Adjusting for Duration Biases in Sexual Behaviour Data

Jesse Knight[1,2] and Sharmistha Mishra[1,2]

[1]Institute of Medical Science, University of Toronto
[2]MAP Centre for Urban Health Solutions, Unity Health Toronto

June 24, 2023

1 Introduction

Quantifying sexual behaviour is necessary to study the epidemiology of sexually transmitted infections (STI), including to inform inputs for STI transmission modelling [1]. Two important quantities are: the duration of time within a "risk group" such as sex workers (for a given definition) [**?**], and the rate of new partnership formation, possibly stratified by partnership type [**?**].

Our aims are to motivate and discuss bias adjuments for estimating:

1. duration within sex work
2. rate of partnership change

from cross-sectional survey data, considering issues of sampling bias and censoring.

Figure 1: TODO

2 Methods

2.1 Risk Group Duration

2.2 Sexual Partnership Duration

Sexual partnerships are often quantified from cross-sectional survey data. Respondents are typically asked to report their numbers of unique partners (x) during a standardized recall period (ω) — *e.g., "How many different people have you had sex with during the past year?"* Such data can then be used to define a rate of partnership change (Q) and/or a number of currrent partnerships (K).

If partnership duration (δ) is long and the recall period is short — *e.g.,* $\omega \approx 0$ for "*Are you currently in a long-term sexual partnership?*" — the reported partnerships mostly reflect *ongoing* partnerships, and thus $x \approx K$. If partnership duration is short and the recall period is long — *e.g.,* $\delta \approx 0$ for "*How many one-off sexual partners have you had during the past year?*" — the reported partnerships mostly reflect *complete* partnerships, and thus $x/\omega \approx Q$. However, if partnership duration and recall period are similar in length, the reported partnerships reflect a mixture of tail-ends, complete, and ongoing partnerships, and thus x overestimates K, but x/ω also overestimates Q.

One solution to this problem would be to ask respondents about *new* partnerships (x') during the recall period, which can give an unbiased estimate of the change rate as: $Q = x'/\omega$. However, for surveys do not include these data, an unbiased estimator for Q and K that uses *all* reported partnerships in the recall period (x) would be useful.

We take the partnership duration to be fixed and known. We start with a similar assumption as in § 2.1: that survey timing is effectively random with respect to partnership duration. Then, if either end of the recall period would capture an ongoing partnership, the intersection point would be, on average, at the partnership mid-point. Thus, the recall period is effectively extended by half the partnership duration $\delta/2$ on each end, and δ overall, as illustrated in

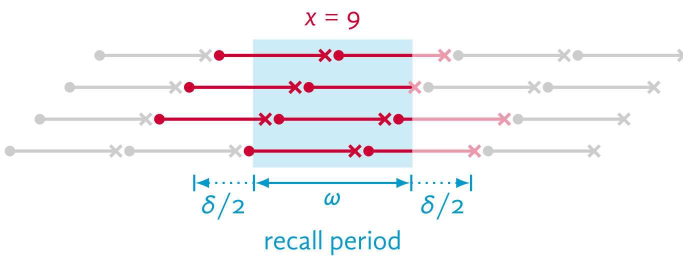

Figure 2: TODO

Figure 2. As such, we can define Q and K as:

$$Q = \frac{x}{\omega + \delta} \tag{1}$$

$$K = \frac{x\delta}{\omega + \delta} = Q\delta \tag{2}$$

As an example, Figure 2 also illustrates a recall period of $\omega = 1$ year, for which $x = 9$ partnerships are reported, having durations of $\delta = 9$ months. Thus, we can compute $Q = 9/(1 + 0.75) = 5.14$ partnerships per year and $K = 5.14(0.75) = 3.86$ current partners; these ars slight underestimates of the true values $Q = 5.33$, $K = 4$, due to randomness in the exact "location" of the recall period.

3 Results

4 Discussion

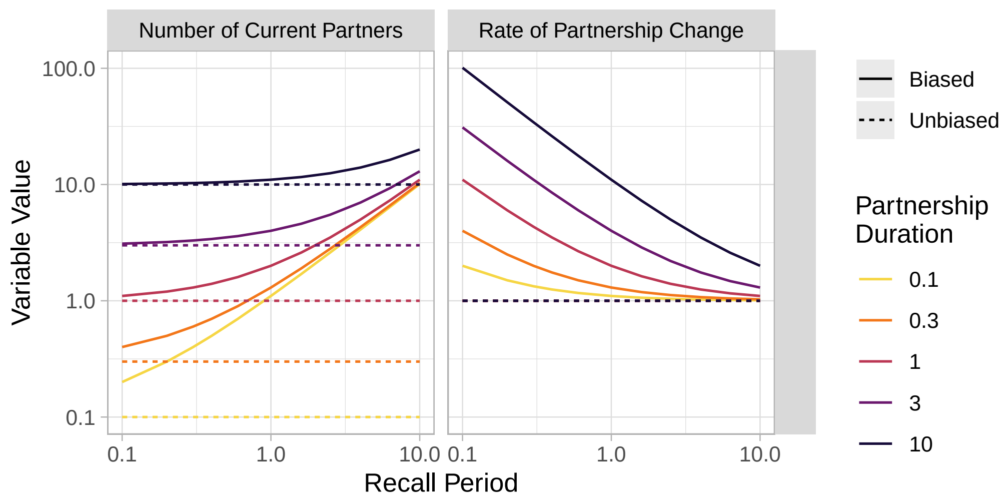

Figure 3: Biased vs unbiased estimates of: number of current partners and rate of partnership change, for different recall periods and partnership durations. Figures A.1 and A.2 also show 95% CI for biased and unbiased estimates from 1000 simulated surveys with $N = 10, 100, 1000$.

References

[1] K. A. Fenton et al. "Measuring sexual behaviour: Methodological challenges in survey research". In: *Sexually Transmitted Infections* 77.2 (Apr. 1, 2001), pp. 84–92. pmid: 11287683. URL: https://doi.org/10.1136/sti.77.2.84.

A Supplement

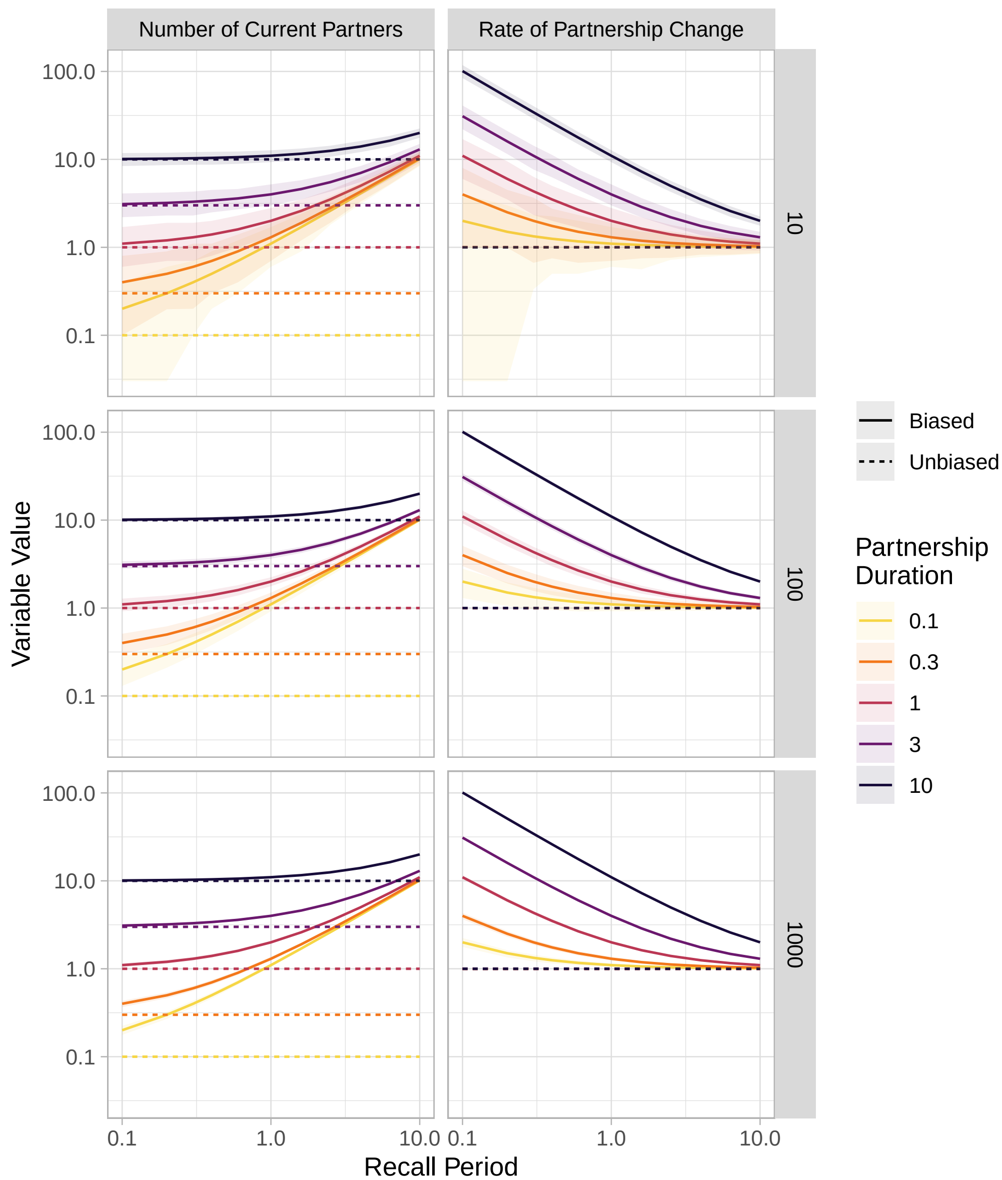

Figure A.1: Biased vs unbiased estimates of: number of current partners and rate of partnership change, for different recall periods and partnership durations. Ribbons show 95% CI for biased estimates from 1000 simulated surveys with $N = 10, 100, 1000$.

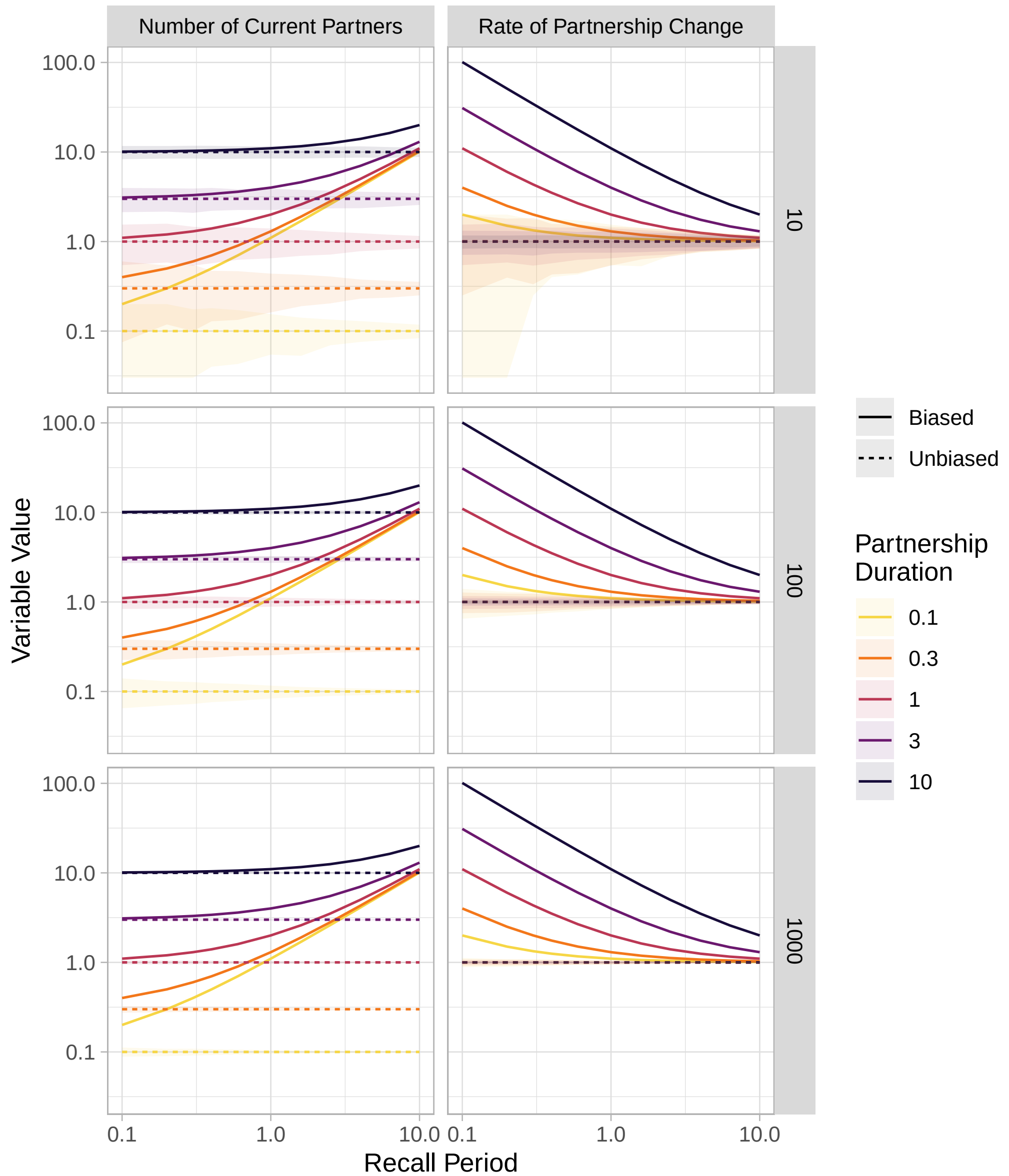

Figure A.2: Biased vs unbiased estimates of: number of current partners and rate of partnership change, for different recall periods and partnership durations. Ribbons show 95% CI for unbiased estimates from 1000 simulated surveys with $N = 10, 100, 1000$.