


# Privacy and Rationality



Alessandro Acquisti and Jens Grossklags

Heinz School, CMU and SIMS, UC Berkeley

[acquisti@andrew.cmu.edu](mailto:acquisti@andrew.cmu.edu)

[jensg@sims.berkeley.edu](mailto:jensg@sims.berkeley.edu)

# Open questions in the economics of privacy

---

- What explains the attitudes/behavior dichotomy?
  - Do people care about privacy?
- Can individuals protect themselves?
  - Should they?

# GMail debate

---

*It is true that there are potential costs of using Gmail for email storage [...] The question is whether consumers should have the right to make that choice and balance the tradeoffs, or whether it will be preemptively denied to them by privacy fundamentalists out to deny consumers that choice.*

-- Declan McCullagh

# Privacy and rationality

---

- Forward looking agent, utility maximizer, Bayesian updater, perfectly informed
  - Theory
  - Empirical studies
  - Explanation of attitudes/behavior dichotomy
    - “It's rational anyway” (Syverson 2003)
    - “Only when matters to them” (Shostack 2003)

# Alternative view

---

- Highlight difference between market for *personal information* and market for *privacy*
- Check for knowledge of risks and protection, strategic behavior, risk attitudes, behavioral distortions
- Theory:
  - Acquisti *ACM EC 04*
- Empirical approach
  - Acquisti and Grossklags *WEIS 04*

# The Two Markets *of* Privacy

---

- Privacy issues actually originate from two different markets
  - Market for personal information
  - Market for privacy
- Related, but not identical
- Confusion leads to inconsistencies
  - Different rules, attitudes, considerations
    - Public vs. private
    - Selling vs. buying
    - Specific vs. generic
    - Value for other people vs. damage to oneself
    - Positive lump sum vs. negative annuity

# Privacy trade-offs

---

- Protect:
  - Immediate costs or loss of immediate benefits
  - Future (uncertain) benefits
- Do not protect:
  - Immediate benefits
  - Future (uncertain) costs

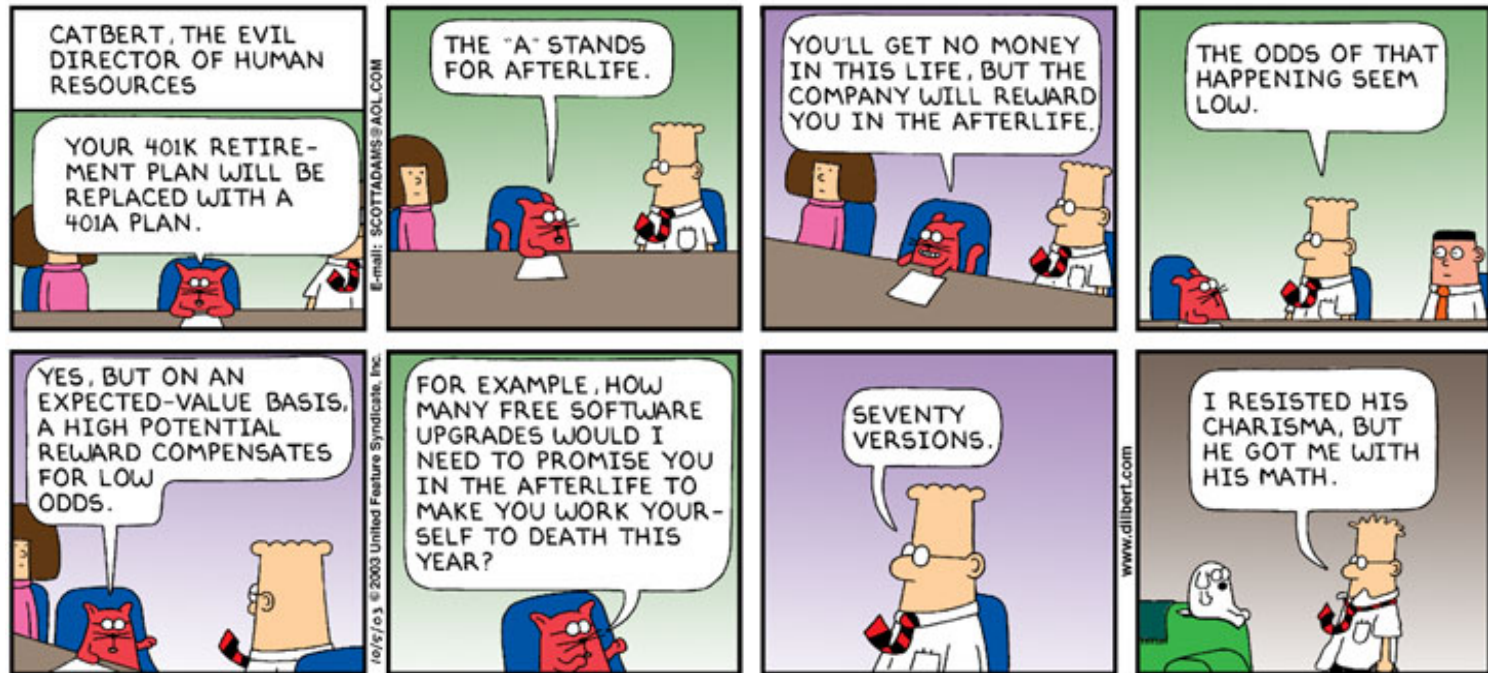
# Why is this problematic?

---

- Incomplete information
- Bounded rationality
- Psychological/behavioral distortions
  - Complacency towards large risks
  - Inability to deal with prolonged accumulation of small risks
  - Coherent arbitrariness
  - Hyperbolic discounting



# Dilbert and expected values



# Hyperbolic discounting

---

$$U^t = \sum_{\tau=t}^T \delta^\tau u_\tau,$$

$$U_t(u_t, u_{t+1}, \dots, u_T) = \delta^t u_t + \beta \sum_{\tau=t+1}^T \delta^\tau u_\tau$$

# Time consistency vs. time inconsistency

---

- Survey time vs. decision time
- Immediate benefits vs. future costs

	Period 1	Period 2	Period 3	Period 4
Benefits from selling period 1	2	0	0	0
Costs from selling period 1	0	1	1	1
Benefits from selling period 3	0	0	2	0
Costs from selling period 3	0	0	0	1

# Survey & experiment

---

- Phase One: pilot
- Phase Two: 100 questions, 119 subjects from CMU list
- Paid, online survey (CMU Berkman Fund)
- Goals
  - Contrast three sets of data
    - *Privacy attitudes*
    - *Privacy behavior*
    - *Market characteristics and psychological distortions*
  - Test rationality assumption
  - Explain behavior and dichotomy
- Phase Three: experiment

# Categories

---

## 1. Demographics and IT usage

## 2. Knowledge of privacy risks

- E.g., when you are releasing personal information during an ecommerce transaction, how likely do you consider the following outcomes? [...]

## 3. Knowledge of protection

- E.g., what technology or service may allow you to browse the Internet and be warned if a website has a privacy policy that is incompatible with your preferences?

## 4. Attitudes towards privacy (generic)

- E.g., how concerned are you about threats to your personal privacy in today's information society?

## 5. Attitudes towards privacy (specific)

- E.g., when visiting information-collecting web sites, how comfortable are you about providing the following types of information? [...]

# Categories

---

## 6. Risk neutrality/aversion (unframed)

- E.g., imagine that you are offered some money or the possibility to participate in a lottery. Do you prefer: *100\$ for certain* OR *.90 prob of getting 0\$ and .10 prob of getting 200\$?*

## 7. Strategic/unstrategic behavior (unframed)

- E.g., beauty contest, bargaining behavior

## 8. Hyperbolic discounting (unframed)

- E.g., would you prefer to receive *x dollars now* OR *500 dollars* in [...] days/months/years?

## 9. Buy and sell value for same piece of information

- E.g., suppose a marketing company wants to buy your personal information [...] How much money would you spend to be completely guaranteed that nobody at all could know the following piece of information related to you without your consent? [...]

# Categories

---

## 10. Behavior, past: “Sell” behavior (i.e., give away information)

- E.g., have you ever given away the following pieces of information to other parties for economic advantage? [...]

## 11. Behavior, past: “Buy” behavior (i.e., protect information)

- E.g., have you ever used a technology to protect your privacy?
- Encryption
- Document shredders
- Do-not-call lists
- Credit alerts
- Using answering machine/caller ID to screen calls
- [...]

# Demographics

---

- Age: 19-55 (average: 24)
- Education:
  - Some college (currently enrolled) (31.40%)
  - At least 4-years bachelors degree (30.58%)
  - Some graduate (currently enrolled) (17.36%)
  - Masters degree (10.74%)
  - PhD, JD or MD (5.79%)
- Household income:
  - <15,000 43 (35.54%)
  - 15,000-30,000 26 (21.49%)
  - 30,000-60,000 32 (26.45%)
  - 60,00-90,000 9 (7.44%)
  - 90,000-120,000 3 (2.48%)
  - >120,000 8 (6.61%)



# Demographics

---








- Nationalities:
  - USA 83%
  - China, India, Belgium, Venezuela, ...
- Jobs:
  - Unemployed 13 (10.74%)
  - Student (full-time) 50 (41.32%)
  - Part-time jobs 13 (10.74%)
  - Full time computer related/IT 10 (8.26%)
  - Full time business/finance/management 4 (3.31%)
  - Full time public sector 6 (4.96%)
  - Full time medical 4 (3.31%)
  - Full time education/academics 12 (9.92%)
  - Other 9 (7.44%)

# Preliminary results

---

- The following results are **preliminary and rough** simple tabulations from the data-set we have just received
- Please contact the authors for further, forthcoming analysis of this data





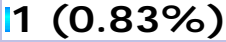


# Privacy attitudes (excerpts)

How important is privacy to you?	
1 - Very important	 73 (60.33%)
2	 31 (25.62%)
3	 9 (7.44%)
4 - Somehow important	 5 (4.13%)
5	 2 (1.65%)
6	 1 (0.83%)
7 - Not important at all	 0 (0.00%)

# Privacy attitudes (excerpts)

Do you think you have enough privacy in today's society?	
Yes	32 (26.89%)
No	87 (73.11%)
How concerned are you about threats to your personal privacy in today's information society?	
1 - Very much	44 (36.36%)
2	21 (17.36%)
3	24 (19.83%)
4 - Somehow	19 (15.70%)
5	10 (8.26%)
6	2 (1.65%)
7 - Not at all	1 (0.83%)

# Privacy attitudes (excerpts)

Has your concern about threats to your personal privacy changed in the last 24 months?	
1 - Much more concerned	 31 (25.62%)
2	 22 (18.18%)
3	 26 (21.49%)
4 - No changes	 41 (33.88%)
5	 1 (0.83%)
6	 0 (0.00%)
7 - Much less concerned	 0 (0.00%)

# Knowledge of privacy risks (excerpts)

**When you are releasing personal information during an ecommerce transaction, how likely do you consider the following outcomes?**

**Attempts to vary price during your next purchase based on your collected data**

1 - Very likely	16 (13.22%)
2 - Quite likely	16 (13.22%)
3 - Somewhat likely	31 (25.62%)
4 - A bit unlikely	34 (28.10%)
5 - Very unlikely	18 (14.88%)
I have no idea	6 (4.96%)

**Use for marketing purposes (e.g., advertising emails)**

1 - Very likely	82 (67.77%)
2 - Quite likely	19 (15.70%)
3 - Somewhat likely	13 (10.74%)
4 - A bit unlikely	3 (2.48%)
5 - Very unlikely	2 (1.65%)
I have no idea	2 (1.65%)

# Knowledge of privacy risks (excerpts)

How likely do you consider the possibility that a 3rd party *can* monitor some details of the following activities you may engage in?

## Using a file sharing client (e.g., Kazaa)

1 - Very likely	70 (57.85%)
2 - Quite likely	22 (18.18%)
3 - Somewhat likely	12 (9.92%)
4 - A bit unlikely	7 (5.79%)
5 - Very unlikely	6 (4.96%)
I have no idea	4 (3.31%)

## Writing a text memo to yourself on a computer connected to the Internet in your organization

1 - Very likely	21 (17.36%)
2 - Quite likely	15 (12.40%)
3 - Somewhat likely	26 (21.49%)
4 - A bit unlikely	34 (28.10%)
5 - Very unlikely	20 (16.53%)
I have no idea	5 (4.13%)

# Knowledge of privacy risks (excerpts)

---

Do you know what Echelon is?	
Yes	■ 15 (12.50%)
No	■ 105 (87.50%)
Do you know what Carnivore is?	
Yes	■ 32 (26.89%)
No	■ 87 (73.11%)
Do you know what Total Information Awareness is?	
Yes	■ 21 (17.50%)
No	■ 99 (82.50%)



# Knowledge of privacy risks (excerpts)

---

**You have completed a credit card purchase with an online merchant. Besides you and the merchant website, who has data about *parts* of your transaction?**

**Nobody: 36.4%**

**Credit card company: 18.7%**

**Hackers: 15%**

*“Nobody, assuming an SSL transaction, without which I would not commit an online transaction using my credit card”*

# Privacy knowledge and overconfidence (excerpts)

---

Can you estimate an interval for which you are 95% sure that it contains the number that correctly answers the following questions?

*Example: Occurrences of identity theft in the US in 2003*

*Solution: lower bound 0.5 Million (complaints with FTC), less conservative estimates: 10 Million*

Rational  
31.9%

Overconfident  
56.3%

Missing  
10.9%

# Privacy risks and bundles (excerpts)

---

<b>Privacy concern</b>	<b><i>Data about offline identity</i></b>	<b><i>Bundled data about offline identity</i></b>
<b><i>Low concern</i></b>	<b>27.30%</b>	<b>6.70%</b>
<b><i>Medium concern</i></b>	<b>26.70%</b>	<b>27.20%</b>
<b><i>High concern</i></b>	<b>20.00%</b>	<b>39.50%</b>
<b><i>Missing data</i></b>	<b>26.00%</b>	<b>26.60%</b>

# Privacy risks and bundles (excerpts)

Imagine that somebody does not know you but knows what your date of birth is, what your sex is, and the zip code where you live. What do you think is the probability that this person can uniquely identify you based on those data?

<10%	29 (23.97%)
11%-25%	26 (21.49%)
26%-50%	28 (23.14%)
51%-75%	13 (10.74%)
76%-90%	7 (5.79%)
>90%	18 (14.88%)
I have no idea	0 (0.00%)

- Sweeney (CMU): 87% of the population of the United States is likely to be uniquely identified by 5-digit ZIP code, birth date, and sex

# Knowledge of privacy protection

(excerpts)

---

- Privacy law:
  - *54% cannot quote a law or describe it*
- OECD Fair information principles:
  - *38% believe they include "litigation against wrongful behavior"*
- Goal: browse anonymously
  - *51% would not know how*
- Goal: browse the Internet with warnings if a website has an incompatible privacy policy
  - *67% would not know how (but most use IE6!)*

# Knowledge of privacy risks and attitude (excerpts)

**Are you informed about the policy regarding monitoring activities of employees/students in your organization?**

	Yes, I am informed	I don't know how such monitoring could take place	I somewhat know ... but don't know the details	There is a policy, but I don't know its details
<b>Low concern</b>	0.00%	0.00%	6.70%	0.00%
<b>Medium concern</b>	0.00%	13.40%	19.60%	6.70%
<b>High concern</b>	19.50%	0.00%	6.70%	20.00%

# “Buy” behavior

---

- 74% adopted some *strategy* or *technology* or otherwise took some particular *action* to protect their privacy:
  - Encryption, PGP
  - Do-not-call list
  - Interrupt purchase
  - Provide fake information
  - [...]
- However, when you look at details, percentages go down...
  - 8% encrypt emails regularly
  - Similar results for shredders, do-not-call lists, caller-IDs, etc.

# “Sell” behavior

---

**When interacting with any party except family and friends (e.g., a merchant or institution), have you ever given away the following pieces of information for a discount or bonus? Or did you receive a better service or recommendation for releasing this information?**

YES (in decreasing order):

- Full name
- Email address
- Home address
- Phone number
- Job title
- Personal interests
- SSN
- Health history



# Attitudes/behavior dichotomy (excerpts)

---

	<b>Have a loyalty card, gave correct info</b>	<b>Have a card, gave fake info</b>	<b>Don't have a card</b>
<b>Low concern</b>	0.00%	6.70%	6.70%
<b>Medium concern</b>	6.70%	13.30%	0.00%
<b>High concern</b>	26.70%	13.30%	0.00%

# Recall of past behavior (excerpts)

---


	<b>Have a loyalty card, gave correct info</b>	<b>Have a card, gave fake info</b>	<b>Don't have a card</b>
<b>Yes, have given identity data</b>	20.00%	0.00%	0.00%
<b>No, have not given identity data</b>	26.70%	26.00%	13.30%

# Password for chocolate?

---

- InfoSec Europe 2004 experiment:
  - 71% of office workers at Liverpool Street Station were willing to reveal their password for a chocolate bar

**Imagine that a person on the street asks you for your access password to your work computer *in exchange* for a chocolate bar. You believe the person does not know you or your workplace. Would you accept the deal?**

Yes  2 (1.67%)

No  118 (98.33%)

- Loewenstein “hot/cold” theory

# "Buy" vs. "sell" price

Is "sell" price higher or lower than "buy" price?

	<b>Social Security Number</b>	<b>Most favorite online user name</b>	<b>Interests outside work/university</b>
<b>Sell &gt; buy</b>	90.00%	76.67%	75.83%
<b>Sell = buy</b>	5.83%	10.00%	10.83%
<b>Sell &lt; buy</b>	0.00%	10.00%	10.00%
<b>Missing</b>	4.17%	3.33%	3.33%

# "Buy" vs. "sell" price

---

## Is "sell" price higher or lower than expected loss?

Sell > expected loss	71.43%
Expected loss > sell	7.56%
Missing	21.01%

## Is "buy" price higher or lower than expected loss?

Buy > expected loss	39.50%
Expected loss > Buy	36.13%
Missing	24.37%

# Does hyperbolic discounting impact privacy attitude?

	<b>Time consistent</b>	<b>Hyperbolic discounting</b>
<b>Low concern</b>	50.00%	54.9%
<b>Medium concern</b>	43.50%	35.3%
<b>High concern</b>	6.50%	9.8%




- Noisy – confounding factors
  - Risk aversion
    - Strong correlation between risk aversion and privacy attitudes
  - Uncertainty
  - Size
- Experiment phase

# Does risk aversion impact privacy attitude?

---

	<b>Risk averse</b>	<b>Risk neutral</b>
<b>Low concern</b>	23.07%	76.92%
<b>Medium concern</b>	30.23%	69.76%
<b>High concern</b>	39.68%	60.31%

# So... who should protect your privacy?

Do you think that privacy should be protected by:	
Government (through legislation)	 65 (53.72%)
Each user by herself (through technology)	 18 (14.88%)
Companies and industry (through self-regulation)	1 (0.83%)
Everybody (warranted naturally through behavioral norms)	 37 (30.58%)
Nobody (should not be especially protected)	0 (0.00%)



# Conclusions (preliminary)

---

- Theory
  - Time inconsistencies may lead to under-protection and over-release of personal information
  - Genuinely privacy concerned individuals may end up not protecting their privacy
  - Large risks accumulate through small steps
  - Not knowing the risk is *not* always the issue
- Evidence
  - Evidence of overconfidence, incorrect assessment of own behavior, incomplete information about risks and protection, attitudes/behavior dichotomy, buy/sell dichotomy
  - Rationality model not appropriate to describe individual privacy behavior
    - Incomplete information
    - Bounded rationality
    - Psychological distortions