ADVERSE SELECTION AND NON-TAKE INFERENCE
Robert Oliver, University of California, Berkeley, California
Aush Thaker, InfoCentricity, Inc., Novato, California

1. What is Adverse Selection? How to define, measure?
2. Traditional Measure: Shift in Score Distributions for Takes
3. Definition of No Adverse Selection
4. Conservation of Counts and Role of Non-Take Inference
5. Differential Risk and Response Elasticities: Bayes’ Rule and “No Free Lunch”
6. Revising Risk Scores with woe on Take Rates
7. Summary
TRADITIONAL VALIDATION OF ADVERSE SELECTION
Comparing Distributions of Scores in Take and Non-Take Populations for Three Price Tiers

Figure 1

Cumulative Distribution Fraction

Default or Late Payment Risk Score

0.1 50 GRAPHFREQ (4 6 8 10) (0 5) (5 5) (14 7 8) TAKE NOT TAKE REC DEV1512
REQUIREMENTS FOR BAD ADVERSE SELECTION (BAS) ?

For behavioral/demographic/financial data $\mathbf{x} \in \mathcal{X}$, loan rate $r \in \mathcal{R}$:

1. $p(B \mid T, \mathbf{x}) > p(B \mid \mathbf{x}) = \Pr\{\text{loan default by borrower} \mid \mathbf{x}\}$

2. $p(T \mid B, \mathbf{x}, r) > p(T \mid \mathbf{x}, r) = \Pr\{\text{borrower Takes loan offer} \mid \mathbf{x}, r\}$

3. \textit{BAS} if \textbf{cdf} of Risk Scores of Takes dominates Non-Takes

4. \textit{BAS} \iff $\frac{\partial p(B \mid T, \mathbf{x}, r)}{\partial r} > 0$ \{Phillips & Raffard (2009)\}

5. Other ....
EQUALITY OF BAYES FACTORS ALWAYS HOLDS:  

\[ \frac{p(B \mid T, x, r)}{p(B \mid x, r)} = \frac{p(T \mid B, x, r)}{p(T \mid x, r)} \]

DEFINITION OF NO ADVERSE SELECTION, NAS:  

\[ \frac{p(B \mid T, x, r)}{p(B \mid x, r)} = \frac{p(T \mid B, x, r)}{p(T \mid x, r)} = 1 \quad \text{G,} \, B \perp T, N, r \mid x \]

DEFINITION OF BAD ADVERSE SELECTION, BAS:  

\[ \frac{p(B \mid T, x, r)}{p(B \mid x, r)} = \frac{p(T \mid B, x, r)}{p(T \mid x, r)} > 1 \quad \text{GAS}: \quad < 1 \]
TWO-STAGE OFFER DECISION MADE BY LENDER
BASED ON RESPONSE AND BASELINE DEFAULT SCORE

Lender Decision:
Offer Rate, \( r \)

Lender Decision:
(Accept/Reject)

Borrower Response
(T/N - Take/Not Take)

Baseline Risk Score

Action-based Response Score

Default Performance
(G/B - Good/Bad)

Utility

G, B \perp r \mid x, T, N

See Figure 2(a)

T: Borrower Takes Lender’s Offer

B: Borrower Defaults on Loan
TWO-STAGE DECISION WITH ACTION-BASED DEFAULT AND RESPONSE SCORE

(Offer-dependent Default Risk with Bad Adverse Selection)

See Figure 2(b)
TAKE, Non-Prepay VAL2012

Baseline Default Risk Score, $s_p(x)$
## COMPARING PREDICTED COUNTS WITH OBSERVED RESPONSES AND GOODS/BADS

<table>
<thead>
<tr>
<th></th>
<th>E[#Goods]</th>
<th>E[#Bads]</th>
<th>Observed #Goods</th>
<th>Observed #Bads</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takes (T)</td>
<td>1530.0</td>
<td>127.8</td>
<td>1508</td>
<td>150</td>
<td>1658</td>
</tr>
<tr>
<td>Non-Takes (N)</td>
<td>3331.0</td>
<td>273.7</td>
<td></td>
<td></td>
<td>3605</td>
</tr>
<tr>
<td>Total Accepts by Lender</td>
<td>4862.0</td>
<td>401.5</td>
<td></td>
<td></td>
<td>5263</td>
</tr>
</tbody>
</table>

Table 1: Predicted and Observed Cell Counts for Goods/Bads, Takes/Not-Takes

(4 6) (2 3) (~4 4) (14 7 8) EXPBADS and EXPADVSEL DEV1612
### COUNTS OF OBSERVED GOODS/BADS AND RESPONSES

<table>
<thead>
<tr>
<th></th>
<th>Goods</th>
<th>Bads</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takes (T)</td>
<td>250</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>Not Takes (N)</td>
<td></td>
<td></td>
<td>1200</td>
</tr>
<tr>
<td>Total Accepts (A)</td>
<td></td>
<td></td>
<td>1500</td>
</tr>
</tbody>
</table>
COUNTS OF OBSERVED GOODS/BADS AND RESPONSES
WITH NON-TAKE INFERENCE COUNT, Z

<table>
<thead>
<tr>
<th></th>
<th>Goods</th>
<th>Bads</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takes (T)</td>
<td>250</td>
<td>50</td>
<td>300</td>
</tr>
<tr>
<td>Not Takes (N)</td>
<td>1200-Z</td>
<td>Z</td>
<td>1200</td>
</tr>
<tr>
<td>Total Accepts (A)</td>
<td>1450-Z</td>
<td>50+Z</td>
<td>1500</td>
</tr>
</tbody>
</table>

Z is the inferred number of Bads among the Non-Takes

See Table 1
### Reallocation of Adverse Counts, Y, to obtain NAS

<table>
<thead>
<tr>
<th></th>
<th>Goods</th>
<th>Bads</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takes (T)</td>
<td>250</td>
<td>50-Y</td>
<td>300-Y</td>
</tr>
<tr>
<td>Not Takes (N)</td>
<td>1200-Z</td>
<td>Z+Y</td>
<td>1200-Y</td>
</tr>
<tr>
<td>Total Accepts (A)</td>
<td>1450-Z</td>
<td>50+Z</td>
<td>1500</td>
</tr>
</tbody>
</table>

\[
\text{NAS: } \frac{250}{50 - Y} = \frac{1200 - Z}{Z + Y}, \quad Y(Z) \geq 0 \text{ is number of Bad Adverse Selects}
\]
Counts of Observed (Inferred) Goods/Bads
(Action Based Default Score Segment 2-3 in VAL2012)

<table>
<thead>
<tr>
<th>3-5% APR Premium</th>
<th>Observed/Inferred #Goods</th>
<th>Observed/Inferred # Bads/Goods (Adverse Selects)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Prepays</td>
<td>547</td>
<td>36 (Y= - 5.3)</td>
<td>583</td>
</tr>
<tr>
<td>Non-Takes and Prepay s</td>
<td>305 0</td>
<td>235.5 (Z)</td>
<td>328 6</td>
</tr>
<tr>
<td>Accepts</td>
<td>359 7</td>
<td>271.5</td>
<td>386 9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5-7% APR Premium</th>
<th>Observed/Inferred #Goods</th>
<th>Observed/Inferred # Bads/Goods (Adverse Selects)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Prepays</td>
<td>493</td>
<td>76 (Y= +31.8 )</td>
<td>569</td>
</tr>
<tr>
<td>Non-Takes and Prepay s</td>
<td>538 2</td>
<td>450.5 (Z)</td>
<td>583 2</td>
</tr>
<tr>
<td>Accepts</td>
<td>587 5</td>
<td>526.5</td>
<td>640 1</td>
</tr>
</tbody>
</table>

See Table 6

(DATA= (3 5 7) (2 3) (~4 4) (14 7 8) EXTRACTRECORDS VAL2012)(3 7 EXPOBSADVSEL2 DATA)(1/6/11)
DIFFERENTIAL RISK AND RESPONSE ELASTICITIES

Response:

\[ \varepsilon^{(T)} \triangleq \frac{\partial p(T \mid x,r)}{\partial r} \frac{r}{p(T \mid x,r)} = \frac{\partial p}{\partial r} / r \]

\[ \varepsilon_B^{(T)} \triangleq \frac{\partial p(T \mid B,x,r)}{\partial r} \frac{r}{p(T \mid B,x,r)} \]

Risk:

\[ \delta^{(B)} \triangleq \frac{\partial p(B \mid x,r)}{\partial r} \frac{r}{p(B \mid x,r)} \]

\[ \delta_T^{(B)} \triangleq \frac{\partial p(B \mid T,x,r)}{\partial r} \frac{r}{p(B \mid T,x,r)} \]

See Equations 16 through 19
Six Response Elasticities: Three for Takes
Six Risk Elasticities: Three for Bads

\[
E = \begin{bmatrix}
\varepsilon_G^{(T)} & \varepsilon_B^{(T)} & \varepsilon^{(T)} \\
\varepsilon_G^{(N)} & \varepsilon_B^{(N)} & \varepsilon^{(N)}
\end{bmatrix} \quad \Delta = \begin{bmatrix}
\delta_T^{(G)} & \delta_T^{(B)} \\
\delta_N^{(G)} & \delta_N^{(B)} \\
\delta_T^{(G)} & \delta_T^{(B)}
\end{bmatrix}
\]

See Equation 20
No Free Lunch Equation

Use Bayes’ Rule:

\[
p(B \mid T, x, r) = \frac{p(T \mid B, x, r)}{p(T \mid x, r)} \frac{p(B \mid x, r)}{p(B \mid x, r)}
\]

Derivative with respect to loan rates:

\[
\frac{\partial}{\partial r} p(B \mid T, x, r) = \frac{\partial}{\partial r} \frac{p(T \mid B, x, r)p(B \mid x, r)}{p(T \mid x, r)}
\]

\[
= \frac{p(B \mid T, x, r)}{r} \left( \varepsilon_B^{(T)} - \varepsilon_T^{(T)} + \frac{\partial p(B \mid x, r)}{\partial r} \frac{r}{p(B \mid x, r)} \right)
\]

yields the risk-response exchange:

\[
\delta_T^{(B)} - \delta_B^{(B)} = \varepsilon_B^{(T)} - \varepsilon_T^{(T)} \tag{Equation 24}
\]

The change (deviation) in the price-risk elasticity for Bads among Takes equals the change (deviation) in the price-response elasticities for Takes among Bads.
Response Updates to Prior Risk Scores

See Equation 38

\[ s_p(r \mid T) \triangleq \ln \frac{p(G \mid T, x, r)}{p(B \mid T, x, r)} = s_p(r) + \ln \frac{q_G(x, r)}{q_B(x, r)} \]

See Equation 40

\[ \Delta s_p \triangleq s_p(r \mid T) - s_p(r \mid N) = \ln \frac{q_G(x, r)}{q_B(x, r)} - \ln \frac{1 - q_G(x, r)}{1 - q_B(x, r)} \]

May help explain the shift in default score distributions for Take/Not-Take populations in Figure 1
SUMMARY

Equality of Bayes’ Factors shows that Bad Adverse Selection of those who Take are equivalent to Increased Take rates for Bads.

Bayes’ factor of 1 defines NAS. Non-Take Inference is directly incorporated in estimates of number of Adverse Selections. This allows asymmetric information available to borrowers to be reflected in their preferences and differential responses.

Price-risk elasticities act in concert with price-response elasticities; there are six elasticities of each type and four conservation equations for exchange of risk and response elasticities. No Free Lunches.

Posterior Take/Non-Take Risk scores (and differences between differential Response Rates) can be estimated from weight of evidence on Take/Non-Take rates.