Privacy & Behavioral Economics: The Paradox of Control & Other Studies

Alessandro Acquisti

Heinz College/CyLab
Carnegie Mellon University

Paduano Symposium
Friday October 15, 2010
What are the trade-offs associated with protecting and sharing personal data...?

How rationally do we calculate, and make decisions about, those trade-offs...?

What are the consequences of those decisions...?

Can we (and should we) assist those decisions, through technology or policy design...?
Why

- Intellectual curiosity
- Policy intervention, Technology design
- Managerial & Marketing implications
1. From the economics of privacy...
2. ... to the behavioral economics of privacy
3. Three studies
   - The inconsistency of privacy valuations
   - The paradox of control
   - Discounting past information
The economics of privacy

- Protection and revelation of personal data flows involve **tangible and intangible trade-offs for the data subject** as well as the potential **data holder**

- Early 1980s
  - The Chicago school approach (Posner 1978, Stigler 1980, ...)

- Mid 1990s

- After 2000

Implicit assumptions under the ‘canonical’ economic approach

- Individuals have stable preferences over privacy.
- Based on those preferences, they mentally trade-off costs and benefits of sharing and protecting data.
- Based on those trade-offs, they decide rationally what personal information to reveal and what to protect.

Therefore, no need for market (regulatory) intervention.
Those assumptions also influence the current debate on privacy

- Lenard and Rubin 2009: “The competitive online market structure suggests that firms do have incentives to satisfy their customers' privacy preferences and that consumers' behavior in the market reflects their preferences.”

- Schmidt: “If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place.”

- Zuckenberg: “[The] social norm [about privacy] is just something that has evolved over time.”
However: The privacy paradox

- Attitudes about privacy...
  - Ostensibly, privacy concerns cost merchants billions in lost e-tail sales...
    (Jupiter Research).... Significant reason for Internet users to avoid Ecommerce... (P&AB)

- ... versus actual behavior
  - Self-professed privacy attitudes do not predict privacy-conscious behavior
    (Spiekermann et al. 2001, Acquisti & Gross 2006’s Facebook study)

  *Do people really care for privacy?*

  *If they do, can they adequately protect their data?*

  *If they don’t (or can’t), should policy-makers do so on their behalf?*
A rational model of privacy decision making

Should I mention my sexual kinks on MySpace?
A rational model of privacy decision making

Maybe I’ll find a lover... But what about my future job prospects? And what if my parents happen to log on...
A rational model of privacy decision making

\[
\sum p_i \sum \frac{1}{(1+d)^t} u(\text{benefits}_{it}) - \sum q_i \sum \frac{1}{(1+d)^t} u(\text{costs}_{it})
\]
Why privacy decision making can’t be “rational” (in the economic sense)

- (Framing issues)
- (Conflicting needs)

1. Incomplete information
2. Bounded rationality
3. Cognitive/behavioral biases
Hence, the need for a behavioral economics of privacy

- Behavioral economics studies systematic “deviations” from the theoretical behavior of the rational economic agent
  - Many of those deviations have applications to the privacy arena (as well as information security)
- Applying behavioral, experimental economics to the understanding of privacy decision-making helps us:
  1. Understand how people make (sometimes biased) decisions about their personal information
  2. Inform policy and technology design, to ameliorate decision-making
Our methodological approach

- Controlled, randomized experiments (in the lab, in the field, survey-based, non survey-based, ...)
- Dependent variable(s) correlated with (heterogeneous, and otherwise latent and therefore unobservable) privacy concerns
  - Actual behavior vs. Survey responses (self-disclosures)
  - Validation studies vs. Comparative studies
- Many influences: BE and BDR, but also:
  - Survey design (e.g. Schwarz 1999); self-disclosure (e.g. Altman and Taylor 1973); privacy and disclosure (e.g., Margulis 2003); privacy concerns (e.g., Culnan and Armstrong 1999); ...
Some of our results

- Some results (2004-2010)
  - Hyperbolic discounting in privacy valuations...
  - Over-confidence, optimism bias in personal disclosures...
  - Confidentiality assurances inhibit information disclosure...
  - Individuals more likely to disclose sensitive information to unprofessional sites than professional sites...
  - ...
Today: Three studies

1. The inconsistency of privacy valuations
   - About how the endowment effect affects people’s willingness to pay for privacy

2. The paradox of control
   - About the propensity to reveal personal information

3. Discounting past information
   - About the impact on others of one’s personal information

Joint works with Laura Brandimarte, Joachim Vosgerau, Leslie John, George Loewenstein
Three studies

1. The inconsistency of privacy valuations
   - With Leslie John and George Loewenstein
2. The paradox of control
3. Discounting past information
Can mere framing change the valuation of personal data?

Consider:

- Willingness to accept (WTA) money to give away information
  
  *vs.*

- Willingness to pay (WTP) money to protect information

**Hypothesis:**

- People assign different values to their personal information depending on whether they are focusing on protecting it or revealing it
Experimental design

- Mall patrons asked to participate in (decoy) survey
- As payment for participation, subjects were offered gift cards
- We manipulated trade-offs between privacy protection and value of cards
- Subjects *endowed* with either:
  - **$10 Anonymous gift card.** "Your name will not be linked to the transactions completed with the card, and its usage will not be tracked by the researchers."
  - **$12 Trackable gift card.** "Your name will be linked to the transactions completed with the card, and its usage will be tracked by the researchers."
- Subjects asked whether they'd like to *switch* cards
  - From $10 Anonymous to $12 Trackable (WTA)
  - From $12 Trackable to $10 Anonymous (WTP)
WTP vs. WTA: Results

χ²(3) = 30.66, p < 0.0005
Implications

- People’s concerns for privacy (and security) depend, in part, on priming and framing
  - This does not necessarily mean that people don’t care for privacy, or are “irrational”
- Rather, it implies that reliance on “revealed preferences” argument for privacy may lead to sub-optimal outcomes if privacy valuations are inconsistent...
  - People may make disclosure decisions that they stand to later regret
- Vicious circle: if you have less privacy, you value privacy less
Three studies

1. The inconsistency of privacy valuations
2. The paradox of control
   - With Laura Brandimarte and George Loewenstein
3. Discounting past information
Privacy and (the paradox of) control

Control :: Privacy
Privacy and (the paradox of) control

Control :: Privacy

=
Privacy and (the paradox of) control

Control :: Privacy
The paradox of control hypothesis

- Conjecture: When deciding what to reveal about ourselves, we confound control over **publication** of private information with control over **access/use** of that information by others
  - Even though **objective privacy costs derive from access to/use** of information by others, not merely its publication

- Hence: **Users who perceive more [less] control over publication of personal information will disclose more [less] sensitive information** – even though they may have less [more] control over access and use of that information

- Why?
  - Saliency of act of publishing, Overconfidence
  - See Henslin 1967, Langer 1975
Three survey-based randomized experiments

- **Experiment 1:** Reducing perceived control over publication of personal information
  - Mediated vs. unmediated publication

- **Experiment 2:** Reducing perceived control over publication of personal information
  - Certainty vs. probability of publication

- **Experiment 3:** Increasing perceived control over publication of personal information
  - Explicit vs. implicit control
Three survey-based randomized experiments

- Experiment 1: Reducing perceived control over publication of personal information
  - Mediated vs. unmediated publication

- Experiment 2: Reducing perceived control over publication of personal information
  - Certainty vs. probability of publication

- Experiment 3: Increasing perceived control over publication of personal information
  - Explicit vs. implicit control
Experiment 2

Design

- Subjects: 100+ CMU students recruited on campus
-Asked to complete online survey
-Justification for the survey: creation of CMU networking website
-Questions focused on students’ life on and off campus
  - Multiple choice, Yes/No, Rating and open-end questions
  - Included quasi-identifiers + privacy intrusive and non-intrusive questions
    - As rated by 31 subjects independently in a pre-study
Experiment 2

- Dependent variables
  - Response rate: whether subject answered or not (White 2004)
    - Recall: Validation studies vs. Comparison studies
- Explanatory variables
  - Experimental treatment
  - Intrusiveness of questions
  - Demographics
Experiment 2

- Manipulation: Profile automatically published vs. profile published with 50% probability (less control)

  - Condition 1
    “The information you provide will appear on a profile that will be automatically created for you. The profile will be published on a new CMU networking website, which will only be accessible by members of the CMU community, starting at the end of this semester. The data will not be used in any other way. NO QUESTION/FIELD REQUIRES AN ANSWER.”

  - Condition 2
    “The information you provide will appear on a profile that will be automatically created for you. Half of the profiles created for the participants will be randomly picked to be published on a new CMU networking website, which will only be accessible by members of the CMU community, starting at the end of this semester. The data will not be used in any other way. NO QUESTION/FIELD REQUIRES AN ANSWER.”
Hypotheses

- **Paradox of control hypothesis**: Subjects more likely to answer sensitive questions in Condition 1 than in Condition 2, but *no more likely* to answer non-sensitive questions.

- Alternative results that would not support the control paradox hypothesis:
  - *Subjects more likely to answer sensitive and non-sensitive questions in Condition 1 than in Condition 2*.
  - *Subjects no more likely to answer sensitive and non-sensitive questions in Condition 1 than in Condition 2*.
Results

Average Response Rates - Experiment 2

- More Intrusive Questions
- Less Intrusive Questions

Condition 1
Condition 2
## Results (DV: Answer yes/no)

### Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Coeff</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>-.25**</td>
<td>.05</td>
</tr>
<tr>
<td>Intrusive</td>
<td>-.64***</td>
<td>.00</td>
</tr>
<tr>
<td>Treat_Int</td>
<td>-.67***</td>
<td>.00</td>
</tr>
<tr>
<td>Age</td>
<td>-.02</td>
<td>.28</td>
</tr>
<tr>
<td>Male</td>
<td>.20</td>
<td>.10</td>
</tr>
</tbody>
</table>

N= 132  Prob > χ² = .000

RE Probit coefficients of panel regression of response rate on treatment with dummy for most intrusive questions, interaction and demographics
Experiment 3

- **Design**
  - Subjects: 100+ CMU students recruited on campus, March 2010
  - Completed online survey
  - Justification for the survey: study on ethical behaviors
  - Ten Yes/No questions that focused on sensitive behaviors (e.g. drug use, stealing)
    - Included demographics + privacy intrusive and non-intrusive questions
      - As rated by 49 subjects independently in a pre-study
Experiment 3

- **Conditions (reduced)**
  - **Implicit control condition**
    “All answers are voluntary. By answering a question, you agree to give the researchers permission to publish your answer.”
  - **Explicit control condition**
    “All answers are voluntary. In order to give the researchers permission to publish your answer to a question, you will be asked to check the corresponding box in the following page.”
Implicit control condition

Study on Ethical Behavior

IMPORTANT: All answers are voluntary. By answering a question, you agree to give the researchers permission to publish your answer.

1. Are you married?  
2. Have you ever been fired by your employer?  
3. Have you ever stolen anything (e.g., from a shop, a person)?  
4. Have you ever used drugs of any kind (e.g., weed, heroin, meth)?  
5. Have you ever lied about your age?  
6. Have you ever had cosmetic surgery?  
7. Have you ever done any kind of voluntary service?  
8. Have you ever been in a public venue (e.g., nightclub, airport)?  
9. Have you ever made a donation to a non-profit organization?  
10. Do you have any permanent tattoos?
Explicit control condition

Study on Ethical Behavior

IMPORTANT: All answers are voluntary. In order to give the researchers permission to publish your answers to the questions, please check the box below.

☐ Publication permission

1. Are you married?
2. Have you ever been fired by your employer?
3. Have you ever stolen anything (e.g., from a shop, a person)?
4. Have you ever used drugs of any kind (e.g., weed, heroin, crack)?
5. Have you ever lied about your age?
6. Have you ever had cosmetic surgery?
7. Have you ever done any kind of voluntary service?
8. Have you ever been in a public venue (e.g., red carpet of a club, arena)?
9. Have you ever made a donation to a non-profit organization?
10. Do you have any permanent tattoos?

Close
Hypotheses

- **Paradox of control hypothesis**: Subjects more likely to answer AND to allow publication of sensitive questions in the explicit control condition than in the implicit control condition

- Alternative results that would not support the control paradox hypothesis:
  - *Subjects more likely to answer sensitive in the explicit control condition, but not more likely to allow their publication*
  - *Subjects no more likely to answer sensitive in the explicit control condition, nor to allow their publication*
Results

Average Response & Publication Rates - Study 3

- More intrusive Questions
- Less Intrusive Questions

Condition 1
Condition 3
## Results (DV: Publish yes/no)

<table>
<thead>
<tr>
<th></th>
<th>Coefficients and p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>1.92**</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
</tr>
<tr>
<td>Intrusive</td>
<td>-.85**</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
</tr>
<tr>
<td>Treat_Int</td>
<td>-1.21**</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
</tr>
<tr>
<td>Age</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>(.521)</td>
</tr>
<tr>
<td>Male</td>
<td>-.11</td>
</tr>
<tr>
<td></td>
<td>(.593)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>65</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob &gt; χ²</td>
<td>.000</td>
</tr>
</tbody>
</table>

RE Probit coefficients of panel regression of response rate on treatment with dummy for most intrusive questions, interaction and demographics.

* indicates significance at 10% level; ** indicates significance at 5% level
Results

- When someone other than themselves is responsible for the publication, or when the publication itself becomes uncertain – which reduces the probability of access/use by others – people refrain from disclosing

- Perceived less *more* control over publication reduces *

  *increases* revelation of private information

  - Even though objective risks of revelation decrease *increase*
Implications

- This effect tends to be stronger for more intrusive questions
  - It is not the publication of private information *per se* that disturbs people, but the fact that someone else will publish it for them
- Results call into questions OSNs’ arguments of protecting privacy by providing more control to members
  - *Giving more control to users over information publication seems to generate higher willingness to disclose sensitive information*
Three studies

1. The inconsistency of privacy valuations
2. The paradox of control
3. Discounting past information
   - With Laura Brandimarte and Joachim Vosgerau
Research question

- How does information about a person or company’s past, retrieved today, get ‘discounted’?
  - Specifically: does information about a person’s past with negative valence receive more weight in impression formation than information with positive valence?
A differential discounting hypothesis

What the literature focused on (e.g., Brickman et al., 1978):

We introduce the hypothesis of differential discounting:
Hypothesis

- Impact of information with negative valence lasts longer than impact of info with positive valence, not only because of asymmetric effects of valence, but also because of different weights – or discount rates – applied to the two types of info

- This may be due to
  - Mobilization effects (Taylor 1991) and evolutionary theory (Baumeister et al. 2001)
  - Negativity bias (Seligman & Maier 1967)
  - Negative info is more attention grabbing (Pratto & John 1991)
Three survey-based randomized experiments

- We ran three survey-based randomized experiments, manipulating valence of information about third parties provided to subjects and the time to which that information referred.
- Subjects were asked to express a judgment on the person or company they just read about.
- Three experiments:
  - The dictator game
  - The company experiment
  - The wallet experiment
The wallet experiment

- **Hypothetical scenario**: subjects are presented background information about another person, and asked to express a judgment about her
  - **Baseline condition**: only baseline information is provided
  - **Treatment conditions**: manipulation of valence and time:
    1. We add to the baseline info one detail with either positive or negative valence
    2. And, we vary the time to which that detail refers
Here is some background information about Mr. A. Please review this information, and be ready to answer the questions below and in the next page.

Mr. A was born in San Diego, California, where he attended elementary and middle school. As a child, he obtained his social security number and received the standard DPT vaccination.

When he was 16 years old, he moved to Sacramento, California, with his family. He attended high school there and got his driving license.

After graduation he moved to Houston, Texas.
Here is some background information about Mr. A. Please review this information, and be ready to answer the questions below and in the next page.

Mr. A was born in San Diego, California, where he attended elementary and middle school. As a child, he obtained his social security number and received the standard DPT vaccination.

When he was 16 years old, he moved to Sacramento, California, with his family. He attended high school there and got his driving license.

Just about graduation, he found a lost woman's purse containing $10,000 in cash. He reported the discovery to the police, and the rightful owner retrieved her money.

After graduation he moved to Houston, Texas where he has been living and working for the past 12 months.
Here is some background information about Mr. A. Please review this information, and be ready to answer the questions below and in the next page.

Mr. A was born in San Diego, California, where he attended elementary and middle school. As a child, he obtained his social security number and received the standard DPT vaccination.

When he was 16 years old, he moved to Sacramento, California, with his family. He attended high school there and got his driving license.

*Just about graduation, he found a lost woman's purse containing $10,000 in cash. He did not report the discovery to the police, and the rightful owner did not retrieve her money.*

After graduation he moved to Houston, Texas where he has been living and working for the past 12 months.
Here is some background information about Mr. A. Please review this information, and be ready to answer the questions below and in the next page.

Mr. A was born in San Diego, California, where he attended elementary and middle school. As a child, he obtained his social security number and received the standard DPT vaccination.

When he was 16 years old, he moved to Sacramento, California, with his family. He attended high school there and got his driving license.

Just about graduation, he found a lost woman's purse containing $10,000 in cash. He reported the discovery to the police, and the rightful owner retrieved her money.

After graduation he moved to Houston, Texas where he has been living and working for the past 5 years.
Here is some background information about Mr. A. Please review this information, and be ready to answer the questions below and in the next page.

Mr. A was born in San Diego, California, where he attended elementary and middle school. As a child, he obtained his social security number and received the standard DPT vaccination.

When he was 16 years old, he moved to Sacramento, California, with his family. He attended high school there and got his driving license.

Just about graduation, he found a lost woman's purse containing $10,000 in cash. He did not report the discovery to the police, and the rightful owner did not retrieve her money.

After graduation he moved to Houston, Texas where he has been living and working for the past 5 years.
Experimental conditions

- One baseline condition
- 2x2 treatment conditions:

<table>
<thead>
<tr>
<th>Reported wallet, 5 years ago</th>
<th>Reported wallet, 12 months ago</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did not report wallet, 5 years ago</td>
<td>Did not report wallet, 12 months ago</td>
</tr>
</tbody>
</table>
Dependent variables:

- How much subjects liked the person described
- How much subjects would have liked to work with her (Interpersonal Judgment Scale, Byrne 1961)
Figure 6. Average level of liking and trust indices across conditions in Experiment 2
Bad is not just stronger than good...

.... It is also discounted differently than good

Implications: future impact of information revealed today
People’s concerns for privacy (and security) depend, in part, on priming and framing.

- This does not necessarily mean that people don’t care for privacy, or are “irrational,” or make wrong decisions about privacy.

- Rather, it implies that reliance on “revealed preferences” argument for privacy may lead to sub-optimal outcomes if privacy valuations are inconsistent...
  
  - People may make disclosure decisions that they stand to later regret.
  - Risks greatly magnified in online information revelation.
Overall implications of these privacy studies

- “Choice & notification” privacy model may be outdated
- Implications for policy-making & the debate on privacy regulation
  - Consider: Chicago School approach vs. privacy advocates
  - “Nudging” privacy?
For more info

- Google: economics privacy
- Visit: http://www.heinz.cmu.edu/~acquisti/economics-privacy.htm
- Email: acquisti@andrew.cmu.edu