Using sampled social network data to estimate adult death rates

Dennis M. Feehan
UC Berkeley

The challenge: measuring mortality on a survey

Adult deaths are challenging to measure with a survey

- We can’t sample and interview dead people
- Death is a rare event

The challenge: measuring mortality on a survey

Adult deaths are challenging to measure with a survey

- We can’t sample and interview dead people
- Death is a rare event

Leading approach to overcoming these challenges: the sibling method

Sibling survival

Sibling survival method: ask respondents to list their siblings, when they were born, and whether or not they died

Sibling survival

Sibling survival method: ask respondents to list their siblings, when they were born, and whether or not they died

Good because

- We learn about people we don’t interview
- We learn about more than one person from each respondent

Sibling survival

But there are also challenges with sibling survival

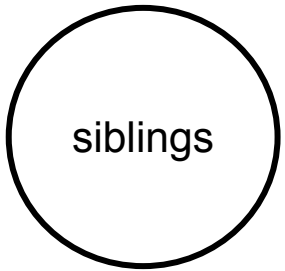
- We don’t learn about enough siblings per interview to produce precise death rate estimates
- Considerable disagreement about how data should be analyzed
- Not well suited to some situations -- disasters, short timeframes, specific geographical areas, etc

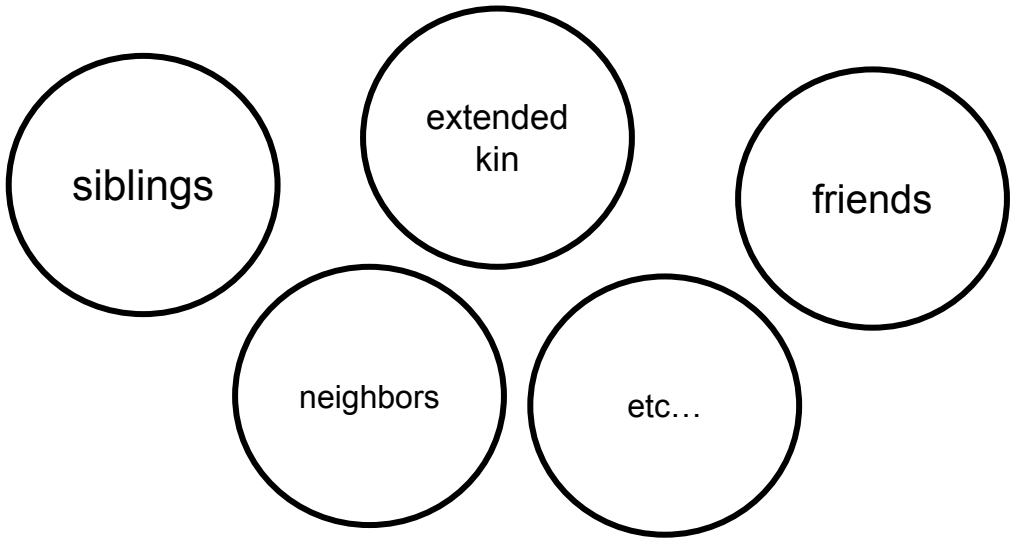
Sibling survival

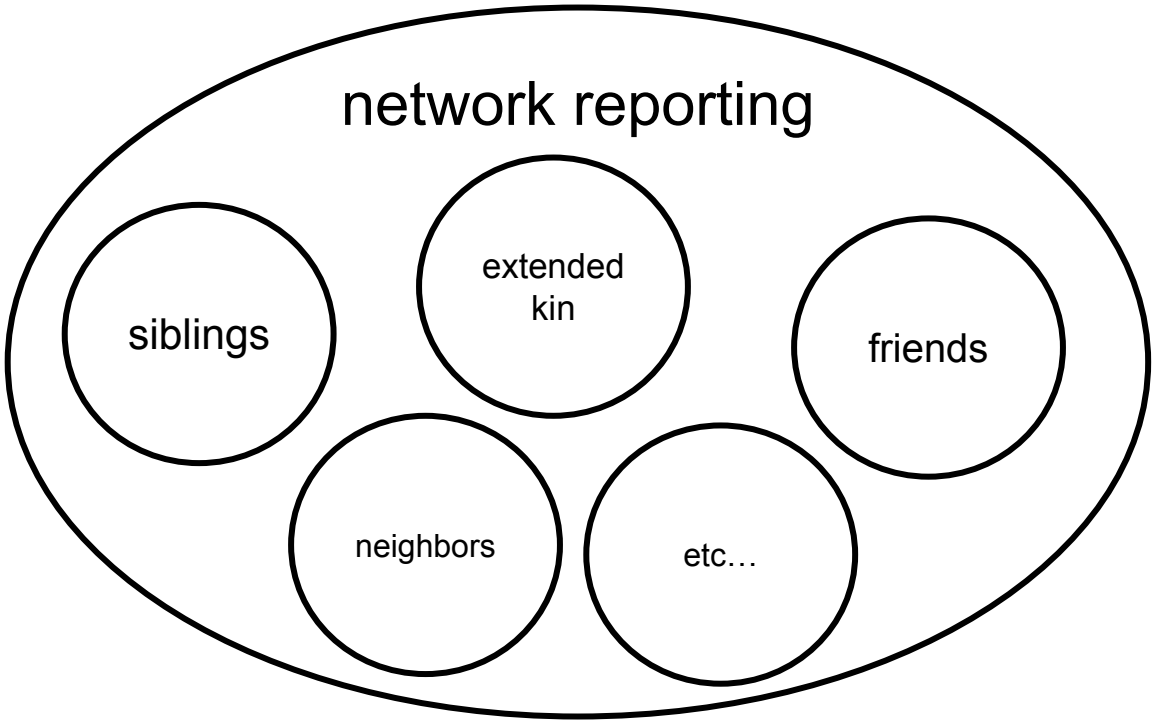
But there are also challenges with sibling survival

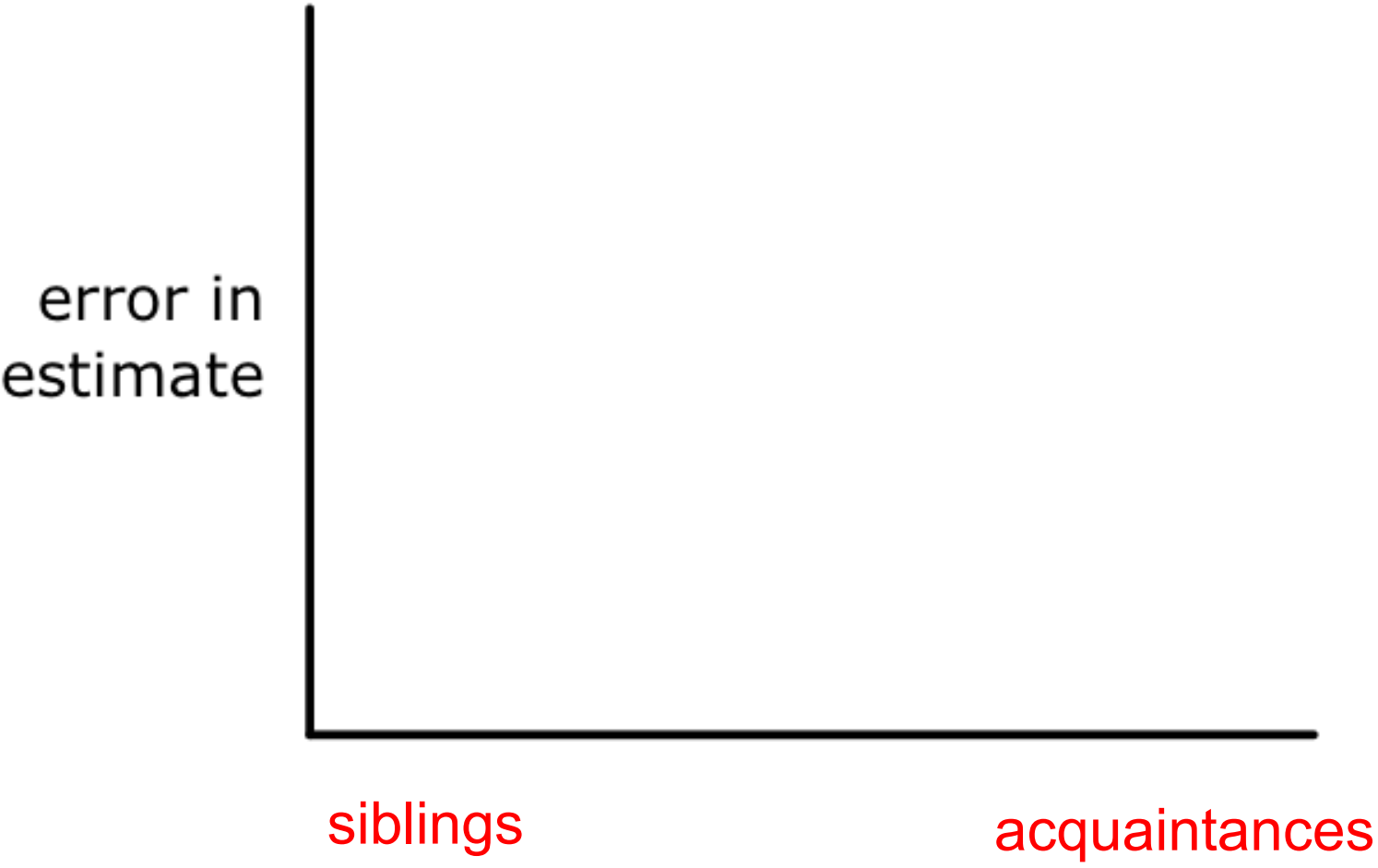
- We don’t learn about enough siblings per interview to produce precise death rate estimates
- Considerable disagreement about how data should be analyzed
- Not well suited to some situations -- disasters, short timeframes, specific geographical areas, etc

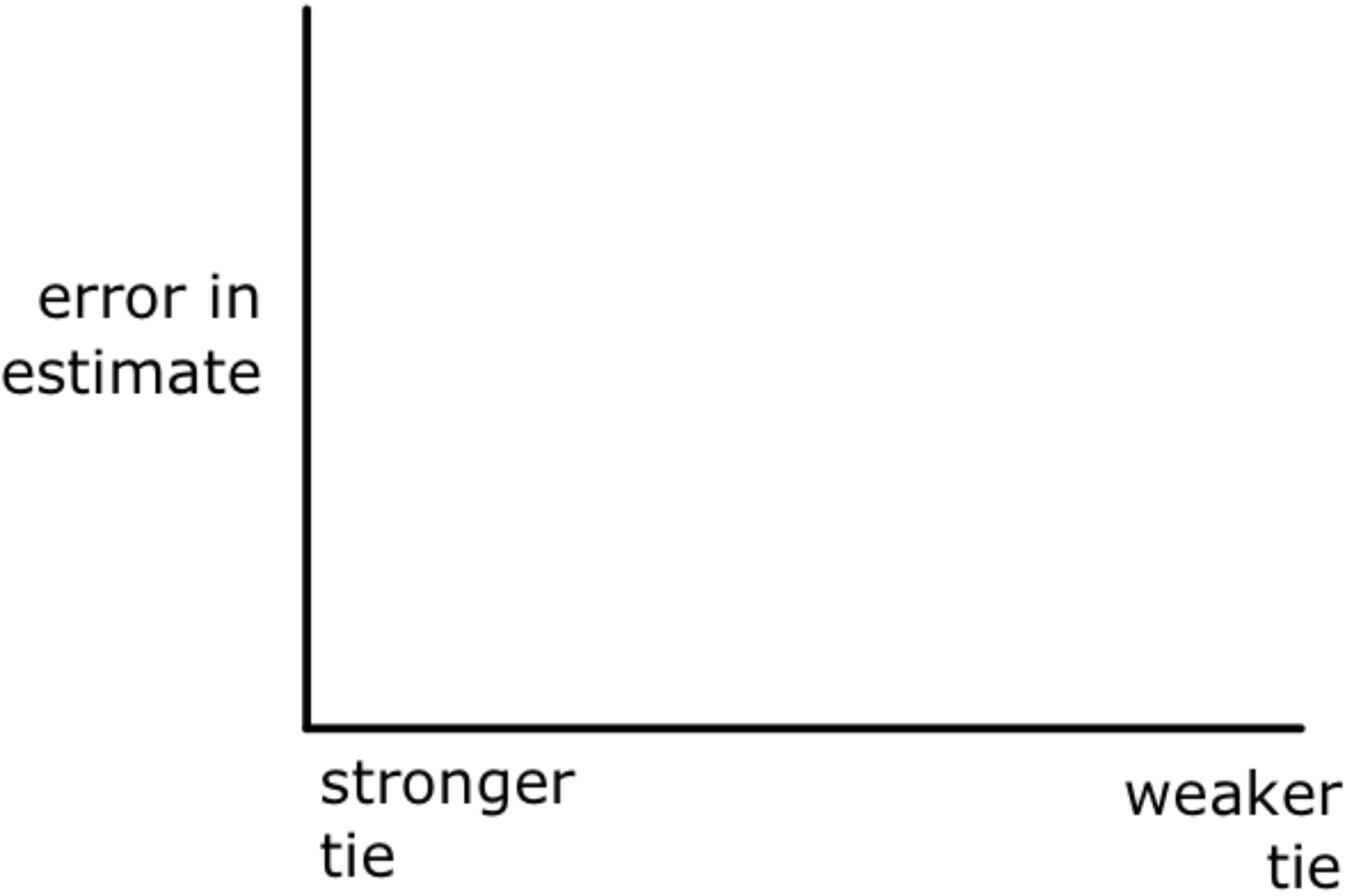
What about going beyond sibship and asking about other types of social relationships?

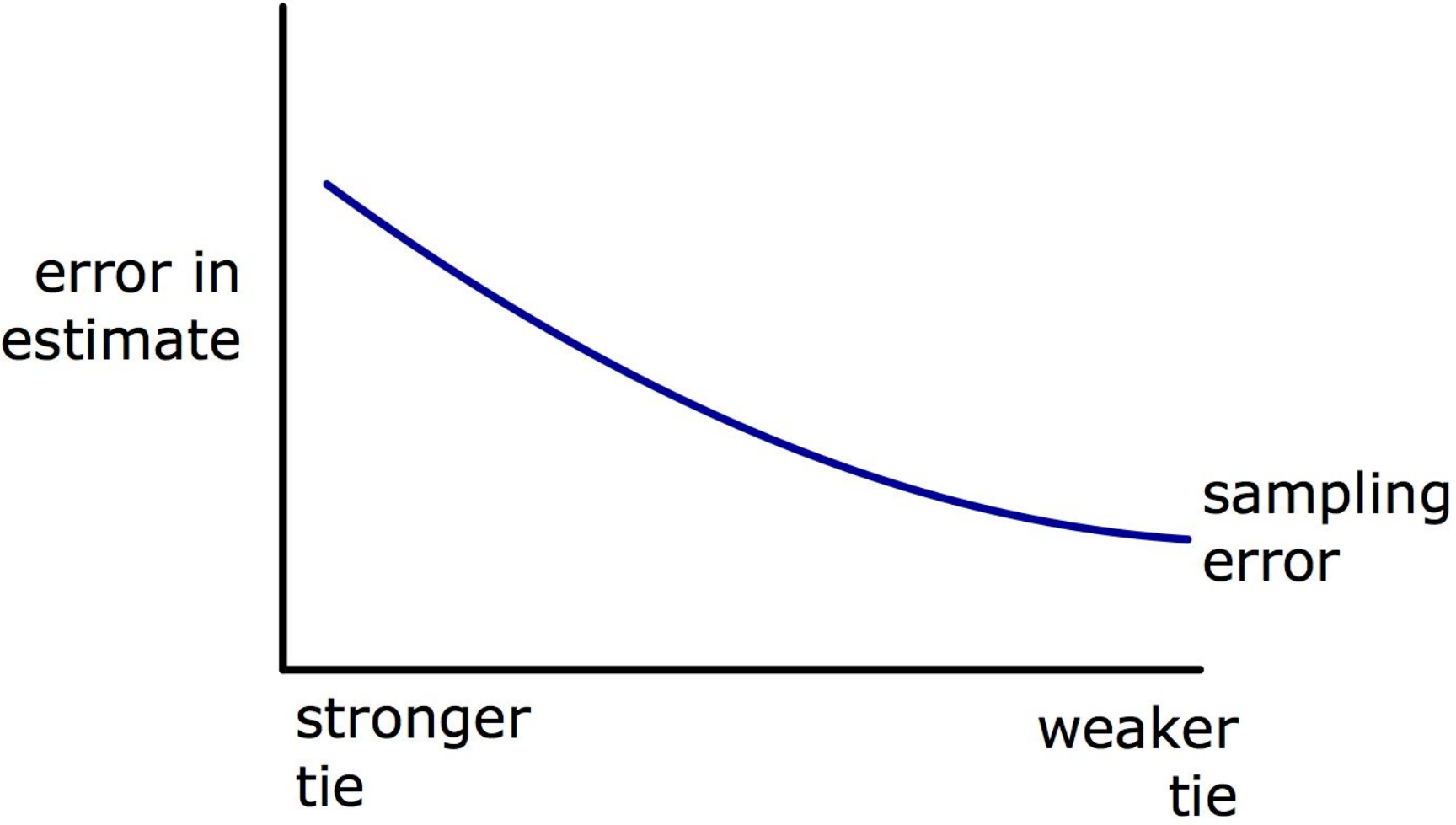

siblings

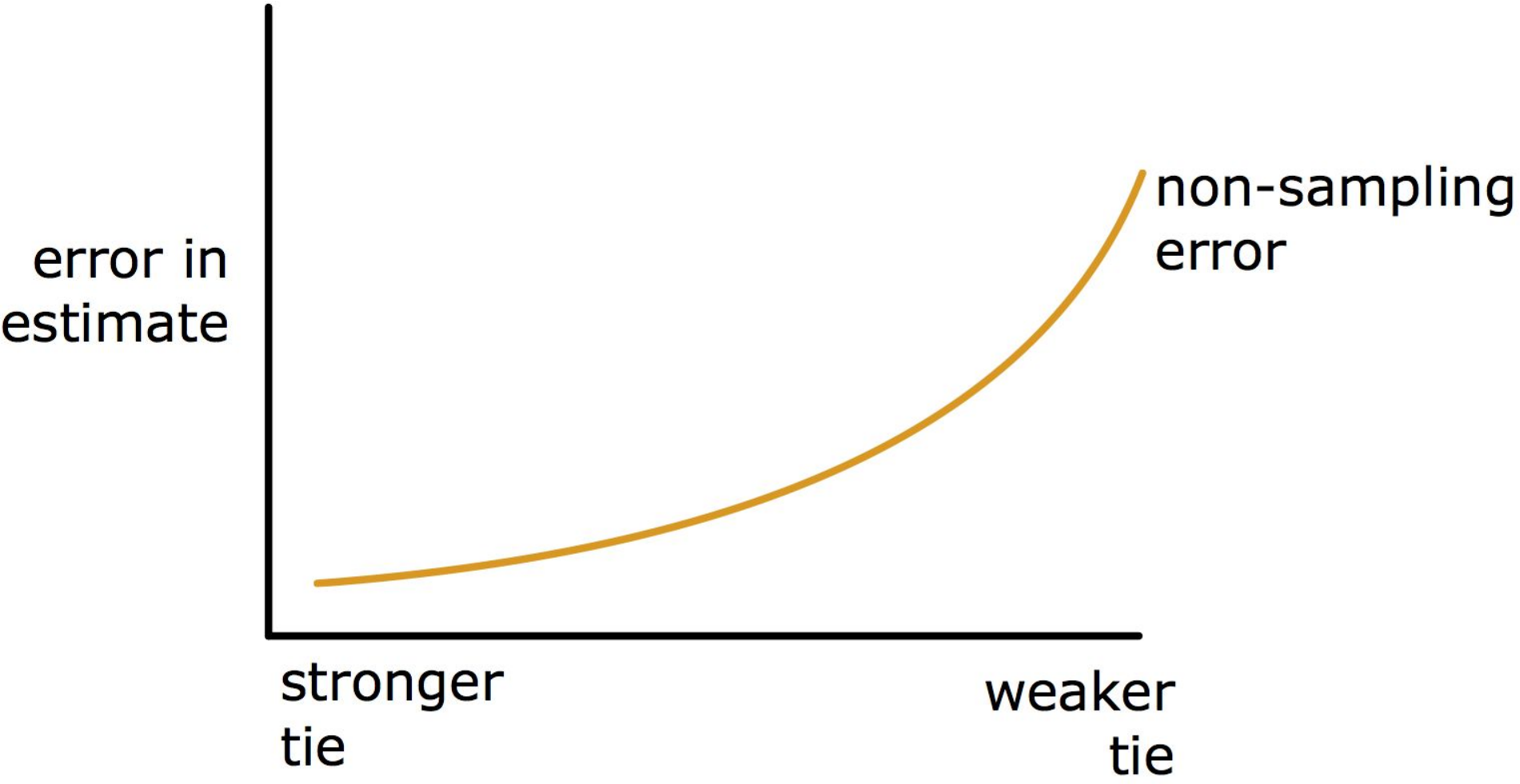

siblings
extended kin
friends
neighbors
etc…

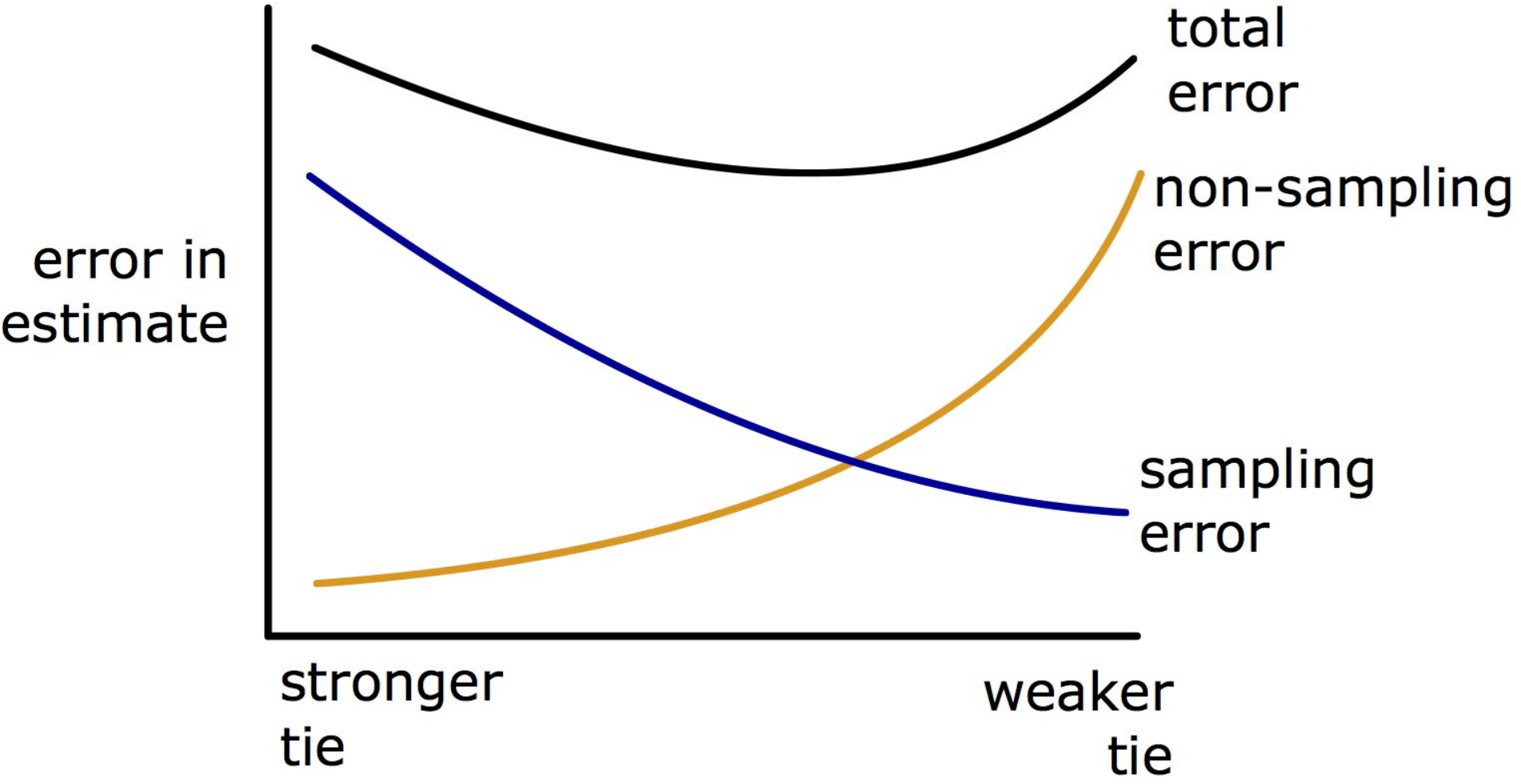

network reporting
siblings
extended kin
friends
neighbors
etc…

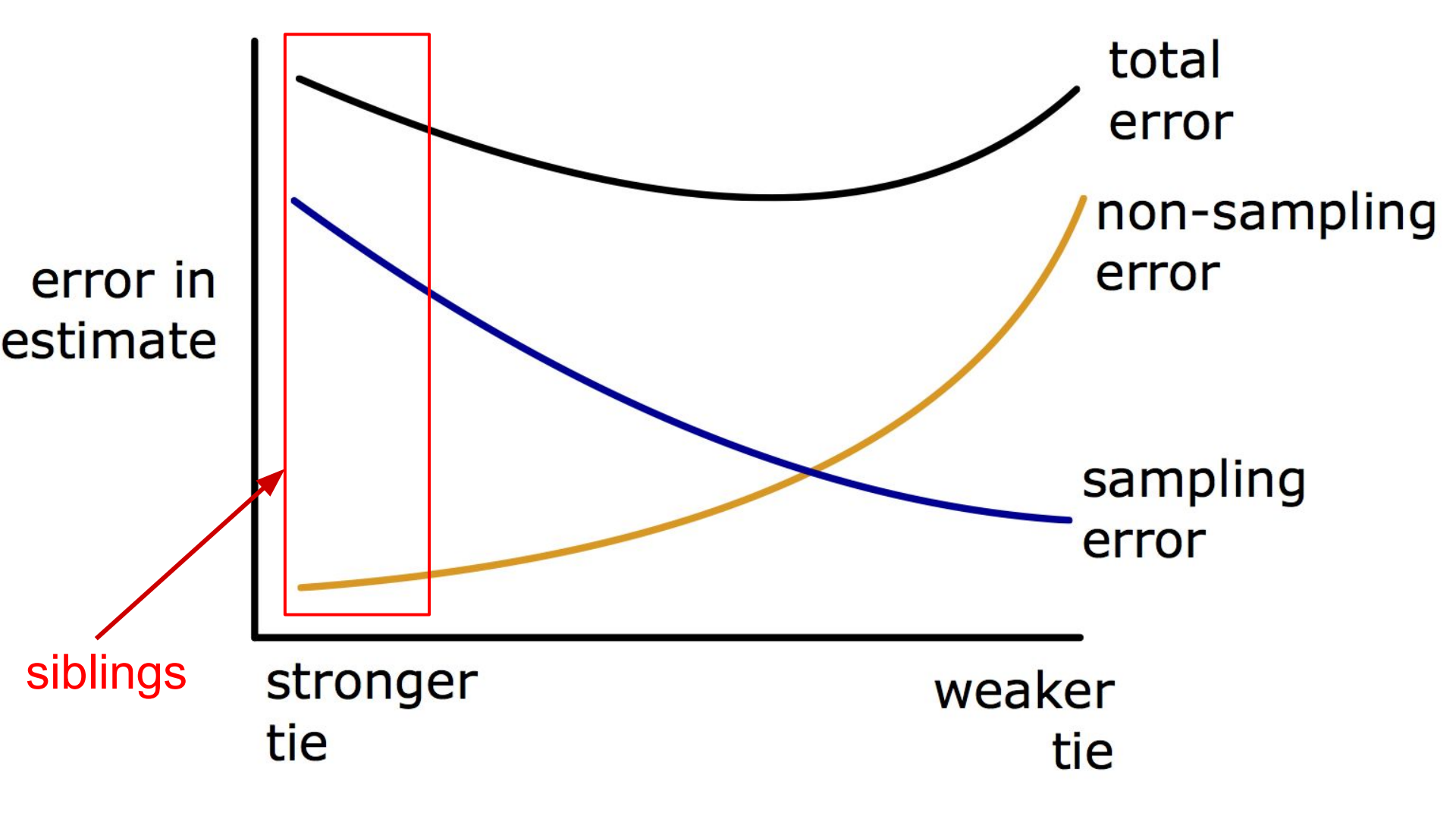

error in
estimate
siblings
acquaintances

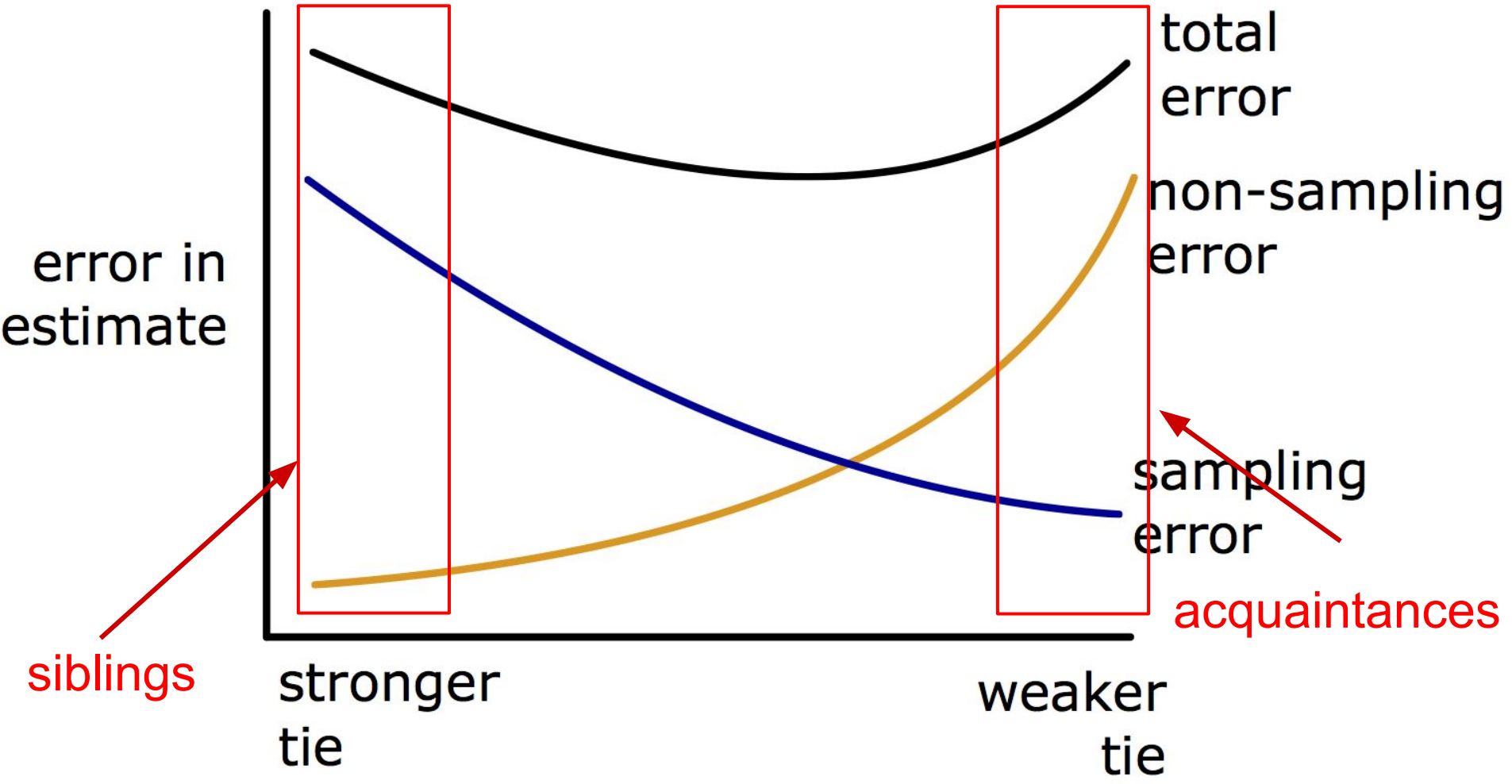

error in
estimate
stronger
tie
weaker
tie


error in estimate
sampling error
stronger tie
weaker tie


error in estimate
non-sampling error
stronger tie
weaker tie


total
error
non-sampling
error
error in
estimate
sampling
error
stronger
tie
weaker
tie


total
error
non-sampling
error
error in
estimate
sampling
error
siblings
stronger
tie
weaker
tie


total error
non-sampling error
error in estimate
sampling error
siblings
acquaintances
stronger tie
weaker tie

Data: household survey in Rwanda

Map source: Wikipedia

Data: household survey in Rwanda

- Intended to mimic a Demographic and Health Survey
- Stratified, two-stage cluster sample of approximately 5,000 Rwandans aged 15 and over (oversampled Kigali)

Survey experiment in Rwanda

Acquaintance Network

- People of all ages who live in Rwanda
- People the respondent knows by sight and by name, and who know the respondent by sight and by name
- People the respondent has had some contact with -- either in person, over the phone, or on the computer -- in the previous 12 months

Meal Network

Survey experiment in Rwanda

Acquaintance Network

- People of all ages who live in Rwanda
- People the respondent knows by sight and by name, and who know the respondent by sight and by name
- People the respondent has had some contact with -- either in person, over the phone, or on the computer -- in the previous 12 months

Meal Network

- People of all ages who live in Rwanda
- People the respondent knows by sight and by name, and who know the respondent by sight and by name
- People you have shared a meal or drink with in the past 12 months. These could be family members, friends, co- workers, or neighbors. You should include meals or drinks taken at any location, such as at home, at work, or in a restaurant.

Recap

Sibling survival

- respondents report about deaths among their siblings
- expect relatively little information per interview
- … but expect respondents to be reasonably well-informed about their siblings

Network survival

- respondents report about deaths among people in their network
- expect more information per interview - and more for acquaintances than for meals
- … but less clear how well-informed respondents will be about network members

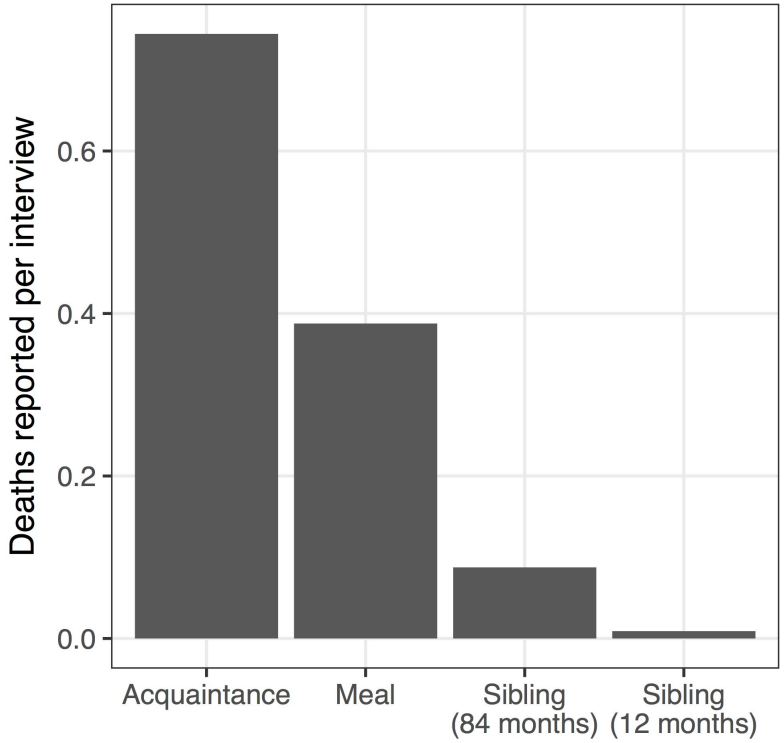
Comparison estimates: Rwanda DHS sibling survival

Sibling method results from Rwanda 2010-11 DHS

- Based on interviews with 13,761 women who were asked to report on their siblings
- The sibling estimates of death rates are based on the 7-year period before the interviews
  (the network results are for 1 year before the interview)

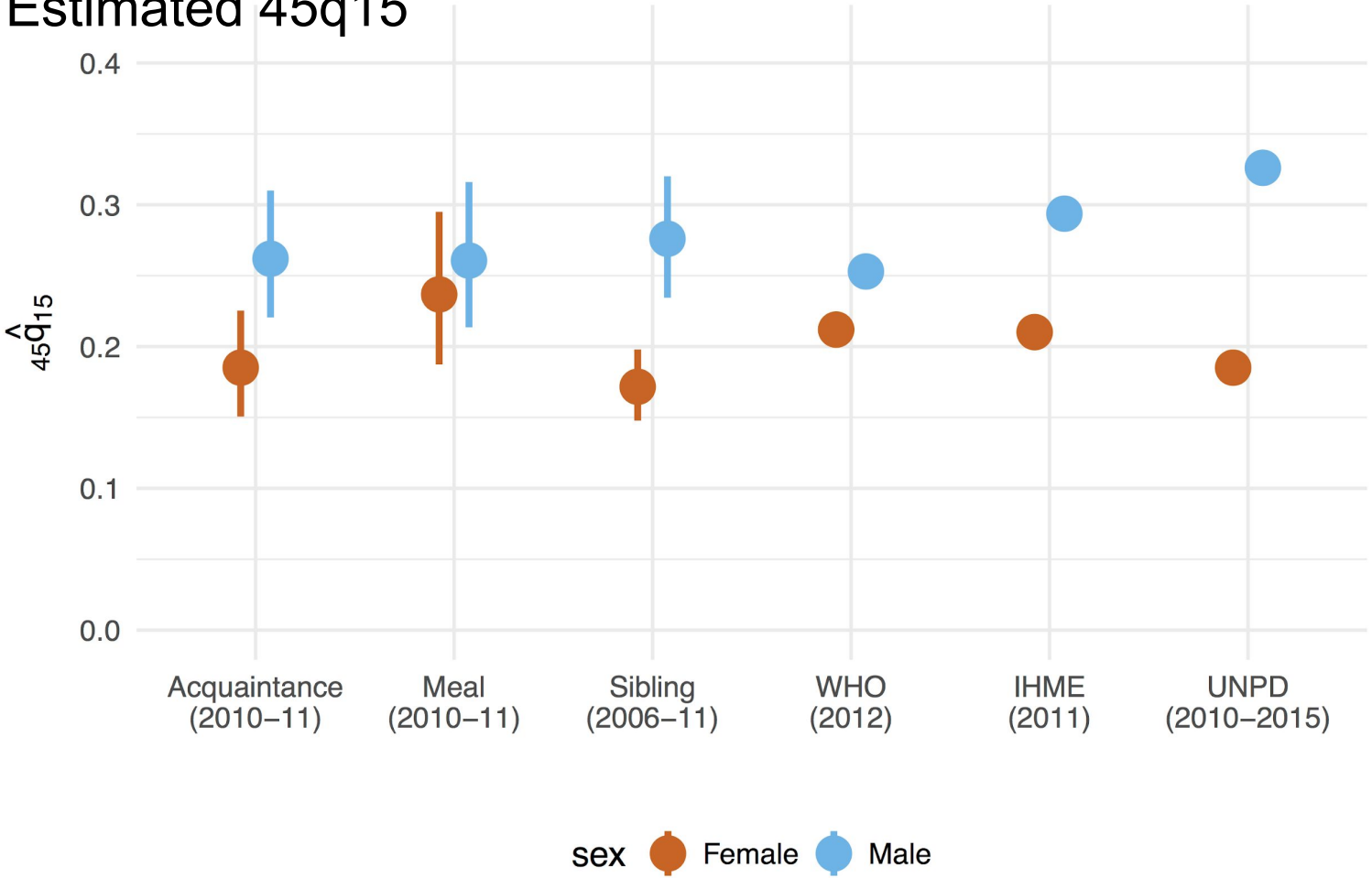
Deaths per interview

Deaths per interview



Estimated 45q15

Estimated 45q15



Summary of Rwanda empirical results

- A network survival study is feasible on a Demographic and Health Survey
- We learned about more deaths from each interview using the network methods
- The estimated age-specific death rates are roughly similar for the sibling method and for the meal and acquaintance tie definitions (especially for males)

Other projects underway

- using network reports in combination with an online sample
  - (idea: we can quickly and easily reach people online, but still learn about people we can’t directly interview — deaths, outmigrants, people not online, people without a cell phone, etc)
- improve degree estimation
- large study in 27 Brazilian cities - compare network survival, sibling, and models
- better understand how to produce estimates from sibling histories

What I haven’t talked about and where this could go

- The study design in Rwanda also provides a template that can be used to embed experiments in data collection, with the goal of improving the method over time
  - Ideal situation: experiment to converge on tie definitions that make the most sense in a given setting. (This may not be the same everywhere!)
  - Papers also have a way to produce blended estimates using data from both arms of the experiment -- so these experiments need not mean that only half the sample gets used for the actual estimates

What I haven't talked about and where this could go

- the papers develop a framework for sensitivity analysis
  - answers the question: how do reporting errors/other factors affect estimates?
  - reveals what we could potentially try to measure in order to adjust estimates produced by this method (see work by Helleringer and colleagues on measuring reporting errors for the sibling method)
  - these quantities could be measured in an HDSS site or other setting with rich, accurate data collection

$$D_\alpha = \underbrace{\left(\frac{y_{F,D_\alpha}}{\bar{d}_{F_\alpha,F}}\right)}_{\substack{\text{network} \\ \text{survival} \\ \text{estimand}}} \times \underbrace{\frac{1}{\bar{d}_{D_\alpha,F}/\bar{d}_{F_\alpha,F}}}_{\substack{\text{degree ratio} \\ \delta_{F,\alpha}}} \times \underbrace{\frac{1}{\bar{v}_{D_\alpha,F}/\bar{d}_{D_\alpha,F}}}_{\substack{\text{true positive rate} \\ \tau_{F,\alpha}}} \times \underbrace{\frac{y^+_{F,D_\alpha}}{y_{F,D_\alpha}}}_{\substack{\text{precision} \\ \eta_{F,\alpha}}}.$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{adjustment factors}}$$

Where this could go

Frequent surveys producing timely estimates (maybe through SMS/phone, maybe in person)

combined with

Less frequent, higher effort data collected in a setting like an HDSS site that can be used to gather information needed to adjust the rapid estimates via the sensitivity framework

Where this could go

All of this could be tuned, over time, to use the tie definition that produces the best information in a given setting

It could produce district-level estimates if the survey is designed with that goal in mind

Can estimate other quantities with network reports also; for example, out-migration

Perhaps it could also locate deaths that could be the target of a VA?

Thanks!

My website http://www.dennisfeehan.org has more information and links to papers and data.

Thanks!

My website http://www.dennisfeehan.org has more information and links to papers and data.

Thanks to my collaborators on several related projects: Matthew J. Salganik (Princeton), Mary Mahy (UNAIDS), Aline Umubyeyi (U. of Rwanda), Wolfgang Hladik (CDC)