Big Data and Insurance Public Policy
The Griffith Insurance Education Foundation, an affiliate of The Institutes, is a 501(c)(3) non-profit, non-partisan, and non-advocative educational organization dedicated to the teaching and study of insurance and risk management.

In keeping with the non-partisan, non-advocative mission of The Griffith Foundation, we will keep our comments and contributions to today's program unbiased and purely educational.
Big Data and Insurance Public Policy

Dr. Dani Bauer
Hickman-Larson Chair in Actuarial Science
Department of Risk and Insurance
Wisconsin School of Business
University of Wisconsin-Madison
daniel.bauer@wisc.edu

J. Tyler Leverty, PhD
Gerald D. Stephens CPCU Distinguished Chair
Department of Risk and Insurance
Wisconsin School of Business
University of Wisconsin-Madison
 ty-leverty@wisc.edu
Outline

- What is Big Data?
- Why is Big Data Important to Insurance Markets?
- What are the Promises of Big Data for the Insurance Industry?
- What are the Potential Public Policy Concerns?
What is Big Data?

- Brave “new” world:
  - Netflix knows what movies you want to watch; Amazon what products you want to buy
  - Target’s Andrew Pole predicts pregnancy before father knows
  - Cars drive themselves...

- How does this work?:
  - Same way actuaries have priced car insurance for decades – predicting car accidents based on individual information using regression
    - **BUT**: More data, fancier methods, bigger computers, ...
    - E.g., think about “features” and “response variables” for self-driving cars
What is Big Data?

- Brave “new” Netflix knows what products you want
- Target’s A
cars drive

- How does the
- Same way car accidents

- BUT: Machines
- E.g., this
- Google
- This
- NOT
- This

In what products
M. knows
Self-driving cars
Underwriting

• The process of Selecting and Classifying exposures:

  – Selection: determination of whether to issue insurance

  – Classification: determination of terms, conditions, and premium
Implications of Heterogeneous Buyers

- What if there are 2 groups of buyers, equal number of each?
  
  (1) Low Risk

<table>
<thead>
<tr>
<th>Possible Loss</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0</td>
<td>0.95</td>
</tr>
<tr>
<td>$10,000</td>
<td>0.05</td>
</tr>
</tbody>
</table>

  \[ E(\text{Claim Cost}) = $500 \]

(2) High Risk

<table>
<thead>
<tr>
<th>Possible Loss</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0</td>
<td>0.90</td>
</tr>
<tr>
<td>$10,000</td>
<td>0.10</td>
</tr>
</tbody>
</table>

  \[ E(\text{Claim Cost}) = $1,000 \]
Underwriting

- Initial Scenario:
  - The is only 1 insurer, Equal Treatment Ins. Co.
  - Premium for everyone = $750
  - Does Equal Treatment cover its costs?
Underwriting

- Now Selective Ins. Co. enters the market
  - If Equal Treatment continues to charge $750, how does Selective set prices to maximize profits,
    - Premium to High Risk =
    - Premium to Low Risk =
  - What happens to Equal Treatment?
Why is Big Data Important to Insurance Markets?

• Insurer will be able to predict expected loss on an individual basis
  – Everyone pays the “right” price, no one overpays
  – Still a need for insurance (perfectly predict chance, but not occurrence)
  – “First best” economic outcome, theoretically desirable...

• Efficiency gains, technology enables for better products
  – Ingenie: A box in the car of young drivers, information and advice
  – Lemonade: Return money to causes you care about
But: Changes of the insurance landscape

THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION?

Carl Benedikt Frey† and Michael A. Osborne‡

September 17, 2013

<table>
<thead>
<tr>
<th>Computerisable</th>
<th>Probability</th>
<th>Label</th>
<th>SOC code</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.0028</td>
<td>29-1125</td>
<td></td>
<td>Recreational Therapists</td>
</tr>
<tr>
<td>2.</td>
<td>0.0003</td>
<td>49-1011</td>
<td></td>
<td>First-Line Supervisors of Mechanics, Installers, and Repairers</td>
</tr>
<tr>
<td>...</td>
<td>0.21</td>
<td>15-2011</td>
<td></td>
<td>Actuaries</td>
</tr>
<tr>
<td>565.</td>
<td>0.92</td>
<td>41-3021</td>
<td></td>
<td>Insurance Sales Agents</td>
</tr>
<tr>
<td>698.</td>
<td>0.99</td>
<td>13-2053</td>
<td></td>
<td>Insurance Underwriters</td>
</tr>
<tr>
<td>701.</td>
<td>0.99</td>
<td>23-2093</td>
<td></td>
<td>Title Examiners, Abstractors, and Searchers</td>
</tr>
<tr>
<td>702.</td>
<td>0.99</td>
<td>41-9041</td>
<td></td>
<td>Telemarketers</td>
</tr>
</tbody>
</table>
But: Potential Public Policy Concerns (1)

- Downsides of Perfect Risk Classification
  - Insurance before or behind the veil: Bad rates based on genetic information? What happens to the “uninsurable”? 
  - Repercussions on long term contracts: More information may preclude commitment 

- Privacy:
  - Pay-as-you-drive: do I want the insurer to know where I drive and when I speed?
  - Is it good that Target knows about pregnancy first?
Moral concerns

- Redlining: Correlation is not causation, but an algorithm will not care

- Algorithms are smart: Proxy for features that are explicitly excluded from underwriting (race, gender, etc.) by using available information

  ...Reinforce preexisting inequality...
Questions