Estimating Welfare in Insurance Markets using Variation in Prices

Liran Einav\textsuperscript{1}  Amy Finkelstein\textsuperscript{2}  Mark R. Cullen\textsuperscript{3}

\textsuperscript{1}Stanford and NBER
\textsuperscript{2}MIT and NBER
\textsuperscript{3}Yale School of Medicine

November, 2008
Motivation

- Classic theory on adverse selection emphasizes private market inefficiency and potential for welfare improving government policy.
- Recent large literature on detecting asymmetric information; little on its welfare consequences.
- We propose (and implement) a simple and general framework for empirical welfare analysis in selection markets.
  - Rely on standard consumer and producer theory.
  - Key feature of selection markets: firms’ costs depend on which consumers purchase their products (“endogenous” costs).
  - Pricing variation used to trace out demand curve can also be used to trace out endogenous cost curve.
  - With demand and cost, welfare analysis is simple and familiar.
- Application: market for employer-provided health insurance in U.S.
Application: employer provided health insurance in U.S.

- Setting: Individual-level data from a large private employer in U.S.
  - Plausibly exogenous variation in prices

- Results
  - Detect statistically-significant adverse selection
  - Efficiency cost: ~$10 per employee (annual) or ~3% of surplus at stake from efficient pricing
  - Limited scope for welfare improvements through public policy
    - Note: results specific to context; no reason they generalize

- Highlights importance of moving beyond detection of market failures to quantifying their welfare implications
Only two contracts: $H$ (full coverage) and $L$ (no coverage)

- Easy to extend to other or more contracts
- $p = p_H - p_L$ is the relative price of contract $H$

Note: take non-price characteristics of insurance contracts as given

- As in Akerlof (1970) compared to Rothschild and Stiglitz (1976)
- Empirically relevant – often observably different individuals offered same menu of contract, just at different prices

Individuals defined by a vector of attributes $\zeta_i \sim G(\zeta)$, and have to choose a contract $H$ or $L$

- $\pi(\zeta_i)$ is willingness to pay for $H$ (i.e., $v_H(\zeta_i, \pi(\zeta_i)) = v_L(\zeta_i)$)
- $c(\zeta_i)$ is expected insurable costs given $H$
  - could depend on coverage chosen (i.e. moral hazard)
Theory: Demand, Supply, and Equilibrium

- **Demand:**
  \[ D(p) = \Pr (\pi(\zeta_i) \geq p) \]

- **Supply:**
  - \( N \geq 2 \) identical risk neutral insurance providers, who set prices in a Nash Equilibrium (a-la Bertrand)
  - Average cost (AC):
    \[ AC(p) = E (c(\zeta) | \pi(\zeta) \geq p) \]
  - Marginal cost (MC):
    \[ MC(p) = E (c(\zeta) | \pi(\zeta) = p) \]

- Additional (standard) assumptions \( \rightarrow \) Equilibrium exists, unique, and given by the lowest break-even price:
  \[ p^* = \min \{ p : p = AC(p) \} \]

- Welfare: Social surplus from allocating \( H \) to individual \( i \) is
  \[ TS(\zeta_i) = \pi(\zeta_i) - c(\zeta_i) \]
Theory: Welfare definitions

- Welfare: Social surplus from allocating $H$ to individual $i$ is
  \[ TS(\zeta_i) = \pi(\zeta_i) - c(\zeta_i) \]

- First best allocation: individual $i$ purchases insurance if and only if
  \[ \pi(\zeta_i) \geq c(\zeta_i) \]

- Constrained efficient allocation: maximizes social welfare subject to the constraint that price is the only instrument available for screening.
  - Constrained efficient: individual $i$ purchases insurance if and only if
    \[ \pi(\zeta_i) \geq \mathbb{E}(c(\tilde{\zeta})|\pi(\tilde{\zeta}) = \pi(\zeta_i)) \]
The welfare cost of adverse selection

Price

Quantity

Demand curve

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008
The welfare cost of adverse selection

Price

Demand curve

Quantity

MC curve
The welfare cost of adverse selection

Price

Demand curve

MC curve

Quantity

$P_{\text{eff}}$

$Q_{\text{eff}}$

$E$
The welfare cost of adverse selection

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008
The welfare cost of adverse selection

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008

11 / 39
The welfare cost of adverse selection

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008 12 / 39
The welfare cost of advantageous selection

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008
Sufficient statistics: Demand and Cost Curves

- Graphical analyses illustrate that demand and cost curves are sufficient statistics for welfare analysis.

- This is the essence of our empirical approach: estimate demand and cost curves but remain agnostic as to the primitives that give rise to them.

- As long as we are willing to use revealed choices for welfare analysis, precise source of selection (i.e., the $\zeta$) not germane for welfare analysis of efficiency cost of adverse selection or welfare consequences of policies that change the equilibrium price.

  - e.g., selection could be driven by heterogeneity in risk aversion, heterogeneity in private information, etc.

  - similarly, source of cost curve (such as any role of moral hazard) not germane to analyzing efficiency consequences of selection that occurs, given cost curve.
As analysis made clear, sufficient statistics for welfare analysis are:

- the demand curve $D(p)$
- the average cost curve $AC(p)$
- the marginal cost curve $MC(p)$

Good variation of $p$ & quantity data → estimate $D(p)$

The *same* variation in $p$ & cost data → estimate $AC(p)$ using sample who endogenously choose $H$

From $D(p)$ and $AC(p)$ we can back out $MC(p)$:

$$MC(p) = \frac{\partial (AC(p) \cdot D(p))}{\partial D(p)} = \left( \frac{\partial D(p)}{\partial p} \right)^{-1} \frac{\partial (AC(p) \cdot D(p))}{\partial p}$$

- Note: The slope of $MC(p)$ also provides direct test of selection

Conceptually, variation in $p$ identifies all curves non-parameterically. In practice, likely that need to make functional form assumptions.

- Here structure could be useful to guide functional form
- But graphs highlight which parts of curves are important to “get right”
Example: Illustrating the basic idea

- Individual $i$ defined by $(\pi_i, c_i)$:
  - $\pi_i$ - willingness to pay for insurance
  - $c_i$ - expected costs to the insurer
- Drawn uniformly from $(\pi, c) \in \{(2, 1), (4, 3), (6, 5)\}$
  - Note adverse selection: higher $\pi$ associated with higher $c$
  - Competitive equilibrium ($p = AC$): $p = 4$
  - Efficient allocation: $p \leq 2$ (everyone buys)
- Data: we assume random price variation
  - Data is given by $(p, Q, AC) = \{(2, 1, 3), (4, 2/3, 4), (6, 1/3, 5)\}$
  - Competitive equilibrium is directly observed
  - Can back out marginal cost ($c$) from demand and $AC$, e.g.
    
    $$c(\pi = 2) = \frac{AC(p = 2)Q(p = 2) - AC(4)Q(4)}{Q(p = 2) - Q(p = 4)} = \frac{3 \cdot 1 - 4 \cdot \frac{2}{3}}{1 - \frac{2}{3}} = 1$$

- With $c$ and $\pi$ thus recovered, can compute social welfare $(\pi - c(\pi))$
  from each allocation, and compare, for example, efficient and equilibrium allocations
Comment: moral hazard

- Does not change analysis; complicates presentation
- Since costs are a function of insurance coverage, useful to define $c^H \geq c^L$
  - $c^j$ is expected cost of insurance coverage when behavior is as under $j$ coverage
  - correspondingly two average cost curves ($AC^H$ and $AC^L$) and two marginal cost curves ($MC^H$ and $MC^L$)
- To explicitly recognize moral hazard in preceding analysis, replace $c$, $AC$, and $MC$ with superscript "$H$"
  - Recall that cost curve estimated on sample of individuals who endogenously choose $H$

Intuition

- From firm perspective: only behavior of insured individuals matters; $c^L$ doesn’t matter
- From individual perspective, $c^L$ relevant only via effect on WTP ($\pi$)
- (Aside: framework allows us to test for and quantify moral hazard)
Summary: Attractions and limitations

- Several appealing features:
  - Model demand and costs, but not primitives behind them
  - Extremely simple to implement
  - In principle broadly applicable
  - Bonus: direct test of selection (shape of cost curve)

- Main limitations:
  - Requires good price variation, which isn’t always easy to find
  - Cannot evaluate welfare from contracts not observed
    - Requires structural primitives underlying demand and cost curves
    - Recent attempts to estimate in selection markets (e.g. Einav, Finkelstein and Schrimpf 2007; Lustig 2008)

- Familiar trade-off:
  - Product-space vs. characteristic-space approaches to differentiated-product demand estimation
Part II: Empirical application

- Data and environment
- Empirical constructs and relation to theory
- Pricing variation
- Baseline results
- Robustness
Individual-level data from 2004 on U.S.-based employees of a large multi-national aluminum manufacturer

- New health insurance options introduced for 2004

Data include:

- The menu of health insurance options available to each employee
- The premium associated with each option
- Employee’s choices
- Employee’s (and dependents) subsequent medical expenditure
- Rich demographics – everything price setter likely to observe
Sample restrictions

- Limit to salaried employees (~1/3 of total employees)
  - good pricing variation

- To more closely follow the theoretical framework, further restrict to the two-thirds who chose the two modal options
  - Two PPOs that vary only in financial details (deductibles and out-of-pocket maximum)
    - Note: $H$ (less consumer cost sharing) and $L$ (more consumer cost sharing) are both partial coverages

- Unlikely that sample selection introduces biases

- Baseline specification further limits to those who chose family coverage (3,779 employees)
  - Throughout take coverage tier (single, family, etc.) as given
  - Results similar when we use all coverage tiers
## Summary statistics – representativeness

<table>
<thead>
<tr>
<th></th>
<th>2004 Company Data</th>
<th>March 2005 CPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All employees</td>
<td>Only salaried</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>workers (2)</td>
</tr>
<tr>
<td></td>
<td>All full time</td>
<td>Only in</td>
</tr>
<tr>
<td></td>
<td>workers (6)</td>
<td>manufacturing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>White collar</td>
</tr>
<tr>
<td></td>
<td></td>
<td>workers in</td>
</tr>
<tr>
<td></td>
<td></td>
<td>manufacturing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(8)</td>
</tr>
<tr>
<td>Number of Individuals</td>
<td>36,814</td>
<td>11,178</td>
</tr>
<tr>
<td>Fraction Male</td>
<td>0.78</td>
<td>0.70</td>
</tr>
<tr>
<td>Fraction White</td>
<td>0.77</td>
<td>0.82</td>
</tr>
<tr>
<td>Fraction unionized</td>
<td>0.33</td>
<td>0.14</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>44.24</td>
<td>42.13</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>9.86</td>
<td>11.45</td>
</tr>
<tr>
<td>Median</td>
<td>45</td>
<td>42</td>
</tr>
<tr>
<td>Tenure with company (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>13.23</td>
<td>n/a</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>10.28</td>
<td>n/a</td>
</tr>
<tr>
<td>Median</td>
<td>11</td>
<td>n/a</td>
</tr>
<tr>
<td>Annual Salary (current $US)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>53,103</td>
<td>46,195</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>47,642</td>
<td>45,435</td>
</tr>
<tr>
<td>Median</td>
<td>47,283</td>
<td>35,000</td>
</tr>
</tbody>
</table>

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008 22 / 39
## Summary statistics – sample cuts

<table>
<thead>
<tr>
<th></th>
<th>All employees</th>
<th>Only salaried workers</th>
<th>Only salaried workers with new benefit design</th>
<th>Col. (3) limited to only workers who chose High or Low</th>
<th>Col. (4) limited to workers with family coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Individuals</strong></td>
<td>36,814</td>
<td>11,964</td>
<td>11,325</td>
<td>7,263</td>
<td>3,779</td>
</tr>
<tr>
<td><strong>Fraction Male</strong></td>
<td>0.78</td>
<td>0.73</td>
<td>0.73</td>
<td>0.77</td>
<td>0.86</td>
</tr>
<tr>
<td><strong>Fraction White</strong></td>
<td>0.77</td>
<td>0.87</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td><strong>Fraction unionized</strong></td>
<td>0.33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>44.24</td>
<td>44.51</td>
<td>44.50</td>
<td>45.17</td>
<td>42.66</td>
</tr>
<tr>
<td>Median</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>46</td>
<td>43</td>
</tr>
<tr>
<td><strong>Tenure with company (years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>13.23</td>
<td>13.26</td>
<td>13.23</td>
<td>13.69</td>
<td>12.70</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>10.28</td>
<td>9.95</td>
<td>9.96</td>
<td>10.01</td>
<td>8.93</td>
</tr>
<tr>
<td>Median</td>
<td>11</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td><strong>Annual Salary (current $US)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>53,103</td>
<td>71,622</td>
<td>72,821</td>
<td>74,017</td>
<td>80,999</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>47,642</td>
<td>77,936</td>
<td>79,373</td>
<td>91,530</td>
<td>112,790</td>
</tr>
<tr>
<td>Median</td>
<td>47,283</td>
<td>60,484</td>
<td>61,433</td>
<td>61,822</td>
<td>66,335</td>
</tr>
</tbody>
</table>
Empirical constructs

- $p_i = p_i^H - p_i^L$ where $p_i^j$ is employee $i$'s annual contribution for coverage $j$
- $D_i = 1$ if $i$ chose $H$; $D_i = 0$ if $i$ chose $L$
- $m_i$ is employee $i$'s vector of medical cost during 2004
- $c(m_i; j)$ is the insurer's cost of covering $m_i$ under coverage $j$
- $c_i = c(m_i; H) - c(m_i; L)$ is the incremental insurer's costs from covering $i$ with $H$ vs. $L$ (holding behavior $m_i$ fixed)
Left hand figure shows in network rules:
- **L**: $500 deductible; 10% coinsurance; $5500 out of pocket max
- **H**: $0 deductible; 10% coinsurance; $5000 out of pocket max

Right hand figure graphs $c_i$
- $c_i = c(m_i; H) - c(m_i; L)$ is the incremental insurer’s costs from covering $i$ with $H$ vs. $L$ (holding behavior $m_i$ fixed)
Graphical illustration of costs \( c \) (out of network)

- **Left hand figure shows out of network rules**
  - \( L \): $1000 deductible; 30% coinsurance; $11000 out of pocket max
  - \( H \): $500 deductible; 30% coinsurance; $10000 out of pocket max

- **Right hand figure graphs** \( c_i \)
  - \( c_i = c(m_i; H) - c_L(m_i; L) \) is the incremental insurer’s costs from covering \( i \) with \( H \) (holding behavior \( m_i \) fixed)
Empirical distribution of $c$

Incremental insurer cost ($US$)

0%
0
(0,100]
0.8% 1.5% 1.9% 2.1%
6.6% 7.1%
68.0%
5.6% 2.0% 1.4% 3.0%
0%
20%
40%
60%
80%
(100,225)
225
(225,350]
(350,450)
450
(450,650]
(650,800)
800
(800,1150)
Incremental insurer cost ($US$)
Empirical distribution of c for H vs. L coverage

Chose High Coverage
Chose Low Coverage

Incremental insurer cost ($US)

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008 28 / 39
We estimate (using OLS):

\[ D_i = \alpha + \beta p_i + \epsilon_i \text{ for everyone} \]
\[ c_i = \gamma + \delta p_i + u_i \text{ for those who chose } H \]

recall \( c_i = c(m_i; H) - c(m_i; L) \)

- Marginal cost derived from these without additional estimation
- Will show robustness to other functional forms
Price variation

- Key to approach is exogenous variation in $p_i = p_i^H - p_i^L$.
- Provided by company business structure
  - Company is decentralized, partitioned into “business units” (BUs)
    - approximately 40; organized by functionality; very autonomous
    - multiple BUs in same state (even same plant)
  - BU presidents choose from 6 pricing menus proposed by headquarters
    - can choose separately by job site and worker type (hourly vs. salary)
    - for family coverage, $p_i$ ranges from $384 to $659
- Focus on salary workers throughout:
  - A priori pricing variation seemed more likely exogenous / driven by idiosyncratic aspects of BU president
    - accountants, paralegals, metallurgists, and administrative assistants may face different prices because they are affiliated with "primary metals" instead of “rigid packaging”
    - nature of hourly work however may differ (e.g., potroom operator vs. furnace operator)
  - Born out by data: prices are not correlated with observables of salaried workers (see Table 2) but are for hourly workers
Exogeneity of prices with respect to observables

<table>
<thead>
<tr>
<th></th>
<th>Faced lowest relative price (2,939 workers)</th>
<th>Faced higher relative prices (840 workers)</th>
<th>Difference</th>
<th>Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Mean)</td>
<td>42.74</td>
<td>42.40</td>
<td>0.33</td>
<td>-0.245</td>
<td>0.31</td>
</tr>
<tr>
<td>Tenure (Mean)</td>
<td>13.02</td>
<td>11.63</td>
<td>1.39</td>
<td>-0.565</td>
<td>0.08</td>
</tr>
<tr>
<td>Fraction Male</td>
<td>0.862</td>
<td>0.852</td>
<td>0.009</td>
<td>1.268</td>
<td>0.79</td>
</tr>
<tr>
<td>Fraction White</td>
<td>0.874</td>
<td>0.825</td>
<td>0.049</td>
<td>-6.998</td>
<td>0.40</td>
</tr>
<tr>
<td>Log(Annual Salary) (Mean)</td>
<td>11.16</td>
<td>11.05</td>
<td>0.11</td>
<td>-8.612</td>
<td>0.17</td>
</tr>
<tr>
<td>Spouse Age (Mean)</td>
<td>41.37</td>
<td>41.05</td>
<td>0.32</td>
<td>-0.200</td>
<td>0.41</td>
</tr>
<tr>
<td>Number of covered family members (Mean)</td>
<td>4.14</td>
<td>4.07</td>
<td>0.07</td>
<td>-1.400</td>
<td>0.36</td>
</tr>
<tr>
<td>Age of youngest covered child (Mean)</td>
<td>9.81</td>
<td>9.41</td>
<td>0.40</td>
<td>-0.3</td>
<td>0.26</td>
</tr>
<tr>
<td>Log(2003 Medical Spending + 1)(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>8.13</td>
<td>7.79</td>
<td>0.32</td>
<td>-2.100</td>
<td>0.15</td>
</tr>
<tr>
<td>In most common 2003 plan</td>
<td>8.21</td>
<td>8.08</td>
<td>0.13</td>
<td>-1.700</td>
<td>0.08</td>
</tr>
</tbody>
</table>

\(^a\) Log(2003 Medical Spending + 1)
### Raw data with basic findings

<table>
<thead>
<tr>
<th>Relative Price</th>
<th>Number of Obs.</th>
<th>Fraction chose High Coverage</th>
<th>Their Average Incremental Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>$384</td>
<td>2,939</td>
<td>0.67</td>
<td>$451.40</td>
</tr>
<tr>
<td>$466</td>
<td>67</td>
<td>0.66</td>
<td>$499.32</td>
</tr>
<tr>
<td>$489</td>
<td>7</td>
<td>0.43</td>
<td>$661.27</td>
</tr>
<tr>
<td>$495</td>
<td>526</td>
<td>0.64</td>
<td>$458.60</td>
</tr>
<tr>
<td>$570</td>
<td>199</td>
<td>0.46</td>
<td>$492.59</td>
</tr>
<tr>
<td>$659</td>
<td>41</td>
<td>0.49</td>
<td>$489.05</td>
</tr>
</tbody>
</table>

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)  
Estimating Welfare in Insurance Markets  
November 2008
### Results: estimates

<table>
<thead>
<tr>
<th></th>
<th>1 if chose High (both High and Low) (1)</th>
<th>Incremental Cost (only High) (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Price of High ($US)</td>
<td>-0.00070 (0.00032) [0.034]</td>
<td>0.15524 (0.06388) [0.021]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.940 (0.123) [0.000]</td>
<td>391.690 (26.789) [0.000]</td>
</tr>
<tr>
<td>Mean Dependent Variable</td>
<td>0.652</td>
<td>455.341</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>3,779</td>
<td>2,465</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.008</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Standard errors (in parentheses) clustered on state
p-values in [square brackets]
The welfare cost of adverse selection

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008 34 / 39
Results: graphical illustration

Einav, Finkelstein, and Cullen (Stanford and NBER, MIT and NBER, Yale School of Medicine)

Estimating Welfare in Insurance Markets

November 2008 35 / 39

(Q_{eqm}=0.617, P_{eqm}=463.5)

(Q_{eff}=0.756, P_{eff}=263.9)

CDE=$9.55
Estimated demand and cost curves can also provide benchmarks to help provide context.

Preferred benchmark:
- Welfare cost of price subsidy required to achieve efficient price – i.e. 
  \[ \lambda(P_{eq} - P_{eff})Q_{eff} \] – is about 5 times welfare gain from moving from adverse selection equilibrium to efficient price.

Other benchmarks:
- Welfare cost of mandatory coverage by \( H \) is about 3 times equilibrium welfare cost of adverse selection.
- Welfare cost of adverse selection \(~3\%\) of total surplus at stake from efficient pricing.
Robustness

- Explore robustness to:
  - functional form
  - accounting for tax subsidies
  - alternative samples

- Welfare estimates are quite robust. e.g.
  - Welfare gain from price subsidy that achieves efficient price always substantially below welfare cost of price subsidy
  - Welfare loss from adverse selection always lower than welfare loss from mandatory coverage by $H$

- Also examine concern about sample selection (limited to those who choose $H$ or $L$)
  - find that relative price of $H$ vs. $L$ does not predict whether employee chooses one of the other options
  - suggests sample selection unlikely to be important for demand estimates (and of course is irrelevant for cost estimates)
### Potential sample selection

| Dependent variable: | "Outside good" does not include "opt out" | "Outside good" does include "opt out"
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Family coverage tier only</td>
<td>All coverage tiers</td>
</tr>
<tr>
<td>Relative price</td>
<td>-0.0000093 (0.00035) [0.98]</td>
<td>-0.000021 (0.00040) [0.96]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.287 (0.1580) [0.08]</td>
<td>0.292 (0.1150) [0.02]</td>
</tr>
<tr>
<td>Mean dependent variable</td>
<td>0.283</td>
<td>0.300</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>5,271</td>
<td>10,386</td>
</tr>
</tbody>
</table>

**Note:** 1 if "outside good" was chosen, 0 otherwise.
Summary

- Propose an approach to quantify importance of adverse selection
- Many attractive features:
  - Not required to make assumptions about underlying primitives (preferences, information, etc.)
  - Simple to implement
  - Likely applicable across in many insurance settings
  - Provides direct test of selection (as distinct from moral hazard)
- Approach requires good variation in prices
  - Many likely sources: regulation, tax policy, firm experimentation, etc.
- Caveats:
  - Requires using revealed choices for welfare analysis
  - Approach not useful for welfare analysis of new contracts; requires that we model structural primitives explicitly.
- Application: employer provided health insurance market
  - Detect adverse selection but not obviously remediable by public policy
  - This underscores the original motivation to quantify adverse selection (rather than simply test for it)