Public Opinion on Specific Potential Uses of Administrative Data

Aleia Clark Fobia, U.S. Census Bureau

FCSM. College Park, MD

October 15, 2024

Background

- Generally expanding use of administrative data by national statistical agencies

- Opinion studies on use of administrative data in the US Federal Statistical System has been ongoing

- Analytic/theoretical frameworks often tacit and focus mainly on privacy concerns
  - Posits use of administrative data as a change that threatens privacy
  - What are people concerned about in terms of the use of admin data?

Contextual Integrity

- What is contextual integrity? Nissenbaum (2009; 2011)
  - Appropriate information flows-what information is appropriate to reveal in a given context
  - Distributional norms-transfer of information from one party to another
- How does thinking about privacy from the perspective of contextual integrity shift the conversation around the use of administrative data?

Research Questions

- In what ways, if any, does the use of administrative data challenge appropriate information flows in the context of government data collection?

- In what ways, if any, does the use of administrative data challenge distributional norms in the context of government data collection?

Data and methodology

1. Two-year, multi-mode research project designed to measure public perceptions regarding a national demographic frame (RTI and Census)
   - Nationally representative web survey using an address-based probability panel (May 2022 – February 2023); and 10,000 complete responses; Qualtrics; Ipsos Knowledge panel
   - Focus groups with targeted groups and individuals who participated in the web survey (March 2023 – September 2023).
   - 3 groups conducted virtually with web survey participants, 9 in-person groups in Washington, D.C., Illinois, North Carolina, California- 12 groups 91 participants

2. Mixed-methods research project focused on public opinions on specific uses of administrative data at the Census Bureau (RTI and Census)
   - Focus groups followed by development of survey instrument. Findings from focus groups informed survey instrument. 20 cognitive interviews to test instrument. 10 in-person focus groups.
   - Instrument fielded in May 2024 on the Census Household Panel-9,700 responses

Appropriate information flows

- Appropriate data
- Control over data collection
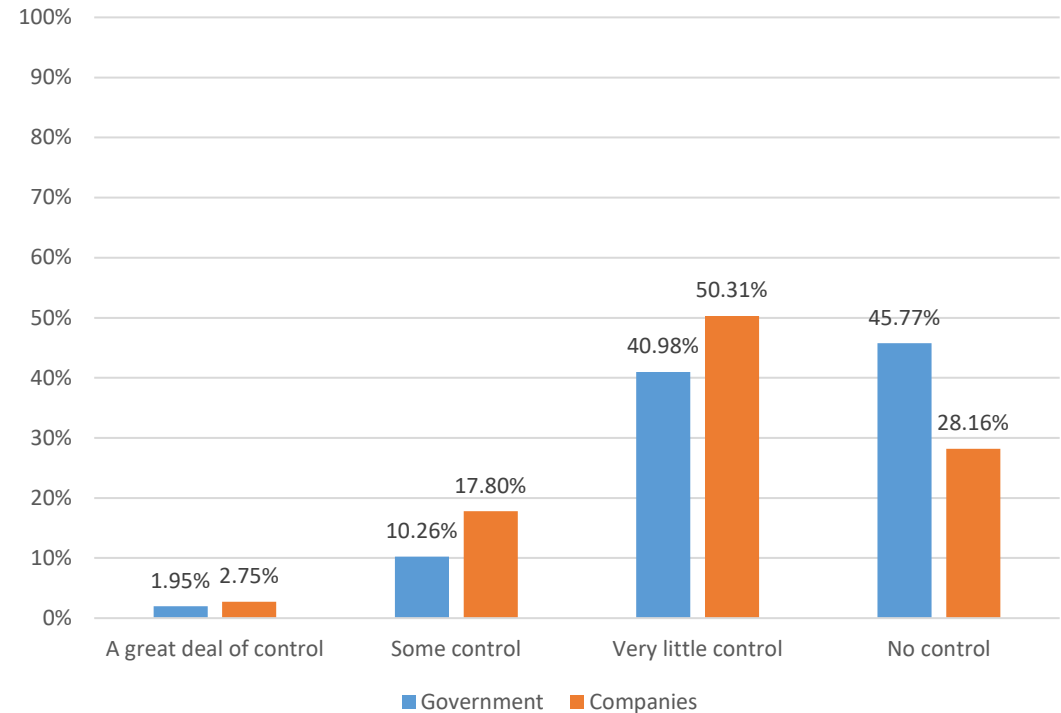- Accuracy and Information flows
- Burden and control

Appropriate data

- Willing to share certain kinds of data but more hesitant about other data.
  - Willing to share publicly available data, such as voter registration or home ownership; demographic information such as age, gender, or race; and number of people in their household.
  - Less willing to share income, social security number, and specific employment information such as employer name
  - Spanish-speaking participants were particularly worried about sharing any personal identifying information like names, addresses, and dates of birth.
- Clear boundaries around the appropriateness of certain types of data
  - "[The government does] not have to know if I have good or bad credit score, of if I've paid my car. There is information that is strictly mine […], they need to limit it to what they need. They don't have to have access, that is my privacy." (Cognitive interviews)

Control over data collection

- Generally, not much control- less with government
  - "[There are] both benefits and risks and no way to avoid the government getting this information." (Cognitive Interview)
  - Refusing to answer questions or entire surveys or answering untruthfully to questions as a form of control.

- Giving data is required to function in society both private sector and government
  - "You need to provide data to do everyday life things, like get a job with a social security number or drive a car."
  - "On some websites I do have control, but some there is no option of turning that off and you have to accept the cookie settings in order to use the website."

How much control do you feel you have, if any, over the data that companies/the government collects about you?

	A great deal of control	Some control	Very little control	No control
Government	1.95%	10.26%	40.98%	45.77%
Companies	2.75%	17.80%	50.31%	28.16%

Government ■ Companies ■

Information flows

- Use of administrative records for non-responders
- Primary concerns related to the accuracy of administrative data.
- Participants were concerned that administrative data might not accurately capture moves, changes to household composition (e.g., deaths, births), and changes to gender identity.
- Some felt strongly that everyone should have the option to complete the census.
- A few were concerned about the potential for bias for who gets to fill out their census if administrative records are used to decide who gets the option to complete the census.
- Some participants thought that using records data was invasive, had concerns about how the data would be used, preferred to fill out the census themselves, and had more concerns about data accuracy than other groups.
- Participants from one focus group preferred the idea of providing administrative data to households and asking for corrections.

Burden and control

- Framing experiment in a question about support for using administrative data
  - This database will mean that you are asked to fill out government surveys less often because your information is already available.
  - This database will help complete survey responses when information is missing.
- Interpreted as decreasing control over data
  - These three participants pointed out that filling out surveys less often could hurt you in that the information on file may be outdated. In the words of one participant [FG12 – Older AA], "It's good to be able to re-up information with the government on where you are at in life…your current status."
- Removes respondents right to refuse data collection
  - "People may be skipping for a reason, and you're taking that right away from them to not answer."
- Some saw as a way to increase accuracy
- Burden is normal and expected and removing it takes away some control in the process of data collection.
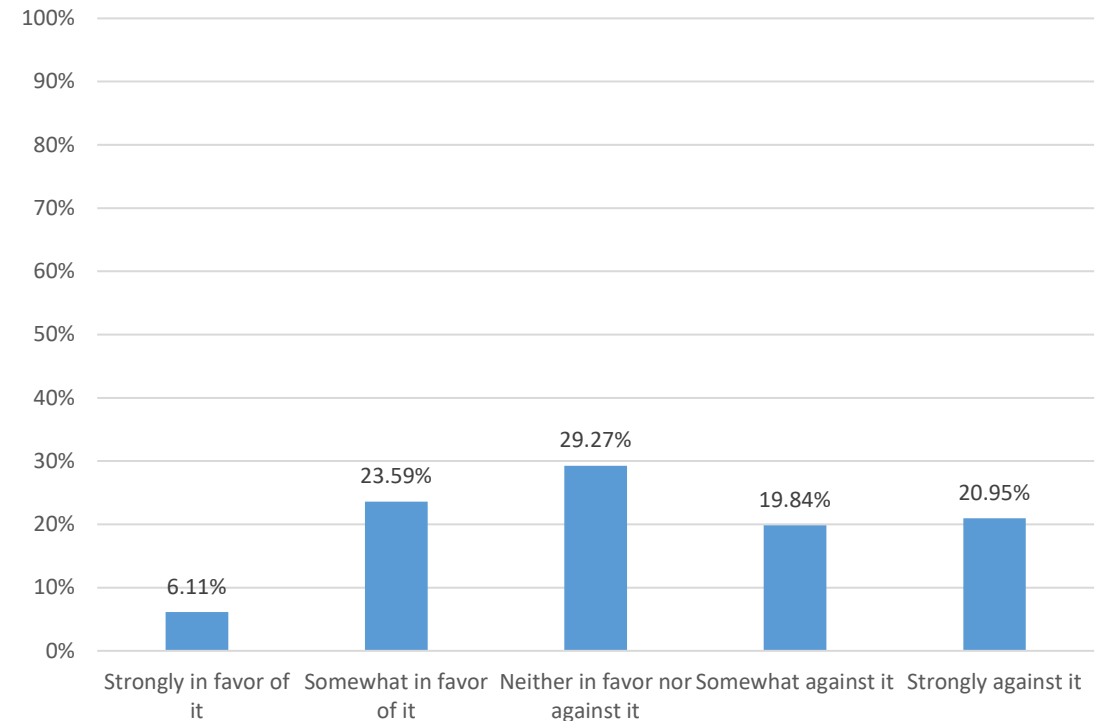
Distributional norms

- Data use expectations
- Boundaries of information flow
  - Private company data
  - Admin data vs data from neighbors/landlords
- Norms of consent to distribute
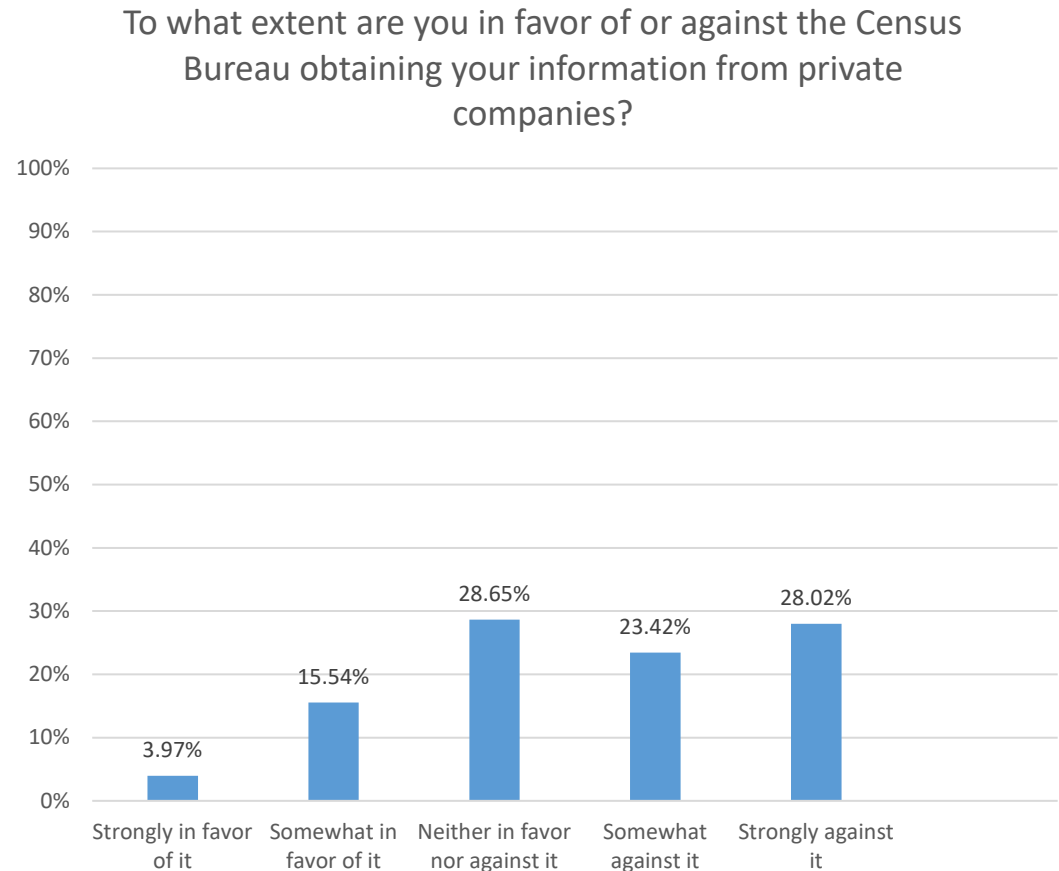
Data use expectations

- Respondents expect to know how data will be used
  - "I would want to know what the new purposes are, more transparency. I would be more in favor if I knew more about how it's being used.
  - "I don't know what 'new purposes' means. That's vague. When things are vague, it makes me uncomfortable. Could mean more taxes. Could mean better healthcare. So, not knowing what that is, is terrifying because I don't trust the government at all."

- Shared data is tied to the perceived use of that data

To what extent are you in favor of or against the information you provided for one purpose being used by the government in additional ways?

Strongly in favor of it	Somewhat in favor of it	Neither in favor nor against it	Somewhat against it	Strongly against it
6.11%	23.59%	29.27%	19.84%	20.95%

Boundaries: Data from private companies

- Government getting data from private companies violates distributional norms
  - "I think it's better that the government shares information among themselves, but I'm a little uncomfortable with private companies being a part of it." (Cognitive interview)
  - "It means government buying data from private companies, it's just another privacy sink where we're losing more of our privacy to government, and they've figured out a way to do it legally." (Focus Group)

To what extent are you in favor of or against the Census Bureau obtaining your information from private companies?

Response	Percentage
Strongly in favor of it	3.97%
Somewhat in favor of it	15.54%
Neither in favor nor against it	28.65%
Somewhat against it	23.42%
Strongly against it	28.02%

United States®
Census
Bureau

Boundaries: Admin data vs data from neighbors/landlords

- Asked focus group participants about information from neighbors or landlords versus use of admin data
- Many thought that their neighbors/landlords would provide less accurate data compared to administrative records. There were more negative reactions towards landlords compared to neighbors.

  - Participants believed that neighbors/landlords would not know enough about their household to answer the census on their behalf.
  - They did not want their neighbors/landlord making assumptions about their gender and race, or assuming they knew information that they did not.
  - Some worried that landlords might not know the number of people in the home because renters sometimes hide how many people stay at a residence.
  - A few were worried about purposefully false reporting.
  - Some mentioned they would not be comfortable sharing a neighbor's details.

United States® Census Bureau

Norms of consent to distribute

- Respondents mentioned consent consistently throughout
  - Two participants thought that individuals should provide consent for sharing with other agencies and that it should be their choice about which agencies can access their data. (Cognitive interviews)
  - Four participants brought up the question of whether someone would be able to explicitly consent to additional uses of their data at the point of providing the information. (Cognitive interviews)

Conclusion/discussion

- Contextual integrity shifts emphasis to how use of administrative data does or does not violate appropriate information flows and distributional norms
  - Clear boundaries around use of data from private companies, appropriate types of data
  - Norms around knowledge/communication of data use, consent, control
- Critiques of contextual integrity (McDonald and Forte 2020)
  - Norms are not created by all groups equally
  - Norms are elements of social control rather than cultural agreement
- Use of administrative data is not going away- findings are instructive for how to message around their use

Thank you!

Aleia Clark Fobia

aleia.yvonne.clark.fobia@census.gov