

The Economics and Behavioral Economics of Privacy

Alessandro Acquisti

Heinz College/CyLab
Carnegie Mellon University

December 5, 2010

Background

- Laurea: Economics (Rome)
- Masters: Economics (Trinity College); Econometrics and Mathematical Economics (LSE)
- PhD: Information Systems (UC Berkeley)

What

- *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 and Marketing implications

Overview

1. The economics of privacy
2. The behavioral economics of privacy
3. Privacy in online social networks

Overview

1. **The economics of privacy**
2. The behavioral economics of privacy
3. Privacy in online social networks

The economics of privacy

- *Protection & revelation of personal data flows involve **tangible and intangible** trade-offs for the **data subject** as well as the potential **data holder***
- Some of our theoretical studies
 - Microeconomics
 - Game theory

When are compulsory data breach disclosures optimal?

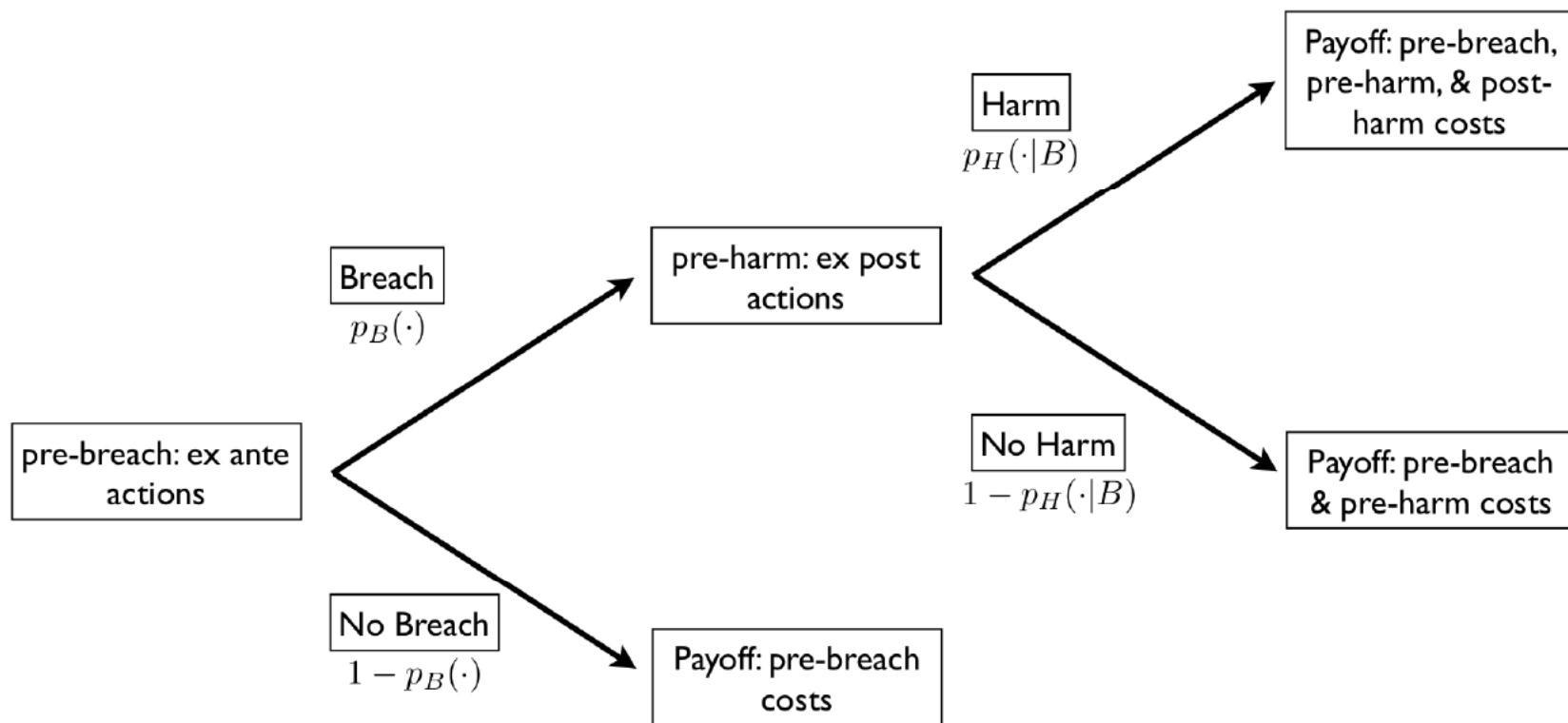


Figure 3.1 Sequence of events

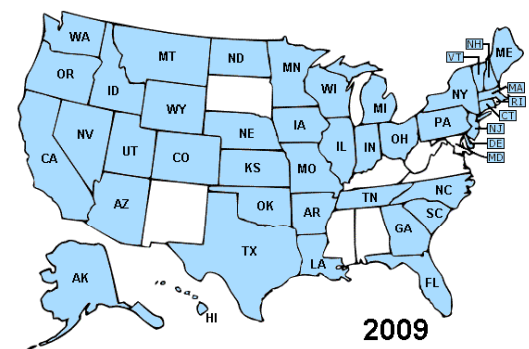
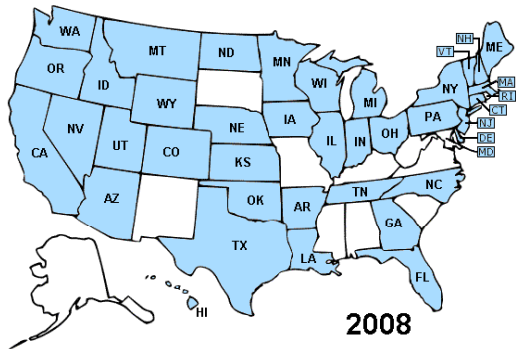
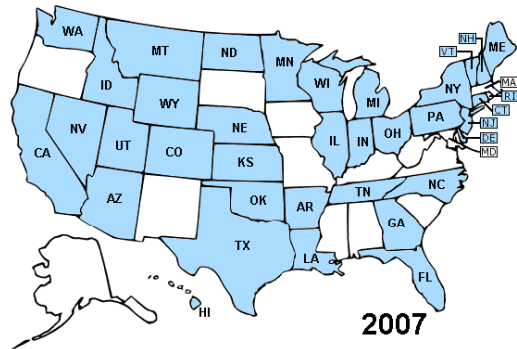
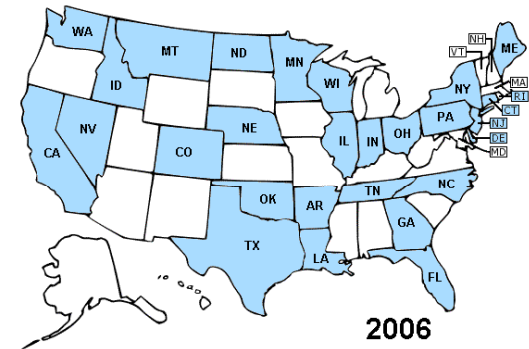
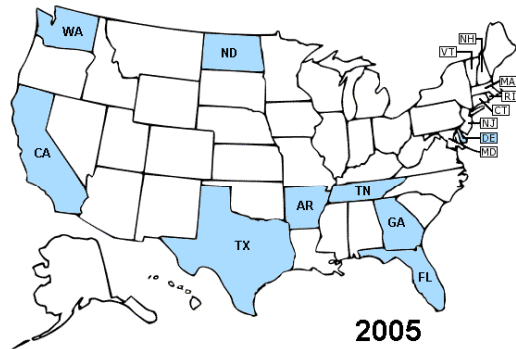
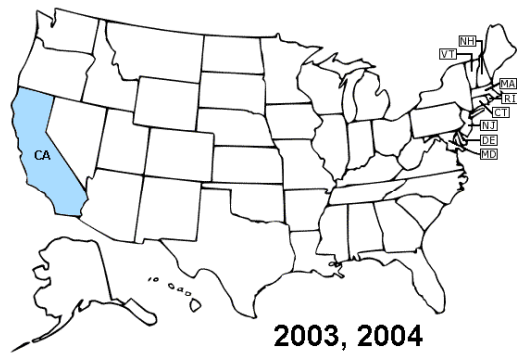
The economics of privacy

- *Protection & revelation of personal data flows involve **tangible and intangible** trade-offs for the **data subject** as well as the potential **data holder***
- Some of our theoretical studies
 - Microeconomics
 - Game theory

The economics of privacy

- *Protection & revelation of personal data flows involve **tangible and intangible** trade-offs for the **data subject** as well as the potential **data holder***
- Some of our theoretical studies
 - Microeconomics
 - Game theory
- Some of our empirical studies
 - Econometric regressions
 - Panel data, Difference-in-difference models, ...

Do data breach disclosure laws reduce identity theft?



Do data breach disclosure laws reduce identity theft?

- Panel data analysis
 - State/year level data about identity theft
 - Different states enacted data breach disclosure laws at different times
- We find that data breach disclosure laws are reducing identity theft by about 6%

Guns, Privacy, and Crime

- Does *online* information affect *offline* crimes?

THE COMMERCIAL APPEAL
Memphis, Tennessee

News Sports Business Opinion Entertainment Lifestyle Media

Today Local State Nation World Weather Obituaries Politics Education Columns

Tennessee Handgun Carry Permits

[« Return to Data Center](#)

Residents of Tennessee can apply for a handgun carry permit with the [Department of Public Safety](#).

Last Name	
First Name	
City	
ZIP	
Search	

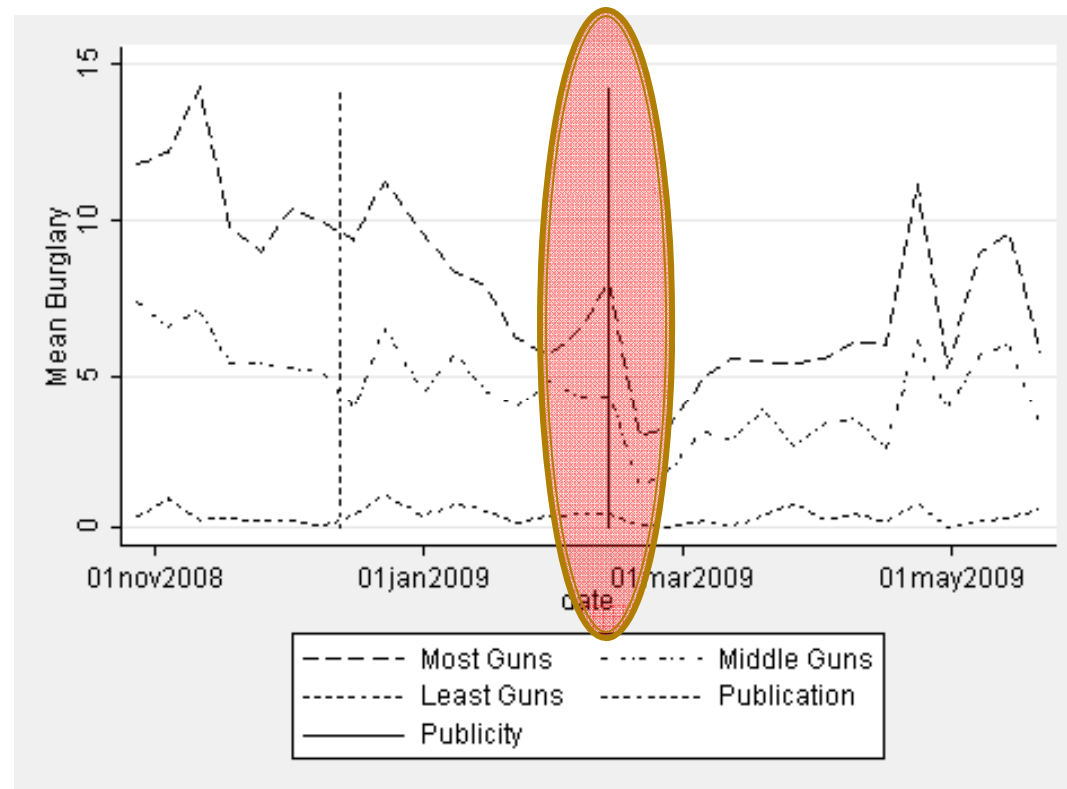
Last Name ▲	First Name	City	ZIP
	JAMES	MEMPHIS	38104
	RICHARD	MEMPHIS	38104
	KATHLEEN	MEMPHIS	38104
	PAUL	MEMPHIS	38104
	REBECCA	MEMPHIS	38104
	MICHAEL	MEMPHIS	38104
	FAYEZ	MEMPHIS	38104
	AHMED	MEMPHIS	38104
	MICHAEL	MEMPHIS	38104
	EDITH	MEMPHIS	38104
	REINALDO	MEMPHIS	38104
	JOE	MEMPHIS	38104
	ROBERT	MEMPHIS	38104
	BRADLY	MEMPHIS	38104
	CATHY	MEMPHIS	38104
	TEJEROLE	MEMPHIS	38104
	LEWIS	MEMPHIS	38104
	GREGORY	MEMPHIS	38104
	MARY	MEMPHIS	38104
	JAMES	MEMPHIS	38104
	JASON	MEMPHIS	38104
	SALIM	MEMPHIS	38104
	MASTIN	MEMPHIS	38104
	MERRITT	MEMPHIS	38104
	JUSTIN	MEMPHIS	38104

1 2 3 4 5 6 7 8 9 10 ► ►►

Records 1-25 of 631

Guns, Privacy, and Crime

- Difference-in-difference model: testing whether crime rates affected post-publicization as function of gun permit density by zip code



Overview

1. The economics of privacy
2. **The behavioral economics of privacy**
3. Privacy in online social networks

The behavioral economics of privacy

- Privacy attitudes/behavior dichotomy
- Hurdles which hamper consumers' privacy decision making
 1. Incomplete information
 2. Bounded rationality
 3. Cognitive/behavioral biases
- Behavioral economics has provided evidence for several systematic “deviations” from the theoretical rational behavior of economic agents
 - Many of those deviations have applications to privacy decision making (as well as information security)

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...
 - ...

The inconsistency of privacy valuations

- 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

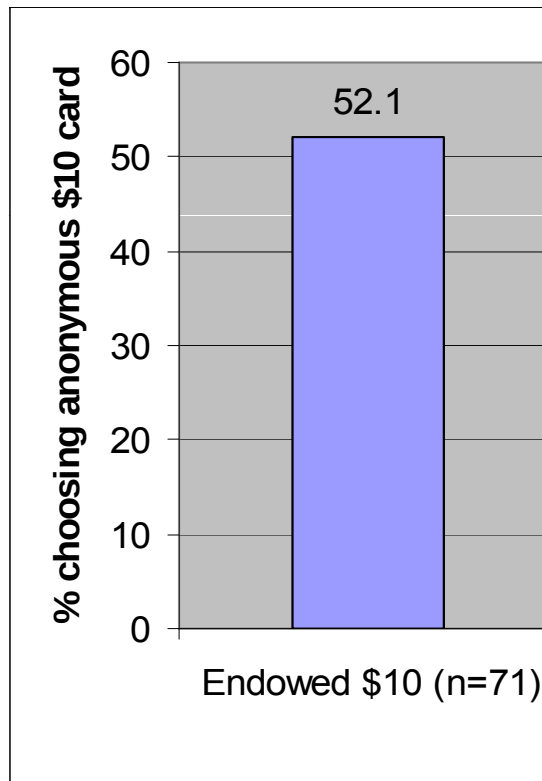
- Hypothesis:

- 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)

Willingness to Pay for Privacy vs. Willingness to Accept for Data



Implications of this stream of studies

- 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
- Rather: Reliance on "revealed preferences" argument for privacy may lead to sub-optimal outcomes
 - People may make disclosure decisions that they stand to later regret
- Creates opportunity for 'nudging' interventions (NSF Large Award, Google Grant)

Overview

1. The economics of privacy
2. The behavioral economics of privacy
3. **Privacy in online social networks**

Privacy in online social networks

- We started the empirical research on privacy and security in online social networks
 - June 2005: mined CMU network (PET 2005)
 - February 2006: mining + survey of CMU Facebook members and non-members (WPES 2006)
- Found evidence of mismatch between attitudes and behavior, mismatch between knowledge and risks, optimism bias, ...

Predicting SSNs from public data

- We reverse-engineered SSN issuance patterns, showing that they were significantly less random than previously predicted
- We found that mere knowledge of an individual's DOB and State of birth is sufficient to predict that individual's SSN

Quick Search

My Profile edit
My Friends
My Groups
My Parties
My Messages
My Account
My Privacy

paid announcement:
CNU students
Needed!
We Need CNU
Undergrads For A
Survey
On Internet Usage,
The First 100
Students
To Participate Will
Receive a Free 40 GB
Ipad

announcement: tell us,
announce something

Search for: [Name] [Email] [Phone] [Address] [City] [State] [Zip] [Country]

Information

Account Info
Name: Dan Bolonis [add to friends]
Member Since: June 23, 2005
Last Update: September 13, 2005

Basic Info
Sex: Male
Residence: Carnegie Mellon '09
School: Student
Status: Student
Concentration: Mechanical
Hometown: Pittsburgh, PA
Home Town: Middlebury, NJ 07748
High School: Middlebury - South High School '05

Contact Info
School Email: dlboloni@andrew.cmu.edu
Screenname: Final

Other Geographies
Dan has friends in...
Monmouth (1) Pittsburgh (1)

Friends at Carnegie Mellon
Dan has 100 Carnegie Mellon friends...

Favorite Books:
Favorite Movies:
Favorite Quotes:

Other Schools
Dan has friends at...
ASU (1) Monmouth (1)
American (2) Montclair (3)
Assumption (1) NJIT (1)
BU (1) NYU (2)

+

Net Detective - Specialty Search - Find out the truth about anyone! - Netscape

http://www.standard.netdetective.net/sp_search_ss.php?first_name=john&middle_initial=

Net Detective - Specialty Search

Find out the truth about anyone! Login Get Advertisement Vacation My Account Help

Did you know? The Social Security Death Index (SSDI) is generated from the US Social Security Administration's Death Master File.

Search By: Unclaimed Money | S.S. Death Index

First Name Last Name State

SSN Validate Search

SS Death Index Search Results

Subject: JONES, JOHN

State: FL Total Entries: 211 (Returned: 1 - 20)

Next >

#	Name	SSN	Last Residence Zip	Last Residence State	Date of Birth	Date of Death
1	JONES, JOHN C.	312-22-2212	31212	FL	1911-08-03	1993-09-27
2	JONES, JOHN W.	344-36-4946	34422	FL	1927-02-26	2000-03-31
3	JONES, JOHN E.	331-14-9791	32719	FL	1927-04-04	1998-02-14
4	JONES, JOHN W.	196-20-7542	31763	FL	1905-12-20	1996-12-27
5	JONES, JOHN E.	262-14-6236	33541	FL	1918-09-03	2000-03-13
6	JONES, JOHN T.	343-20-3999	34609	FL	1914-11-10	1999-06-04
7	JONES, JOHN H.	383-19-3213	33309	FL	1923-02-22	1994-01-18
8	JONES, JOHN B.	389-12-7415	33611	FL	1924-06-26	1999-09-26
9	JONES, JOHN O.	396-19-2308	33904	FL	1922-01-07	1989-04-06
10	JONES, JOHN W.	1001-04-1901	33701	FL	1913-11-18	1980-10-09

= SSN

Individual Re-Identification via Online Social Networks' Images

Unidentified
Database:
Match.com Photo



Identified Database:
Facebook Photo



One (or two) examples

1. The paradox of control
2. Discounting past information

Joint works with Laura Brandimarte, Joachim Vosgerau, and George Loewenstein

Two examples

1. The paradox of control
2. 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."*

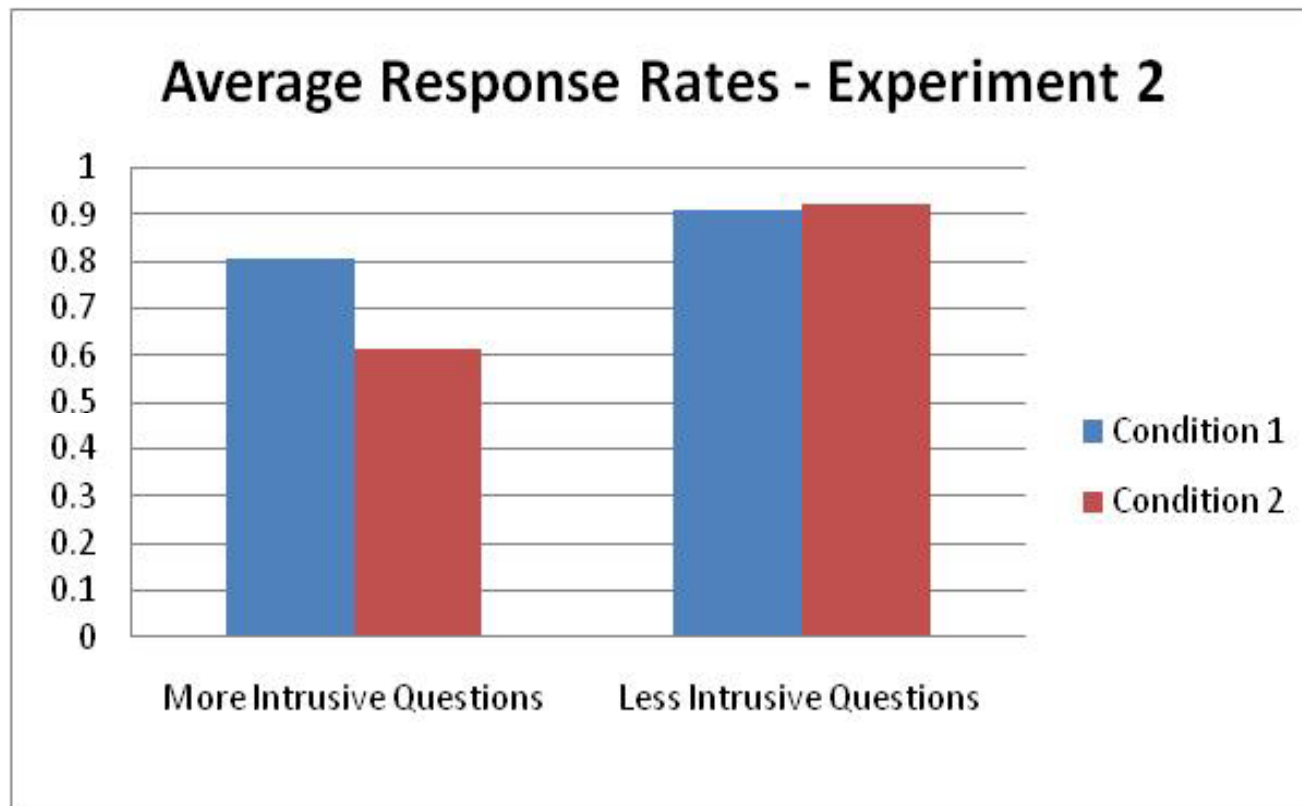
- Condition 2

*"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



Results (DV: Answer yes/no)

Table 2.

	<i>Coeff</i>	<i>P-value</i>
Treatment	-.25**	.05
Intrusive	-.64***	.00
Treat_Int	-.67***	.00
Age	-.02	.28
Male	.20	.10
	N= 132	Prob > χ^2 = .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

qualtrics.com*

Study on Ethical Behavior

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

	Yes	No
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 had sex in a public venue (e.g.: restroom of a club, airplane)?	☐	☐
9. Have you ever made a donation to a non-profit organization?	☐	☐
10. Do you have any permanent tatoos?	☐	☐

Close

Explicit control condition

qualtrics.com

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

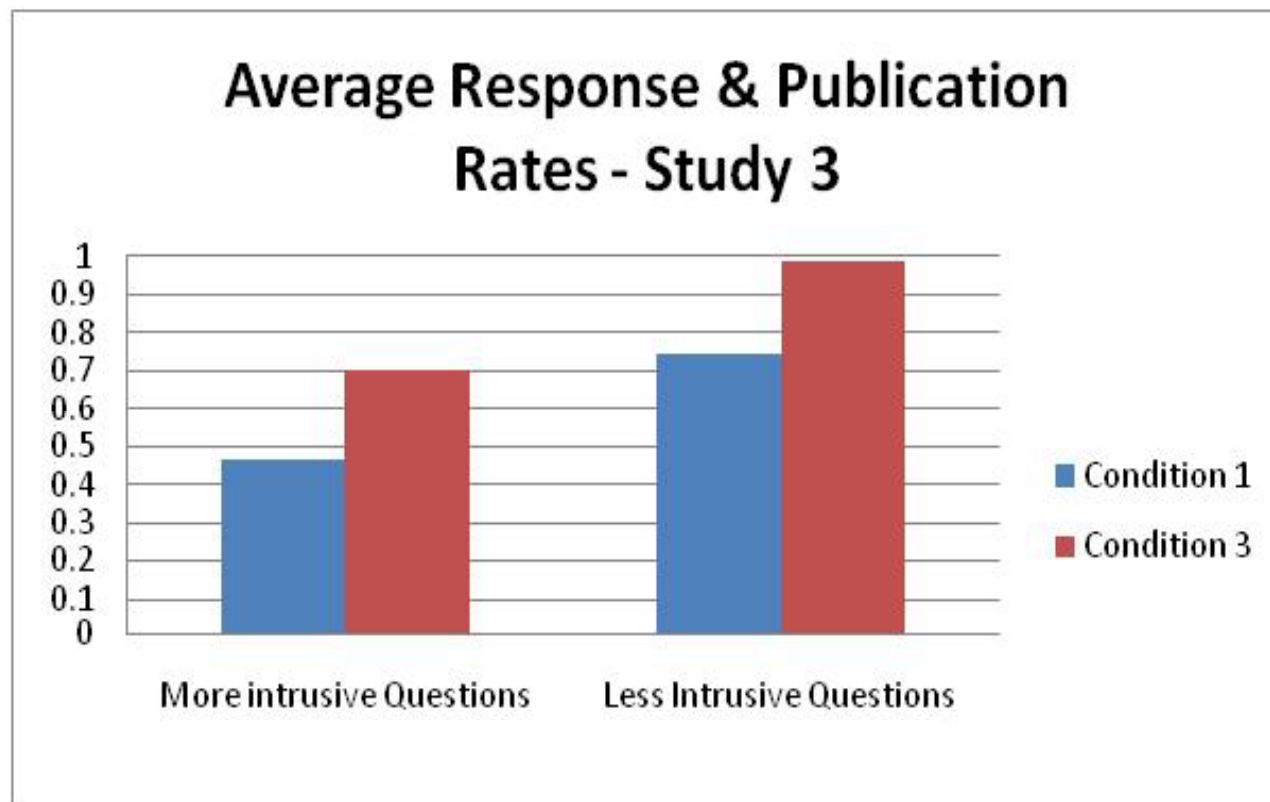
	Yes	No
1. Are you married?	☐	☐
2. Have you ever been fired by your employer?	☐	☐
3. Have you ever stolen anything (e.g.: from a shop, a person)?	☐	☐
4. Have you ever used drugs of any kind (e.g.: weed, heroin, crack)?	☐	☐
5. Have you ever lied about your age?	☐	☐
6. Have you ever had cosmetic surgery?	☐	☐
7. Have you ever done any kind of voluntary service?	☐	☐
8. Have you ever had sex in a public venue (e.g.: restroom of a club, airplane)?	☐	☐
9. Have you ever made a donation to a non-profit organization?	☐	☐
10. Do you have any permanent tatoos?	☐	☐

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



Results (DV: Publish yes/no)

	<i>Coefficients and p-values</i>
Treatment	1.92** (.000)
Intrusive	-.85** (.000)
Treat_Int	-1.21** (.002)
Age	.03 (.521)
Male	-.11 (.593)
N	65
Prob > χ^2	.000

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*

Two examples

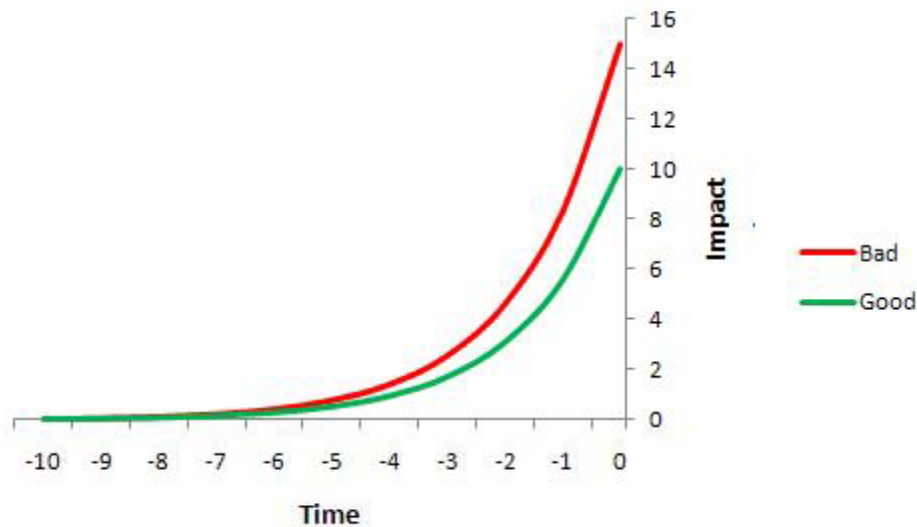
1. The paradox of control
2. **Discounting past information**

Research question

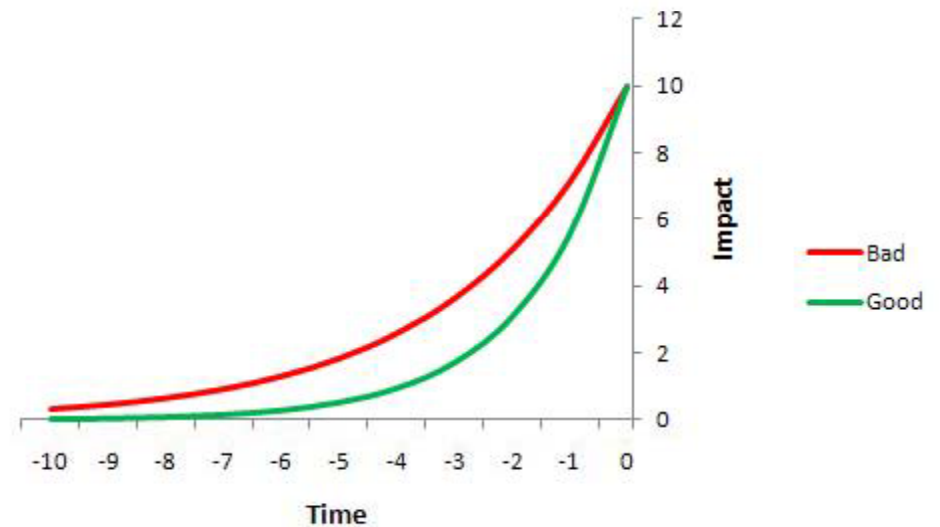
- How does information about a person or company's past, retrieved today, get 'discounted'?
- Specifically: does information about a person's past with **negative valence** receive more weight in impression formation than information with **positive valence**?

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

Baseline condition

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.

Treatment conditions (positive/recent)

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 .***

Treatment conditions (negative/recent)

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 .***

Treatment conditions (positive/old)

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.

Treatment conditions (negative/old)

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:

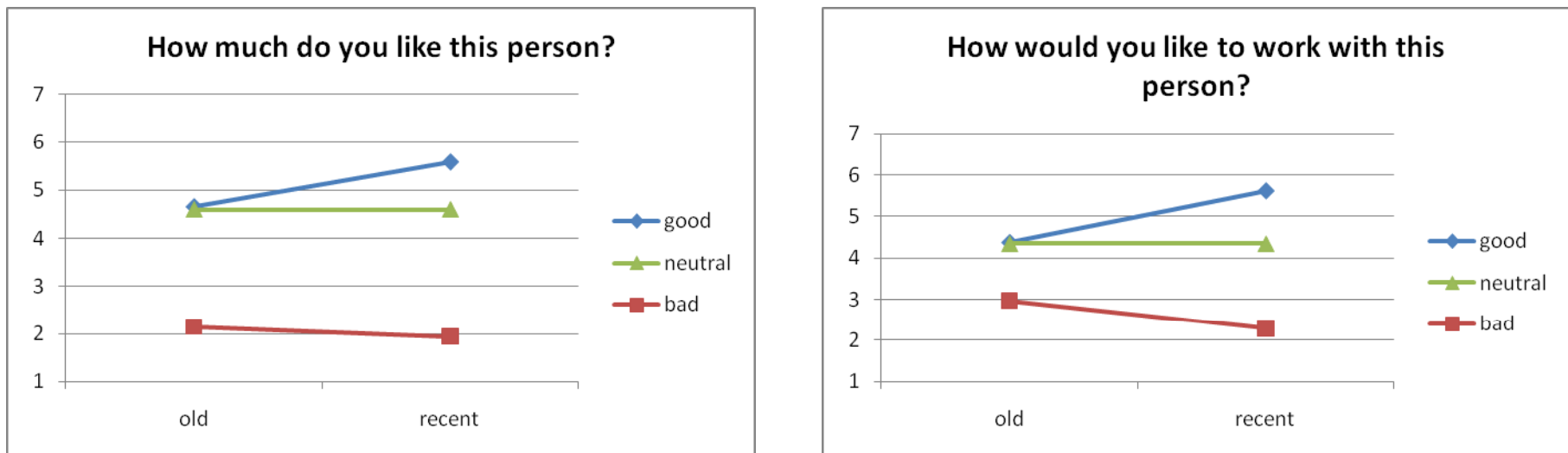
Reported wallet, 5 years ago	Reported wallet, 12 months ago
Did not report wallet, 5 years ago	Did not report wallet, 12 months ago

Dependent variables

- Dependent variables:
 - How much subjects liked the person described
 - How much subjects would have liked to work with her (Interpersonal Judgment Scale, Byrne 1961)

Results

Figure 6. Average level of liking and trust indices across conditions in Experiment 2



Implications

- Bad is not just stronger than good...
- It is also discounted differently than good
- Implications: future impact of information revealed today

Overall implications of these privacy studies

- 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