

# W4 Ethics

## Privacy and Confidentiality

### Privacy and Confidentiality

- Privacy: Includes the famous right to be left alone, and the ability to share information selectively but not publicly (U.S. White House 2014)
- Confidentiality: means preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information (McCallister, Grance, and Scarfone 2010)

### Challenge

- How to balance the risk of providing access with the associated utility?

## Linkage Consent

### Informed Consent

- Prior to linkage, respondent consent is usually required
  - In Germany, this is law (Federal Data Protection Act, 2013, Part I, Section 4; Code of Social Law X, 2013, Section 75)
- The purpose of the consent process is to ensure respondents are informed about:
  - Which data sources will be linked
  - Intended uses of the linked data
  - possible benefits (and risks, if any)
  - responsibility of ensuring data confidentiality
  - voluntary nature of request

### Methods of Obtaining Linkage Consent

- Active consent (opt-in)
  - respondent must take explicit action
  - e.g. verbal agreement, signature, provision of unique identifier (e.g. social security number)
  - typically requested during survey interview
- Passive consent (opt-out)
  - respondent only takes action if they do not want to link
  - must inform interviewer or submit opt-out card
  - otherwise consent is implicit
- Controversy over which procedure is more “informed”
  - most surveys use active consent procedure

### Linkage Consent Rates

- Consent rates vary from study-to-study
  - range: 39 to 97% (da Silva 2012)
  - range: 24 to 89% (Sakshaug and Kreuter, 2012)
- Some evidence that consent rates are decreasing (in the U.S.)
  - National health Interview Survey (1993-2005): 85 to 50%
  - Survey of Income and Program Participation (1996-2004): 88 to 65%
  - current population survey (1994-2003): 90 to 76%

### Why do some people agree/refused to link?

- Reasons for agreement
  - being helpful with the research
  - trust in survey/fieldwork agency
  - having a clear understanding of hwy
  - other (positive) considerations
- Reasons for refusal
  - concerns about sharing of confidential data
  - other (negative) considerations

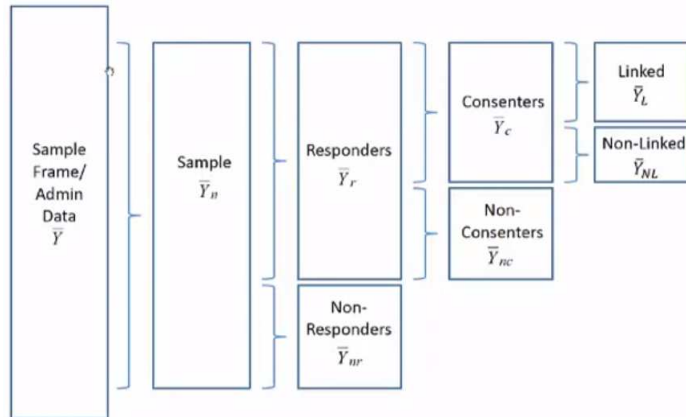
### Consequences of Linkage Non-Consent

- Variance
  - non-consent reduces analytic sample size
  - increases variance of linked-data estimates

- Bias

- people who consent may be systematically different from those who don't based on key survey and administrative variables
- may adversely impact linked-data estimates
- several indicators of bias in the literature

#### Conceptual Pathway to Linkage



#### Correlates of Consent

- Interview/er characteristics

- consent obtained from other household members
- interviewer experience
- interviewer willingness to consent to other (hypothetical) linkages (+) e.g. income tax return, medical data, school records etc

- respondent characteristics

- demographics (e.g. age, foreign citizenship, education)
- income
- benefit receipt
- attitudes toward privacy/confidentiality
- health status
- resistance towards interview (e.g. item missing data, initial refusal)

Study	Consent rate (%)	Survey variables					
		Female	Age	Education	Income/ earnings	Net worth	# Health conditions
Bates, 2005	24.0 – 63.4	↓	↓	↓	↓	--	--
Dahlhamer & Cox, 2007	60.7 – 66.2	--	↓	--	--	--	↑
Banks et al., 2005	77.8 – 81.5	↓	↓	--	↑	↓	--
Halder & Solon, 2000	74.9	↑	--	--	↑	↓	↓
Jenkins et al., 2006	58.5 – 77.4	--	↑	--	--	--	--
Knies et al., 2012	41.0	↓	↓	↑	--	--	↑
Saichuaug et al., 2012	67.8	--	--	↑	--	↓	--
Young et al., 2001	49.4	--	↑	↑	--	--	↓

#### Bias in Survey Estimates

- Consumer Expenditure Quarterly Interview Survey (CEQ): Linkage consent rate about 80%

	Respondent Mean	Consenting Units	Difference
Family income	\$ 50,939.00	\$ 52,869.00	\$ 1,930.00**
Vehicle cost	\$ 599.59	\$ 619.14	\$ 19.55
Property taxes	\$ 454.15	\$ 429.12	– \$ 25.02**
Property value	\$ 247,216.00	\$ 243,507.00	– \$ 3,709.00
Rental value	\$ 1,378.03	\$ 1,351.92	– \$ 26.11**

### Bias in Administrative Estimates

- Bias in linked-administrative variables is also a concern

#### IAB PASS Study (welfare recipient sample)

Variable	Nonresponse Bias	Measurement Bias	Linkage Consent Bias
Age	0.1	0.03	-0.3*
Foreign citizen (%)	-5.6*	-2.5*	-0.9*
Welfare receipt (%)	3.2*	-7.1*	-0.3
Disability (%)	0.4	6.0*	0.01
Employed (%)	1.0	-0.6	0.3
Income (30 days)	-71.4*	394.5*	1.7

Non consent is present, but relatively small compared to other error sources

### Optimizing Linkage Consent

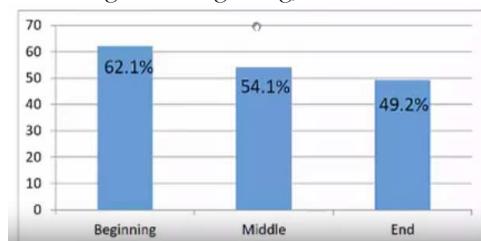
- Recent efforts to overcome the effects of non-consent have largely focused on increasing the consent rate
  - placement
  - wording/framing
  - re-asking for consent among prior refusers (panel surveys)
  - active v.s. passive consent

### Placement of Consent Question

- historically linkage consent question has been asked at the end of interview: conventional wisdom is that interviewer-respondent rapport reaches peak at the end
  - however, relationship between rapport and linkage consent is mixed
    - Jenkins (2006): positive effect
    - Sala (2012): negative effect
- Recent (experimental) evidence suggests end-placement is suboptimal compared to:
  - asking in the context of topic-related items (Sala, Knies, and Burton 2014)
  - asking at the beginning of the interview (Sakshaug, Tutz, and Kreuter 2013)

### Placement in an establishment survey

- N = 4,222 responding establishments in Germany
- Asking for consent to link establishment's register information from Federal employment agency
- 1/3 assigned to beginning, middle and end condition



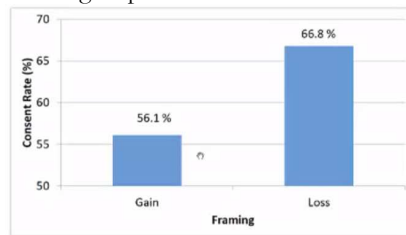
### Wording of the consent question

- Surveys have some flexibility in scripting the consent question
  - exact wording varies across studies
- often the benefits of linkage are emphasized to respondents
  - e.g. saves time, reduces costs and burden, improves data accuracy
- However, empirical support for this strategy is mixed
  - no effect on consent rates (pascale 2011, Sakshaug, Tutz and Kreuter 2013): telephone survey
  - positive effect of time-saving argument (Sakshaug and Kreuter, 2014) Web Survey

### Loss Framing

- Instead of emphasizing the positive benefits of linkage, emphasize the negative consequences of not linking one's data (Based on the tenets of Prospect Theory; Kahneman and Tversky 1979 1984)
- Game frame: the information you have provided so far would be a lot more valuable to us if we could link it to...
- Loss frame: the information you have provided so far would be much less valuable to use we can't link it to ....

## Framing Experiment



- Respondents in the loss framing group were more likely to consent than those in the gain framing group (Kreuter Sakshaug, and Tourangeau 2015)

## Interaction: Placement v.s. Framing (Kreuter 2015)

Phone	Beginning	End	Total n
Gain	90.8	78.7	598
Loss	90.5	81.2	610
Total n	613	595	1208

Web	Beginning	End	Total
Gain	82.6	62.4	520
Loss	86.3	75.4	489
Total	511	498	1009

## Key Points Summary

- Interviewers' own willingness to consent to record linkage is positively correlated with respondents' willingness to consent
- Confidentiality DOESN NOT MEAN removing all identifying records from a data source. Confidentiality means private information is protected. However those with access have the ability to identify individuals.
- A reduced linkage consent rate can increase bias and variance of analysis
- FALSE STATEMENT: Respondents willingness to consent to a linkage request increases on average if the consent question is asked at the end of the survey.
- Opt-in request (meaning a respondent has to actively agree to the record linkage) usually results in higher consent rates