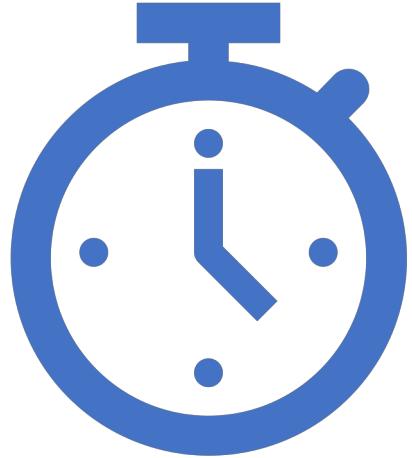


Emily Pollock
NMG
2 May 2019

Relationship Duration & Survival Analysis





Quick Guide to Censoring



Relationship Duration in NSFG



Guide to Censoring & Sampling Bias

Right Censoring

- Type 1
 - Event is only observed if it occurs prior to a set time
 - e.g. study ends and patient has not experienced a cardiac event post-surgery, or
 - We have the start date of a relationship, but it is still active on the day of interview
- Type 2
 - Study ends after X failures
 - Not usually present in social science data
- If ignored, results in bias towards shorter time-to-event / duration

We have a *minimum* duration:

The time between surgery and study end or loss-to-follow-up



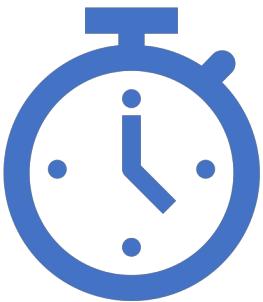
Guide to Censoring & Sampling Bias

Interval / Left Censoring

- Interval
 - event occurs between observations
 - e.g. a patient seroconverts between HIV tests
- Left*:
 - Special case of interval censoring
 - Event occurred prior to observation period at undetermined time
 - e.g. when did you first have sex? “I don’t know, but I have had sex”
- Both results in bias towards longer time-to-event / duration

We have a *maximum* duration:

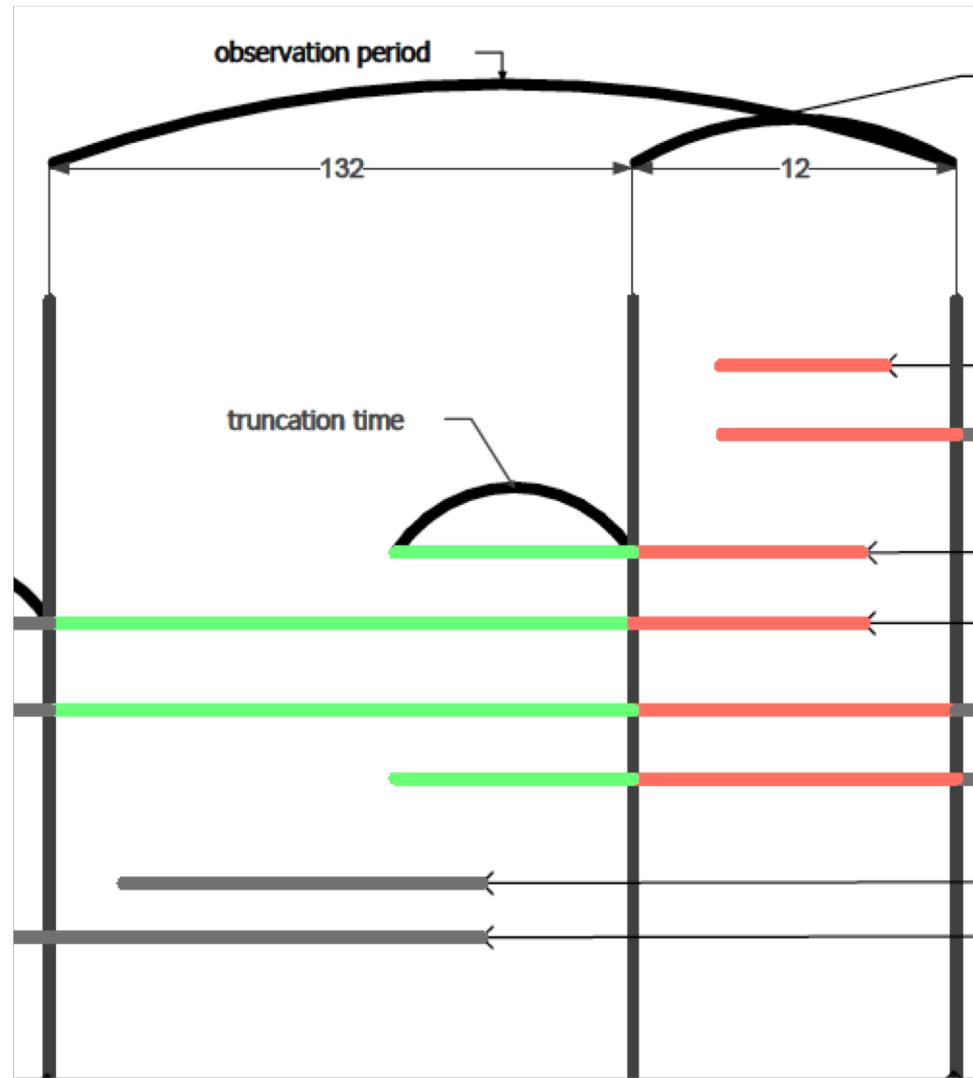
time between HIV tests, and the time between interview day and perhaps an assumption about average onset of sexual activity

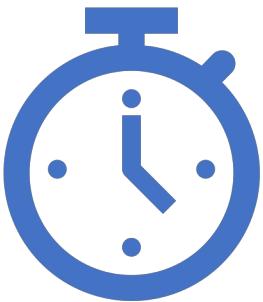


Guide to Censoring & Sampling Bias

Left Truncation

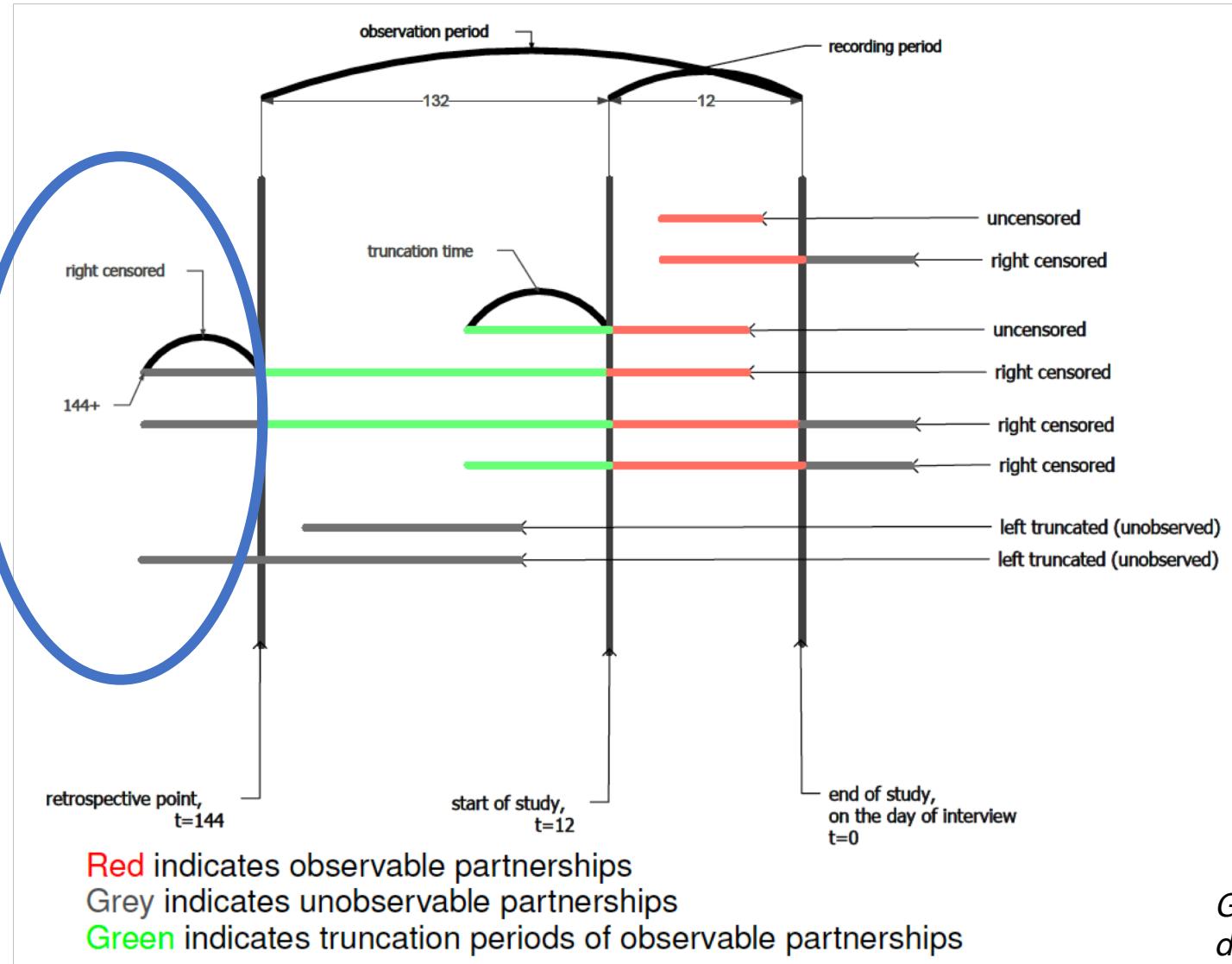
- ‘delayed time entry’
- Bias towards longer duration





Guide to Censoring & Sampling Bias

not present in
NSFG data



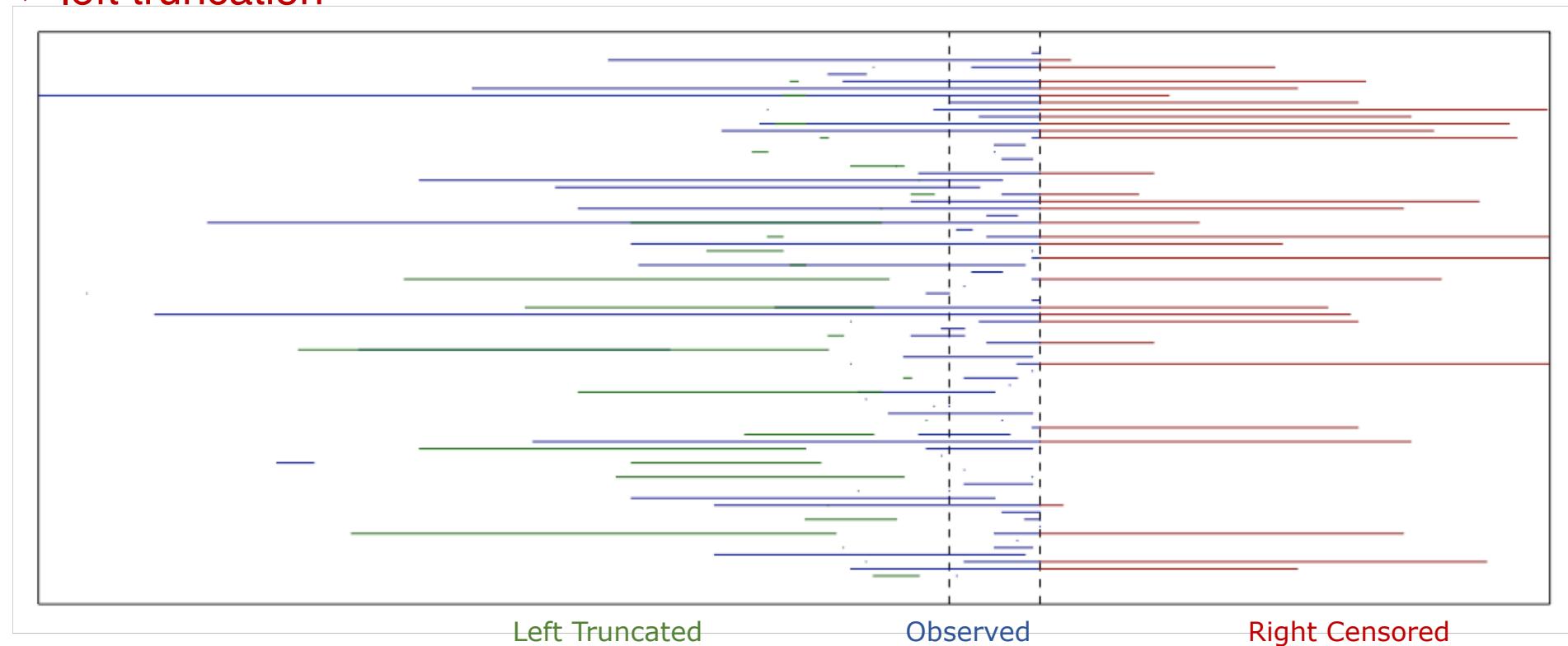


Censoring in NSFG

Describe up to 3 sexual partners in last 12 months (cross-sectional egocentric data)
→ right censoring

Report start and end dates (retrospective data)

→ left truncation



Relationship Duration in NSFG





Motivation

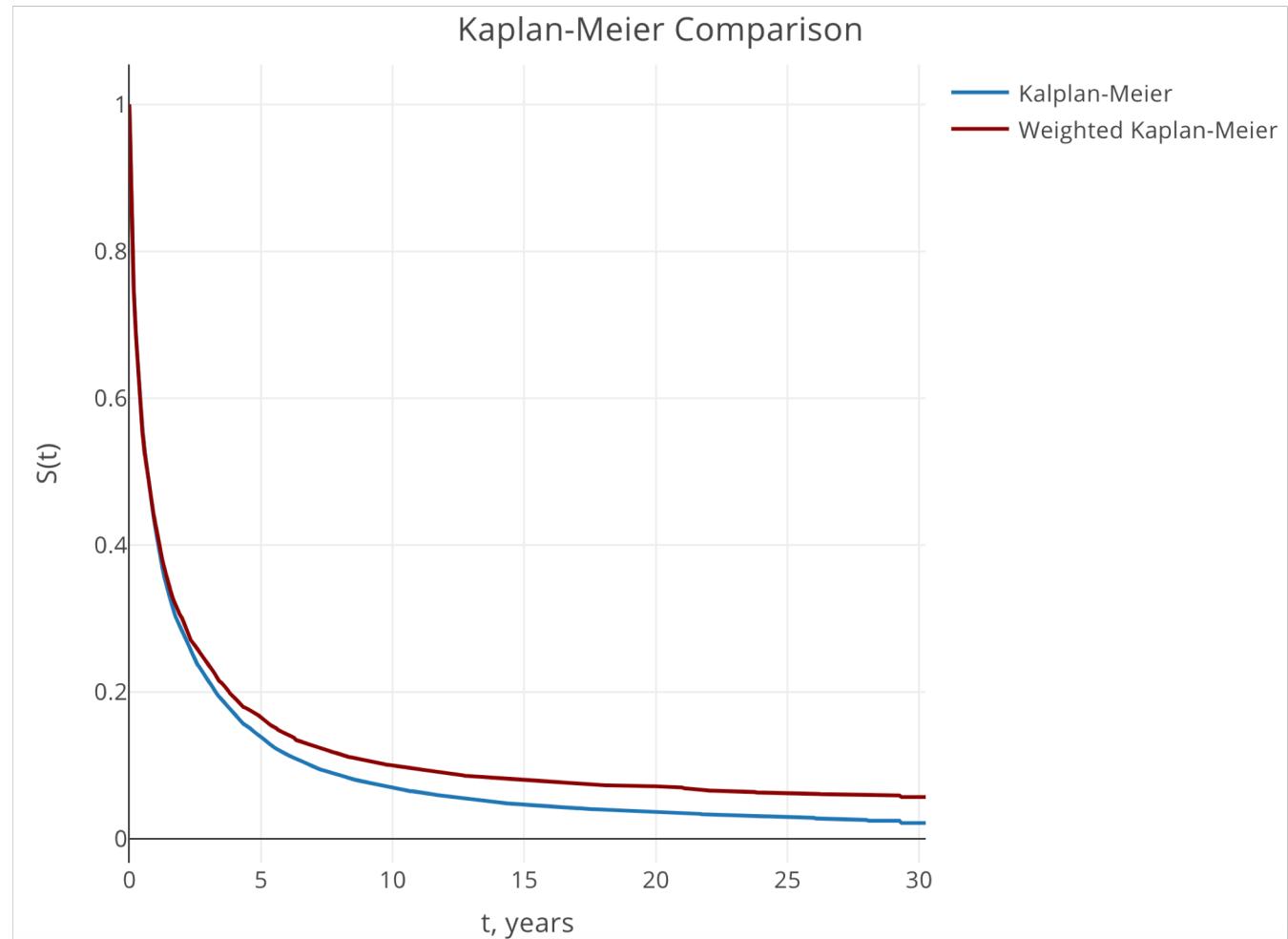
What are the patterns of relationship duration by age, and is there a reasonably simple framework by which we can represent them in our ERGMs?

- Is an exponential distribution, either homogenous across all nodes or heterogenous (homogenous within groups)
- How does ego age influence the rate of both formation and dissolution?
- How does the survey design influence how we understand patterns by age?



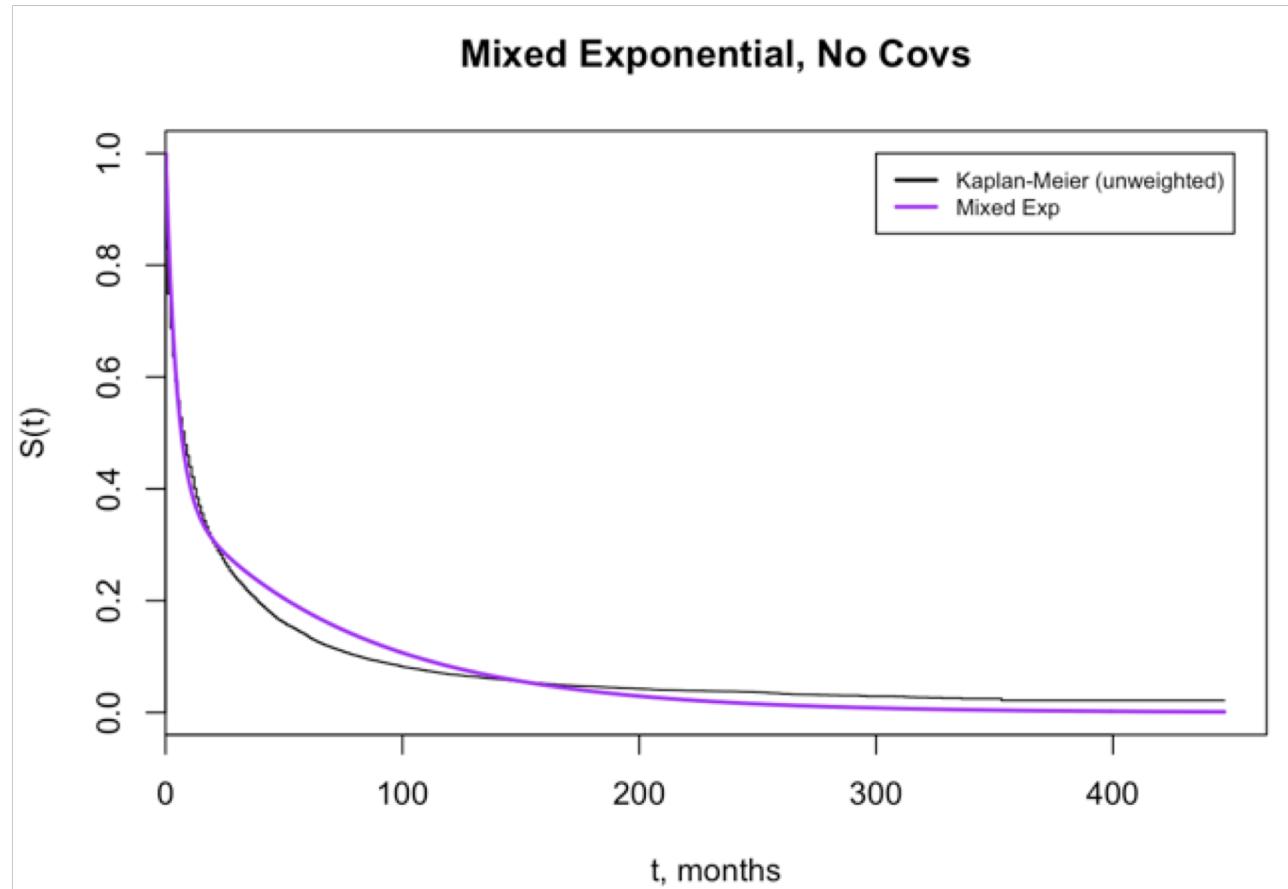
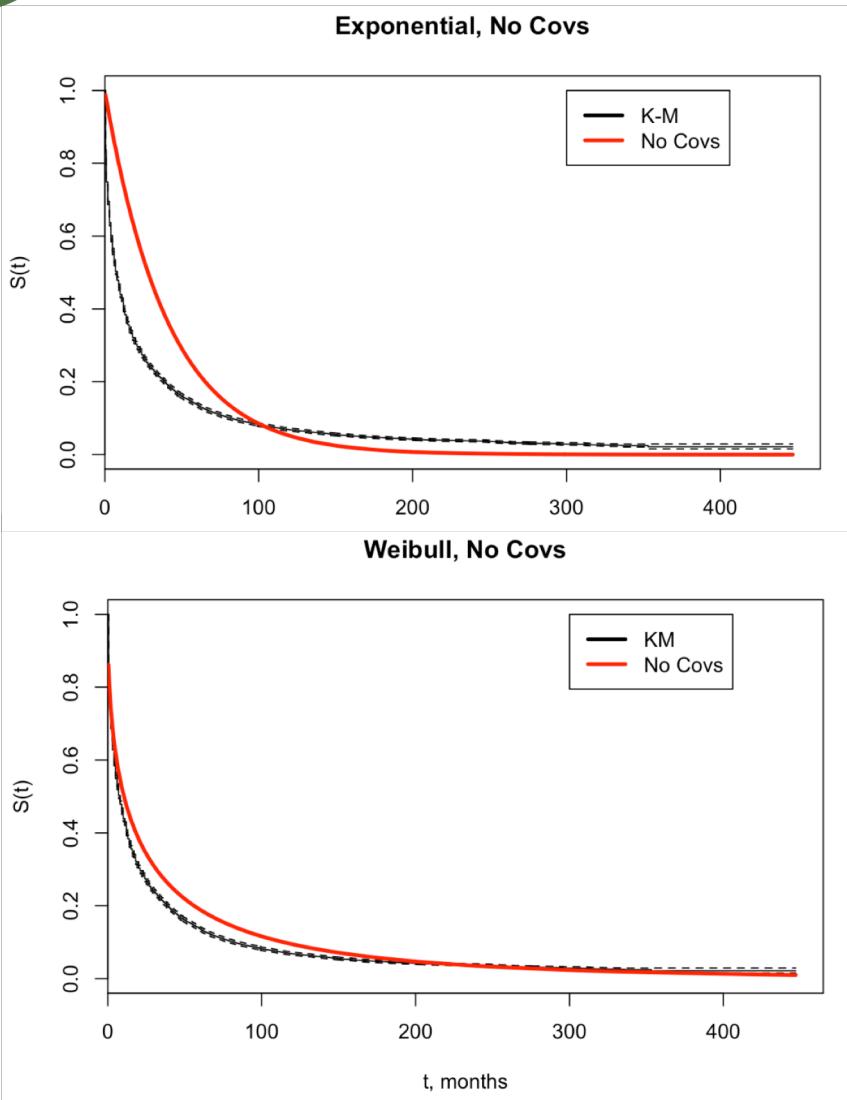
Modified Kaplan-Meier

- Kaplan-Meier Curves as comparison
- Weighted to NSFG weights vs Unweighted





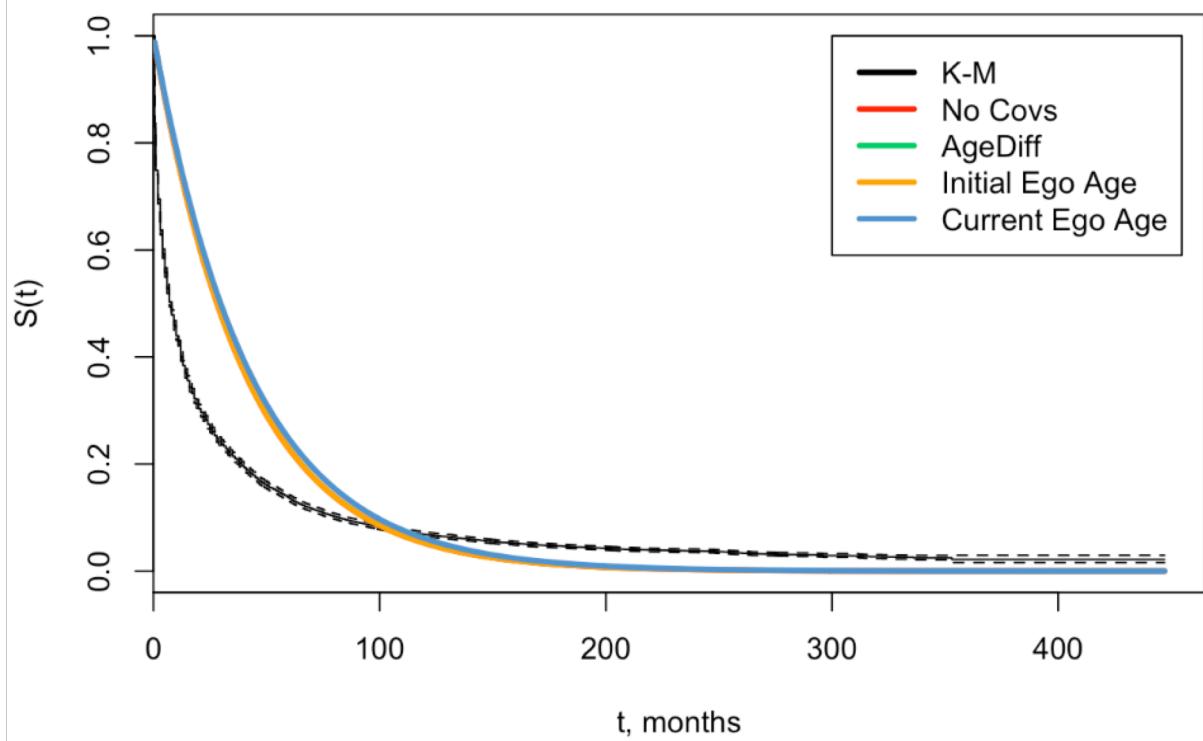
Duration-Only Models



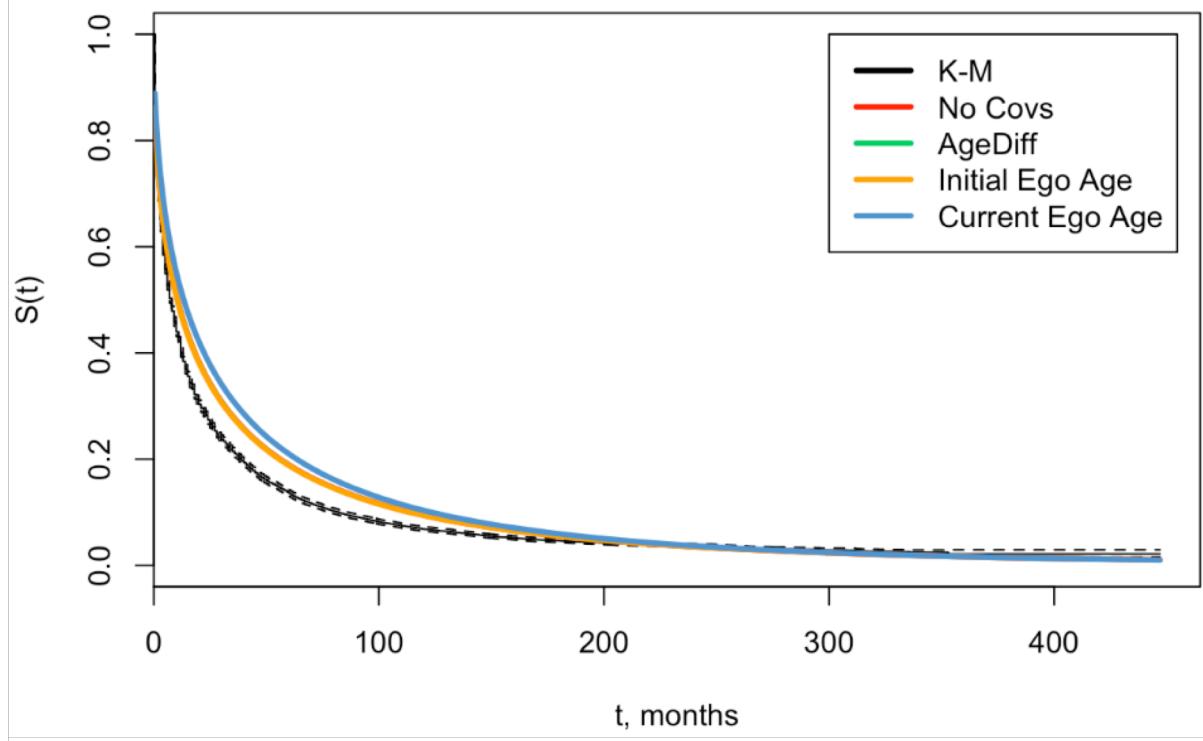


Age Models

Basic Age Dynamics, Exp Dist

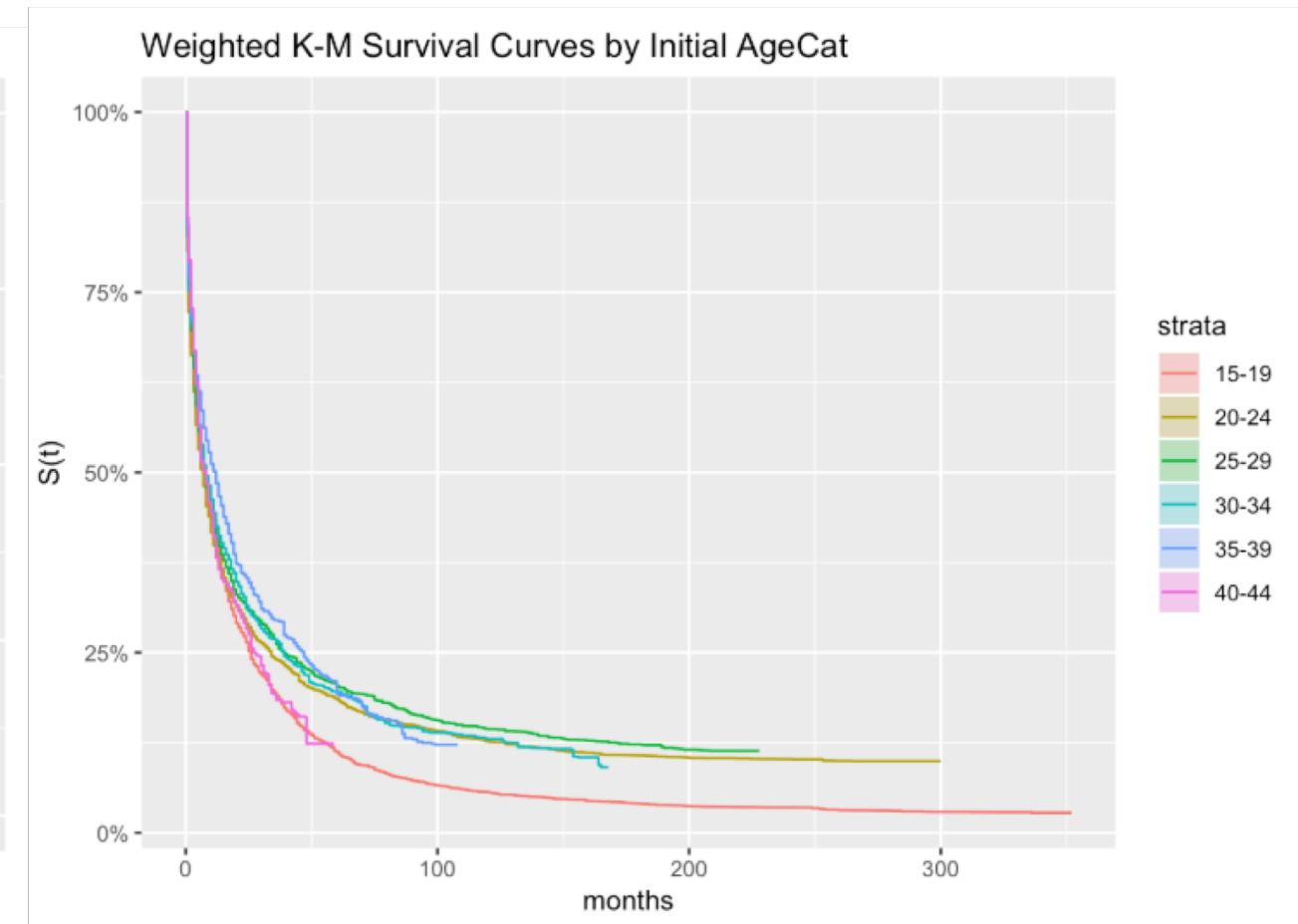
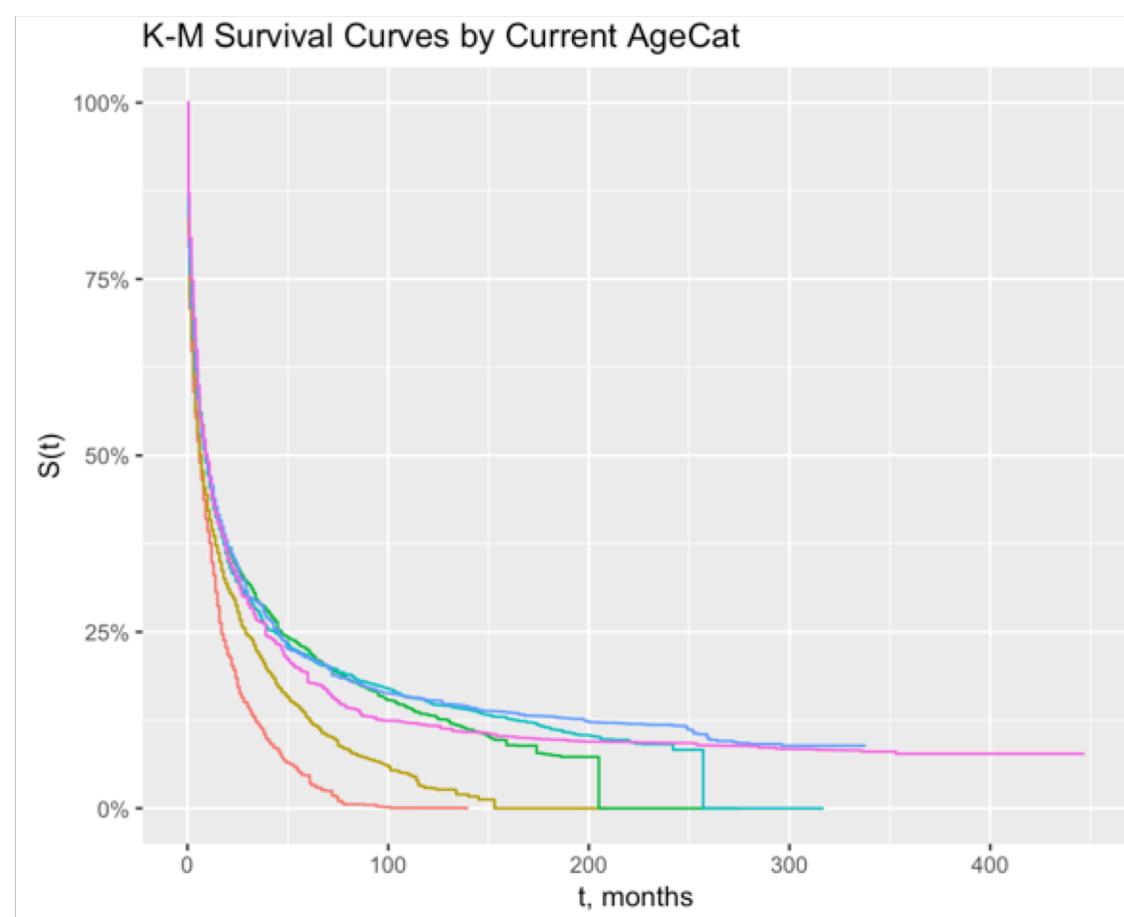


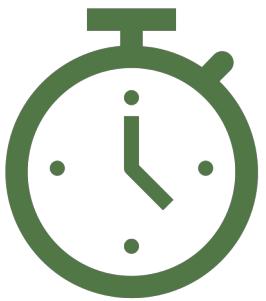
Basic Age Dynamics, Weibull Dist



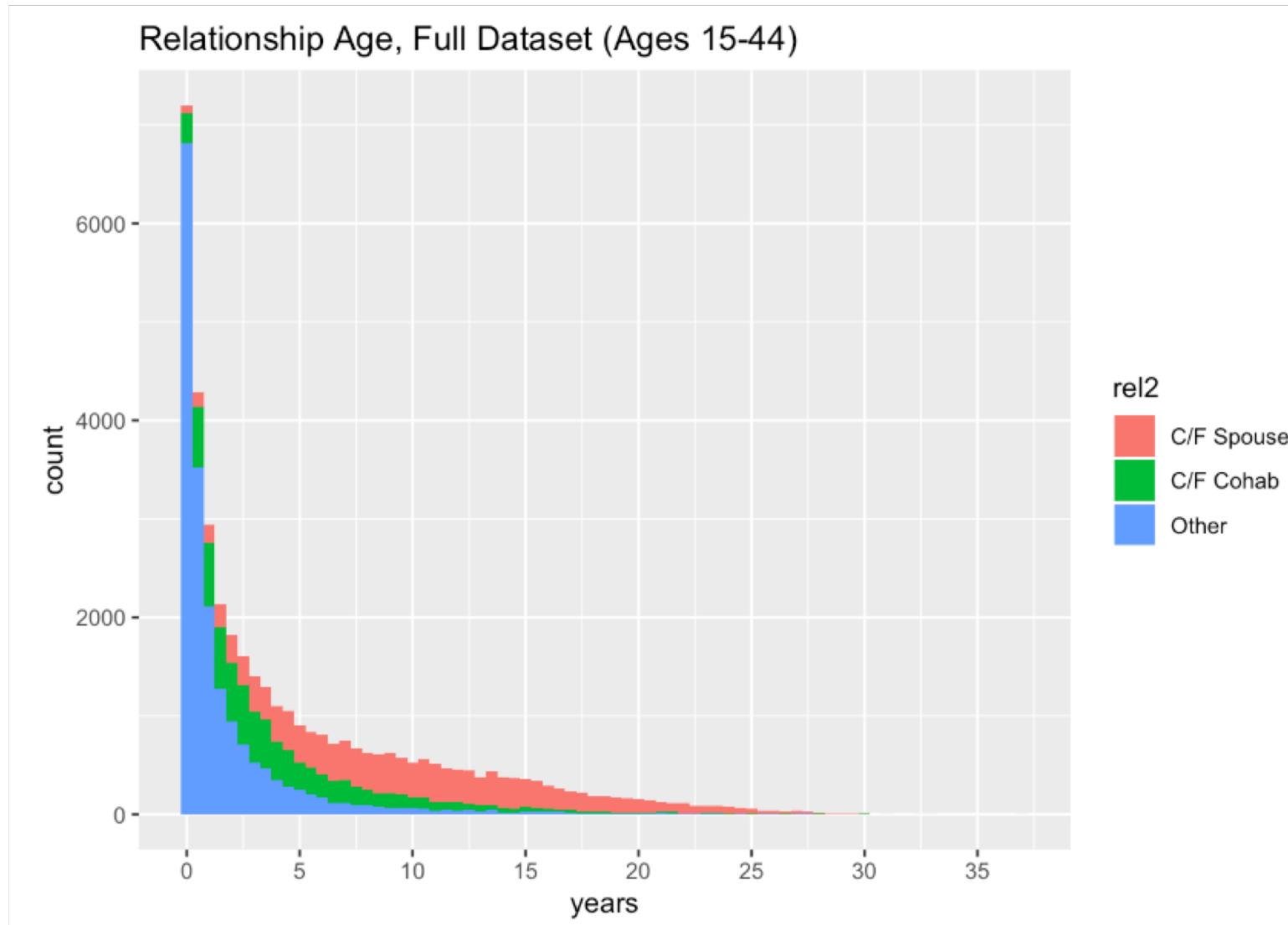


Gaining Some Intuition



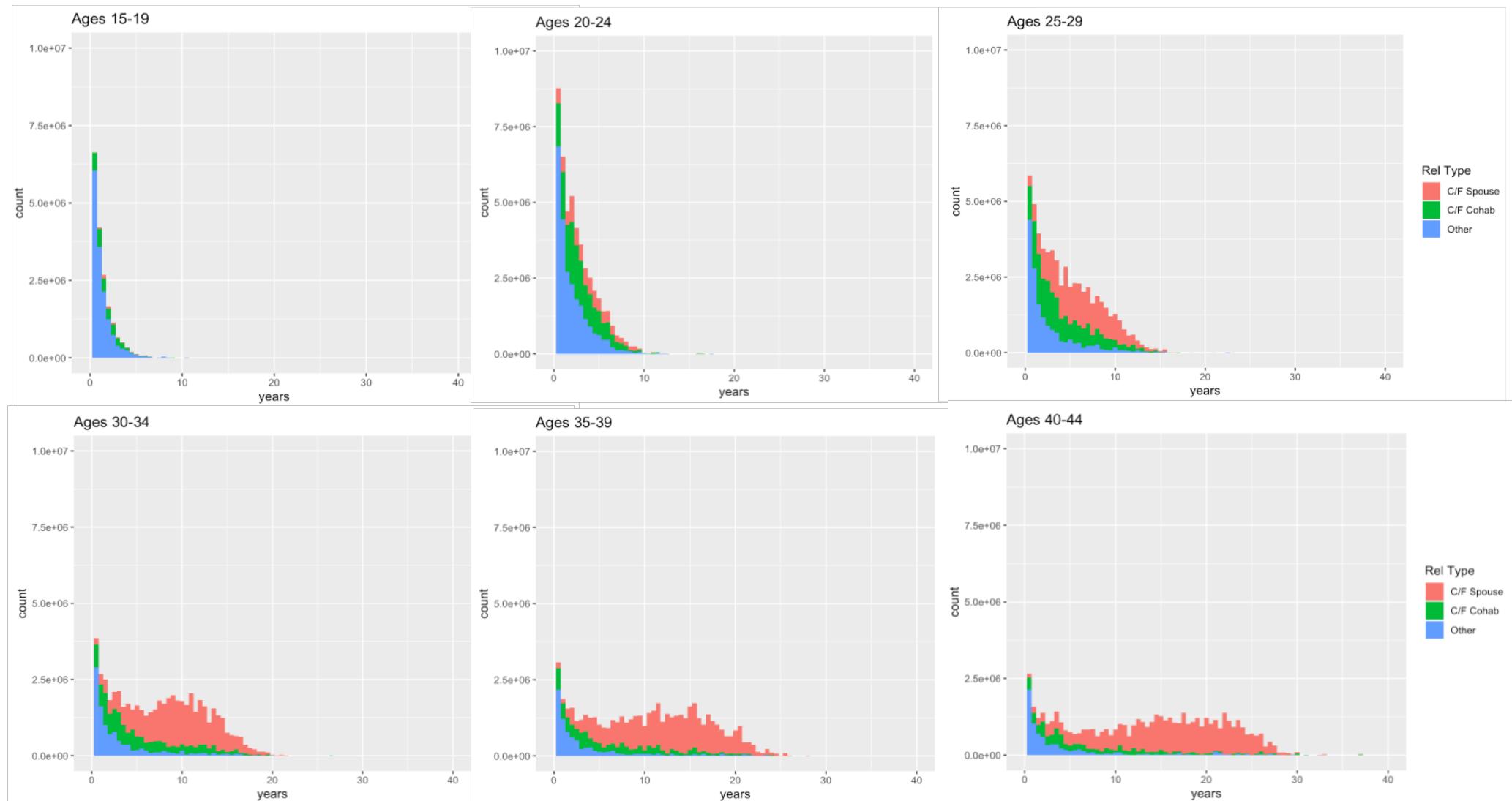


Gaining Some Intuition



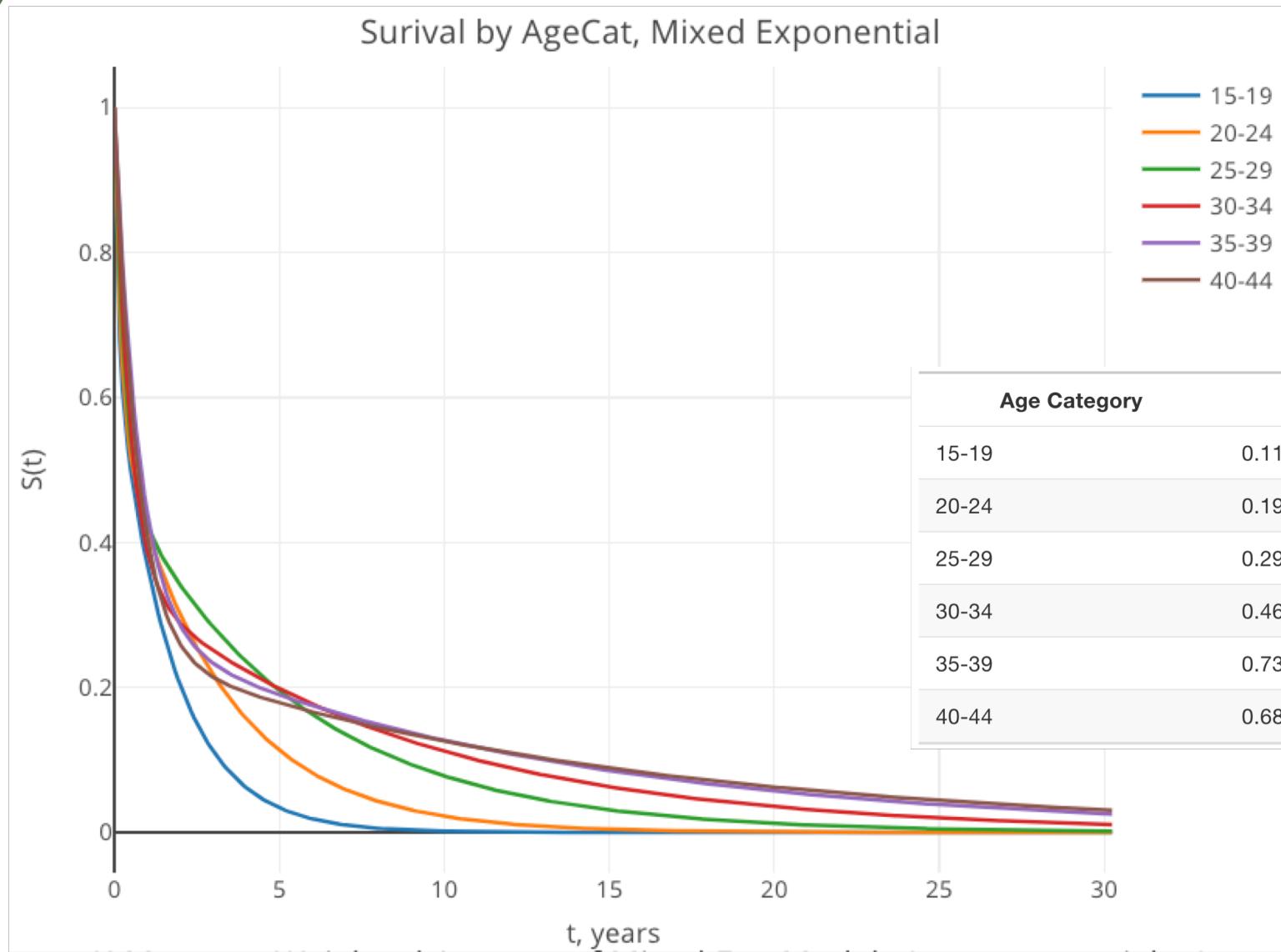


Gaining Some Intuition





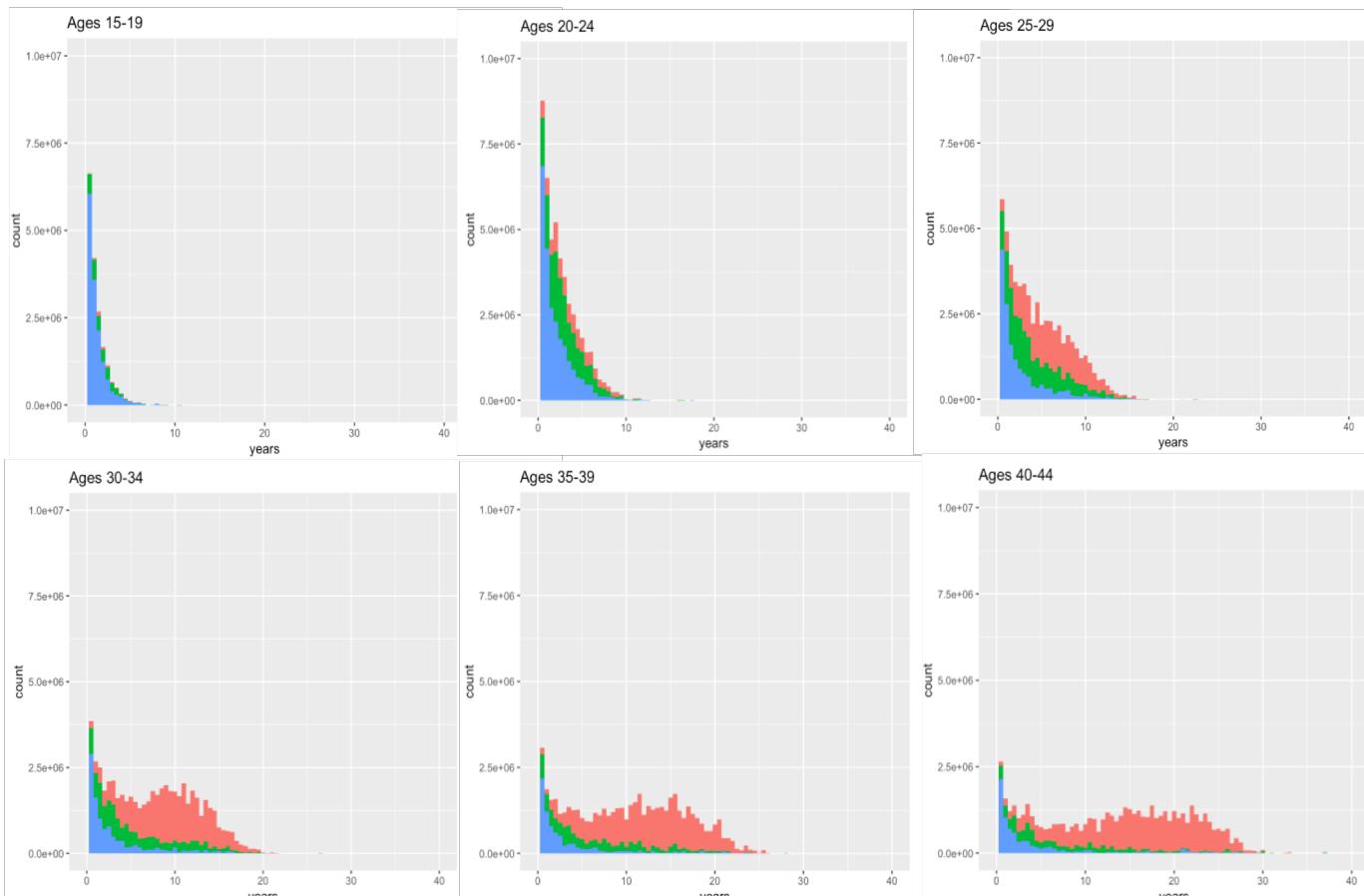
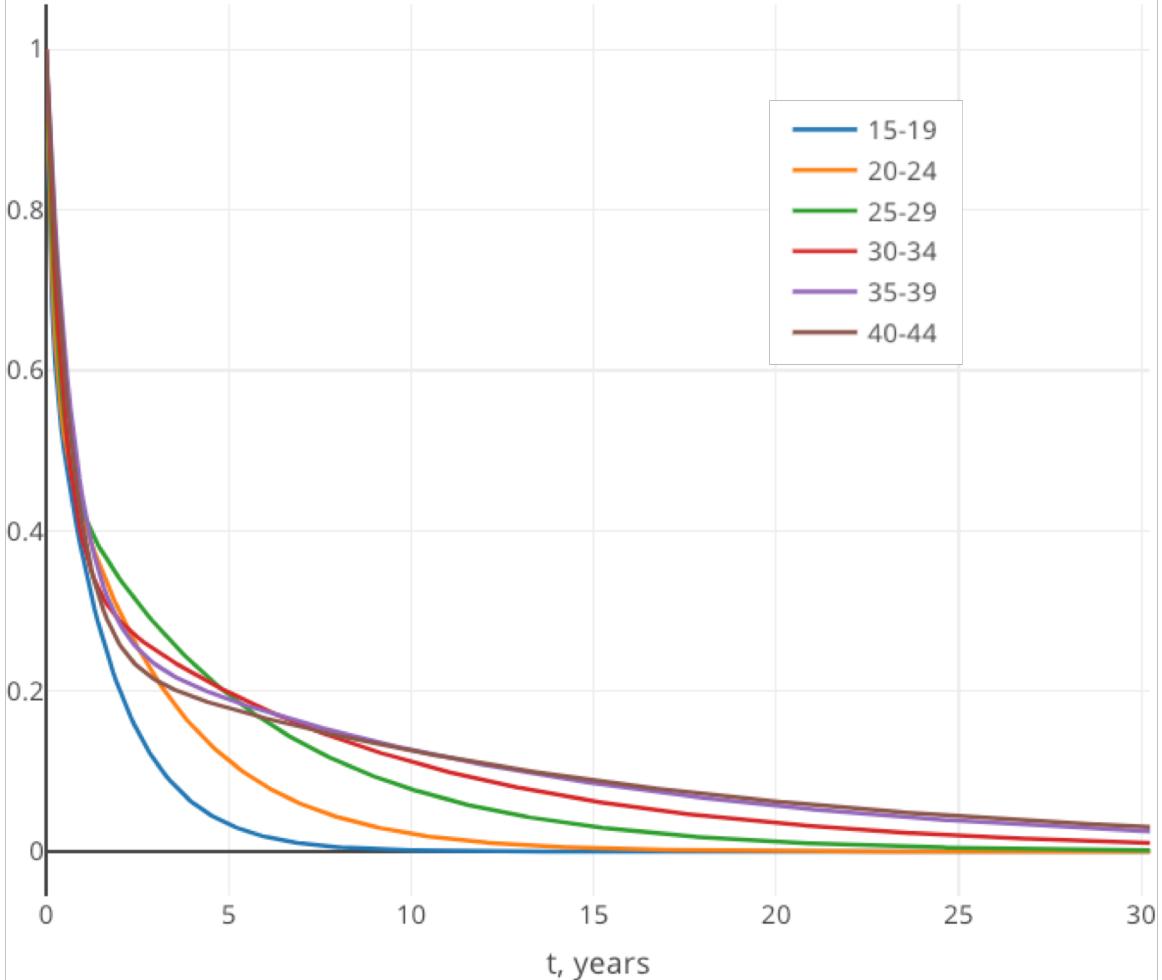
Mixture Models





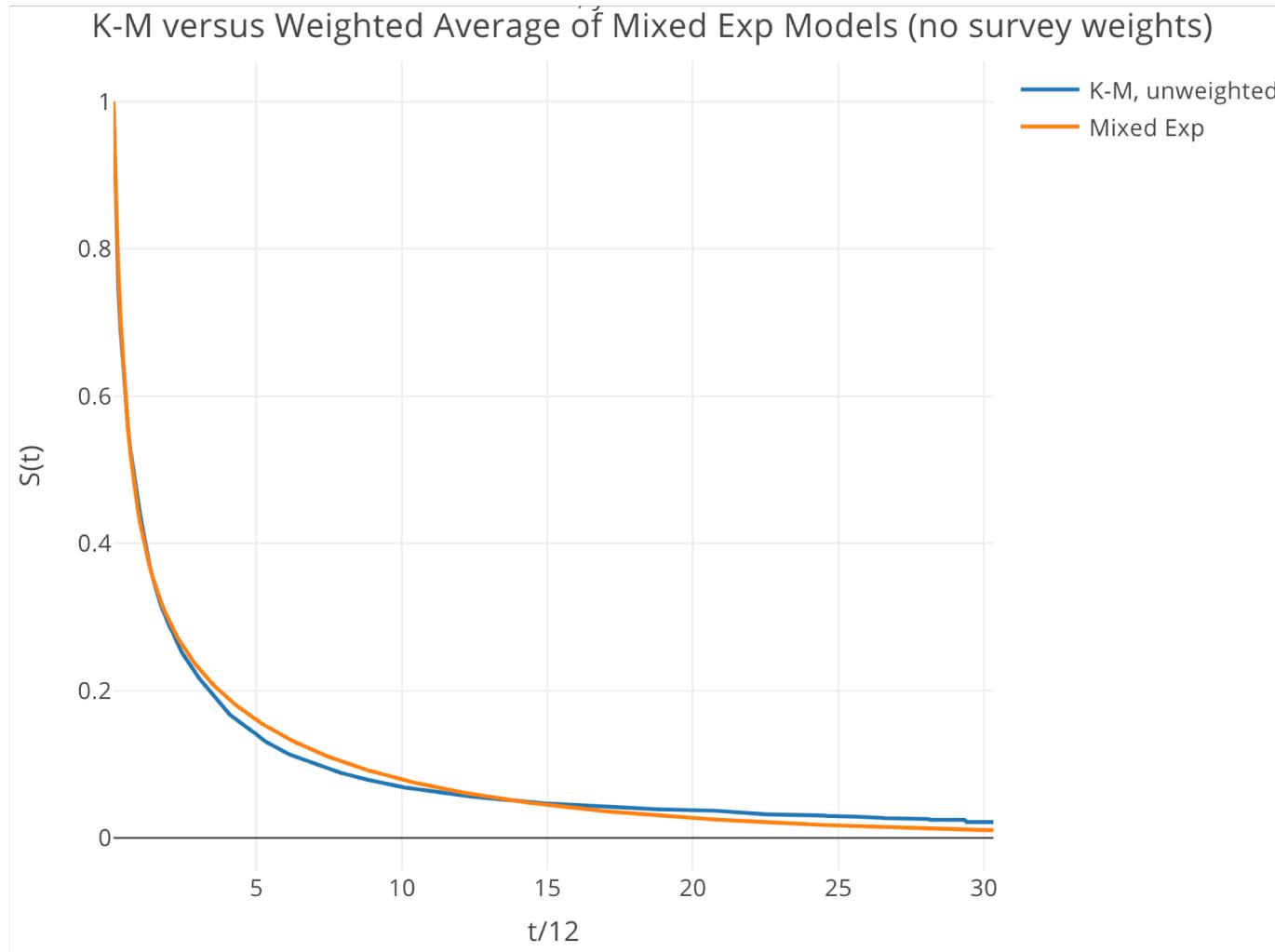
Mixture Models

Survival by AgeCat, Mixed Exponential



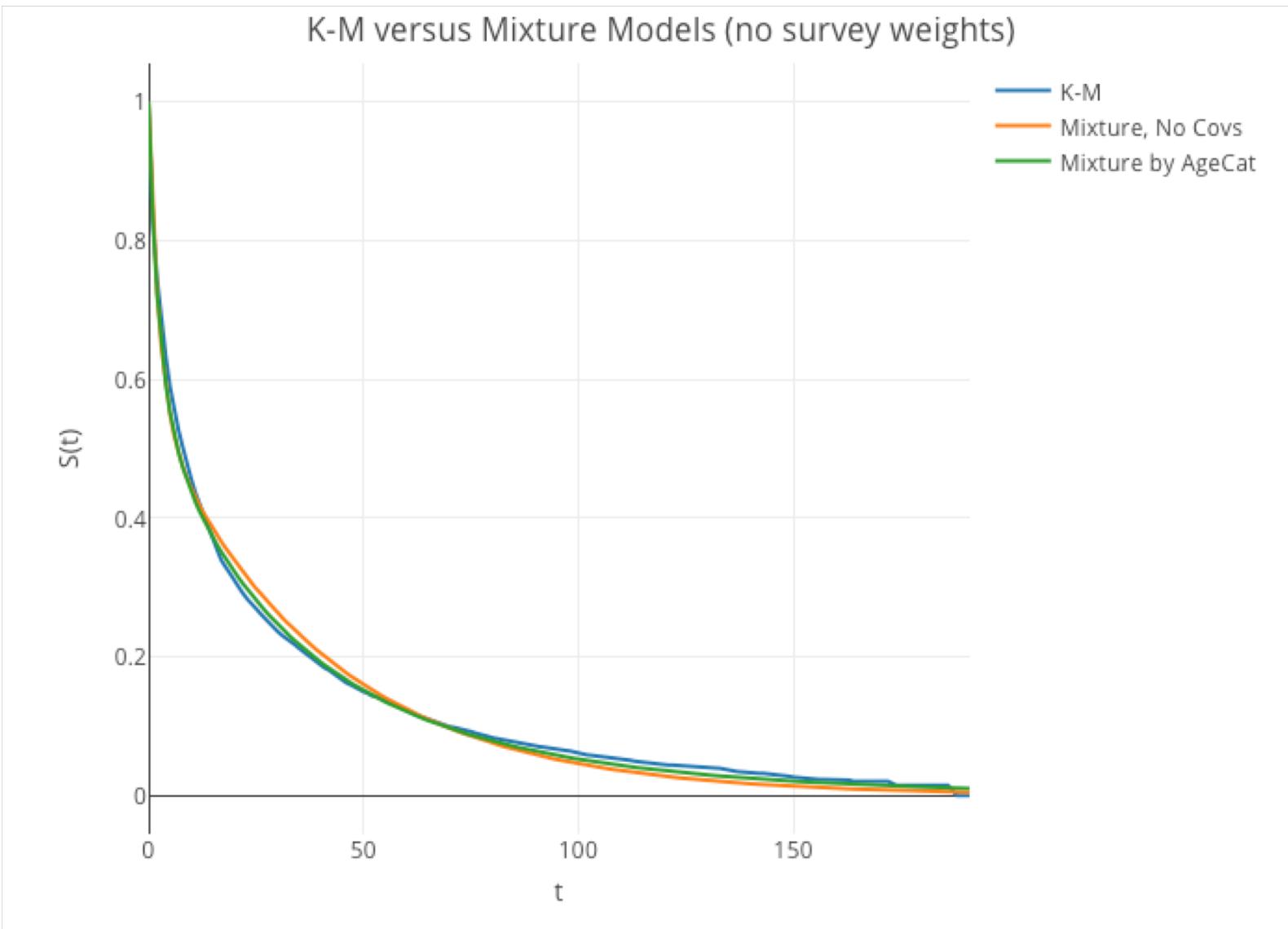


Mixture Models





15-29-year-olds





Next Steps

- Part 1 (EPT Model): use insights from mixture models to model dissolution among the 15-29 age group
- Part 2:
 - Look into family demography literature for intuition and methods for better framework for generating very complex age dynamics (particularly at ages 30+)
 - Simulation modeling
 - Set up different scenarios based on different assumptions about relationship duration, formation rates over time, and degree distribution
 - Run sim, do a “look back” at the data in the same way NSFG asked respondents to report their relationships