Outline

1. Reference Dependence: Mergers
2. Reference Dependence: Non-Bunching Papers
3. Reference Dependence: Labor Supply
4. Reference Dependence: Employment and Effort
5. Reference Dependence: Domestic Violence
6. Reference Dependence: Insurance
1 Reference Dependence: Mergers

• Baker, Pan, Wurgler (JFE 2012)

• On the appearance, very different set-up:
  – Firm A (Acquirer)
  – Firm T (Target)

• After negotiation, Firm A announces a price $P$ for merger with Firm T
  – Price $P$ typically at a 20-50 percent premium over current price
  – About 70 percent of mergers go through at price proposed
  – Comparison price for $P$ often used is highest price in previous 52 weeks, $P_{52}$
  – Example of how Cablevision (Target) trumpets deal
Figure 1. Slide from Cablevision Presentation to Shareholders, October 24, 2007. The management of Cablevision recommended acceptance of a $36.26 per share cash bid from the Dolan family. The slide compares this bid price to various recent prices including 52-week highs.

Valuation Achieved

Market Premia

- 179% higher than the lowest price during the 52-week period ended October 6, 2006
- 49% higher than the 52-week high during the period ended October 6, 2006
- 30% higher than the average closing price for the 180 days prior to the $36.26 May 2007 offer
- 10% higher than the 5-year and 52-week high prior to May 2, 2007
- Proposal

- $13.00*
- $24.26
- $27.90
- $32.86
- $36.26

* Adjusted to reflect payment of $1.00/share special dividend.
• Assume that Firm T chooses price $P$, and A decides accept reject

• As a function of price $P$, probability $p(P)$ that deal is accepted (depends on perception of values of synergy of A)

• If deal rejected, go back to outside value $\bar{U}$

• Then maximization problem is same as for housing sale:

\[
\max_P p(P)U(P) + (1 - p(P))\bar{U}
\]

• Can assume T reference-dependent with respect to

\[
v(P|P_0) = \begin{cases} 
P - P_{52} & \text{if } P \geq P_{52}; \\
\lambda (P - P_{52}) & \text{if } P < P_{52},
\end{cases}
\]
• Obtain same predictions as in housing market
• (This neglects possible reference dependence of A)
• Baker, Pan, and Wurgler (2009): Test reference dependence in mergers
  – Test 1: Is there bunching around $P_{52}$? (GM did not do this)
  – Test 2: Is there effect of $P_{52}$ on price offered?
  – Test 3: Is there effect on probability of acceptance?
  – Test 4: What do investors think? Use returns at announcement
• Test 1: Offer price $P$ around $P_{52}$
  – Some bunching, missing left tail of distribution
Notice that this does not tell us how the missing left tail occurs:

- Firms in left tail raise price to $P_{52}$?
- Firms in left tail wait for merger until 12 months after past peak, so $P_{52}$ is higher?
- Preliminary negotiations break down for firms in left tail

Would be useful to compare characteristics of firms to right and left of $P_{52}$
• Test 2: Kernel regression of price offered $P$ (Renormalized by price 30 days before, $P_{-30}$, to avoid heterosked.) on $P_{52}$:

$$100 \times \frac{P - P_{-30}}{P_{-30}} = \alpha + \beta \left[ 100 \times \frac{P_{52} - P_{-30}}{P_{-30}} \right] + \varepsilon$$
• Test 3: Probability of final acquisition is higher when offer price is above $P_{52}$ (Skip)

• Test 4: What do investors think of the effect of $P_{52}$?
  
  – Holding constant current price, investors should think that the higher $P_{52}$, the more expensive the Target is to acquire
  
  – Standard methodology to examine this:
    
    * 3-day stock returns around merger announcement: $CAR_{t-1,t+1}$
    * This assumes investor rationality
    
    * Notice that merger announcements are typically kept top secret until last minute $\rightarrow$ On announcement day, often big impact
Regression (Columns 3 and 5):

\[ CAR_{t-1,t+1} = \alpha + \beta \frac{P}{P_{-30}} + \varepsilon \]

where \( P/P_{-30} \) is instrumented with \( P_{52}/P_{-30} \)

Table 8. Mergers and Acquisitions: Market Reaction. Ordinary and two-stage least squares regressions of the 3-day CAR of the bidder on the offer premium.

<table>
<thead>
<tr>
<th>Offer Premium:</th>
<th>OLS 1</th>
<th>OLS 2</th>
<th>IV 3</th>
<th>OLS 4</th>
<th>IV 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offer Premium</td>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>-.0186***</td>
<td>-.0204***</td>
<td>-.215***</td>
<td>-.0443***</td>
<td>-.253***</td>
</tr>
<tr>
<td>(t)</td>
<td>(-2.64)</td>
<td>(-2.74)</td>
<td>(-3.48)</td>
<td>(-4.21)</td>
<td>(-4.39)</td>
</tr>
</tbody>
</table>

Results very supportive of reference dependence hypothesis – Also alternative anchoring story
2 Reference Dependence: Non-Bunching

• Previous papers had bunching implication
  – Some papers test for bunching (mergers, tax evasion, marathon running)
  – Some papers do not test it... but should! (housing)

• For bunching test, need
  – Reference point $r$ obvious enough to people AND researcher (house purchase price, zero taxes, round number goal)
  – Effort can be altered to get to reference point
• Next set of papers, these conditions do not apply:
  – Reference point $r$ not clear (labor supply, effort and crime, job search)
  – Choice is not about effort (domestic violence, insurance)

• Identification in these papers typically relies on variants of:
  – Loss aversion induces higher marginal utility of income to left of reference point
  – Identify comparing when to the left of reference point, versus to the right
  – Still need some model about reference point (more later on this)
3 Reference Dependence: Labor Supply

• Does reference dependence affect work/leisure decision?

• Framework:
  – effort $h$ (no. of hours)
  – hourly wage $w$
  – Returns of effort: $Y = w \times h$
  – Linear utility $U(Y) = Y$
  – Cost of effort $c(h) = \theta h^2 / 2$ convex within a day

• Standard model: Agents maximize

$$U(Y) - c(h) = wh - \frac{\theta h^2}{2}$$
• (Assumption that each day is orthogonal to other days – see below)

• Reference dependence: Threshold $T$ of earnings agent wants to achieve

• Loss aversion for outcomes below threshold:

$$U = \begin{cases} 
wh - T & \text{if } wh \geq T \\
\lambda (wh - T) & \text{if } wh < T
\end{cases}$$

with $\lambda > 1$ loss aversion coefficient

• Referent-dependent agent maximizes

$$wh - T - \frac{\theta h^2}{2} \quad \text{if } h \geq T/w$$

$$\lambda (wh - T) - \frac{\theta h^2}{2} \quad \text{if } h < T/w$$
• Derivative with respect to $h$:

$$w - \theta h \text{ if } h \geq T/w$$

$$\lambda w - \theta h \text{ if } h < T/w$$

1. Case 1 ($\lambda w - \theta T/w < 0$).

   - Optimum at $h^* = \lambda w/\theta < T/w$
2. Case 2 \((\lambda w - \theta T/w > 0 > w - \theta T/w)\)
   
   – Optimum at \(h^* = T/w\)

3. Case 3 \((w - \theta T/w > 0)\)
   
   – Optimum at \(h^* = w/\theta > T/w\)
• **Standard theory** \((\lambda = 1)\).

• Interior maximum: \(h^* = w/\theta\) (Cases 1 or 3)

• Labor supply

• Combine with labor demand: \(h^* = a - bw\), with \(a > 0, b > 0\).
• Optimum:

\[ L^S = \frac{w^*}{\theta} = a - bw^* = L^D \]

or

\[ w^* = \frac{a}{b + 1/\theta} \]

and

\[ h^* = \frac{a}{b\theta + 1} \]

• Comparative statics with respect to \( a \) (labor demand shock): \( a \uparrow \rightarrow h^* \uparrow \) and \( w^* \uparrow \)

• On low-demand days (low \( w \)) work less hard \( \rightarrow \) Save effort for high-demand days
• Model with reference dependence ($\lambda > 1$):
  
  – Case 1 or 3 still exist

  – BUT: Case 2. Kink at $h^* = T/w$ for $\lambda > 1$

  – Combine Labor supply with labor demand: $h^* = a - bw$, with $a > 0, b > 0$. 
• Case 2: Optimum:

\[ L^S = T/w^* = a - bw^* = L^D \]

and

\[ w^* = \frac{a + \sqrt{a^2 - 4Tb}}{2b} \]

• Comparative statics with respect to \( a \) (labor demand shock):
  - \( a \uparrow \rightarrow h^* \uparrow \) and \( w^* \uparrow \) (Cases 1 or 3)
  - \( a \uparrow \rightarrow h^* \downarrow \) and \( w^* \uparrow \) (Case 2)

• Case 2: On low-demand days (low \( w \)) need to work harder to achieve reference point \( T \rightarrow \) Work harder \( \rightarrow \) Opposite to standard theory

• (Neglected negligible wealth effects)
Camerer, Babcock, Loewenstein, and Thaler (QJE 1997)

- Data on daily labor supply of New York City cab drivers
  - 70 Trip sheets, 13 drivers (TRIP data)
  - 1044 summaries of trip sheets, 484 drivers, dates: 10/29-11/5, 1990 (TLC1)
  - 712 summaries of trip sheets, 11/1-11/3, 1988 (TLC2)

- Notice data feature: Many drivers, few days in sample
• Analysis in paper neglects wealth effects: Higher wage today $\rightarrow$ Higher lifetime income

• Justification:
  – Correlation of wages across days close to zero
  – Each day can be considered in isolation
  – $\rightarrow$ Wealth effects of wage changes are very small

• Test:
  – Assume variation across days driven by $\Delta a$ (labor demand shifter)
  – Do hours worked $h$ and $w$ co-vary positively (standard model) or negatively?
- Raw evidence
• Estimated Equation:

\[ \log(h_{i,t}) = \alpha + \beta \log\left(\frac{Y_{i,t}}{h_{i,t}}\right) + X_{i,t}\Gamma + \varepsilon_{i,t}. \]

• Estimates of \( \hat{\beta} \):
  
  - \( \hat{\beta} = -0.186 \) (s.e. 129) – TRIP with driver f.e.
  
  - \( \hat{\beta} = -0.618 \) (s.e. .051) – TLC1 with driver f.e.
  
  - \( \hat{\beta} = -0.355 \) (s.e. .051) – TLC2

• Estimate is not consistent with prediction of standard model

• Indirect support for income targeting
• Issues with paper:

• Economic issue 1. Reference-dependent model does not predict (log-) linear, negative relation

• What happens if reference income is stochastic? (Koszegi-Rabin, 2006)
- Econometric issue 1. Division bias in regressing hours on log wages
- Wages is not directly observed — Computed at $Y_{i,t}/h_{i,t}$
- Assume $h_{i,t}$ measured with noise: $\tilde{h}_{i,t} = h_{i,t} \ast \phi_{i,t}$. Then,
  \[
  \log(\tilde{h}_{i,t}) = \alpha + \beta \log\left(\frac{Y_{i,t}}{\tilde{h}_{i,t}}\right) + \varepsilon_{i,t}.
  \]
  becomes
  \[
  \log(h_{i,t}) + \log(\phi_{i,t}) = \alpha + \beta \left[\log(Y_{i,t}) - \log(h_{i,t})\right] - \beta \log(\phi_{i,t}) + \varepsilon_{i,t}.
  \]
- Downward bias in estimate of $\hat{\beta}$
- Response: instrument wage using other workers’ wage on same day
• IV Estimates:

<table>
<thead>
<tr>
<th>Sample</th>
<th>TRIP</th>
<th>TLC1</th>
<th>TLC2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log hourly wage</td>
<td>-.319</td>
<td>-.313</td>
<td>-.926</td>
</tr>
<tr>
<td></td>
<td>(.298)</td>
<td>(.236)</td>
<td>(.259)</td>
</tr>
<tr>
<td>High temperature</td>
<td>-.000</td>
<td>.002</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.002)</td>
<td>(.002)</td>
</tr>
</tbody>
</table>

• Notice: First stage not very strong (and few days in sample)
Econometric issue 2. Are the authors really capturing demand shocks or supply shocks?

- Assume $\theta$ (disutility of effort) varies across days.
- Even in standard model we expect negative correlation of $h_{i,t}$ and $w_{i,t}$

- Camerer et al. argue for plausibility of shocks due to $a$ rather than $\theta$
• Farber (JPE, 2005)
• Re-Estimate Labor Supply of Cab Drivers on new data
• Address Econometric Issue 1
• Data:
  – Daily summary not available (unlike in Camerer et al.)
  – Notice: Few drivers, many days in sample
• First, replication of Camerer et al. (1997)

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>4.012</td>
<td>3.924</td>
<td>3.778</td>
</tr>
<tr>
<td></td>
<td>(.349)</td>
<td>(.379)</td>
<td>(.381)</td>
</tr>
<tr>
<td>Log(wage)</td>
<td>-.688</td>
<td>-.685</td>
<td>-.637</td>
</tr>
<tr>
<td></td>
<td>(.111)</td>
<td>(.114)</td>
<td>(.115)</td>
</tr>
<tr>
<td>Day shift</td>
<td>...</td>
<td>.011</td>
<td>.134</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.040)</td>
<td>(.062)</td>
</tr>
<tr>
<td>Minimum temperature &lt; 30</td>
<td>...</td>
<td>.126</td>
<td>.024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.053)</td>
<td>(.058)</td>
</tr>
<tr>
<td>Maximum temperature ≥ 80</td>
<td>...</td>
<td>.041</td>
<td>.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.055)</td>
<td>(.064)</td>
</tr>
<tr>
<td>Rainfall</td>
<td>...</td>
<td>-.022</td>
<td>-.054</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.073)</td>
<td>(.071)</td>
</tr>
<tr>
<td>Snowfall</td>
<td>...</td>
<td>-.096</td>
<td>-.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.036)</td>
<td>(.035)</td>
</tr>
<tr>
<td>Driver effects</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Day-of-week effects</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.068</td>
<td>.098</td>
<td>.198</td>
</tr>
</tbody>
</table>

• Farber (2005) however cannot replicate the IV specification (too few drivers on a given day)
• Key specification: Hazard model that does not suffer from division bias
  – Dependent variable is dummy \( \text{Stop}_{i,t} = 1 \) if driver \( i \) stops at hour \( t \):
    \[
    \text{Stop}_{i,t} = \Phi \left( \alpha + \beta \gamma Y_{i,t} + \beta_h h_{i,t} + \Gamma X_{i,t} \right)
    \]
  – Control for hours worked so far \( (h_{i,t}) \) and other controls \( X_{i,t} \)
• Does a higher earned income \( Y_{i,t} \) increase probability of stopping \( (\beta > 0) \)?

<table>
<thead>
<tr>
<th>Variable</th>
<th>( X^* )</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total hours</td>
<td>8.0</td>
<td>.013</td>
<td>.087</td>
<td>.111</td>
<td>.010</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.012)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Waiting hours</td>
<td>2.5</td>
<td>.010</td>
<td>-.005</td>
<td>.001</td>
<td>.004</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.012)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Break hours</td>
<td>.5</td>
<td>.006</td>
<td>-.015</td>
<td>-.003</td>
<td>-.001</td>
<td>-.002</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.011)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Shift income ( \div ) 100</td>
<td>1.5</td>
<td>.053</td>
<td>.036</td>
<td>.014</td>
<td>.016</td>
<td>.011</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.030)</td>
<td>(0.015)</td>
<td>(0.016)</td>
<td>(0.015)</td>
<td></td>
</tr>
</tbody>
</table>

| Driver (21)         | no       | yes  | yes  | yes  | yes  | yes  |
| Day of week (7)     | no       | no   | yes  | yes  | yes  | yes  |
| Hour of day (19)    | 2:00 p.m.| no   | no   | yes  | yes  | yes  |
| Log likelihood      | -2,092.9 | -1,965.0 | -1,789.5 | -1,784.7 | -1,767.6 |

Note.—The sample includes 18,461 trips in 584 shifts for 21 drivers. Probit estimates are normalized to reflect the marginal effect at \( X^* \) of \( X \) on the probability of stopping. The normalized probit estimate is \( \beta \cdot \phi(X^* \gamma) \), where \( \phi(\cdot) \) is the standard normal density. The values of \( X^* \) chosen for the fixed effects are equally weighted for each day of the week and for each driver. The hours from 5:00 a.m. to 10:00 a.m. have a common fixed effect. The evaluation point is after 5.5 driving hours, 2.5 waiting hours, and 0.5 break hour in a dry hour on a day with moderate temperatures in midtown Manhattan at 2:00 p.m. Robust standard errors accounting for clustering by shift are reported in parentheses.
• Positive, but not significant effect of $Y_{i,t}$ on probability of stopping:
  
  – 10 percent increase in $Y$ ($15) → 1.6 percent increase in stopping prob. (.225 pctg. pts. increase in stopping prob. out of average 14 pctg. pts.) → .16 elasticity

  – Cannot reject large effect: 10 pct. increase in $Y$ increase stopping prob. by 6 percent

• Qualitatively consistent with income targeting

• Also notice:
  
  – Failure to reject standard model is not the same as rejecting alternative model (reference dependence)

  – Alternative model is not spelled out
• Final step in Farber (2005): Re-analysis of Camerer et al. (1997) data with hazard model
  – Use only TRIP data (small part of sample)
  – No significant evidence of effect of past income $Y$
  – However: Cannot reject large positive effect

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>Driver</th>
<th>4</th>
<th>10</th>
<th>16</th>
<th>18</th>
<th>20</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours</td>
<td></td>
<td>.073</td>
<td>.056</td>
<td>.043</td>
<td>.010</td>
<td>.195</td>
<td>.198</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.060)</td>
<td>(.047)</td>
<td>(.015)</td>
<td>(.007)</td>
<td>(.045)</td>
<td>(.030)</td>
</tr>
<tr>
<td>Income+100</td>
<td></td>
<td>.178</td>
<td>.089</td>
<td>.064</td>
<td>.048</td>
<td>-.160</td>
<td>-.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.167)</td>
<td>(.059)</td>
<td>(.041)</td>
<td>(.020)</td>
<td>(.123)</td>
<td>(.150)</td>
</tr>
<tr>
<td>Number of shifts</td>
<td>40</td>
<td>45</td>
<td>70</td>
<td>72</td>
<td>46</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Number of trips</td>
<td>884</td>
<td>912</td>
<td>1,754</td>
<td>2,023</td>
<td>1,125</td>
<td>882</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-124.1</td>
<td>-116.0</td>
<td>-221.1</td>
<td>-260.6</td>
<td>-123.4</td>
<td>-116.9</td>
<td></td>
</tr>
</tbody>
</table>
• Farber (2005) cannot address the Econometric Issue 2: Is it Supply or Demand that Varies

• **Fehr and Goette (AER 2007).** Experiments on Bike Messengers

• Use explicit randomization to deal with Econometric Issues 1 and 2

• Combination of:
  – *Experiment 1.* Field Experiment shifting wage and
  – *Experiment 2.* Lab Experiment (relate to evidence on loss aversion)...
  – ... on the same subjects

• Slides courtesy of Lorenz Goette
• Other work:

• **Farber (AER 2008)** goes beyond Farber (JPE, 2005) and attempts to estimate model of labor supply with loss-aversion

  – Estimate loss-aversion \( \delta \)

  – Estimate (stochastic) reference point \( T \)

• Same data as Farber (2005)

• Results:

  – significant loss aversion \( \delta \)

  – however, large variation in \( T \) mitigates effect of loss-aversion
- $\delta$ is loss-aversion parameter

- Reference point: mean $\theta$ and variance $\sigma^2$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$ (contprob)</td>
<td>-0.691</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>(0.243)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\theta}$ (mean ref inc)</td>
<td>159.02</td>
<td>206.71</td>
<td>250.86</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>(4.99)</td>
<td>(7.99)</td>
<td>(16.47)</td>
<td></td>
</tr>
<tr>
<td>$\hat{\delta}$ (cont increment)</td>
<td>3.40</td>
<td>5.35</td>
<td>4.85</td>
<td>5.38</td>
</tr>
<tr>
<td></td>
<td>(0.279)</td>
<td>(0.573)</td>
<td>(0.711)</td>
<td>(0.545)</td>
</tr>
<tr>
<td>$\hat{\sigma}^2$ (ref inc var)</td>
<td>3199.4</td>
<td>10440.0</td>
<td>15944.3</td>
<td>8236.2</td>
</tr>
<tr>
<td></td>
<td>(294.0)</td>
<td>(1660.7)</td>
<td>(3652.1)</td>
<td>(1222.2)</td>
</tr>
</tbody>
</table>

| Driver $\theta_i$ (15) | No | No | No | Yes |
| Vars in Cont Prob       |    |    |    |     |
| Driver FE's (14)        | No | No | Yes| No  |
| Accum Hours (7)         | No | Yes| Yes| Yes |
| Weather (4)             | No | Yes| Yes| Yes |
| Day Shift and End (2)   | No | Yes| Yes| Yes |
| Location (1)            | No | Yes| Yes| Yes |
| Day-of-Week (6)         | No | Yes| Yes| Yes |
| Hour-of-Day (18)        | No | Yes| Yes| Yes |
| Log(L)                   | -1867.8 | -1631.6 | -1572.8 | -1606.0 |
| NumberParms              | 4  | 43  | 57  | 57  |
• **Crawford and Meng (AER 2011)**

• Re-estimates the Farber paper allowing for two dimensions of reference dependence:
  - Hours (loss if work more hours than $\bar{h}$)
  - Income (loss if earn less than $\bar{Y}$)

• Re-estimates Farber (2005) data for:
  - Wage above average (income likely to bind)
  - Wages below average (hours likely to bind)

• Perhaps, reconciling Camerer et al. (1997) and Farber (2005)
  - $w > w^e$: hours binding $\rightarrow$ hours explain stopping
  - $w < w^e$: income binding $\rightarrow$ income explains stopping
### Table 1: Probability of Stopping: Probit Model with Linear Effect

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) Pooled data</th>
<th></th>
<th>(2) Pooled data</th>
<th></th>
<th>(3) Pooled data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$v^g &gt; w^g$</td>
<td>$v^g \leq w^g$</td>
<td>$v^g &gt; w^g$</td>
<td>$v^g \leq w^g$</td>
<td>$v^g &gt; w^g$</td>
<td>$v^g \leq w^g$</td>
</tr>
<tr>
<td>Total hours</td>
<td>.013</td>
<td>.005</td>
<td>.016</td>
<td>.010</td>
<td>.003</td>
<td>.011</td>
</tr>
<tr>
<td></td>
<td>(.009)**</td>
<td>(.007)**</td>
<td>(.003)**</td>
<td>(.008)**</td>
<td>(.006)**</td>
<td>(.005)**</td>
</tr>
<tr>
<td>Waiting hours</td>
<td>.010</td>
<td>.007</td>
<td>.016</td>
<td>.001</td>
<td>.001</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>(.003)**</td>
<td>(.007)</td>
<td>(.001)**</td>
<td>(.009)</td>
<td>(.012)</td>
<td>(.004)</td>
</tr>
<tr>
<td>Break hours</td>
<td>.006</td>
<td>.005</td>
<td>.004</td>
<td>-.003</td>
<td>-.006</td>
<td>-.003</td>
</tr>
<tr>
<td></td>
<td>(.003)**</td>
<td>(.001)**</td>
<td>(.008)</td>
<td>(.006)</td>
<td>(.009)</td>
<td>(.004)</td>
</tr>
<tr>
<td>Income/100</td>
<td>.053</td>
<td>.076</td>
<td>.055</td>
<td>.013</td>
<td>.045</td>
<td>.009</td>
</tr>
<tr>
<td></td>
<td>(.000)**</td>
<td>(.001)**</td>
<td>(.007)**</td>
<td>(.010)</td>
<td>(.019)**</td>
<td>(.024)</td>
</tr>
<tr>
<td>Min temp&lt;30</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Max temp&gt;80</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hourly rain</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Daily snow</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Location dummies</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Driver dummies</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Day of week</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hour of day</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-2039.2</td>
<td>-1148.4</td>
<td>-882.6</td>
<td>-1789.5</td>
<td>-1003.8</td>
<td>-753.4</td>
</tr>
<tr>
<td>Pseudo R2</td>
<td>0.1516</td>
<td>0.1555</td>
<td>0.1533</td>
<td>0.2555</td>
<td>0.2618</td>
<td>0.2773</td>
</tr>
<tr>
<td>Observation</td>
<td>13461</td>
<td>7936</td>
<td>5525</td>
<td>13461</td>
<td>7936</td>
<td>5525</td>
</tr>
</tbody>
</table>
4 Reference Dependence: Employment and Effort

- Back to labor markets: Do reference points affect performance?

- Mas (QJE 2006) examines police performance

- Exploits quasi-random variation in pay due to arbitration

- Background
  - 60 days for negotiation of police contract \(\rightarrow\) If undecided, arbitration
  - 9 percent of police labor contracts decided with final offer arbitration
• Framework:

  - pay is \( w \times (1 + r) \)
  
  - union proposes \( r_u \), employer proposes \( r_e \), arbitrator prefers \( r_a \)
  
  - arbitrator chooses \( r_e \) if \( |r_e - r_a| \leq |r_u - r_a| \)
  
  - \( P(r_e, r_u) \) is probability that arbitrator chooses \( r_e \)
  
  - Distribution of \( r_a \) is common knowledge (cdf \( F' \))
  
  - Assume \( r_e \leq r_a \leq r_u \) \( \Rightarrow \) Then

\[
P = P\left( r_a - r_e \leq r_u - r_a \right) = P\left( r_a \leq \frac{r_u + r_e}{2} \right) = F\left( \frac{r_u + r_e}{2} \right)
\]
• Nash Equilibrium:

- If $r_a$ is certain, Hotelling game: convergence of $r_e$ and $r_u$ to $r_a$

- Employer’s problem:

$$\max_{r_e} PU (w (1 + r_e)) + (1 - P) U (w (1 + r^*_u))$$

- Notice: $U' < 0$

- First order condition (assume $r_u \geq r_e$):

$$\frac{P'}{2} [U (w (1 + r^*_e)) - U (w (1 + r^*_u))] + PU' (w (1 + r^*_e)) w = 0$$

- $r^*_e = r^*_u$ cannot be solution \(\Rightarrow\) Lower $r_e$ and increase utility ($U' < 0$)
− Union’s problem: maximizes

\[ \max_{r_u} PV \left( w (1 + r_u^*) \right) + (1 - P) V \left( w (1 + r_u) \right) \]

− Notice: \( V' > 0 \)

− First order condition for union:

\[
\frac{P'}{2} [V (w (1 + r_u^*)) - V (w (1 + r_u^*))] + (1 - P) V' (w (1 + r_u^*)) w = 0
\]

− To simplify, assume \( U(x) = -bx \) and \( V(x) = bx \)

− This implies

\[
V (w (1 + r_u^*)) - V (w (1 + r_u^*)) = -U (w (1 + r_e^*)) - U (w (1 + r_u^*)) \Rightarrow
\]

\[
-bP^*w = - (1 - P^*) bw
\]
– Result: $P^* = 1/2$

• Prediction (i) in Mas (2006): “If disputing parties are equally risk-averse, the winner in arbitration is determined by a coin toss.”

• Therefore, as-if random assignment of winner

• Use to study impact of pay on police effort

• Data:
  – 383 arbitration cases in New Jersey, 1978-1995
  – Observe offers submitted $r_e$, $r_u$, and ruling $\bar{r}_a$
  – Match to UCR crime clearance data (number of crimes solved by arrest)
• Compare summary statistics of cases when employer and when police wins

• Estimated $\hat{P} = .344 \neq 1/2 \rightarrow$ Unions more risk-averse than employers

• No systematic difference between Union and Employer cases except for $r_e$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arbitrator rules for employer</strong></td>
<td>0.344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Final Offer: Employer</strong></td>
<td>6.11</td>
<td>6.44</td>
<td>5.94</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>[1.65]</td>
<td>[1.54]</td>
<td>[1.68]</td>
<td>(0.18)</td>
</tr>
<tr>
<td><strong>Final Offer: Union</strong></td>
<td>7.65</td>
<td>7.87</td>
<td>7.54</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>[1.71]</td>
<td>[2.03]</td>
<td>[1.51]</td>
<td>(0.18)</td>
</tr>
<tr>
<td><strong>Population</strong></td>
<td>21,345</td>
<td>22,893</td>
<td>20,534</td>
<td>2,358</td>
</tr>
<tr>
<td></td>
<td>[33,463]</td>
<td>[34,561]</td>
<td>[32,915]</td>
<td>(3,598)</td>
</tr>
<tr>
<td><strong>Contract length</strong></td>
<td>2.09</td>
<td>2.09</td>
<td>2.09</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>[0.66]</td>
<td>[0.64]</td>
<td>[0.66]</td>
<td>(0.071)</td>
</tr>
<tr>
<td><strong>Size of bargaining unit</strong></td>
<td>42.58</td>
<td>41.36</td>
<td>43.22</td>
<td>-1.85</td>
</tr>
<tr>
<td></td>
<td>[97.34]</td>
<td>[53.33]</td>
<td>[113.84]</td>
<td>(15.66)</td>
</tr>
<tr>
<td><strong>Arbitration year</strong></td>
<td>85.56</td>
<td>85.85</td>
<td>85.41</td>
<td>0.436</td>
</tr>
<tr>
<td></td>
<td>[4.75]</td>
<td>[5.10]</td>
<td>[4.56]</td>
<td>(0.510)</td>
</tr>
<tr>
<td><strong>Clearances per 100,000 capita</strong></td>
<td>120.31</td>
<td>122.28</td>
<td>118.57</td>
<td>3.71</td>
</tr>
<tr>
<td></td>
<td>[106.65]</td>
<td>[108.76]</td>
<td>[104.35]</td>
<td>(9.46)</td>
</tr>
</tbody>
</table>
- Graphical evidence of effect of ruling on crime clearance rate

- Significant effect on clearance rate for one year after ruling

- Estimate of the cumulated difference between Employer and Union cities on clearance rates and crime
- Arbitration leads to an average increase of 15 clearances out of 100,000 each month
- Effects on crime rate more imprecise

<table>
<thead>
<tr>
<th></th>
<th>All crime</th>
<th>Violent crime</th>
<th>Property crime</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Constant</td>
<td>612.18</td>
<td>150.26</td>
<td>461.81</td>
</tr>
<tr>
<td>(63.98)</td>
<td>(23.23)</td>
<td>(42.00)</td>
<td></td>
</tr>
<tr>
<td>Post-arbitration</td>
<td>26.86</td>
<td>24.68</td>
<td>19.19</td>
</tr>
<tr>
<td>× Employer win</td>
<td>(25.29)</td>
<td>(14.68)</td>
<td>(18.17)</td>
</tr>
<tr>
<td>Post-arbitration</td>
<td>7.64</td>
<td>6.68</td>
<td>0.170</td>
</tr>
<tr>
<td>× Union win</td>
<td>(16.24)</td>
<td>(11.42)</td>
<td>(11.68)</td>
</tr>
<tr>
<td>Row 3 – Row 2</td>
<td>-19.21</td>
<td>-18.01</td>
<td>-19.02</td>
</tr>
<tr>
<td>(30.06)</td>
<td>(19.12)</td>
<td>(9.56)</td>
<td>(21.60)</td>
</tr>
<tr>
<td>Employer Win (Yes = 1)</td>
<td>-31.81</td>
<td>-20.43</td>
<td>-11.35</td>
</tr>
<tr>
<td>(84.42)</td>
<td>(27.57)</td>
<td>(59.50)</td>
<td></td>
</tr>
</tbody>
</table>

Fixed-effects? Yes

Mean of the dependent variable

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>444.03</td>
<td>95.49</td>
<td>98.26</td>
<td>348.45</td>
</tr>
<tr>
<td>[364.23]</td>
<td>[363.76]</td>
<td>[292.10]</td>
<td>[1865.8]</td>
</tr>
</tbody>
</table>

Sample size

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>9.528</td>
</tr>
</tbody>
</table>

$R^2$

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
</tr>
</tbody>
</table>


Do reference points matter?

Plot impact on clearances rates \((12, -12)\) as a function of \(\bar{r}_a - \left(\frac{r_e + r_u}{2}\right)\)

Figure V
Estimated expected change in clearances conditional on the deviation of the award from the average of the offers
• Effect of loss is larger than effect of gain

<table>
<thead>
<tr>
<th>Table VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneous effects of arbitration decisions on clearances by loss size, award, and deviation from the expected offer; -12 to +12 month event time window</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5) Police lose</th>
<th>(6) Police win</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(2.31)</td>
<td>(9.58)</td>
<td>(8.45)</td>
<td>(4.76)</td>
<td>(3.14)</td>
<td>(4.17)</td>
</tr>
<tr>
<td>Post-Arbitration × Award</td>
<td>1.23</td>
<td>-1.00</td>
<td>0.23</td>
<td>-0.83</td>
<td>-0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.16)</td>
<td>(0.98)</td>
<td>(1.89)</td>
<td>(4.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Arbitration × Loss size</td>
<td>-10.31</td>
<td>-10.93</td>
<td>-10.31</td>
<td>-10.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.59)</td>
<td>(1.89)</td>
<td>(4.54)</td>
<td>(4.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Arbitration × Union win</td>
<td>13.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Arbitration × (expected award-award)</td>
<td>-17.72</td>
<td>2.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.94)</td>
<td>(4.13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Arbitration × p(loss size)²</td>
<td>Included</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample Size</td>
<td>59,137</td>
<td>59,137</td>
<td>59,137</td>
<td>59,137</td>
<td>52,857</td>
<td>55,879</td>
</tr>
<tr>
<td>R²</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.60</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Standard errors, clustered on the intersection of arbitration window and city, are in parentheses. Standard deviations are in brackets. Observations are municipality × month cells. The sample is weighted by population size in 1976. The dependent variable is clearances per 100,000 capita. Loss size is defined as the union demand (percent increase on previous wage) less the arbitrator's award. Amongst cities that underwent arbitration, the mean loss size is 0.489 with a standard deviation of 0.953. The expected award is the mathematical expectation of the award given the union and employer offers and the predicted probability of an employer win. The predicted probability of an employer win is estimated with a probit model using as predictors year of arbitration dummies, the average of the final offers, log population, and the length of the contract. See text for details. The samples in models (1)-(4) consist of the 12 months before to the 12 months after arbitration, for jurisdictions that underwent arbitration, as well as all jurisdictions that never underwent arbitration for all months between 1978 and 1996. The sample in model (5) consists of cities where the union lost in arbitration and the comparison group of non-arbitrating cities. The sample in model (6) consists of cities where the union won in arbitration and the comparison group of non-arbitrating cities. All models include month × year effects (252), arbitration window effects (383), and city effects (432). Author's calculation based on NJ PERC arbitration cases matched to monthly municipal clearance rates at the jurisdiction level from FBI Uniform Crime Reports.
• Column (3): Effect of a gain relative to \((r_e + r_u)/2\) is not significant; effect of a loss is

• Columns (5) and (6): Predict expected award \(\hat{r}_a\) using covariates, then compute \(\bar{r}_a - \hat{r}_a\)
  
  – \(\bar{r}_a - \hat{r}_a\) does not matter if union wins
  
  – \(\bar{r}_a - \hat{r}_a\) matters a lot if union loses

• Assume policeman maximizes

\[
\max_e [\bar{U} + U(w)] e - \theta \frac{e^2}{2}
\]

where

\[
U(w) = \begin{cases} 
  w - \hat{w} & \text{if } w \geq \hat{w} \\
  \lambda (w - \hat{w}) & \text{if } w < \hat{w}
\end{cases}
\]
• Reduced form of reciprocity model where altruism towards the city is a function of how nice the city was to the policemen ($\bar{U} + U (w)$)

• F.o.c.:

$$\bar{U} + U (w) - \theta e = 0$$

Then

$$e^* (w) = \frac{\bar{U}}{\theta} + \frac{1}{\theta} U (w)$$

• It implies that we would estimate

$$Clearances = \alpha + \beta (\bar{r}_a - \hat{r}_a) + \gamma (\bar{r}_a - \hat{r}_a) \mathbf{1} (\bar{r}_a - \hat{r}_a < 0) + \varepsilon$$

with $\beta > 0$ (also in standard model) and $\gamma > 0$ (not in standard model)
• Compare to observed pattern

• Close to predictions of model
5 Reference Dependence: Domestic Violence

• Consider a man in conflicted relationship with the spouse

• What is the effect of an event such as the local football team losing or winning a game?

• With probability $h$ the man loses control and becomes violent
  
  – Assume $h = h(u)$ with $h' < 0$ and $u$ the underlying utility
  
  – Denote by $p$ the ex-ante expectation that the team wins
  
  – Denote by $u(W)$ and $u(L)$ the consumption utility of a loss
  
  – Using a Koszegi-Rabin specification, then ex-post the utility from a win
is

\[ U (W|p) = u(W) \text{ [consumption utility]} + p [0] + (1 - p) \eta [u(W) - u(L)] \text{ [gain-loss utility]} \]

- Similarly, the utility from a loss is

\[ U (L|p) = u(L) + (1 - p) [0] - \lambda p \eta [u(W) - u(L)] \]

- Implication:

\[ \frac{\partial U (L|p)}{\partial p} = -\lambda \eta [u(W) - u(L)] < 0 \]

- The more a win is expected, the more a loss is painful \(\rightarrow\) the more likely it is to trigger violence

- The (positive) effect of a gain is higher the more unexpected (lower \(p\))
• Card and Dahl (QJE 2011) test these predictions using a data set of:
  – Domestic violence (NIBRS)
  – Football matches by State
  – Expected win probability from Las Vegas predicted point spread

• Separate matches into
  – Predicted win (+3 points of spread)
  – Predicted close
  – Predicted loss (-3 points)
Table 4. Emotional Shocks from Football Games and Male-on-Female Intimate Partner Violence Occurring at Home, Poisson Regressions.

<table>
<thead>
<tr>
<th>Coefficient Estimates</th>
<th>Intimate Partner Violence, Male on Female, at Home</th>
<th>Baseline Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss * Predicted Win (Upset Loss)</td>
<td>.083 (.026)</td>
<td>.077 (.027)</td>
<td>.080 (.027)</td>
<td>.074 (.028)</td>
<td>.076 (.028)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss * Predicted Close (Close Loss)</td>
<td>.031 (.023)</td>
<td>.034 (.024)</td>
<td>.036 (.024)</td>
<td>.024 (.025)</td>
<td>.026 (.025)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Win * Predicted Loss (Upset Win)</td>
<td>-.002 (.027)</td>
<td>.011 (.027)</td>
<td>.021 (.028)</td>
<td>.013 (.029)</td>
<td>.011 (.029)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted Win</td>
<td>-.004 (.022)</td>
<td>-.019 (.032)</td>
<td>-.015 (.032)</td>
<td>.000 (.033)</td>
<td>-.068 (.044)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted Close</td>
<td>-.012 (.023)</td>
<td>-.017 (.032)</td>
<td>-.016 (.032)</td>
<td>-.007 (.034)</td>
<td>-.074 (.044)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted Loss</td>
<td>-.000 (.022)</td>
<td>-.004 (.031)</td>
<td>-.011 (.031)</td>
<td>.006 (.033)</td>
<td>-.057 (.042)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-game Day</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nielsen Rating</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>.009 (.004)</td>
<td>---</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Municipality fixed effects | X | X | X | X | X |
Year, week, & holiday dummies | X | X | X | X | X |
Weather variables | X | X | X | X | X |
Nielsen Data Sub-sample | X | X | X | X | X |

Number of Municipalities | 765 | 765 | 765 | 749 | 749 |
Observations | 77,520 | 77,520 | 77,520 | 71,798 | 71,798 |
• Findings:
  1. Unexpected loss increase domestic violence
  2. No effect of expected loss
  3. No effect of unexpected win, if anything increases violence

• Findings 1-2 consistent with ref. dep. and 3 partially consistent (given that violence is a function is very negative utility)

• Other findings:
  – Effect is larger for more important games
  – Effect disappears within a few hours of game end $\rightarrow$ Emotions are transient
  – No effect on violence of females on males
6 Reference Dependence: Insurance

- Much of the laboratory evidence on prospect theory is on risk taking.

- Field evidence considered so far (mostly) does not involve risk:
  - House Sale
  - Merger Offer

- Field evidence on risk taking?

- Sydnor (AEJ Applied, 2010) on deductible choice in the life insurance industry

- Menu Choice as identification strategy as in DellaVigna and Malmendier (2006)

- Slides courtesy of Justin Sydnor
Dataset

- 50,000 Homeowners-Insurance Policies
  - 12% were new customers
- Single western state
- One recent year (post 2000)
- Observe
  - Policy characteristics including deductible
    - 1000, 500, 250, 100
  - Full available deductible-premium menu
  - Claims filed and payouts by company
Features of Contracts

- Standard homeowners-insurance policies (no renters, condominiums)
- Contracts differ only by deductible
- Deductible is *per claim*
- No experience rating
  - Though underwriting practices not clear
- Sold through agents
  - Paid commission
  - No “default” deductible
- Regulated state
## Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Full Sample</th>
<th>1000</th>
<th>500</th>
<th>250</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insured home value</td>
<td>206,917</td>
<td>266,461</td>
<td>205,026</td>
<td>180,895</td>
<td>164,485</td>
</tr>
<tr>
<td></td>
<td>(91,178)</td>
<td>(127,773)</td>
<td>(81,834)</td>
<td>(65,089)</td>
<td>(53,808)</td>
</tr>
<tr>
<td>Yearly premium paid</td>
<td>719.80</td>
<td>798.60</td>
<td>715.60</td>
<td>687.19</td>
<td>709.78</td>
</tr>
<tr>
<td></td>
<td>(312.76)</td>
<td>(405.78)</td>
<td>(300.39)</td>
<td>(267.82)</td>
<td>(269.34)</td>
</tr>
<tr>
<td>Number of years insured by the company</td>
<td>8.4</td>
<td>5.1</td>
<td>5.8</td>
<td>13.5</td>
<td>12.8</td>
</tr>
<tr>
<td></td>
<td>(7.1)</td>
<td>(5.6)</td>
<td>(5.2)</td>
<td>(7.0)</td>
<td>(6.7)</td>
</tr>
<tr>
<td>Average age of H.H. members</td>
<td>53.7</td>
<td>50.1</td>
<td>50.5</td>
<td>59.8</td>
<td>66.6</td>
</tr>
<tr>
<td></td>
<td>(15.8)</td>
<td>(14.5)</td>
<td>(14.9)</td>
<td>(15.9)</td>
<td>(15.5)</td>
</tr>
<tr>
<td>Number of paid claims in sample year (claim rate)</td>
<td>0.042</td>
<td>0.025</td>
<td>0.043</td>
<td>0.049</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.17)</td>
<td>(0.22)</td>
<td>(0.23)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>N</td>
<td>49,992</td>
<td>8,525</td>
<td>23,782</td>
<td>17,536</td>
<td>149</td>
</tr>
<tr>
<td>Percent of sample</td>
<td>100%</td>
<td>17.05%</td>
<td>47.57%</td>
<td>35.08%</td>
<td>0.30%</td>
</tr>
</tbody>
</table>

* Means with standard errors in parentheses.
Deductible Pricing

- \( X_i = \) matrix of policy characteristics
- \( f(X_i) = \) "base premium"
  - Approx. linear in home value
- Premium for deductible \( D \)
  - \( P_i^D = \delta_D f(X_i) \)
- Premium differences
  - \( \Delta P_i = \Delta \delta f(X_i) \)
- \( \Rightarrow \) Premium differences depend on base premiums (insured home value).
### Premium-Deductible Menu

<table>
<thead>
<tr>
<th>Available Deductible</th>
<th>Full Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>$615.82</td>
</tr>
<tr>
<td>(292.59)</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>+99.91</td>
</tr>
<tr>
<td>(45.82)</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>+86.59</td>
</tr>
<tr>
<td>(39.71)</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>+133.22</td>
</tr>
<tr>
<td>(61.09)</td>
<td></td>
</tr>
</tbody>
</table>

**Risk Neutral Claim Rates?**

- 100/500 = 20%
- 87/250 = 35%
- 133/150 = 89%

* Means with standard deviations in parentheses
The curves in the upper graphs are fan locally-weighted kernel regressions using a quartic kernel. The dashed lines give 95% confidence intervals calculated using a bootstrap procedure with 200 repetitions. The range for additional premium covers 98% of the available data.

The graph in the upper left gives the fraction that chose either the $250 or $500 deductibles versus the additional premium an individual faced to move from a $1000 to the $500 deductible. The graph in the upper right represents the average expected savings from switching to the $1000 deductible for customers facing a given premium difference. The potential savings is calculated at the individual level and then the kernel regressions are run. Because they filed no claims, for most customers this measure is simply the premium reductions they would have seen with the $1000 deductible. For the roughly 4% of customers who filed claims the potential savings is typically negative.

What if the x-axis were insured home value?
The graph in the upper left gives the fraction that chose either the $250 or $500 deductibles as a function of the insured home value. The graph in the upper right represents the average expected savings from switching to the $1000 deductible for customers who chose one of the lower deductibles. The potential savings is calculated at the individual level and then the kernel regressions are run. Because they filed no claims, for most customers this measure is simply the premium reductions they would have seen with the $1000 deductible. For the roughly 4% of customers who filed claims the potential savings is typically negative.

The curves in the upper graphs are fitted locally-weighted kernel regressions using a quartic kernel. The dashed lines give 95% confidence intervals calculated using a bootstrap procedure with 200 repetitions.

The range for insured home value covers 99% of the available data.
## Potential Savings with 1000 Ded

### Claim rate?

### Value of lower deductible?

### Additional premium?

### Potential savings?

<table>
<thead>
<tr>
<th>Chosen Deductible</th>
<th>Number of claims per policy</th>
<th>Increase in out-of-pocket payments per claim with a $1000 deductible</th>
<th>Increase in out-of-pocket payments per policy with a $1000 deductible</th>
<th>Reduction in yearly premium per policy with $1000 deductible</th>
<th>Savings per policy with $1000 deductible</th>
</tr>
</thead>
<tbody>
<tr>
<td>$500</td>
<td>0.043</td>
<td>469.86</td>
<td>19.93</td>
<td>99.85</td>
<td>79.93</td>
</tr>
<tr>
<td>N=23,782 (47.6%)</td>
<td>(.0014)</td>
<td>(2.91)</td>
<td>(0.67)</td>
<td>(0.26)</td>
<td>(0.71)</td>
</tr>
<tr>
<td>$250</td>
<td>0.049</td>
<td>651.61</td>
<td>31.98</td>
<td>158.93</td>
<td>126.95</td>
</tr>
<tr>
<td>N=17,536 (35.1%)</td>
<td>(.0018)</td>
<td>(6.59)</td>
<td>(1.20)</td>
<td>(0.45)</td>
<td>(1.28)</td>
</tr>
</tbody>
</table>

Average forgone expected savings for all low-deductible customers: $99.88

* Means with standard errors in parentheses
Back of the Envelope

- **BOE 1:** Buy house at 30, retire at 65, 3% interest rate ⇒ $6,300 expected
  - With 5% Poisson claim rate, only 0.06% chance of losing money

- **BOE 2:** (Very partial equilibrium) 80% of 60 million homeowners could expect to save $100 a year with “high” deductibles ⇒ $4.8 billion per year
Consumer Inertia?

Percent of Customers Holding each Deductible Level

Number of Years Insured with Company

- 0-3
- 3-7
- 7-11
- 11-15
- 15+

Deductible Levels:
- 1000
- 500
- 250
- 100

%
### Look Only at New Customers

<table>
<thead>
<tr>
<th>Chosen Deductible</th>
<th>Number of claims per policy</th>
<th>Increase in out-of-pocket payments per claim with a $1000 deductible</th>
<th>Increase in out-of-pocket payments per policy with a $1000 deductible</th>
<th>Reduction in yearly premium per policy with $1000 deductible</th>
<th>Savings per policy with $1000 deductible</th>
</tr>
</thead>
<tbody>
<tr>
<td>$500</td>
<td>0.037</td>
<td>475.05</td>
<td>17.16</td>
<td>94.53</td>
<td>77.37</td>
</tr>
<tr>
<td>N = 3,424 (54.6%)</td>
<td>(.0035)</td>
<td>(7.96)</td>
<td>(1.66)</td>
<td>(0.55)</td>
<td>(1.74)</td>
</tr>
<tr>
<td>$250</td>
<td>0.057</td>
<td>641.20</td>
<td>35.68</td>
<td>154.90</td>
<td>119.21</td>
</tr>
<tr>
<td>N = 367 (5.9%)</td>
<td>(.0127)</td>
<td>(43.78)</td>
<td>(8.05)</td>
<td>(2.73)</td>
<td>(8.43)</td>
</tr>
</tbody>
</table>

Average forgone expected savings for all low-deductible customers: $81.42
Risk Aversion?

- Simple Standard Model
  - Expected utility of wealth maximization
  - Free borrowing and savings
  - Rational expectations
  - Static, single-period insurance decision
  - No other variation in lifetime wealth

- Consumption maximization:

\[
\max_{c_t} U(c_1, c_2, \ldots, c_T),
\]
\[
s.t. \ c_1 + c_2 + \ldots + c_T = y_1 + y_2 + \ldots y_T.
\]

- (Indirect) utility of wealth maximization

\[
\max_w u(w),
\]
\[
\text{where } \ u(w) = \max_{c_t} U(c_1, c_2, \ldots, c_T),
\]
\[
s.t. \ c_1 + c_2 + \ldots + c_T = y_1 + y_2 + \ldots + y_T = w
\]

\[\Rightarrow \ w \text{ is lifetime wealth}\]
Model of Deductible Choice

- Choice between \((P_L, D_L)\) and \((P_H, D_H)\)
- \(\pi = \) probability of loss
  - Simple case: only one loss
- EU of contract:
  - \(U(P,D,\pi) = \pi u(w-P-D) + (1- \pi)u(w-P)\)
Bounding Risk Aversion

Assume CRRA form for $u$:

$$u(x) = \frac{x^{(1-\rho)}}{(1-\rho)} \quad \text{for } \rho \neq 1, \quad \text{and} \quad u(x) = \ln(x) \quad \text{for } \rho = 1$$

Indifferent between contracts iff:

$$\pi \frac{(w - P_L - D_L)^{(1-\rho)}}{(1-\rho)} + (1-\pi) \frac{(w - P_L)^{(1-\rho)}}{(1-\rho)} = \pi \frac{(w - P_H - D_H)^{(1-\rho)}}{(1-\rho)} + (1-\pi) \frac{(w - P_H)^{(1-\rho)}}{(1-\rho)}$$
Getting the bounds

- Search algorithm at individual level
  - New customers
- Claim rates: Poisson regressions
  - Cap at 5 possible claims for the year
- Lifetime wealth:
  - Conservative: $1 million (40 years at $25k)
  - More conservative: Insured Home Value
## CRRA Bounds

<table>
<thead>
<tr>
<th>Chosen Deductible</th>
<th>Measure of Lifetime Wealth (W): (Insured Home Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W</td>
</tr>
<tr>
<td>$1,000$</td>
<td>256,900</td>
</tr>
<tr>
<td>$1,000$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,135,565</td>
</tr>
<tr>
<td>$500$</td>
<td>190,317</td>
</tr>
<tr>
<td>N = 3,424 (54.6%)</td>
<td>{64,634}</td>
</tr>
<tr>
<td>$250$</td>
<td>166,007</td>
</tr>
<tr>
<td>N = 367 (5.9%)</td>
<td>{57,613}</td>
</tr>
</tbody>
</table>
50-50 gamble:
Lose $1,000/ Gain $10 million
- 99.8% of low-ded customers would reject

Labor-supply calibrations, consumption-savings behavior $\Rightarrow \rho < 10$
- Gourinchas and Parker (2002) -- 0.5 to 1.4
- Chetty (2005) -- $< 2$
Wrong level of wealth?

- Lifetime wealth inappropriate if borrowing constraints.
- $94 for $500 insurance, 4% claim rate
  - $W = $1 million $\Rightarrow \rho = 2,013$
  - $W = $100k $\Rightarrow \rho = 199$
  - $W = $25k $\Rightarrow \rho = 48$
Prospect Theory

- Kahneman & Tversky (1979, 1992)
- Reference dependence
  - Not final wealth states
- Value function
  - Loss Aversion
  - Concave over gains, convex over losses
- Non-linear probability weighting
Model of Deductible Choice

- Choice between $(P_L, D_L)$ and $(P_H, D_H)$
- $\pi = \text{probability of loss}$
- EU of contract:
  - $U(P, D, \pi) = \pi u(w-P-D) + (1- \pi)u(w-P)$
- PT value:
  - $V(P, D, \pi) = v(-P) + w(\pi)v(-D)$
- Prefer $(P_L, D_L)$ to $(P_H, D_H)$
  - $v(-P_L) - v(-P_H) < w(\pi)[v(-D_H) - v(-D_L)]$
Loss Aversion and Insurance

  - Choice A
    - 25% chance of $200 loss [80%]
    - Sure loss of $50 [20%]
  - Choice B
    - 25% chance of $200 loss [35%]
    - Insurance costing $50 [65%]
No loss aversion in buying

- Novemsky and Kahneman (2005)
  (Also Kahneman, Knetsch & Thaler (1991))
  - Endowment effect experiments
  - Coefficient of loss aversion = 1 for “transaction money”
- Köszegi and Rabin (forthcoming QJE, 2005)
  - Expected payments
- Marginal value of deductible payment > premium payment (2 times)
So we have:

- Prefer \((P_L, D_L)\) to \((P_H, D_H)\):
  \[v(-P_L) - v(-P_H) < w(\pi)[v(-D_H) - v(-D_L)]\]

- Which leads to:
  \[P_L^\beta - P_H^\beta < w(\pi)\lambda[D_H^\beta - D_L^\beta]\]

- Linear value function:
  \[WTP = \Delta P = \frac{w(\pi)\lambda\Delta D}{\Delta P} = 4 \text{ to } 6 \text{ times EV}\]
Parameter values

- Kahneman and Tversky (1992)
  - $\lambda = 2.25$
  - $\beta = 0.88$
- Weighting function
  \[ w(\pi) = \frac{\pi^\gamma}{(\pi^\gamma + (1 - \pi)^\gamma)^{\gamma'}} \]
  - $\gamma = 0.69$
WTP from Model

- Typical new customer with $500 ded
  - Premium with $1000 ded = $572
  - Premium with $500 ded = +$94.53
  - 4% claim rate
- Model predicts WTP = $107
- Would model predict $250 instead?
  - WTP = $166. Cost = $177, so no.
## Choices: Observed vs. Model

<table>
<thead>
<tr>
<th>Chosen Deductible</th>
<th>Predicted Deductible Choice from Prospect Theory NLIB Specification: ( \lambda = 2.25, \gamma = 0.69, \beta = 0.88 )</th>
<th>Predicted Deductible Choice from EU(W) CRRA Utility: ( \rho = 10, W = \text{Insured Home Value} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1,000 \ (N = 2,474 \ (39.5%))$</td>
<td>87.39% 11.88% 0.73% 0.00%</td>
<td>100.00% 0.00% 0.00% 0.00%</td>
</tr>
<tr>
<td>$500 \ (N = 3,424 \ (54.6%))$</td>
<td>18.78% 59.43% 21.79% 0.00%</td>
<td>100.00% 0.00% 0.00% 0.00%</td>
</tr>
<tr>
<td>$250 \ (N = 367 \ (5.9%))$</td>
<td>3.00% 44.41% 52.59% 0.00%</td>
<td>100.00% 0.00% 0.00% 0.00%</td>
</tr>
<tr>
<td>$100 \ (N = 3 \ (0.1%))$</td>
<td>33.33% 66.67% 0.00% 0.00%</td>
<td>100.00% 0.00% 0.00% 0.00%</td>
</tr>
</tbody>
</table>
Conclusions

- (Extreme) aversion to moderate risks is an empirical reality in an important market
- Seemingly anomalous in Standard Model where risk aversion = DMU
- Fits with existing parameter estimates of leading psychology-based alternative model of decision making
Alternative Explanations

- Misestimated probabilities
  - \( \approx 20\% \) for single-digit CRRA
  - Older (age) new customers just as likely

- Liquidity constraints

- Sales agent effects
  - Hard sell?
  - Not giving menu? ($500?, data patterns)
  - Misleading about claim rates?

- Menu effects
• **Barseghyan, Molinari, O’Donoghue, and Teitelbaum (AER 2013)**
  – Micro data for same person on 4,170 households for 2005 or 2006 on
    * home insurance
    * auto collision insurance
    * auto comprehensive insurance

• Estimate a model of reference-dependent preferences with Koszegi-Rabin
  reference points
  – Separate role of loss aversion, curvature of value function, and proba-
    bility weighting

• Key to identification: variation in probability of claim:
  – * home insurance $\rightarrow 0.084$
    * auto collision insurance $\rightarrow 0.069$
    * auto comprehensive insurance $\rightarrow 0.021$
Figure 1: Empirical Density Functions for Predicted Claim Probabilities
• This allows for better identification of probability weighting function

• Main result: Strong evidence from probability weighting, implausible to obtain with standard risk aversion

• Share of probability weighting function

• With probability weighting, realistic demand for low-deductible insurance

• Follow-up work: distinguish probability weighting from probability distortion
Figure 2: Estimated $\Omega(\mu)$ – Model 1
Table 6: Economic Significance

<table>
<thead>
<tr>
<th>Standard risk aversion</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>r=0</td>
<td>r=0.00064</td>
<td>r=0</td>
<td>r=0.00064</td>
<td>r=0.0129</td>
<td></td>
</tr>
<tr>
<td>Probability distortions?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>( \mu )</td>
<td>WTP</td>
<td>WTP</td>
<td>WTP</td>
<td>WTP</td>
<td>WTP</td>
</tr>
<tr>
<td>0.020</td>
<td>10.00</td>
<td>14.12</td>
<td>41.73</td>
<td>57.20</td>
<td>33.76</td>
</tr>
<tr>
<td>0.050</td>
<td>25.00</td>
<td>34.80</td>
<td>55.60</td>
<td>75.28</td>
<td>75.49</td>
</tr>
<tr>
<td>0.075</td>
<td>37.50</td>
<td>51.60</td>
<td>67.30</td>
<td>90.19</td>
<td>104.86</td>
</tr>
<tr>
<td>0.100</td>
<td>50.00</td>
<td>68.03</td>
<td>77.95</td>
<td>103.51</td>
<td>130.76</td>
</tr>
<tr>
<td>0.125</td>
<td>62.50</td>
<td>84.11</td>
<td>86.41</td>
<td>113.92</td>
<td>154.00</td>
</tr>
</tbody>
</table>

Notes: WTP denotes—for a household with claim rate \( \mu \), the utility function in equation (2), and the specified utility parameters—the household's maximum willingness to pay to reduce its deductible from $1000 to $500 when the premium for coverage with a $1000 deductible is $200. Columns (3) and (4) use the probability distortion estimates from Model 1a.
7 Next Lecture

- No lecture next week
- More on Reference-Dependence
- Social Preferences