

All Organizations Were Once New: Choosing Appropriate Statistical Methods to Analyze Data on New Ventures

Tiantian Yang

Howard E. Aldrich

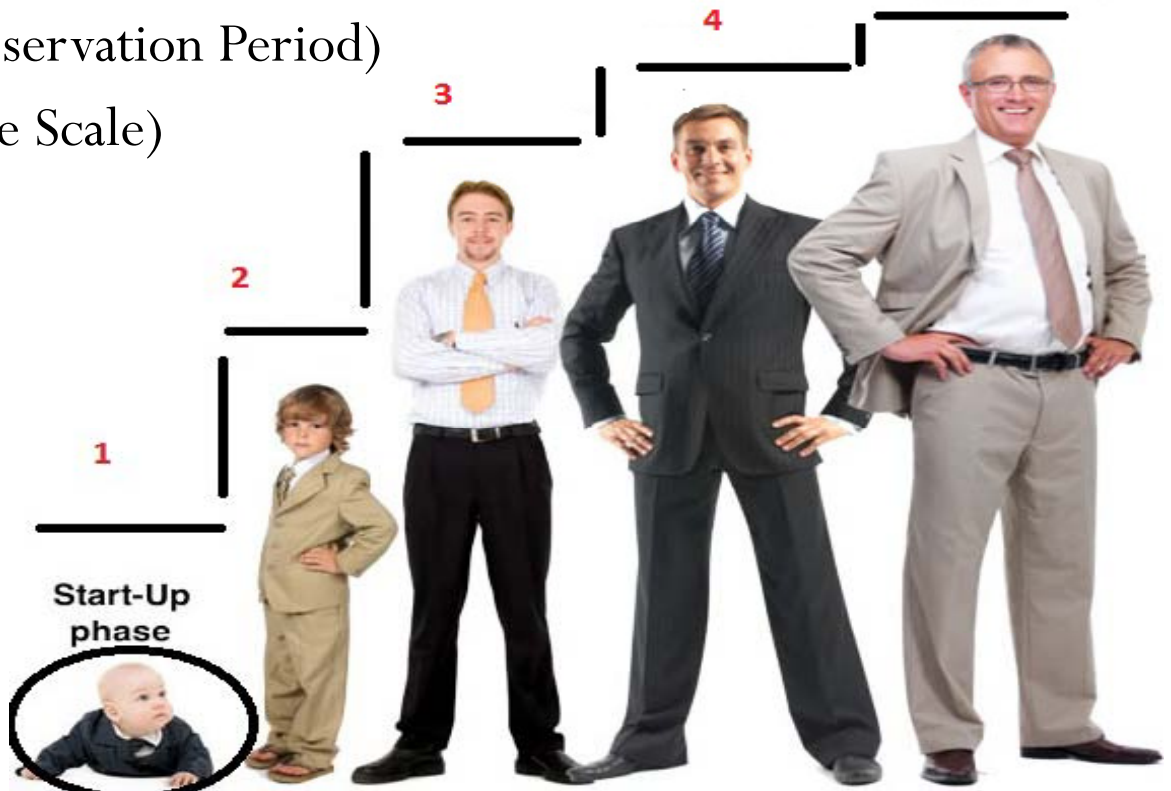
University of North Carolina-Chapel Hill

Global Entrepreneurship: The Latest Research on Business Creation

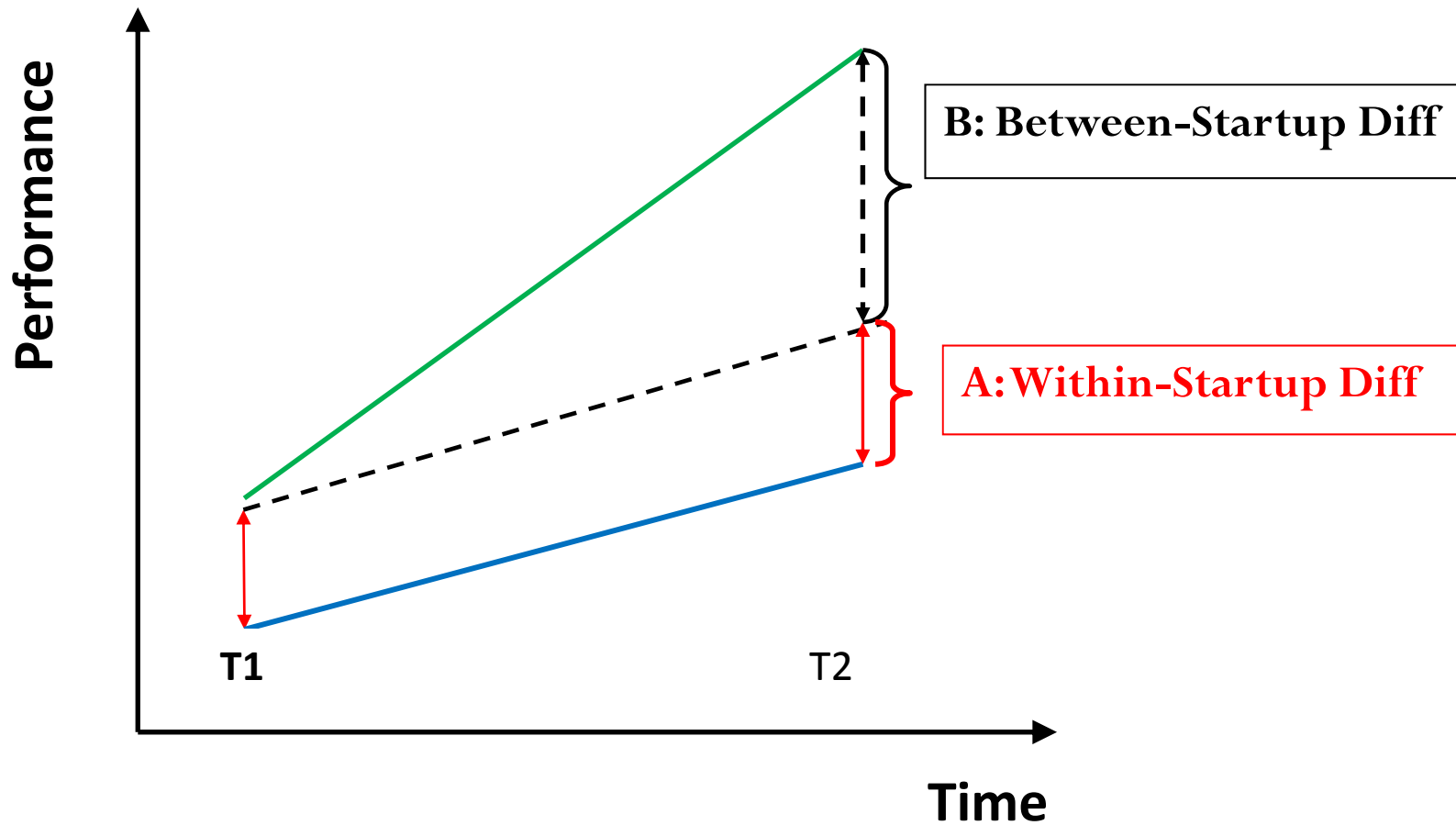
October 2010

Liability of Newness

- Identification of New Firms
 - Emerging Organizations *VS* Established Firms
- Process of Organizational Changes
 - Changes over time (Observation Period)
 - Pace and Rhythm (Time Scale)
- Panel Data
 - PSEDI & PSEDII



Why not Cross-Sectional Data

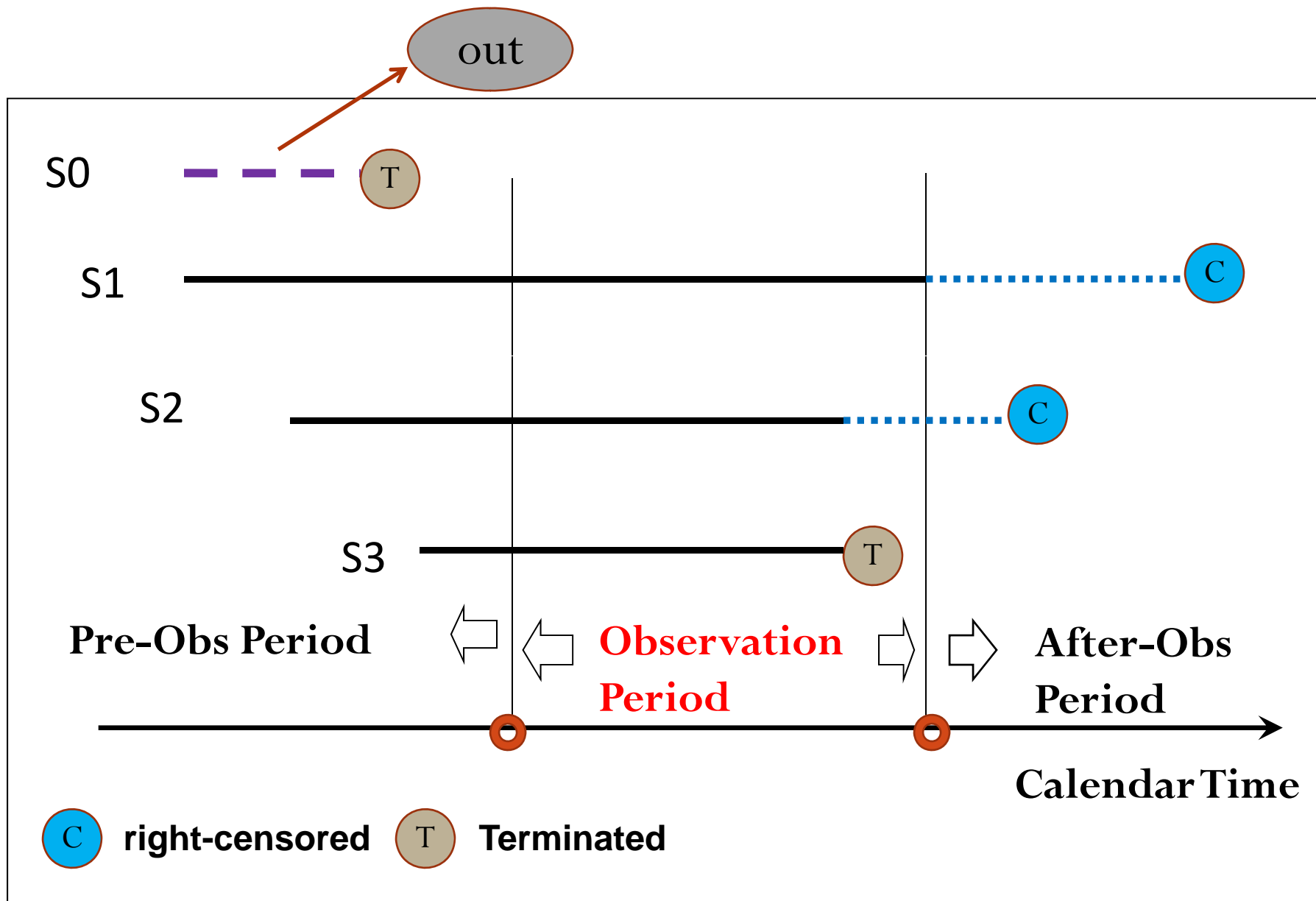


Properties of Panel Data on New Ventures

- New Ventures Have Existed before We Begin Our Observations
 - Difficulties in Identifying New Ventures
- A Study Ends before All New Ventures Die
 - Limited Researcher Resources vs Long Survival Time of Some Firms
- Left-Truncation and Right-Censoring



Figure 1. Left-Truncated Data with Known Start-Time



Left-Truncation

- Left-truncation arises when observation period starts after subjects in a study have already come into existence before the study began
- Left-truncated data are incomplete and include disproportionate numbers of *hardy* subjects. (S0 and S1)
- Data having Left-Truncated Cases lead to *Over-estimation* of Survival Rates

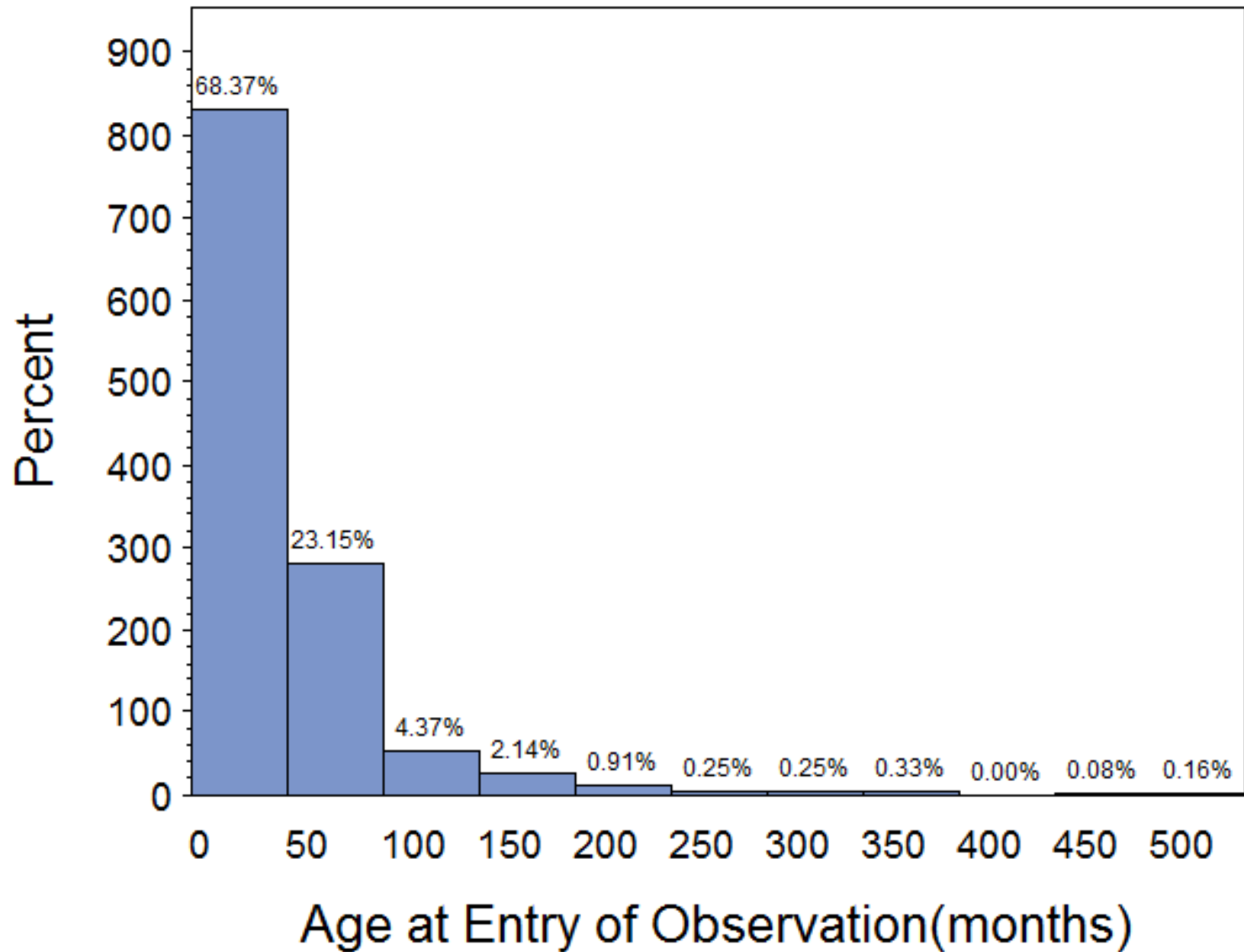
An Example: Using PSEDII

- Target Population: New Ventures
 - nascent entrepreneurs who are active in business creation but have not yet created a profitable firm
- Sampling Procedures
 - Identification of New Ventures
 - Exclude Established Firms
- Data Structure
 - Initial Interview
 - Four Follow-up Interviews

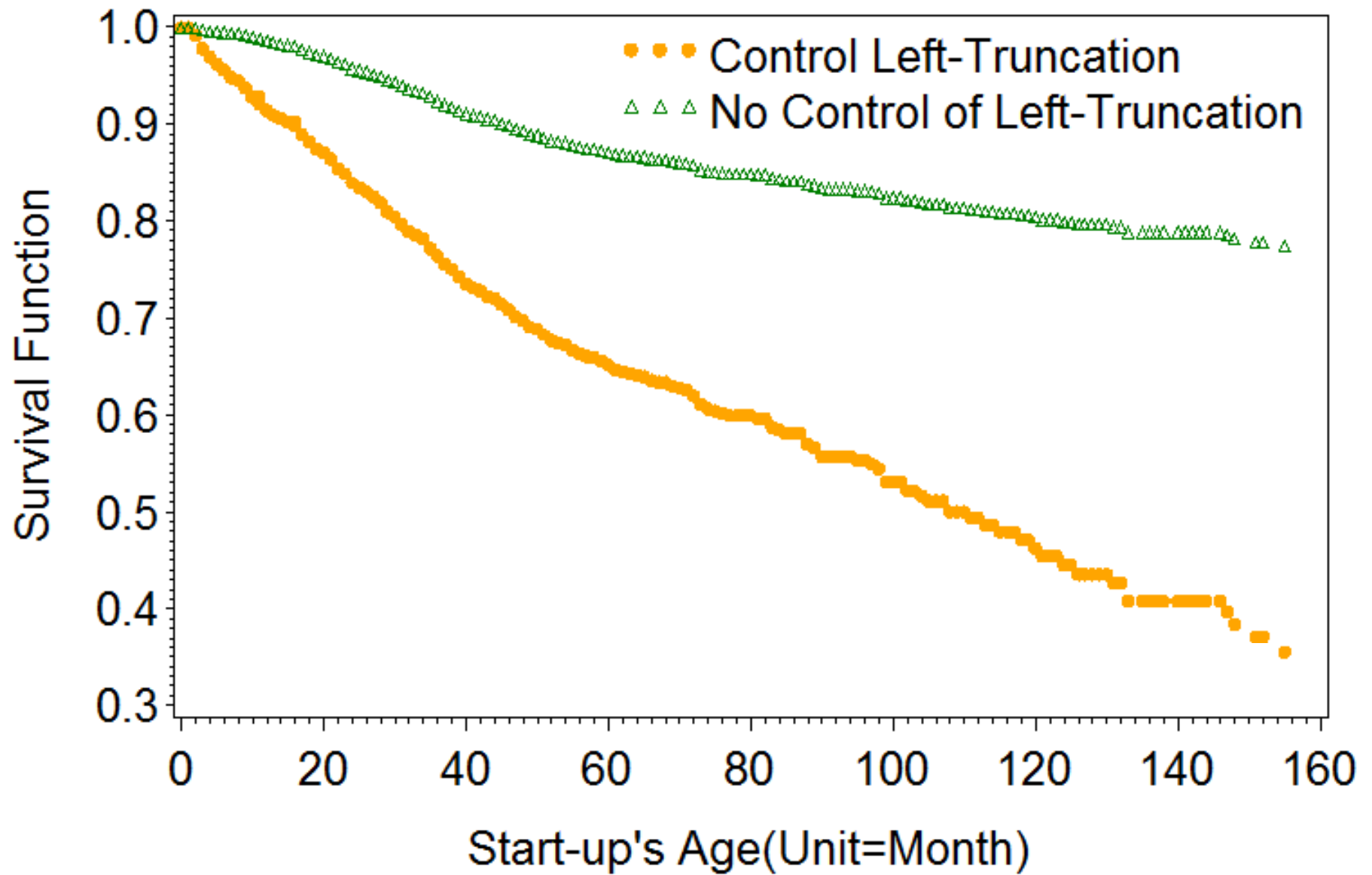
Table 1. The Number of Observations in Waves (PSED II)

Interview Waves	Frequency	Percent
only A	115	9.47%
A, B	278	22.90%
A, C	44	3.62%
A, D	22	1.81%
A, E	14	1.15%
A, B, C	194	15.98%
A,B,D	16	1.32%
A,B,E	9	0.74%
A, C, D	12	0.99%
A, C,E	7	0.58%
A, D, E	5	0.41%
A, B, C, D	98	8.07%
A, B, C, E	26	2.14%
A,B, D, E	9	0.74%
A, C, D, E	23	1.89%
A, B, C, D, E	342	28.17%
Total	1214	100.00%

Start-ups' Ages at Date of Screening Interview



Compare Survival Function: Control for Left-Truncation and No Control

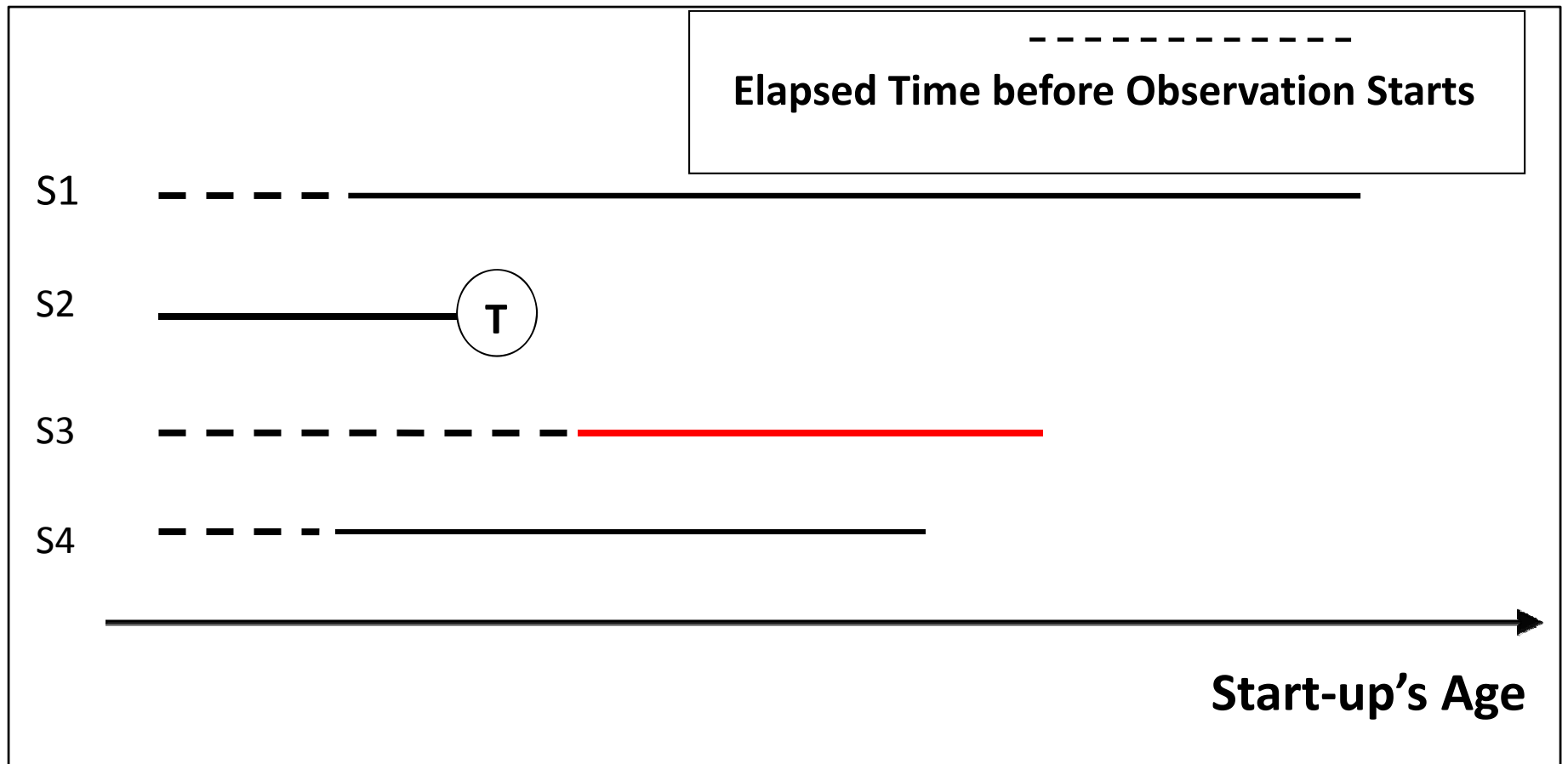


Statistical Solution to Left-Truncation

- Conditional Likelihood Approach in Proportional Hazard Model
- Excludes Subjects from the *Risk Set* When They Have yet not Entered into Our Observation



Risk set for Left Truncated Data (S2 was terminated at age 10 months)



Risk set for Start-up 2 at age 10 months is {S1, S2, S4}, WHY?

Implications

- Data Structures
- Model Assumptions
- Model Assumptions are Satisfied by Our Data
- You MUST take account of left truncation if you want to examine startup “age” in your analysis