tests...that can detect smaller abnormalities, undergo more follow-up testing, and undergo more biopsies, and they may be served by health systems that have a lower threshold for labeling results as abnormal.”

- Welch and Fisher (2017)
Breast Cancer Characteristics Corroborate Treatment Effect Heterogeneity

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
<th>Difference in Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Tumor Size Among Breast Cancers (in mm)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Always Takers</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>Treated Compliers</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Share of Invasive Breast Cancer Among Breast Cancers (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>73</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>(9)</td>
<td>(7)</td>
</tr>
</tbody>
</table>
Suggestive Evidence for All-Cause Mortality

test rejects treatment effect homogeneity: 1{22 > -13} = 1.00
[0.36]

untreated outcome test: 428 - 990 = -562

LATE = -13
(38)

always taker average treatment effect
lower bound = 22
(59)

All-Cause Deaths 20 Years After Enrollment (per 10,000)

$p_c = 0.19$ Always Takers

$p_I = 0.95$ Never Takers

$U_D$ : unobserved net cost of treatment
Procedures Corroborate Treatment Effect
Heterogeneity

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
<th>Difference in Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>(1)</strong> Always Takers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor Size Among Breast Cancers (in mm)</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td><strong>(2)</strong> Treated Compliers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Invasive Breast Cancer Among Breast Cancers (%)</td>
<td>73</td>
<td>75</td>
</tr>
<tr>
<td><strong>(9)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Mastectomy Among Breast Cancers with Mastectomy or Lumpectomy (%)</td>
<td>45</td>
<td>23</td>
</tr>
<tr>
<td><strong>(9)</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• 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 experience a higher level of overdiagnosis

• Robustness
• Conclusions
Results Are Robust Along Many Dimensions

• Alternative outcomes
  – All-cause mortality
  – Breast cancer morality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
test rejects treatment effect homogeneity: $1 \{206 > 58\} = 1.00$ [0.03]

untreated outcome test: $366 - 667 = -301 (119)$

always taker average treatment effect lower bound = 206 (65)

upper bound

LATE = 58 (34)

$U_D$: unobserved net cost of treatment
## Main Specification For Comparison

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Untreated Outcome Test</th>
<th>Always Taker Average Treatment Effect</th>
<th>Local Average Treatment Effect LATE</th>
<th>Lower Bound on Treatment Effect Ratio (2)/(3)</th>
<th>Test Rejects Treatment Effect Homogeneity 1{(2 &gt; (3)}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer incidence</td>
<td>19,505</td>
<td>-301 (119)</td>
<td>206 (65)</td>
<td>58 (34)</td>
<td>3.5 (12)</td>
<td>1.00 [0.03]</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• **Alternative outcomes**
  – All-cause mortality
  – Breast cancer mortality

• **Alternative sample restrictions**
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• **Alternative definitions of mammography**
  – Narrower

• **Alternative follow-up lengths**
## Robust to Alternative Outcomes

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Untreated Outcome Test</td>
<td>Always Taking Average Treatment Effect Lower Bound</td>
<td>Local Average Treatment Effect LATE</td>
<td>Lower Bound on Treatment Effect Ratio (2)/(3)</td>
<td>Test Rejects Treatment Effect Homogeneity 1{(2) &gt; (3)}</td>
</tr>
<tr>
<td><strong>Main Specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer incidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>19,505</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-301</td>
<td></td>
<td>206</td>
<td>58</td>
<td>3.5</td>
<td>1.00</td>
</tr>
<tr>
<td>(119)</td>
<td></td>
<td>(65)</td>
<td>(34)</td>
<td>(12)</td>
<td>[0.03]</td>
</tr>
<tr>
<td><strong>Alternative Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All-cause mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>19,505</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-562</td>
<td></td>
<td>22</td>
<td>-13</td>
<td>-</td>
<td>1.00</td>
</tr>
<tr>
<td>(147)</td>
<td></td>
<td>(59)</td>
<td>(38)</td>
<td>-</td>
<td>[0.36]</td>
</tr>
<tr>
<td>Breast cancer mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>19,505</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-43</td>
<td></td>
<td>30</td>
<td>-12</td>
<td>-</td>
<td>1.00</td>
</tr>
<tr>
<td>(47)</td>
<td></td>
<td>(25)</td>
<td>(13)</td>
<td>-</td>
<td>[0.23]</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• Alternative outcomes
  – All-cause mortality
  – Breast cancer morality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
## Robust to Alternative Sample Restrictions

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Untreated Outcome Test</td>
<td>Always Taker Average Treatment Effect</td>
<td>Lower Bound on Treatment Effect LATE</td>
<td>Lower Bound on Treatment Effect Ratio (2)/(3)</td>
</tr>
<tr>
<td><strong>Main Specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer incidence</td>
<td>19,505</td>
<td>-301</td>
<td>206</td>
<td>58</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>(119)</td>
<td>(65)</td>
<td>(34)</td>
<td>(12)</td>
<td>[0.03]</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>-1,237</td>
<td>309</td>
<td>79</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>(147)</td>
<td>(48)</td>
<td>(43)</td>
<td>(35)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>All participants aged 40-49 at enrollment</td>
<td>50,430</td>
<td>-826</td>
<td>298</td>
<td>69</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>(107)</td>
<td>(40)</td>
<td>(31)</td>
<td>(31)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>All participants aged 50-59 at enrollment</td>
<td>39,405</td>
<td>-1,555</td>
<td>419</td>
<td>39</td>
<td>10.7</td>
</tr>
<tr>
<td></td>
<td>(140)</td>
<td>(49)</td>
<td>(34)</td>
<td>(89)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>All participants</td>
<td>89,835</td>
<td>-1,156</td>
<td>332</td>
<td>55</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>(96)</td>
<td>(31)</td>
<td>(21)</td>
<td>(95)</td>
<td>[0.00]</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• Alternative outcomes
  – All-cause mortality
  – Breast cancer mortality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
Robust to Alternative Definitions of Mammography

<table>
<thead>
<tr>
<th>N</th>
<th>Untreated Outcome Test</th>
<th>Always Taker Average Treatment Effect</th>
<th>Local Average Treatment Effect (LATE)</th>
<th>Lower Bound on Treatment Effect Ratio (2)/(3)</th>
<th>Test Rejects Treatment Effect Homogeneity 1{(2) &gt; (3)}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19,505</td>
<td>206</td>
<td>58</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(119)</td>
<td>(65)</td>
<td>(34)</td>
<td>(12)</td>
</tr>
<tr>
<td>Main Specification</td>
<td>Breast cancer incidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At least two active study period years after enrollment</td>
<td>19,505</td>
<td>239</td>
<td>54</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(95)</td>
<td>(95)</td>
<td>(32)</td>
<td>(16)</td>
</tr>
<tr>
<td></td>
<td>At least three active study period years after enrollment</td>
<td>19,505</td>
<td>167</td>
<td>55</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(73)</td>
<td>(145)</td>
<td>(32)</td>
<td>(9)</td>
</tr>
<tr>
<td></td>
<td>All active study period years after enrollment</td>
<td>19,505</td>
<td>158</td>
<td>64</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(61)</td>
<td>(190)</td>
<td>(38)</td>
<td>(10)</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• Alternative outcomes
  – All-cause mortality
  – Breast cancer mortality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
## Robust to Breast Cancer Incidence 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 Treatment Effect</th>
<th>Average Lower Bound</th>
<th>Local Average Treatment Effect LATE</th>
<th>Lower Bound on Treatment Effect Ratio (2)/(3)</th>
<th>Test Rejects Treatment Effect Homogeneity 1{(2) &gt; (3)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main specification: 20</td>
<td>19,505</td>
<td>-301</td>
<td>206</td>
<td>58</td>
<td>3.5</td>
<td>1.00</td>
<td>[0.03]</td>
</tr>
<tr>
<td>19</td>
<td>19,505</td>
<td>-269</td>
<td>196</td>
<td>52</td>
<td>3.8</td>
<td>1.00</td>
<td>[0.03]</td>
</tr>
<tr>
<td>18</td>
<td>19,505</td>
<td>-311</td>
<td>210</td>
<td>54</td>
<td>3.9</td>
<td>1.00</td>
<td>[0.01]</td>
</tr>
<tr>
<td>17</td>
<td>19,505</td>
<td>-322</td>
<td>214</td>
<td>49</td>
<td>4.1</td>
<td>1.00</td>
<td>[0.01]</td>
</tr>
<tr>
<td>16</td>
<td>19,505</td>
<td>-342</td>
<td>232</td>
<td>56</td>
<td>4.1</td>
<td>1.00</td>
<td>[0.00]</td>
</tr>
<tr>
<td>15</td>
<td>19,505</td>
<td>-381</td>
<td>211</td>
<td>84</td>
<td>2.5</td>
<td>1.00</td>
<td>[0.02]</td>
</tr>
<tr>
<td>14</td>
<td>19,505</td>
<td>-404</td>
<td>201</td>
<td>80</td>
<td>2.5</td>
<td>1.00</td>
<td>[0.03]</td>
</tr>
<tr>
<td>13</td>
<td>19,505</td>
<td>-431</td>
<td>223</td>
<td>75</td>
<td>3.0</td>
<td>1.00</td>
<td>[0.02]</td>
</tr>
<tr>
<td>12</td>
<td>19,505</td>
<td>-443</td>
<td>191</td>
<td>64</td>
<td>3.0</td>
<td>1.00</td>
<td>[0.03]</td>
</tr>
<tr>
<td>11</td>
<td>19,505</td>
<td>-423</td>
<td>195</td>
<td>55</td>
<td>3.5</td>
<td>1.00</td>
<td>[0.01]</td>
</tr>
</tbody>
</table>

(Numbers in parentheses indicate standard errors)
Robust to Breast Cancer Incidence at Alternative Follow-Up Lengths: 1-10

<table>
<thead>
<tr>
<th>Years Since Enrollment</th>
<th>N</th>
<th>(1) Untreated Outcome Test</th>
<th>Always Taking Average Treatment Effect Lower Bound</th>
<th>(3) Local Average Treatment Effect LATE</th>
<th>(4) Lower Bound on Treatment Effect Ratio (2)/(3)</th>
<th>(5) Test Rejects Treatment Effect Homogeneity 1{(2 &gt; 3)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>19,505</td>
<td>-419</td>
<td>200</td>
<td>47</td>
<td>4.2</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(107)</td>
<td>(45)</td>
<td>(22)</td>
<td>(33)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>9</td>
<td>19,505</td>
<td>-413</td>
<td>192</td>
<td>34</td>
<td>5.6</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(103)</td>
<td>(42)</td>
<td>(21)</td>
<td>(32)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>8</td>
<td>19,505</td>
<td>-409</td>
<td>175</td>
<td>35</td>
<td>5.1</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(100)</td>
<td>(40)</td>
<td>(20)</td>
<td>(39)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>7</td>
<td>19,505</td>
<td>-393</td>
<td>177</td>
<td>46</td>
<td>3.9</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(95)</td>
<td>(36)</td>
<td>(17)</td>
<td>(17)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>6</td>
<td>19,505</td>
<td>-412</td>
<td>185</td>
<td>50</td>
<td>3.7</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(95)</td>
<td>(34)</td>
<td>(15)</td>
<td>(8)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>5</td>
<td>19,505</td>
<td>-382</td>
<td>180</td>
<td>45</td>
<td>4.0</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(90)</td>
<td>(32)</td>
<td>(14)</td>
<td>(4)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>4</td>
<td>19,505</td>
<td>-393</td>
<td>152</td>
<td>46</td>
<td>3.3</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(91)</td>
<td>(29)</td>
<td>(13)</td>
<td>(2)</td>
<td>[0.00]</td>
</tr>
<tr>
<td>3</td>
<td>19,505</td>
<td>-354</td>
<td>104</td>
<td>37</td>
<td>2.8</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(85)</td>
<td>(23)</td>
<td>(12)</td>
<td>(4)</td>
<td>[0.01]</td>
</tr>
<tr>
<td>2</td>
<td>19,505</td>
<td>-337</td>
<td>63</td>
<td>25</td>
<td>2.5</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(82)</td>
<td>(18)</td>
<td>(10)</td>
<td>(3)</td>
<td>[0.04]</td>
</tr>
<tr>
<td>1</td>
<td>19,505</td>
<td>-342</td>
<td>35</td>
<td>20</td>
<td>1.8</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(82)</td>
<td>(12)</td>
<td>(7)</td>
<td>(1)</td>
<td>[0.10]</td>
</tr>
</tbody>
</table>
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• 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 experience a higher level of overdiagnosis

• Robustness
• Conclusions
Implications for Guidelines and Future Research

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

“The USPSTF recommends selectively offering or providing this service to individual patients based on professional judgment and patient preferences”
### 2016 USPSTF 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 Humphry, 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>Kopparberg</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>
</tbody>
</table>

Overall ($I^2 = 25\% ; P = 0.230$)

---

**Annals of Internal Medicine**

**Review**

---

54 of 52
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>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>Kopparberg</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>
</tbody>
</table>

Overall ($I^2 = 25\%; P = 0.230$)  
| Relative Risk (95% CI) |
|-------------------------|-------------------------|
| 0.92 (0.75–1.02)        |                         |

- **Favors Screening Group**: 0.25, 1.00, 4.00
- **Relative Risk (95% CI)**: 0.25, 1.00, 4.00
- **Favors Control Group**: 0.25, 1.00, 4.00
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)”
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:

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

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

Assumptions:

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

**A.2.** (Independence) $(U_D, \gamma_T)$ and $(U_D, \gamma_U) \perp Z$
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) \((U_D, \gamma_I)\) 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:

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

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

Assumptions:

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

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

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

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

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:

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

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

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

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

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

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

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

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

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

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

\[ Z = 0 \]

\[ D=1 \]

\[ 0 \leq p \leq p_c \]

\[ p_c < p \leq 1 \]

\[ 0.00 \quad p_c = 0.19 \quad 1.00 \]

Always Takers

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

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

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

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

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

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

\[ p_C = 0.19 \]

Always Takers

\[ p_I = 0.95 \quad 1.00 \]

Never Takers

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

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

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

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

**Diagram:**

- **Z = 1**
  - D = 1
  - 0 ≤ p ≤ p_I
  - p_I < p ≤ 1

- **Z = 0**
  - D = 1
  - 0 ≤ p ≤ p_C
  - p_C < p ≤ 1

- **p_C = 0.19**
  - Always Takers

- **p_I = 0.95**
  - Never Takers

**UD:** unobserved net cost of treatment
First Stage:

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

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

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

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

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

\[ 0.00 \quad p_C = 0.19 \quad \text{Always Takers} \quad \text{Compliers} \quad p_I = 0.95 \quad 1.00 \quad \text{Never Takers} \]

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

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

Second Stage:

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

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

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

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

Second Stage:

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

<table>
<thead>
<tr>
<th>( p_C )</th>
<th>0.19</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_I )</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Always Takers | Compliers | Never Takers
---|---|---
\( U_D \): unobserved net cost of treatment
First stage:

\[ V = V_U + (V_T - V_U)D \quad Z \in \{0, 1\}, \ Z \perp (\nu_U - \nu_T) \]
\[ V_T = \mu_T(Z) + \nu_T \]
\[ V_U = \mu_U(Z) + \nu_U \]
\[ D = 1\{V_U \leq V_T\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]
\[ U_D = F(\nu_U - \nu_T) \]

Second stage:

\[ Y = Y_U + (Y_T - Y_U)D \]
\[ Y_T = \eta_T(U_D, \gamma_T) \quad Z \perp (\gamma_T, \gamma_U) \]
\[ Y_U = \eta_U(U_D, \gamma_U) \]

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, Chang, Lorch, and Small (2014); Black, Joo, Lalonde, Smith, and Taylor (2015); Mogstad, Brinch, and Wiswall (2017); Kowalski (2016, 2018). Also related to Hausman (1978); Heckman (1979); Willis and Rosen (1979); Angrist (2004); Huber (2013).)
First stage:

\[ V = V_U + (V_T - V_U)D \]
\[ V_T = \mu_T(Z) + \nu_T \]
\[ V_U = \mu_U(Z) + \nu_U \]
\[ D = 1\{V_U \leq V_T\} \]
\[ \Rightarrow D = 1\{U_D \leq P(D = 1 \mid Z = z)\} \]

\[ Z \in \{0, 1\}, \ Z \perp (\nu_U - \nu_T) \]
Let \( F \) be the CDF of \( \nu_U - \nu_T \)

\[ U_D = F(\nu_U - \nu_T) \]

Second stage:

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

\[ Z \perp (\gamma_T, \gamma_U) \]

Ancillary Assumption — Weak Monotonicity:

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

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

(Mogstad, Brinch, and Wiswall (2017); Kowalski (2016, 2018). Also related to Heckman (1979); Olsen (1980); Kline and Walters (2018).)
Results Are Robust to Alternative Identification Strategy

• Kim and Lee (2017) – regression discontinuity
Never Takers Die More Than Compliers

- Women more likely to receive mammograms are healthier
- Breast cancer mortality **without** screening (Kim and Lee, 2017)
Always Takers Die More Than Compliers

- Women more likely to receive mammograms are more likely to be harmed by them (under untreated outcome monotonicity)
- Breast cancer mortality with screening (Kim and Lee, 2017)