Privacy Trade-Offs of Geo-Location – General Population Concerns and an Application to the 2020 US Census

Laura Brandimarte
Alessandro Acquisti

How does this affect willingness to disclose
personal information?

- What is the effect of making people aware that their location can be easily identified on willingness to provide further personal information?
  - Will this represent an incentive to disclose personal information, thus increasing response rate to a questionnaire?
  - **Will it raise privacy concerns, and thus backfire?** (Barkhuus & Dey, 2003; Sadeh et al., 2009; Toch et al., 2010)
- Application: US Census 2020 – Are there location privacy concerns specific to the Government or is the Government trusted (Joinson, 2009)?

Context

- NSF Census Research Network (NCRN)
  - 2020 Census
  - Several American universities
  - Census Bureau

Methodology

- Four between-subject randomized experiments

- Manipulations: geo-location awareness, institution requesting data, and salience of privacy

- Dependent variables
  - Perceived intrusiveness of questions
  - Propensity to provide sensitive information
    - Engagement in unethical behaviors (Brandimarte, Acquisti & Loewenstein, 2013; Joinson, Woodley, & Reips, 2007; Phelps et al. 2000; Weisband & Kiesler, 1996)
    - Census-related questions (demographics and living arrangements)

Experiment 1

- Three conditions (we captured location in all of them):
  - Control
  - Geo-Located (Country, State, City, Zip code)
  - Requested Location (Country, State, City, Zip code)
- 7 Census-related questions
- 16 sensitive behavior questions
  - Why did we use them?
- Exit questions about privacy concerns, feeling tracked or monitored

Experiment 1 – Geo-Located Condition

Before we begin with the actual questions, please first verify that you are currently in the following location:

	That's correct	That's incorrect or missing
City: Pittsburgh	○	○
State: PA	○	○
Country: United States	○	○
Zip code: 15213	○	○

Experiment 1 – Requested Location Condition

Before we begin with the actual questions, please tell us your current location.

Current City

Current State

Current Country

Current Zip code

Experiment 1 – Census questions

Where do you currently live?

- House
- Apartment
- Mobile home
- Other - please specify

Do you currently live with other people?

- Yes
- No - I live by myself

Do you own the place where you live?

- Yes
- No - I'm renting
- No - I'm living there without paying rent

Do you have a mortgage or loan?

- Yes
- No

Experiment 1 – Sensitive behavior questions

For each of the following, please indicate how often (if at all) you personally did the described action.

Have you ever:

|	Never	Once	Twice	Sometimes	Often	I prefer not to say
Made up a serious excuse, such as grave illness or death in the family, to get out of doing something?	○	○	○	○	○	○
Let a friend drive after s/he had too had much to drink or had used drugs?	○	○	○	○	○	○
Called in sick when you were not sick?	○	○	○	○	○	○
Tried to gain access to someone else's phone or email without their knowledge or consent?	○	○	○	○	○	○
Had sex with the current partner of a friend?	○	○	○	○	○	○
Lied about your age?	○	○	○	○	○	○
Cheated at sports or games?	○	○	○	○	○	○
Stolen anything worth more than $25?	○	○	○	○	○	○

Experiment 1 - Results

- 403 Mturk workers (37% female, M_{age} = 29.8, SD = 9.4)
- Census-related score: ceiling effect
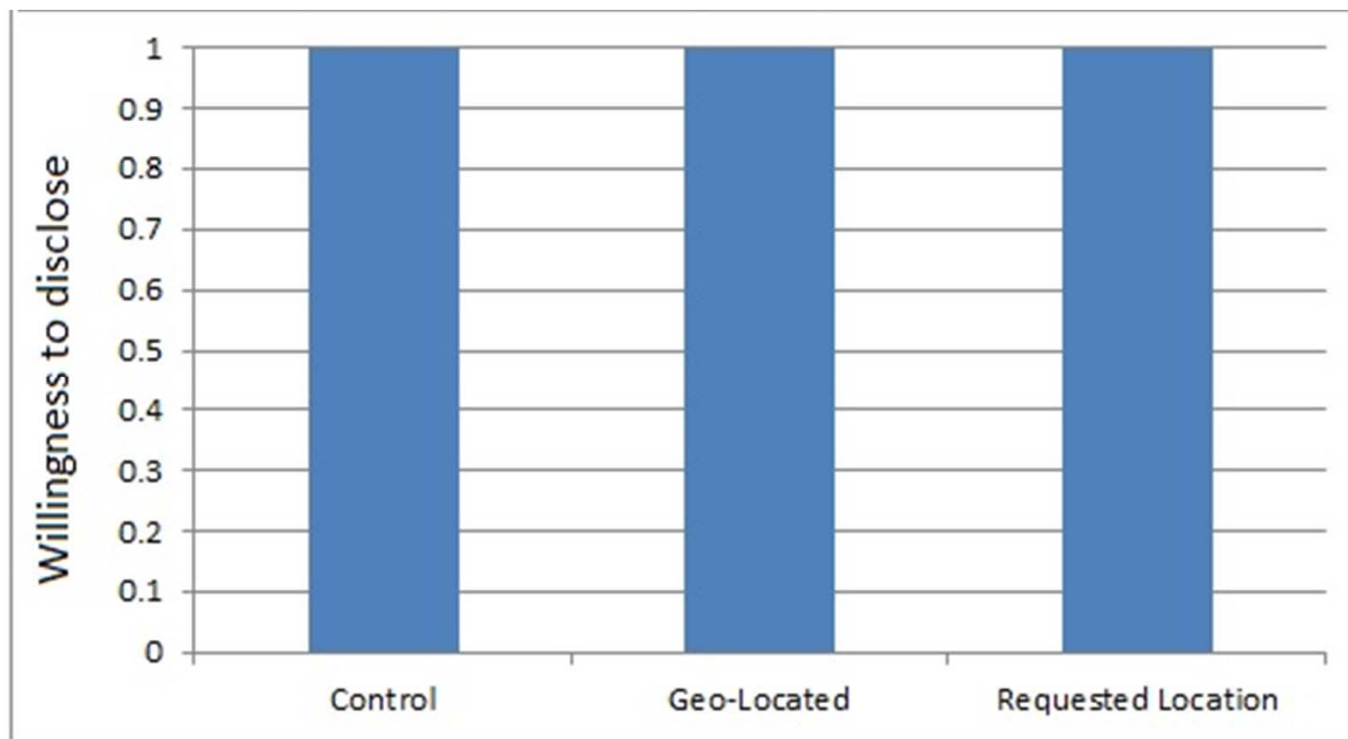
Experiment 1 - Results

- Sensitive behaviors, average disclosure score:
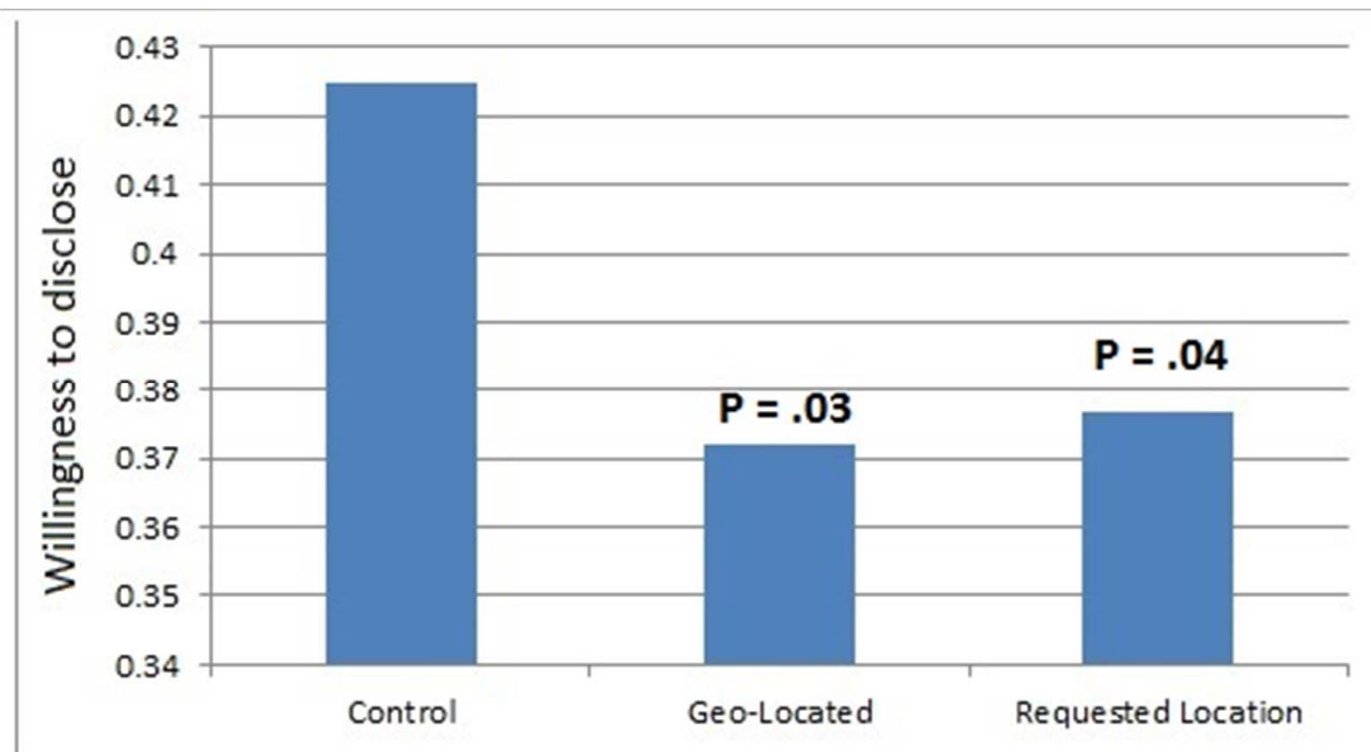
Experiment 1 - Results

- Sensitive behaviors questions: Panel specification, probit estimation

**Table 1. Marginal effects for panel probit estimation – Experiment 1 l. Standard errors in brackets.
** indicates significance at the 1% level; * at 5% level.**

	Disclosure
Control	.138* (.070)
Requested_Location	.006 (.070)
Male	.194** (.061)
Age	-.003 (.003)
White	.213** (.065)
	Number of groups = 403 Wald χ^2 (5) = 26.78

Experiment 2

- N = 694 MTurk Workers (41.2% female, M_{age} = 31.1, SD = 10.6)
- 3x3 between-subjects, manipulating geo-location (Control, Geo-Located, and Requested Location) and Institution (Researchers, Census Bureau, Government)
- 12 Census-related questions
- 16 sensitive behavior questions
- Exit questions

Experiment 2 – Geo-Located Conditions

Before we begin with the actual questions, please first verify that you are currently in the following location.
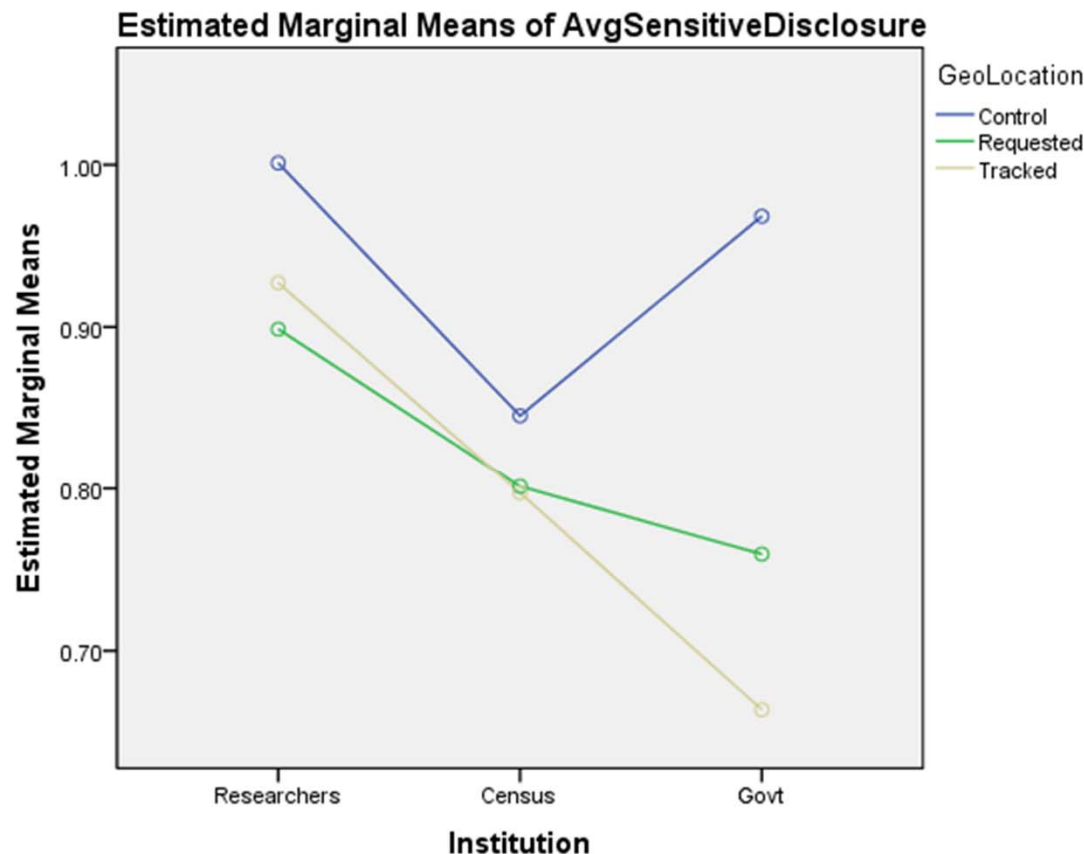
Notice that, for your privacy, we are not showing the full 5-digit Zip code. Please just confirm whether the first two digits (if shown) are correct.

	That's correct	That's incorrect or missing
City: Pittsburgh	○	○
State: PA	○	○
Country: United States	○	○
Zip code: 15***	○	○

Experiment 2 - Results

- **Census questions**: ceiling effect

Experiment 2 - Results

Estimated Marginal Means of AvgSensitiveDisclosure

GeoLocation
— Control
— Requested
— Tracked

y-axis: Estimated Marginal Means — 1.00, 0.90, 0.80, 0.70

x-axis: Institution — Researchers, Census, Govt

Covariates appearing in the model are evaluated at the following values: Gender = 1.4121, Age = 31.1311, White = .7666

- **Census questions**: ceiling effect
- **Sensitive questions**
- Main effect of geo-location $(F(2,682) = 4.165, p < .05)$
  - Higher disclosure rates in the Control conditions that in the two geo-location conditions $(t(685) = 3.22, p = .001)$
  - Requested Location and the Geo-Located condition did not differ from each other $(t(685) = .48, p > .10)$
  - **Effect is strongest for Government institutions (less trusted?)**
- Main effect of type of institution requesting the data $(F(2,682) = 4.493, p < .05)$
  - Higher disclosure rates if Researchers requested info as compared to Census or Government $(t(685) = 3.40, p = .001)$
- No significant interaction

Experiment 3

- N = 603

- Design: 3x2 between-subject

  – Manipulate the alleged entity requesting the data (Governmental institution, Census specifically, Researchers) and the presence of surveillance priming (participants solve anagram containing of either "Snowden" or "Clinton")

- DV: perceived intrusiveness of Census-related questions and sensitive behaviors questions

Experiment 3

To make sure that you are focused on this survey, we ask that you please try to solve the following anagram.

The 7 letters shown below can be used to make the last name of a famous person. These 7 letters are presented here in a random order. Can you think of who the famous person is? Please write your guess in the text box below.

C N T N L I O

Experiment 3

To make sure that you are focused on this survey, we ask that you please try to solve the following anagram.

The 7 letters shown below can be used to make the last name of a famous person. These 7 letters are presented here in a random order. Can you think of who the famous person is? Please write your guess in the text box below.
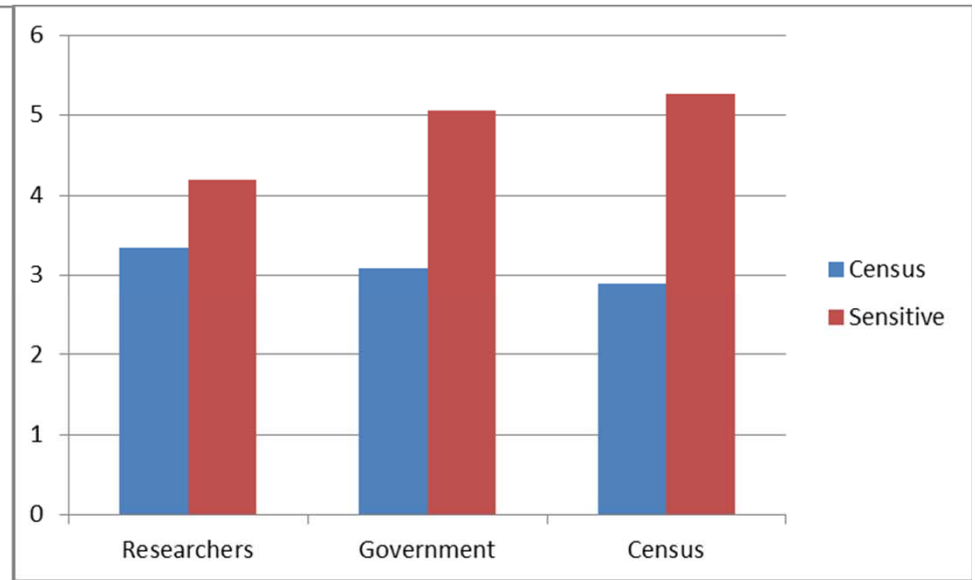
S W D N N O E

Experiment 3 – Results

Perceived Intrusiveness Without Prime
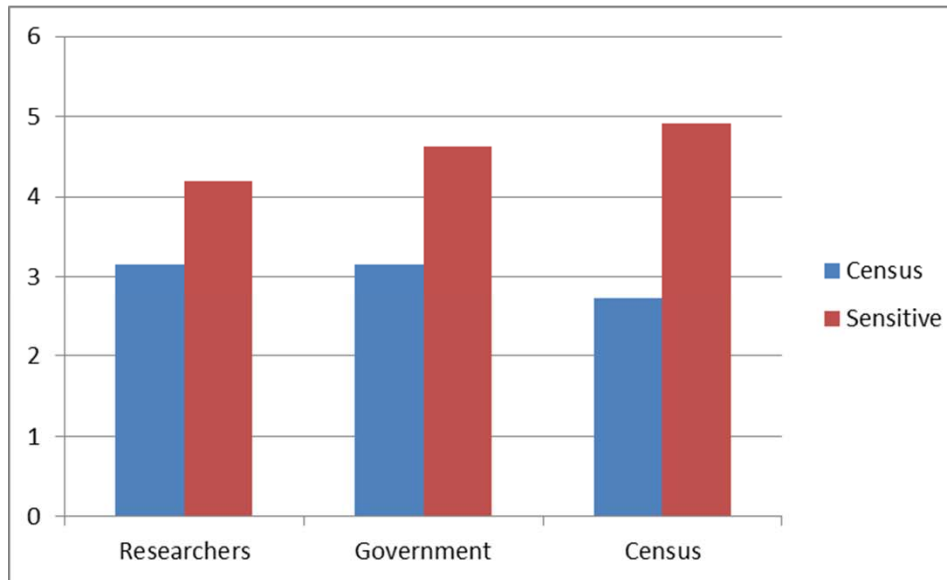
Perceived Intrusiveness With Prime

- Census questions less intrusive than sensitive behaviors questions ($p < .001$)
- Census questions: main effect of institution ($F_{(2, 596)} = 5.476$, $p < .01$), but no effect of priming. No significant interaction
- Sensitive questions: main effect of institution ($F_{(2, 596)} = 15.721$, $p < .001$) and priming ($F_{(2, 596)} = 4.327$, $p < .05$). No significant interaction

Experiment 4

- N = 601 MTurk Workers (43% female, M_{age} = 31.2, SD = 10.3)
- All Geo-Located (saw their City, State, Country and first 2 digits of zip code)
- Design: 3x2 between-subject
  - Manipulate the alleged entity requesting the data (Governmental institution, Census specifically, Researchers) and the presence of surveillance priming (participants solve anagram of either "Snowden" or "Clinton")
- DV: same as Experiment 2: Census-related questions, sensitive behavior questions
- Exit questions about privacy concerns, feeling tracked or monitored

Experiment 4

In the next page, you will be shown a list of various behaviors that people sometimes engage in.

We ask that you read each of those and indicate how often (if at all) you personally did each of the described behaviors.

Please notice that we are collecting this data for our research purposes, but if you feel like some of the questions are intrusive, you always have an option not to answer.

In the next page, you will be shown a list of various behaviors that people sometimes engage in.

We ask that you read each of those and indicate how often (if at all) you personally did each of the described behaviors.

Please notice that we are collecting this data on behalf of the Census Bureau, a US Governmental Institution, for their research purposes, but if you feel like some of the questions are intrusive, you always have an option not to answer.

In the next page, you will be shown a list of various behaviors that people sometimes engage in.

We ask that you read each of those and indicate how often (if at all) you personally did each of the described behaviors.
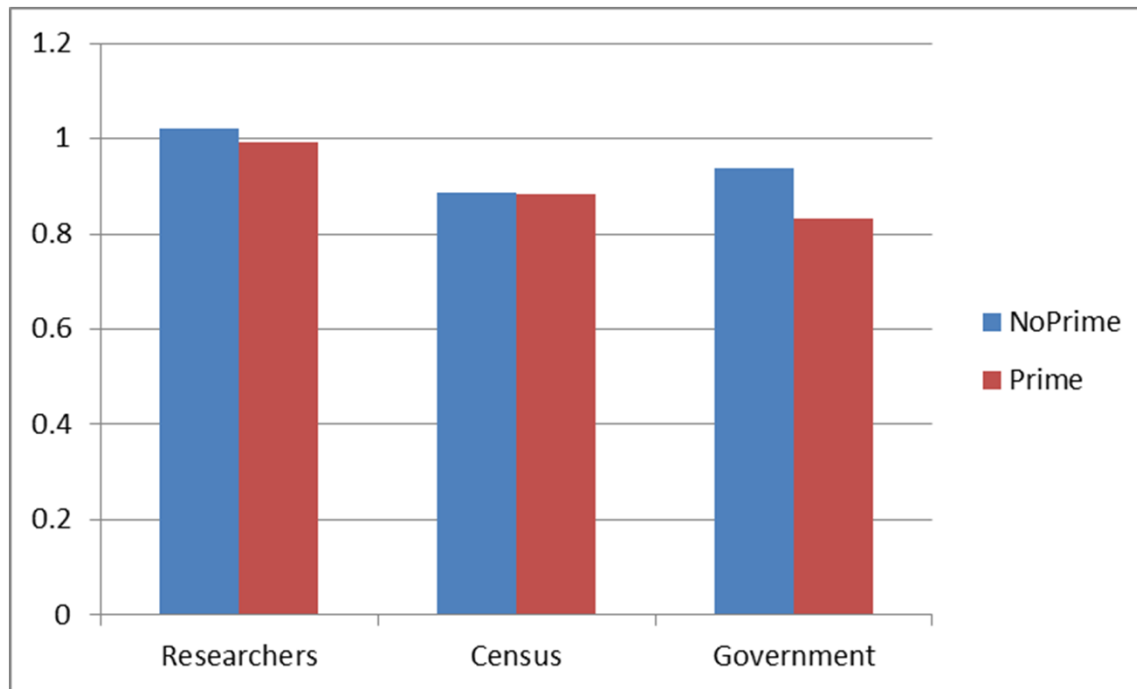
Please notice that we are collecting this data on behalf of the US Government for their research purposes, but if you feel like some of the questions are intrusive, you always have an option not to answer.

Experiment 4 – Results

- **Census questions**: again ceiling effect

Experiment 4 – Results

Sensitive behaviors disclosure scale

- **Census questions**: again ceiling effect
- **Sensitive questions**
- Main effect of institution (F(2,592) = 3.93, p < .05)
  - Higher disclosure to researchers (M = 1.01, SD = .57, t(598) = 2.397, p < .05) than to Census or Government (M = .88, SD = .59)
- No effect of surveillance priming
- No significant interaction

Conclusions

- Our results (always ceiling effect on Census questions) suggest that awareness of geo-location will not affect willingness to disclose non-sensitive information…

- …but it decreases willingness to provide **sensitive information…**

- …and people seem less comfortable disclosing to Census or **Government institutions** than to researchers

- Problem if actual Census forms are **perceived as privacy intrusive**

Conclusions

- This could extend to other (non-location) data (e.g., administrative data from DMV)

- Alternatives to geo-location:
  - Ad campaigns focusing on the completion of the form as a duty
  - Emphasize difference between geo-location and location tracking

Thank you!

lbrandim@andrew.cmu.edu

Questions?