Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias

Kristen Olson
Program in Survey Methodology
University of Michigan
DC-AAPOR Workshop on
Nonresponse Bias in Household Surveys
March 30, 2007

Outline

- Summary of POQ article
  - Motivation
  - Main findings
  - Limitations
- New research related to nonresponse bias

Motivation

- Long time hypothesis about reluctant respondents or otherwise difficult respondents giving error-filled answers.

- Mixed evidence about this
  - Studies based on record checks imply that this is true; studies based on covariance measures not as clear.

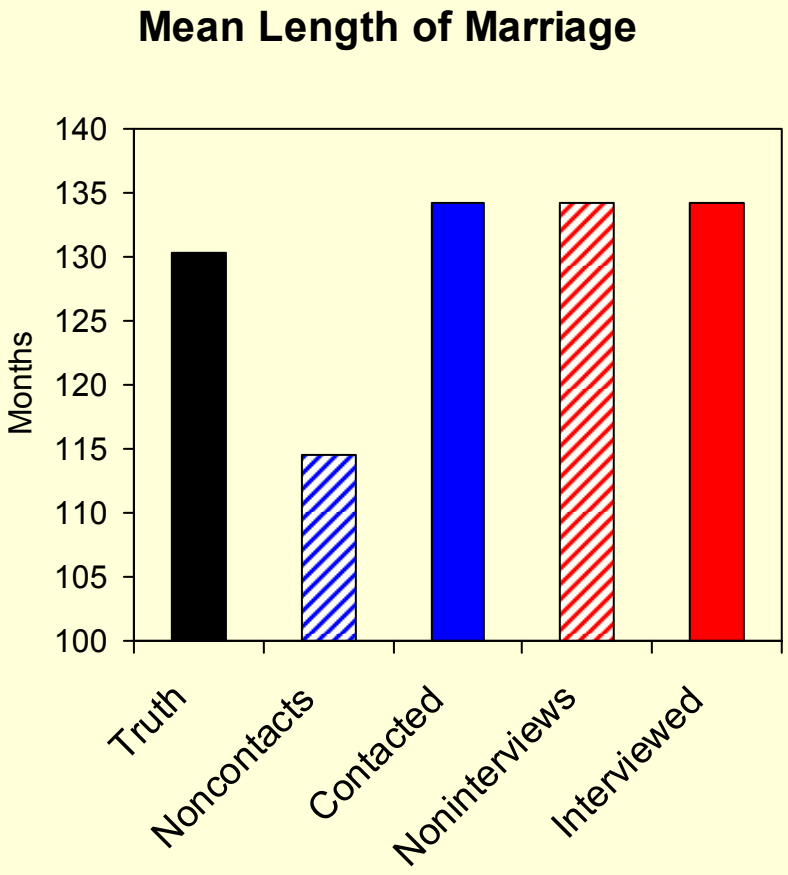
- Most of the studies ignore the reduction of nonresponse error compared to potential increases in measurement error.

- Basic question: Does including reluctant respondents increase the bias of an estimate?

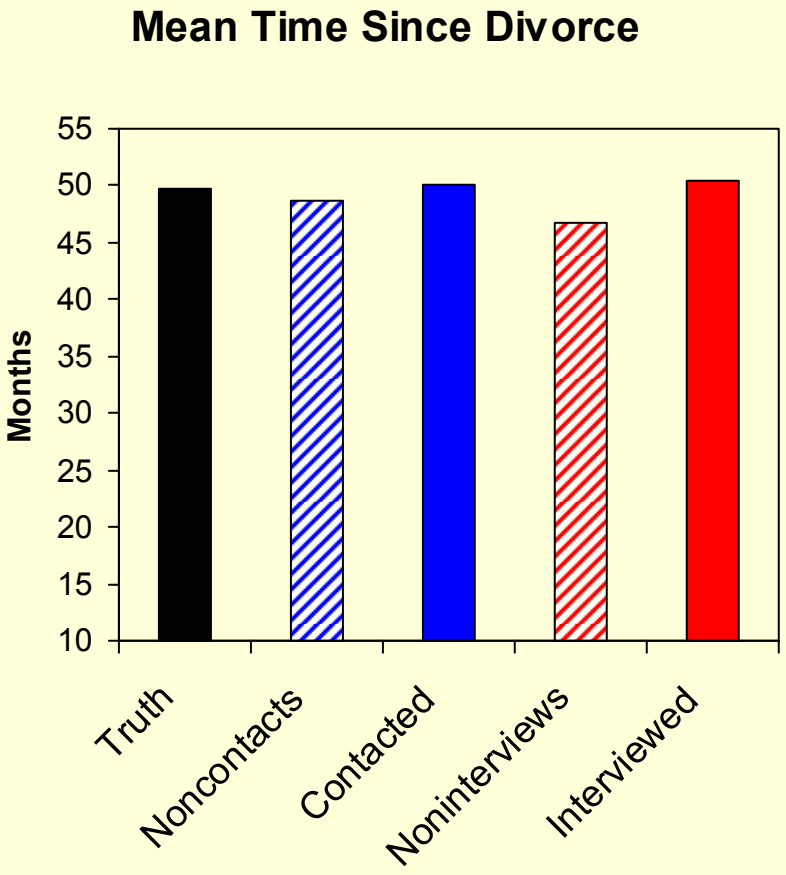
Main Findings

- The relationship between nonresponse propensity and nonresponse bias is both statistic-specific and specific to the type of nonresponse.
- The nonresponse / measurement error trade-off is also statistic specific.
- There is a difference between a variable in a data set and a statistic calculated using that variable.
  - Different functions of a variable may pose different nonresponse /measurement properties.
- In these statistics, using this propensity model, the overall reduction in bias tended to outweighed a net increase in either error source.

Nonresponse bias is statistic specific and varies by type of nonresponse.

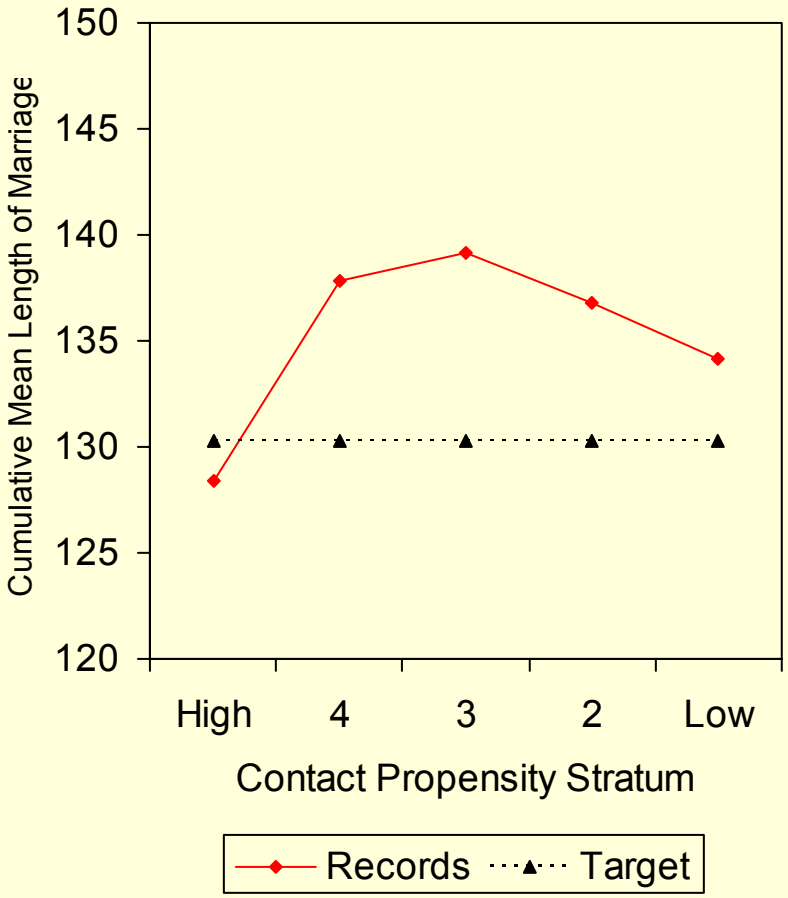


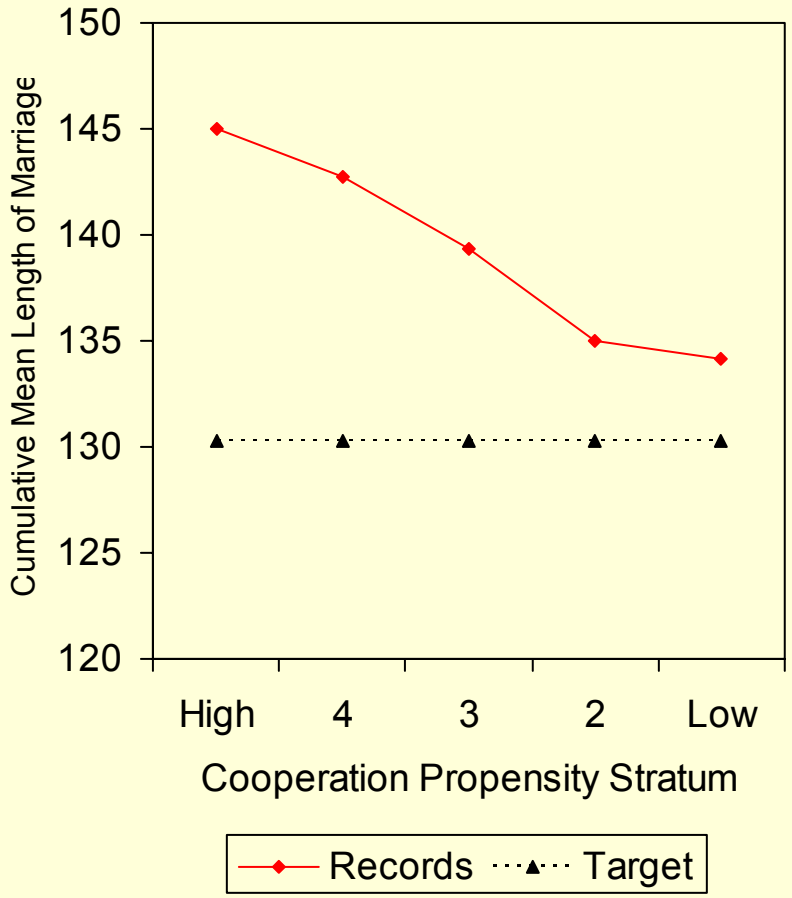
Divorce Date – Marriage Date



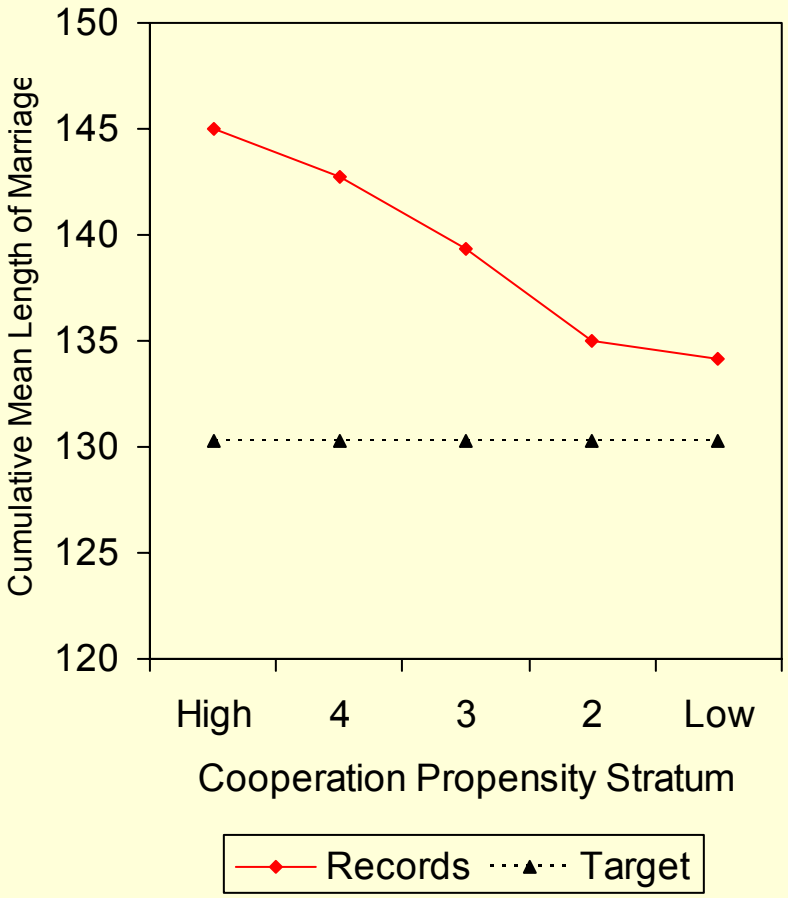
Start of Field Period - Divorce Date

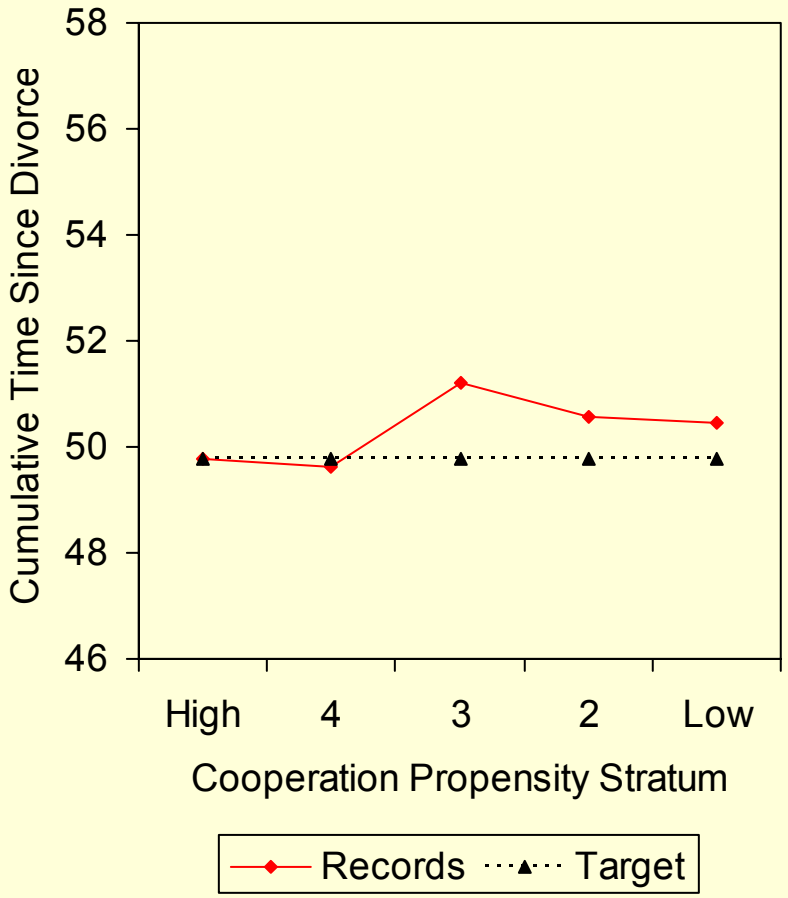
Relationship between nonresponse propensity and nonresponse bias varies across **types of nonresponse**.



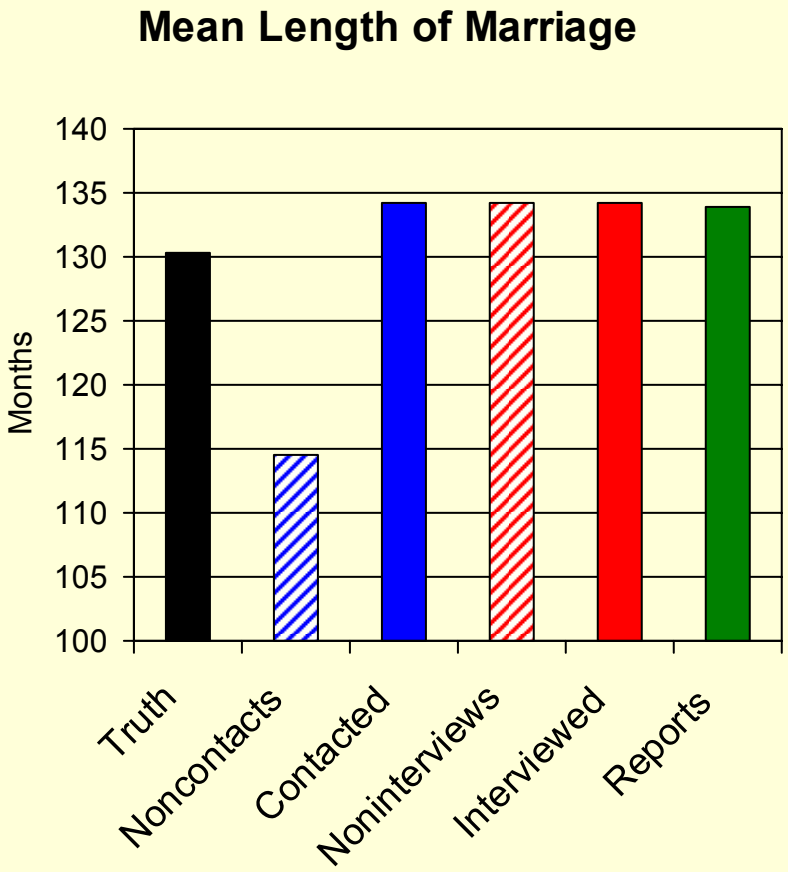


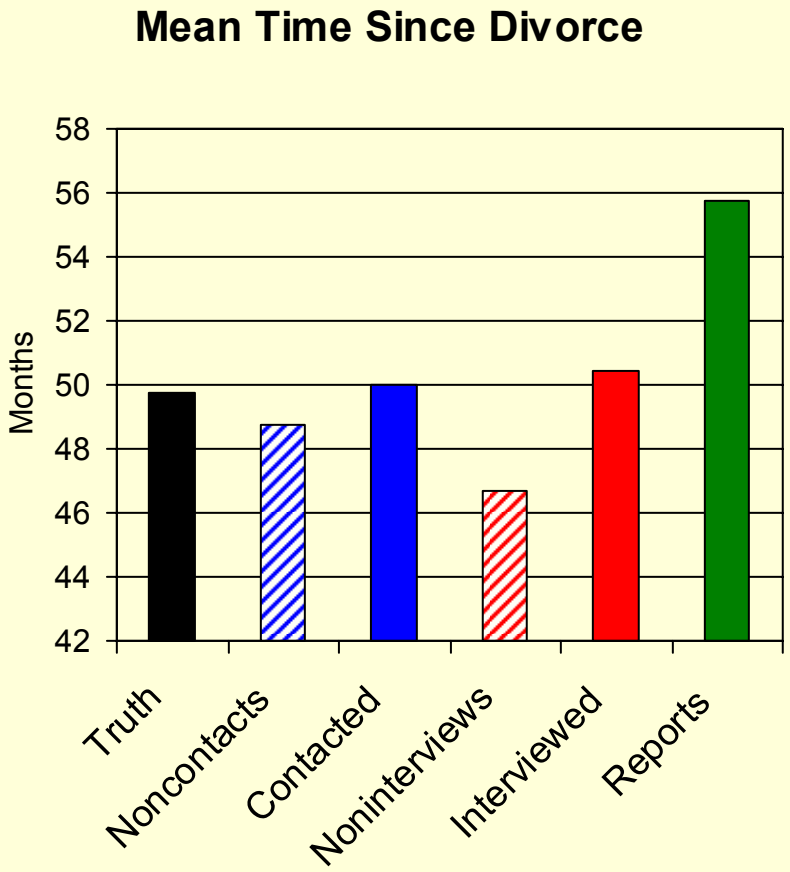
Relationship between nonresponse propensity and nonresponse bias varies **across statistics**.



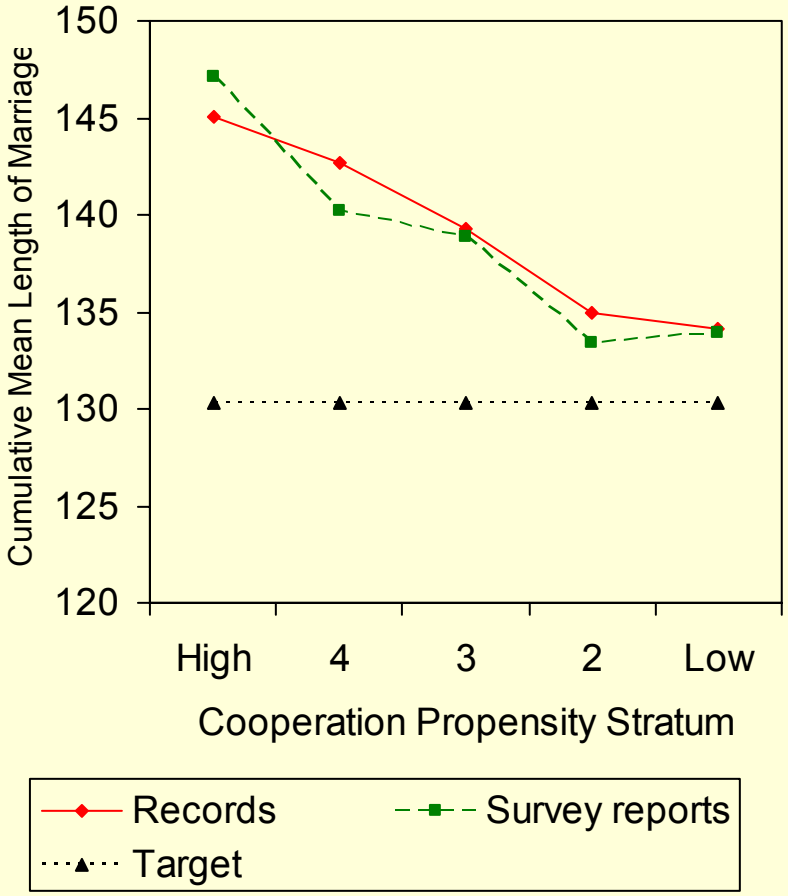


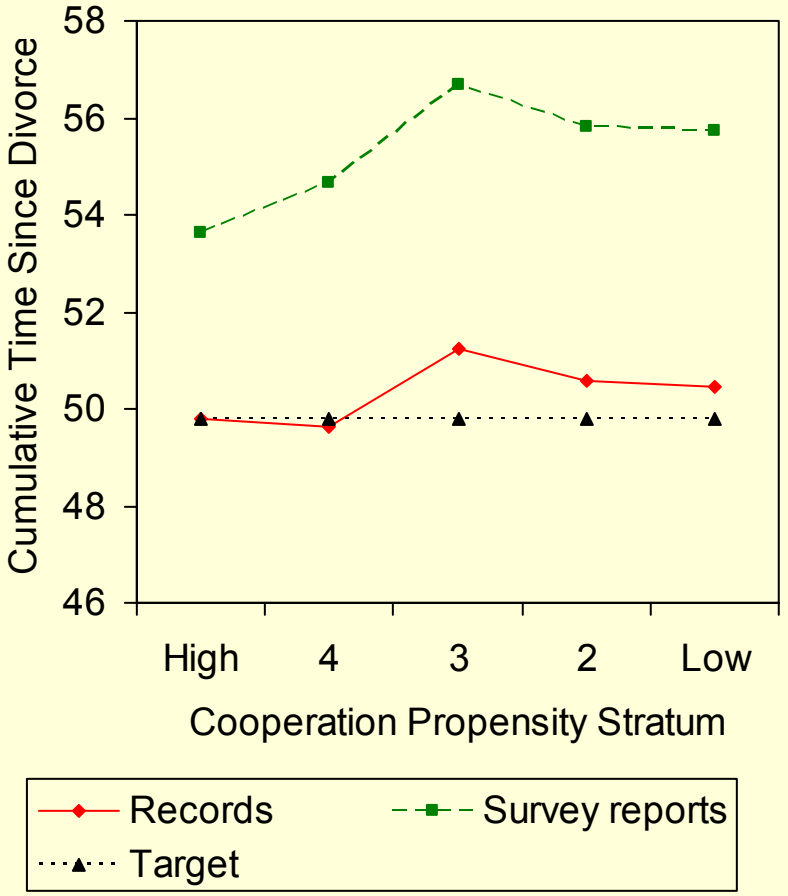
Nonresponse / measurement error properties are statistic-specific.



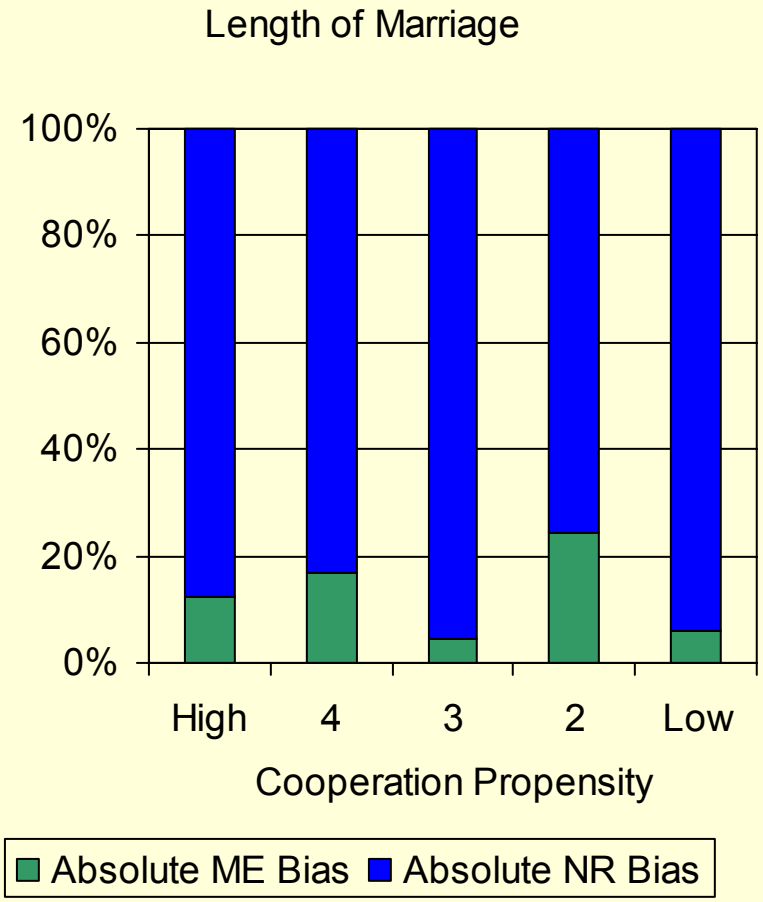


Measurement error bias of a statistic increases for some, but not all, statistics.

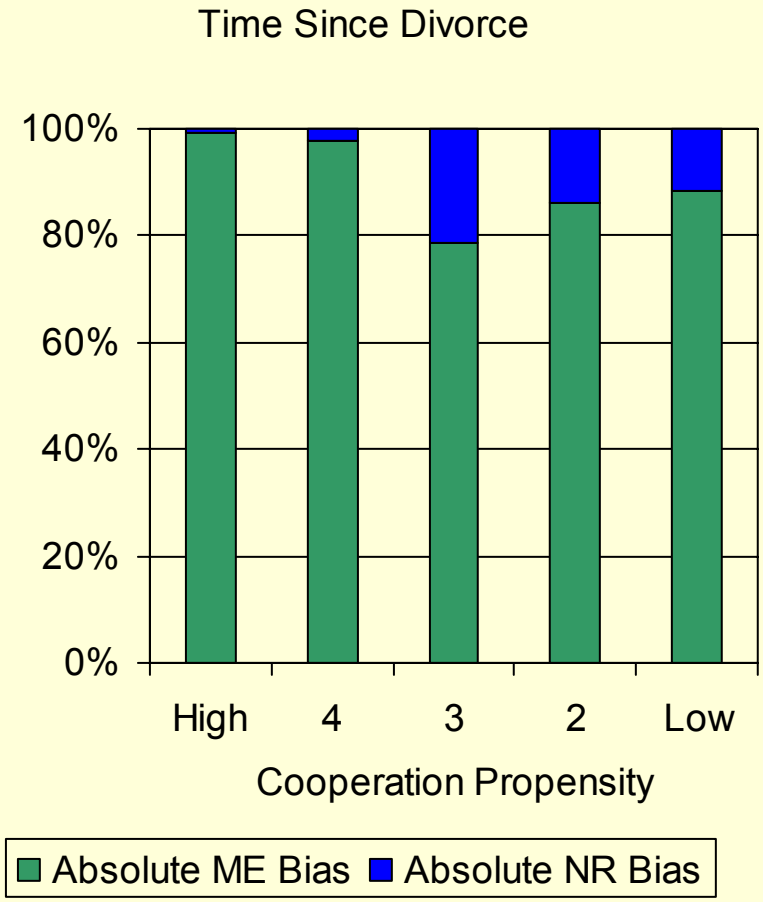




Statistics sharing variables can differ in their nonresponse / measurement error properties.

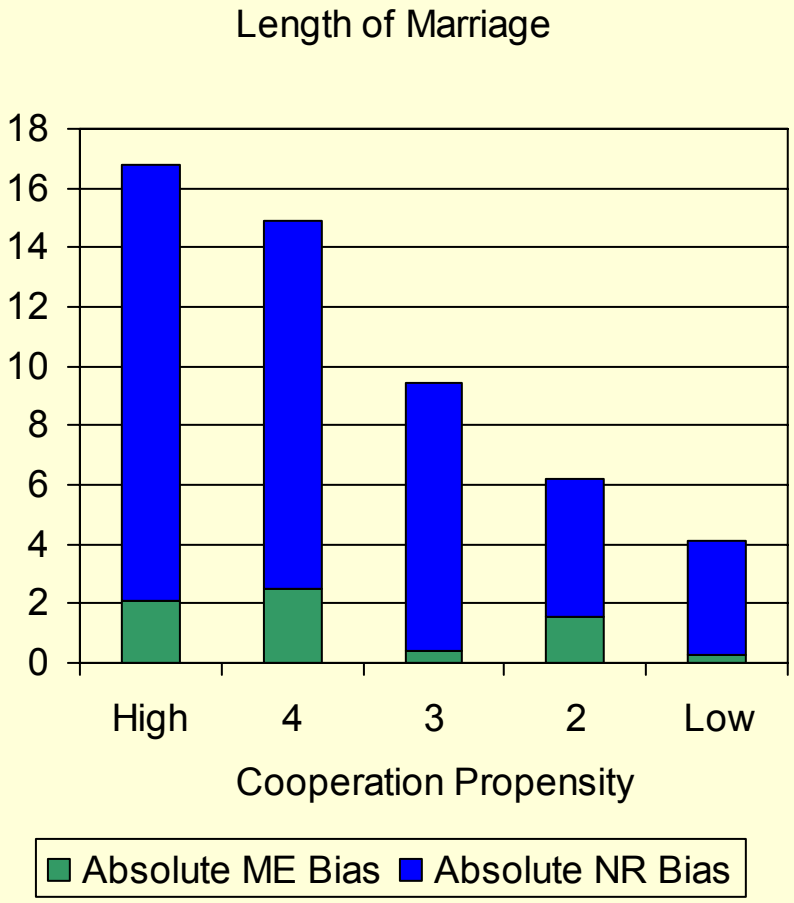


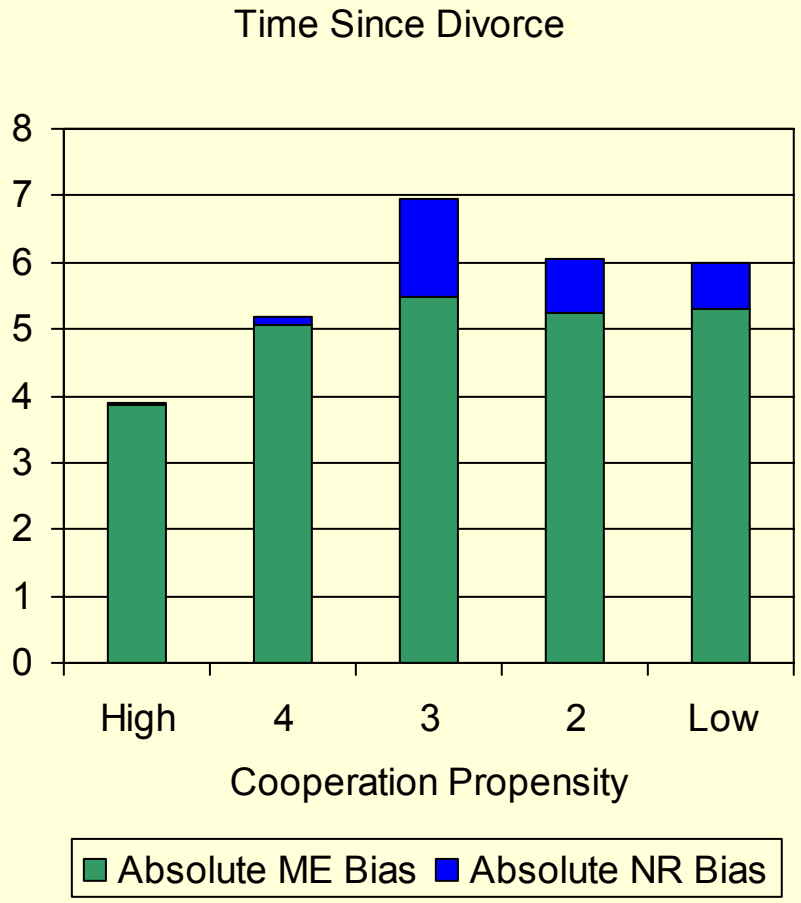
Divorce Date – Marriage Date



Start of Field Period - Divorce Date

Absolute total bias may be reduced.





Main Findings

- The relationship between nonresponse propensity and nonresponse bias is both statistic-specific and specific to the type of nonresponse.
- The nonresponse / measurement error trade-off is also statistic specific.
- There is a difference between a variable in a data set and a statistic calculated using that variable.
  - Different functions of a variable may pose different nonresponse /measurement properties.
- In these statistics, using this propensity model, the overall reduction in bias tended to outweighed a net increase in either error source.

Limitations

- Specification of the propensity model
  - Can we do any better than “it depends?”
- Only looked at the survey mean. What about other statistics?
- Measurement error analysis looked at for the “complete cases.”
  - This confounds a number of processes; these need to be teased apart to really understand what’s happening at the unit level

New Research on Nonresponse Bias

- Examining nonresponse propensity and nonresponse bias in observational studies
  - What does “it depend” on?
    - Who is being asked to participate?
    - The statistics of interest?
    - The recruitment process?
  - Mechanisms and nonresponse bias of a statistic
  - Survey recruitment process and nonresponse bias

Can we gain insights into nonresponse bias by speculating on what mechanisms might affect particular statistics?

- Specify proxy indicators for a mechanism:
  - At home patterns
    - Age, gender, kids, education, ecological measures on commuting time, working at home, mode of transportation to work
  - Social isolation
    - Age, gender, age*gender, kids, ecological measures on single person households, persons living below the poverty line
- Relate to the Y variables
  - Expect at home patterns to be related to length of marriage
    - Age is the common cause: older people are more likely to be at home, and have longer marriages
  - Not clear direction for length of marriage and social isolation
    - Are people who were in longer marriages less isolated because they were part of a longer union? Are they more isolated because their ties are now broken?

Can we gain insights into nonresponse bias by speculating on what mechanisms might affect particular statistics?

	Length of marriage	Months since divorce	Number of marriages	Age at marriage	Age at divorce
Contact Models					
At home patterns	Yes	No	No	Yes	Yes
Access impediments	No	No	Yes	No	No
Cooperation Models					
Social Isolation	Mixed	No	Mixed	Yes	Yes
Social Cohesion	No	No	Yes	Yes	No
Discretionary Time	Yes	No	No	No	Yes
Positive affect toward sponsor	No	No	No	No	No

Can we gain insights into nonresponse bias by speculating on what mechanisms might affect particular statistics? (part 2)

	Length of marriage	Months since divorce	Number of marriages	Age at marriage	Age at divorce
Contact Models					
At home patterns	Yes	No	No	Yes	Yes
Access impediments	No	No	Yes	No	No
Cooperation Models					
Social Isolation	Mixed	No	Mixed	Yes	Yes
Social Cohesion	No	No	Yes	Yes	No
Discretionary Time	Yes	No	No	No	Yes
Positive affect toward sponsor	No	No	No	No	No

New Research on Nonresponse Bias

- Examining nonresponse propensity and nonresponse bias in observational studies
  - What does "it depend" on?
    - Who is being asked to participate?
    - The statistics of interest?
    - The recruitment process?
  - Mechanisms and nonresponse bias of a statistic
  - Survey recruitment process and nonresponse bias

Survey recruitment process and nonresponse bias

- Traditional views of survey nonresponse look at nonresponse at the end of a survey.
  - Deterministic view – nonresponse reflects classes of respondents and nonrespondents.
  - Stochastic view – everyone has a probability of participating in a survey.
- Traditional observational examinations of nonresponse propensity look primarily at respondent or ecological characteristics.

Survey recruitment process and nonresponse bias

- We do many things as survey organizations to gain contact and cooperation.
  - Incentives, mode switches, advance letters, persuasion letters, call timing, etc.
  - Everything that we do is designed to change people's likelihood of survey participation.

Survey recruitment process and nonresponse bias

- Survey participation is a process
  - To fully understand how and why people participate in surveys, need to account for the process
    - What was done to the sample unit, When was it done, Who was it done to
  - Where the process ends affects the final adjustment model.
  - Where the process ends affects the observed nonresponse bias.

- We need a new way of thinking about survey participation.

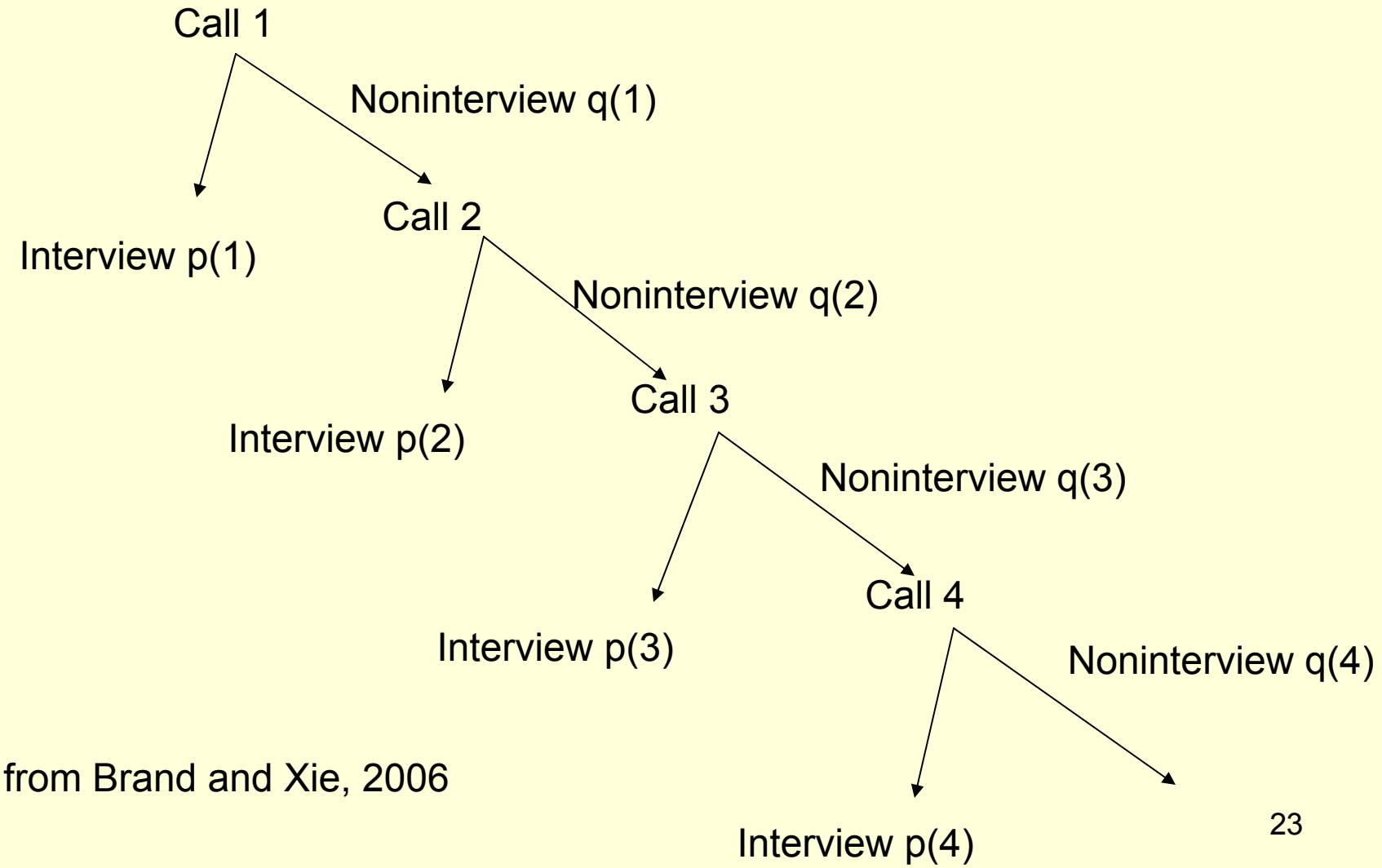
Dynamic Response Propensities

- Sample units have a propensity to respond each time they are attempted.

- Response propensity at time t is determined by sample unit characteristics ***and*** recruitment protocol decisions at time t and at prior times.
  - As additional levels of effort are exerted, the sample unit's propensity changes.
  - Each sample unit has a vector of response propensities.

- Call this vector the sample unit's **"dynamic response propensities**."
  - This is essentially the assumption underlying responsive designs (Groves and Heeringa, 2006).

Dynamic Response Propensities



Adapted from Brand and Xie, 2006

Implications of Dynamic Response Propensities for Nonresponse Bias

- $Bias(\overline{y}_r) = \sigma_{py} \Big/ \overline{p}$ becomes

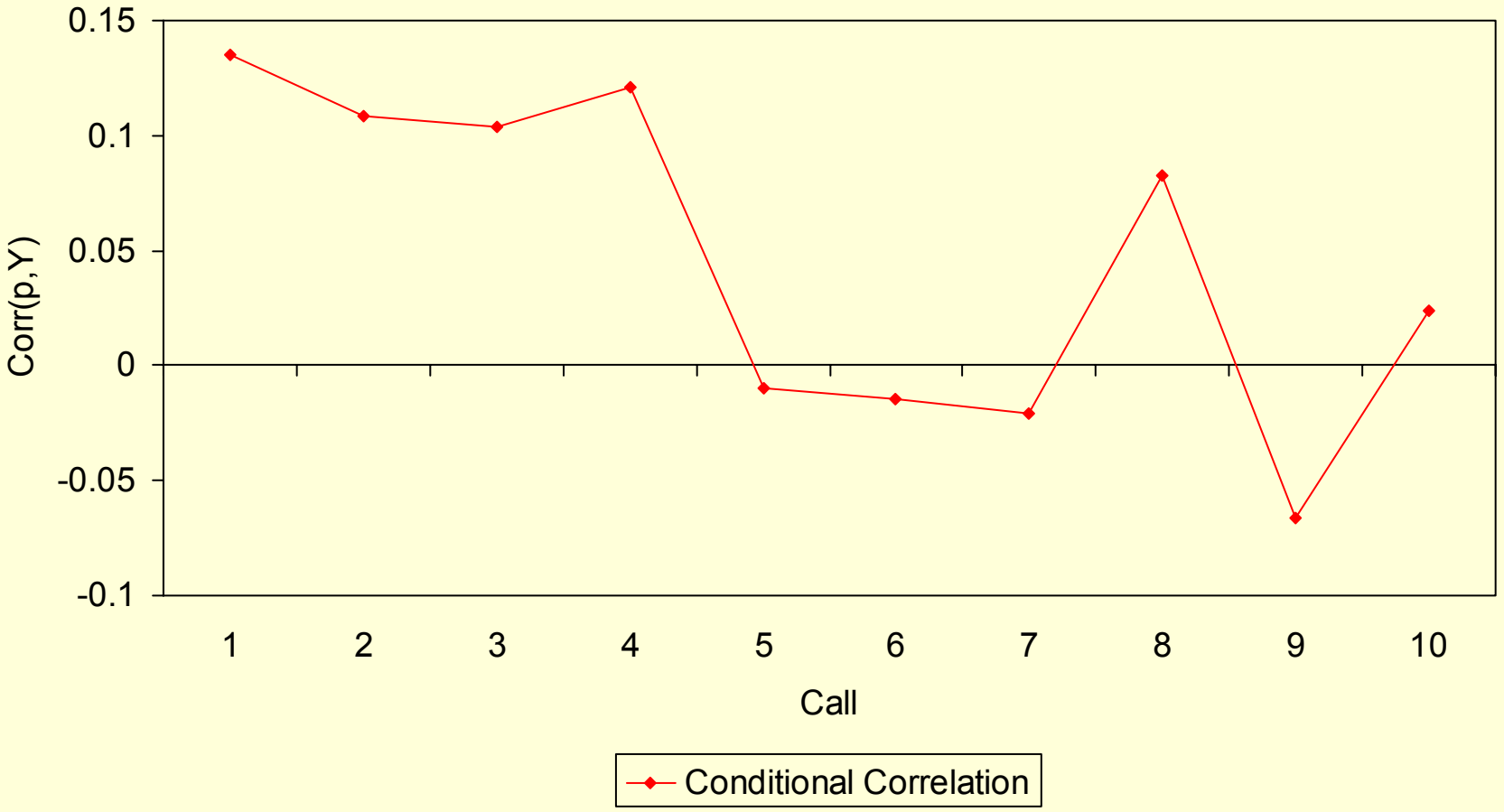
$$Bias(\overline{y}_{rt}) = \sigma_{p_t y} \Big/ \overline{p}_t$$

- That is, the nonresponse bias of the (unadjusted) mean changes over the life course of a survey.
  - Not only because the average response propensity is changing, but because the covariance term is changing

Estimating Dynamic Response Propensities

- How to estimate?
  - Could estimate separate logistic regressions for each call, changing the risk set for each model
  - Instead of lots of logistic regressions, make a few proportionality assumptions (which can be tested).
- Use discrete time hazard model

Correlation between p_t and Length of Marriage



Implications for Measurement Error

- Different mechanisms for survey participation vary in their implications for measurement error.
- Different components of the recruitment process may affect measurement error – does it matter when these components are implemented?

Thanks!

Questions or comments?

olsok@isr.umich.edu