Panel Study of Entrepreneurial Dynamics

Richard Curtin
University of Michigan

Research Design

- Basic Sample Design
  - Used large representative sample to identify nascent entrepreneurs
- Supplemental Samples
  - Women
  - Minorities
  - Mixed gender control group
  - Minority control group
- Survey Waves
  - Telephone and mail interviews
  - Third annual wave completed in 2001
  - Wisconsin conducted wave 1 & 2; Michigan wave 3

Completed Interviews

- Main Sample
  - Wave 1: 446
  - Wave 2: 342 or 77% of eligible wave 1 respondents
  - Wave 3: 256
    - 231 were wave 1 & 2 respondents
      - 93% of eligible (275 = 342 – 67 deaths by wave 2)
    - 25 from wave 1 but lost in wave 2

- Women's Sample
  - Wave 1: 223
  - Wave 2: 159 or 71% of eligible wave 1 respondents
  - Wave 3: 141
    - 119 were wave 1 & 2 respondents
      - 93% of eligible (128 = 159 –31 deaths by wave 2)
    - 22 from wave 1 but lost in wave 2

Completed Interviews

- Minority Sample
  - Wave 1: 161
  - Wave 3: 114 or 71% of eligible wave 1 respondents

- Total Sample of Entrepreneurs
  - Wave 1: 830
  - Wave 2: 501
  - Wave 3: 511
  - Deaths by wave 3: 224

- Control Group Samples (wave 1)
  - Both genders: 223
  - Minority: 208

PSED Weights

- No weights needed if:
  - Equal probability of selection (EPSEM)
  - No differential non response
- PSED needs weights because:
    - Differential selection probabilities
      - Additional samples of women and minorities
    - Differential non response
      - Some types of respondents were harder to reach and maintain in panel

PSED Weights

- Weights
  - Differential selection probabilities
    - Weight by inverse of selection probability
      - If selection probability is 2.0 then weight is 0.5
      - If selection probability is 0.5 then weight is 2.0
  - Differential non response/panel attrition
    - Weight by inverse of response rate
      - If response rate is 0.5 then weight is 2.0
      - If response rate is 2.0 then weight is 0.5
  - Two factors often interact
      - Often more difficult to contact special samples
      - Don’t know impact of differential non response
      - Can combine correction factors

Weight Strategy

- Sample weights needed for:
  - Base screening sample
  - Nascent entrepreneurs
  - Additional sample of women
  - Additional sample of minorities
  - Control groups

- Rather than multiple weights developed integrated weights
  - Enables analysis of all cases across sample types
    - for example, all nascent entrepreneurs in wave 1 regardless of whether in added samples or main sample
  - Can still do separate analysis of subsamples

Calculation of Screener Weights

- Use Census (CPS) data to determine weight targets
  - Used: sex, race, age, and education
    - Too much missing data on income

- Creates multi-way table of distributions:
  - Example

    Male, white, 45-54, college = 5% of actual population
    - If sample = 10% then weight = 0.5
    - If sample = 2.5% then weight = 2.0

- The single set of target distributions correct for both differences in selection and non response

- Outcome: when use weights get same distributions as in CPS data

Calculation of Sample Weights

- Used weighted screener data to determine targets
  - Sex, minority, age, and education
  - Corrects for selection differentials,
    Non response bias, and panel attrition
    - Procedure repeated for all waves separately
  - Results in weights that represents national sample

- Same procedure used for control groups

Pitfalls of Weighting

- Extreme weights
  - How large a weight are you willing to accept
    - Highest three times lowest?
      - Five times? Ten times? Hundred times?
  - Same problems as extreme values in analysis
  - Current weights: Highest three times lowest.

- How much should weights contribute to the variance of a variable?
  - Example: compare two different weights
    - Same estimate of mean or proportion
      But one has much higher variance
      - Variance 1: 95% “true” & 5% due to weights
      - Variance 2: 65% “true” & 35% due to weights
    - New weights are like 1 and old weights like 2

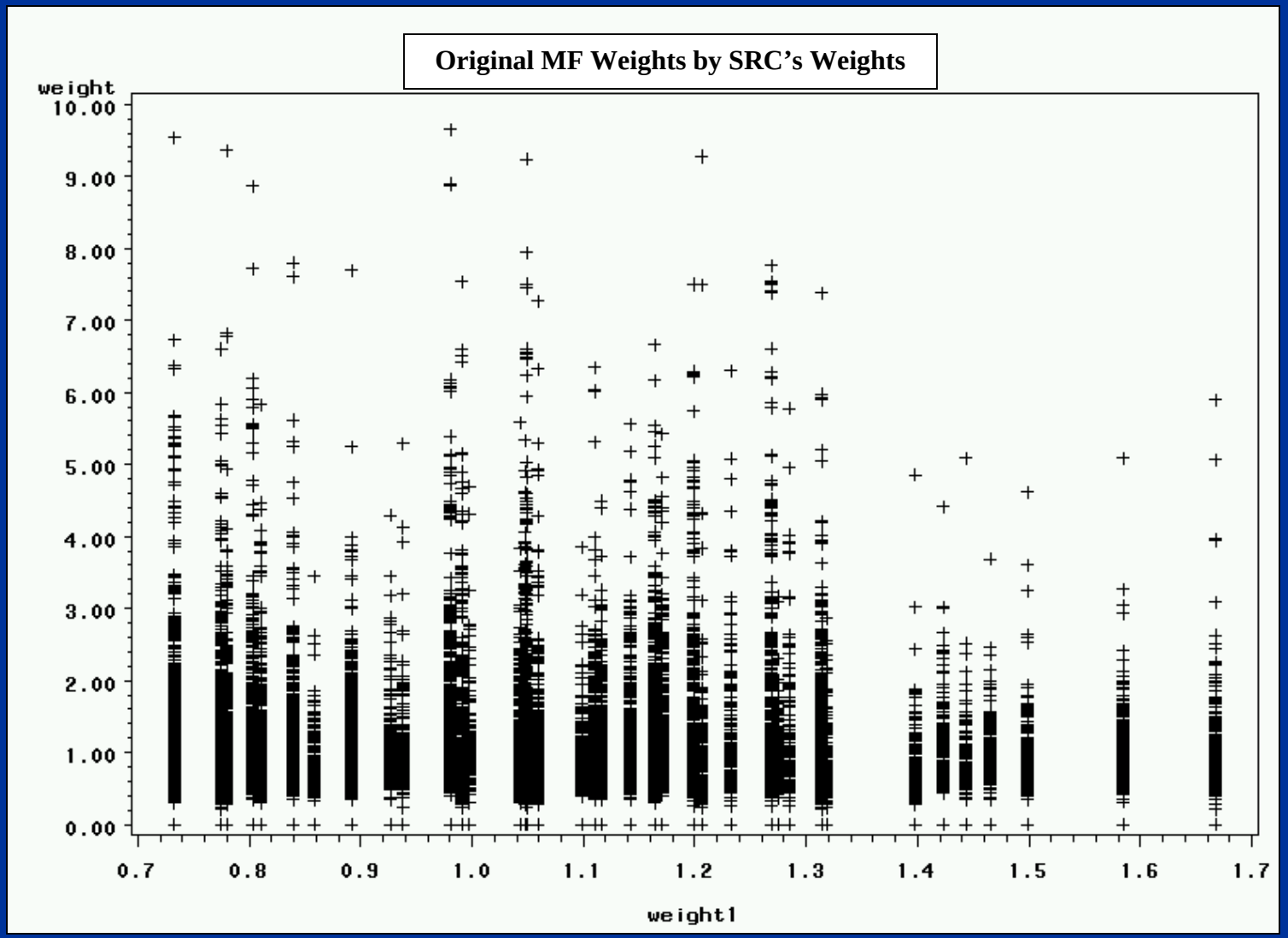
Age Distribution for Entrepreneurs In Screening Sample

- Old and new weights give nearly identical point estimates as the example below indicates

Age	Original	New SRC
18-24	12.62%	12.99%
25-34	28.97%	28.29%
34-44	30.01%	30.19%
45-54	19.49%	19.80%
55 or older	8.88%	8.73%

Original versus New Weights

- Variance of Old weights is much greater than new weights
  Old range from 0.1 to 10.0 (left scale) new from 0.7 to 1.7



Use of Weights in Analysis

- Always use weighted data in analysis of:
  - Nascent entrepreneurs
  - Women and minority samples
  - Control groups
  - Screener sample
- Weight variables integrate all sample types
  - WTW1, WTW2, WTW3
  - WTCG
  - WT_scrn
- Weights for women and minority subsamples
  - If want to do women's analysis, simply select all women in total sample
  - For analysis of minorities, select all minorities

Use Centered Weights in Analysis

- Use of weights always yield correct estimate of mean, but they may not yield correct standard error
  - Need to have mean weight equal to 1.0 in analysis
  - If mean weight within each analysis not equal 1.0 then get bias estimate of standard errors because sum of cases not equal to sum of weights

- How to center weights on 1.0
  - Define analysis subgroup
  - Calculate mean weight in subgroup, call it WTsub
  - Then define NewWT = WTW1 / WTsub

Relative and Absolute Weights

- Centering weights changes absolute value
  not relative value of weights
- Relative value of weights provide accurate estimate of
  means and proportions
- Absolute value of weights provide accurate estimate of
  standard errors of means and proportions
  and hence accurate statistical tests
- Over-samples yield more precise estimates, not less

Documentation

- Project Internet Site
  - http://projects.isr.umich.edu/psed/
  - Complete project documentation
  - Most up-to-date version
- Questionnaires
  - Screening questionnaire
  - Phone interviews from wave 1, 2, 3 (100+ pages each)
  - Mail forms: wave 1 (long) wave 2,3 (short)
  - PDF format; book marked files
- Data File and Codebook
  - Complete SPSS data file (system file and portable)
  - Comprehensive codebooks (450 pages)
  - PDF format; book marked files

Codebook

- Codebook includes
  - All variables in all waves
  - Excludes names and other confidential information
- Variable numbers and questions
  - Uses question number in questionnaire
  - Same numbers as in waves 1 & 2 data files
  - Displays identical questions across waves together
- Codebook order and index
  - Table of contents – grouped by topic
  - Index gives page numbers for every variable
    - Listed in alphabetical order in appendix
    - Can use questionnaire, then look up page number

Codebook Example

WAVE 1	WAVE 2	WAVE 3		
Q120	R577	S577	At what stage of development is the product or serv this (start-up/new firm) will be selling: 1) Complete ready for sale or delivery; 2) A prototype or proced has been tested with customers; 3) A model or proc is being developed; or 4) Still in the idea stage?	
352	183	183	1.	Completed and ready for sale or delivery
154	56	43	2.	Prototype/procedure tested with customers
169	40	25	3.	Model/procedure is being developed
136	12	15	4.	Still in the idea stage
2	4	0	0.	No work has been done on a product or serv
17	3	9	9.	DK; NA

Codebook Items

- Frequencies
  - Sum of frequencies add to the total number of respondents eligible to be asked the question
  - Ineligible respondents coded SPSS system missing
- Eligibility
  - Respondents answer to prior question
  - If question answered in prior wave
  - If respondent decline to participate
  - If respondent was not eligible for subsequent wave
- Frequencies not included for open-ended numerical codes
  - Dollar amounts, dates, etc.

Questionnaire Example

S572b. DATA CHECKPOINT

1. IF YEAR ORGANIZED START-UP **(Q119)** NOT ANSWERED --> **GO TO S573**

2. IF YEAR ORGANIZED START-UP **(Q119)** ANSWERED --> **GO TO S575**

S573. Has a start-up team been organized?

(A start-up team is more than one person that helps to put the firm in place, expecting to share ownership. If both married partners own and operate a business, that is a start-up team.)

1. YES 5. NO 8. DON'T KNOW

GO TO S575

S574. Will a start-up team be organized, or is it not relevant to this business?

1. TEAM WILL BE ORGANIZED 2. NOT RELEVANT TO THIS BUSINESS 8. DON'T KNOW

GO TO S577

Matching Codebook Entries

WAVE 1	WAVE 2	WAVE 3		
Q116	R573	S573	Has a start-up team been organized? (A start-up team is more than one person that helps to p firm in place, expecting to share ownership. If both mar partners own and operate a business, that is a start-up t	
470	33	21	1.	Yes
358	100	79	2.	No
1	0	1	8.	DK
1	0	0	9.	NA

WAVE 1	WAVE 2	WAVE 3		
Q117	R574	S574	Will a start-up team be organized, or is it not relevant to business? (Start-up teams are those active in management and operating the business. Owners that are not managers not part of the start-up team.)	
68	7	5	1.	Team will be organized
281	90	74	2.	Not relevant to this business
10	0	1	8.	DK
1	3	0	9.	NA