Quantitative Data Analysis II

SOC 781

SEM brief overview

Today we will…

- SEM brief overview

Why SEM?

- Enables more complex RQs
  - can handle multiple outcomes (i.e., multivariate modeling)
  - compensates for measurement error in variables

- Four common SEM techniques
  - path analysis
  - confirmatory factor analysis
  - latent variable structural model
  - growth curve

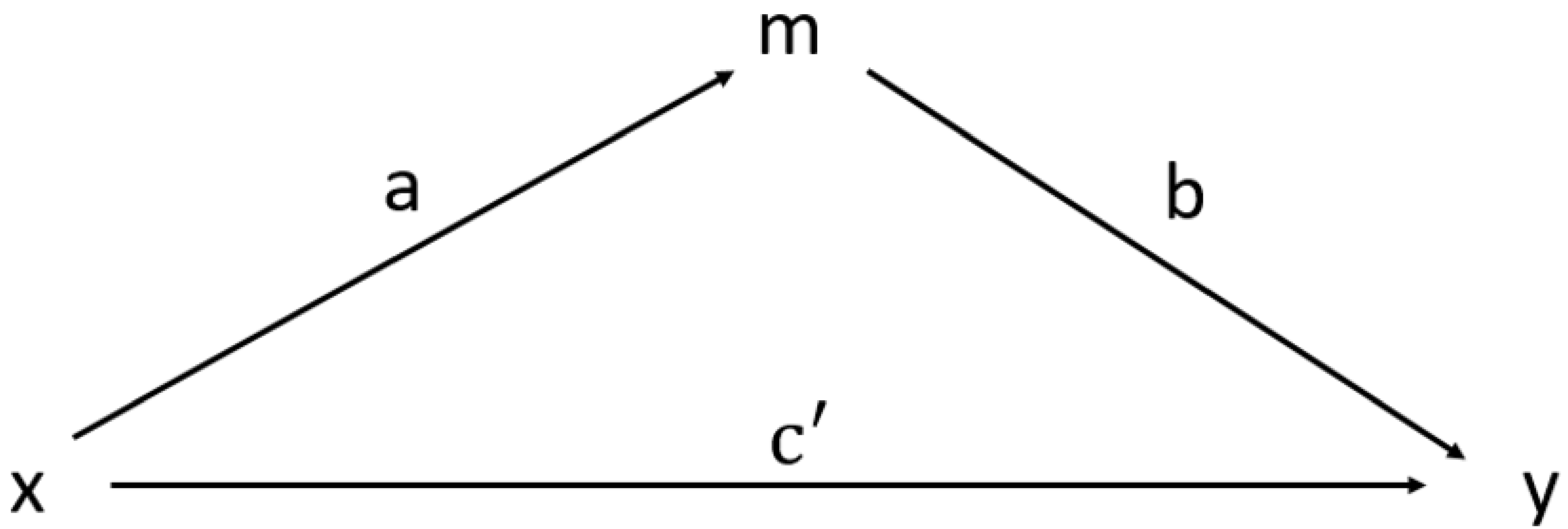
SEM basics

- Relies on covariance matrices
  - matrix algebra
- MANY assumptions
  - model specific
- MANY parameters
  - heavily technical

Path analysis

- Recall mediation techniques for LRM
  - Barron-Kenny approach
    - outdated and limited to continuous outcomes

- *m* is a mediator because it lies "in the path" from *x* to *y*



Path analysis

- Requires decomposition of total effect of *x* on *y*
  - direct and indirect effects through m
- Sobel's method assume normality of sampling distribution
  - typically, not met with samples < 1000
- Bigger problems with BK approach
  - countervailing mediators and confounders
    - sometimes dealt with using bootstrapping
- Path analysis can overcome these limitations
  - because SEMs…

Structural Equation Models

- Use observed variables, x and y, as measures of
  - latent variables ξ and η
- ξ are latent exogenous variables
- η are latent endogenous variables

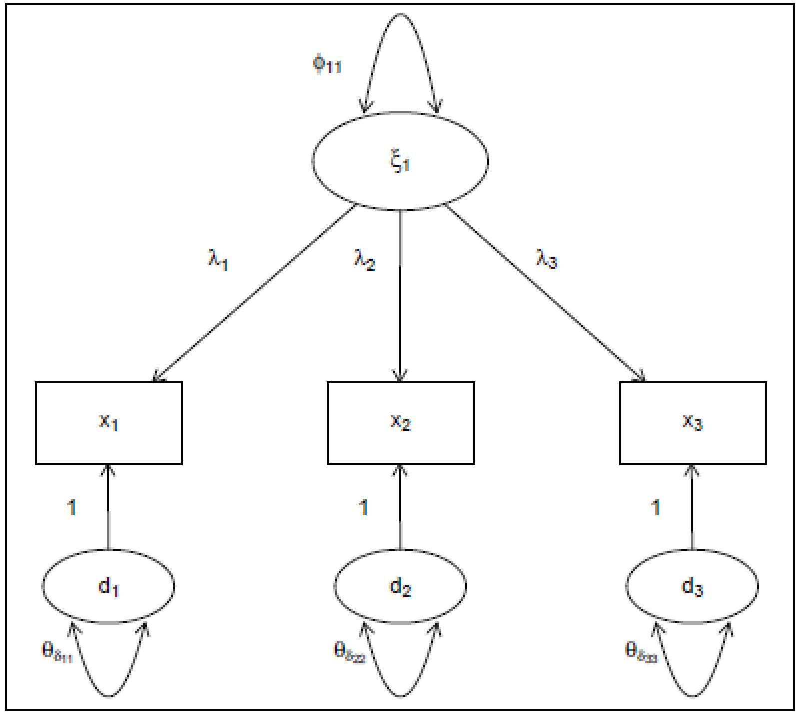
- Measurement models link *x* and *y* to ξ and η

- Theory used to “structure” un/constrained parameters
  - with their own sets of assumptions, but
    - less restrictive that univariate model or sequential univariate models
      - kind of like mlogit vs sequential blrm models

Measurement Models

- Not only useful for addressing error biases
  - also, for reliability: precise/consistent
  - and, for validity: do items measure what is intended?

- latent variable ("factor")
  - inter-item correlation between variables
    - shared variance of a set of variables



Factor Analysis

- Exploratory Factor Analysis (EFA)
  - (1) limiting number of factors to extract, (2) rotation method (orthogonal/oblique), (3) allowing items to load on limited number of factors
- Confirmatory Factor Analysis (CFA)
  - derived from theory
    - pre-specified and tested
- EFA-CFA
  - more of a continuum vs dichotomy

Factor Analysis

- Reliability (ideally $R^2 > 0.75$)
  - function of all arrows to *x*, not just from one factor

- Validity: only possible with multiple items
  - similarity/clusters in standardized loadings
    - distinct clusters suggest distinct factors

Single vs multigroup analysis

- Traditional regression limited to…
  - interaction terms or stratified models
    - limitations: group parameters or ignored shared error variance
- SEM allows
  - all parameters to be freely estimated within each group
  - enables significant difference test

Latent Class Analysis vs Growth Curve Modeling

- LCA and GCM both used to model patterns over time (longitudinal)
  - LCA also used for different purposes (e.g., latent profile analysis)
    - special case of SEM for categorical outcomes
- GCM and HLM approaches are equivalent
  - SEM is multivariate and more flexible but with costs
- LCA and GCM can be combined in same model
  - lots of different (hybridized and extended) approaches

CFA example: domain satisfaction

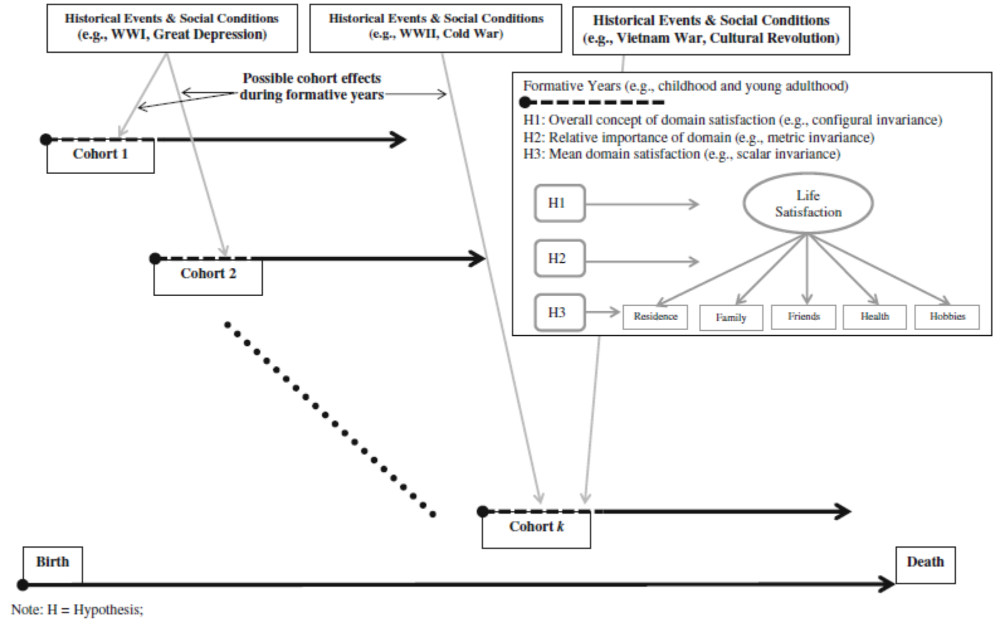


Fig. 1 Simplified life course model of possible cohort effects on domain satisfaction through life experiences during formative years

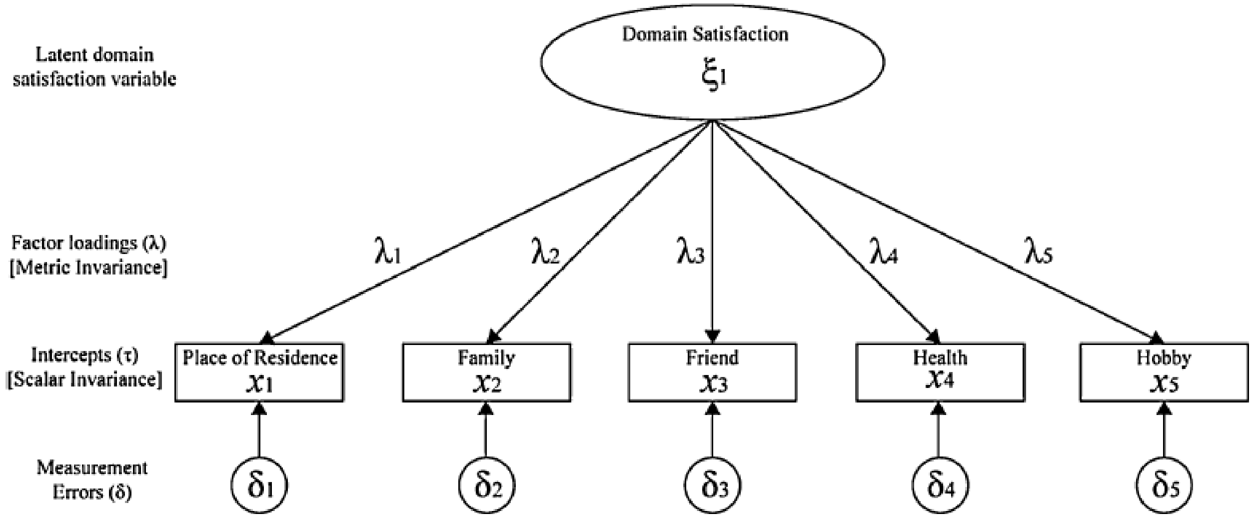
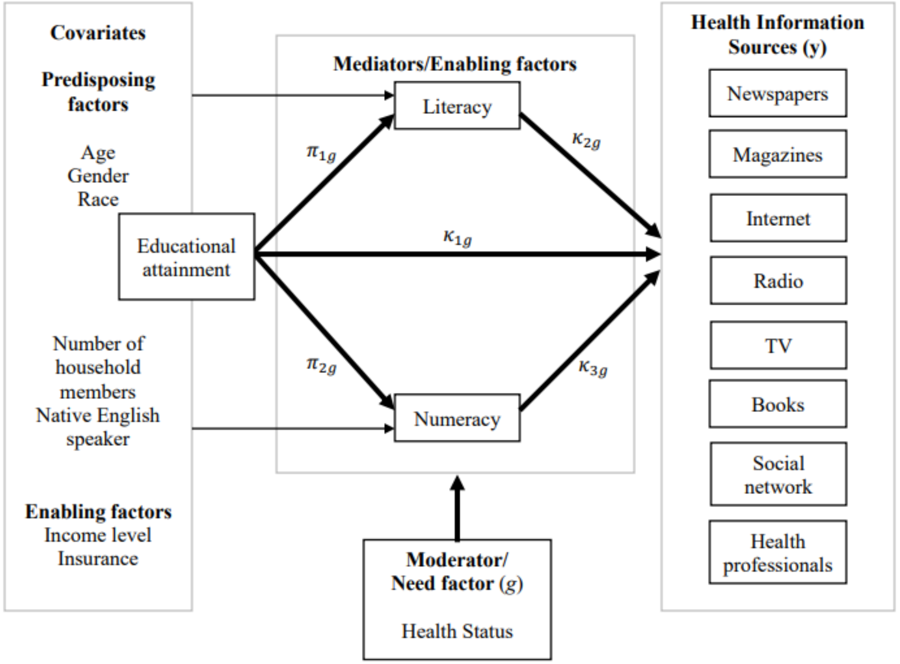


Fig. 2 One factor measurement model of domain satisfaction with five quality of life domain indicators

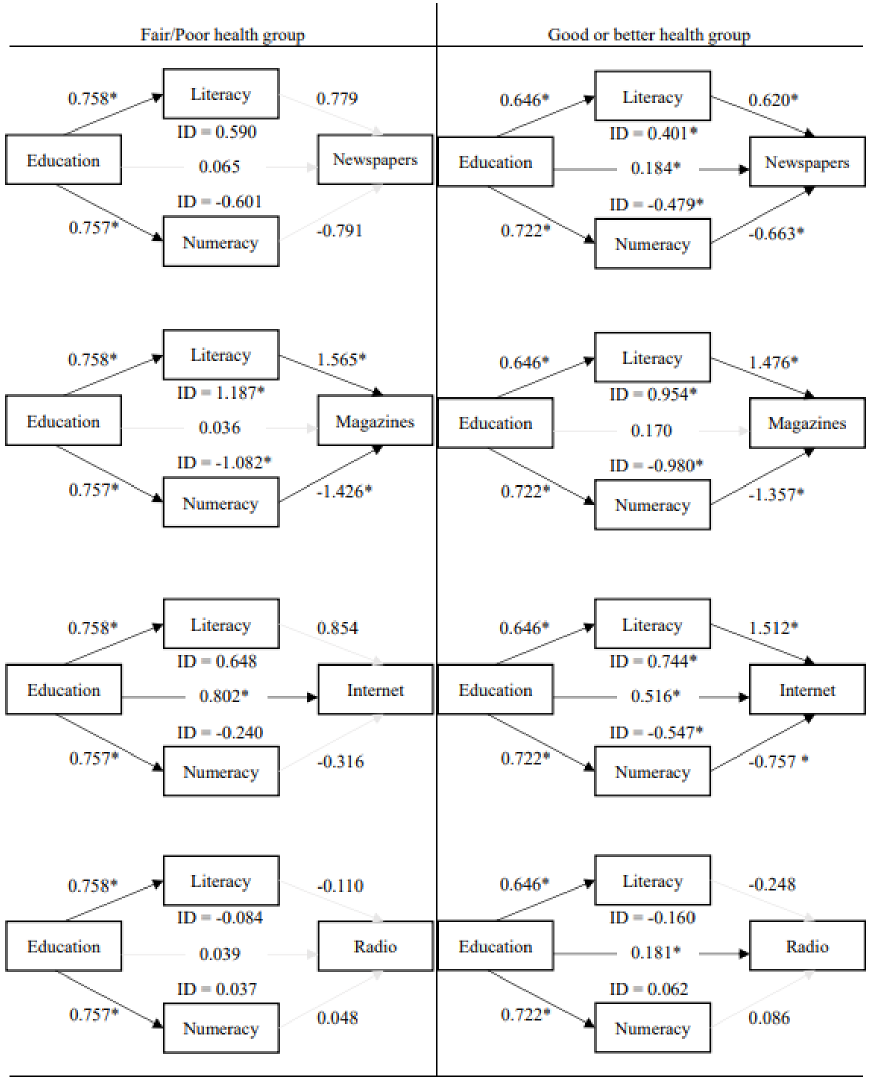
Multi-group Path Analyses: [moderated mediation]

Figure 2: Simplified Path Diagram of the Final Model



Note: Straight line = regression paths; bold lines indicate research questions; all variables in each box with gray lines were allowed to be correlated. For the notation (i.e., Greek letters), see the methods section; (g) groups indicator (fair/poor health vs. good or better health)

Table 3a: Estimated Coefficients from Path Models with a Probit Link Function



*$p < 0.05$; ID = indirect/mediation effect; solid line indicates statistical significance; gray line indicates non-significance; all models were adjusted for covariates, and correlations between literacy and numeracy

Growth Curve Analyses: latent trajectories

ELSEVIER

Social Science & Medicine 66 (2008) 849–861

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Trajectories of functional health: The 'long arm' of childhood health and socioeconomic factors☆

Steven Haas*

Arizona State University, School of Social and Family Dynamics, P.O. Box 873701, Tempe, AZ 85287-3701, USA

Available online 26 December 2007

Table 2

Model fit indices and curve parameters for unconditional latent growth curves of functional limitations under various functional forms (HRS 1994–2002)

	Model fit indices				Growth curve parameters					
					Intercept		Linear term[a]		Quadratic term	
	X^2 (df)	BIC	CFI	RMSEA	Mean	Variance	Mean	Variance	Mean	Variance
Linear model	464.89 (14)	331.67	0.99	0.06	1.87	5.00	0.07	0.04	–	–
Freely-estimated model	407.74 (11)	305.42	0.99	0.06	1.90	5.00	0.59	2.27	–	–
Quadratic model	114.62 (10)	21.60	1.00	0.03	1.92	4.83	0.02	0.14	0.01	0.00

Notes: BIC $= X^2 - (\ln(N) \times \text{df})$.

All models assume constant error variances.

[a] for freely estimated model this represents the total change over the period rather than a linear slope.

Latent Class Analyses: finite-mixture models

Journal of Health and Social Behavior
2016, Vol. 57(3) 407–422

DOI: 10.1177/0022146516660345
jhsb.sagepub.com
SAGE

Life Course Pathways of Economic Hardship and Mobility and Midlife Trajectories of Health

Andrea E. Willson[1] and Kim M. Shuey[1]

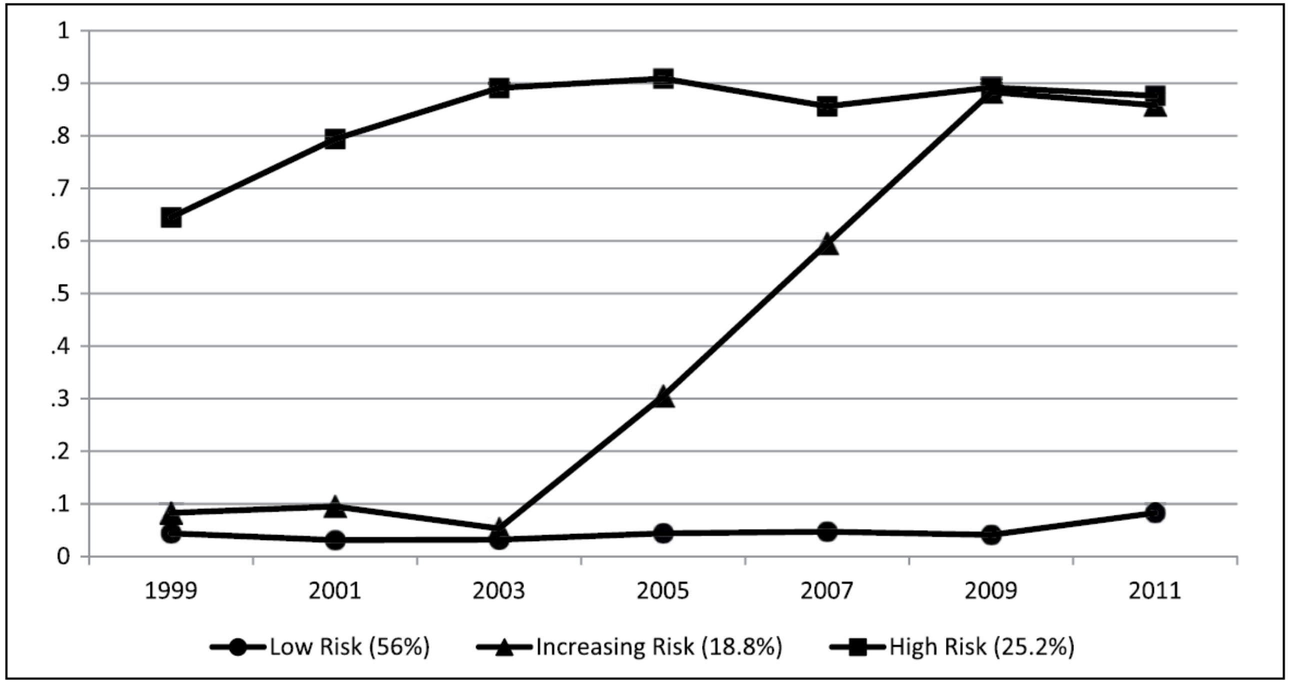


Figure 2. Item-response Probabilities for a Three-class Longitudinal Latent Class Model of Health Risk Trajectories, by Class and Year (Panel Study of Income Dynamics, 1999–2011).

Hybridized and Extended: [survival analyses]

The Gerontologist
Cite journal as: The Gerontologist Vol. 54, No. 2, 221–231
doi:10.1093/geront/gns164

Advance Access publication January 9, 2013

Body Mass Trajectories and Mortality Among Older Adults: A Joint Growth Mixture–Discrete-Time Survival Analysis

Anna Zajacova, PhD[1,*] and Jennifer Ailshire, PhD[2,3]

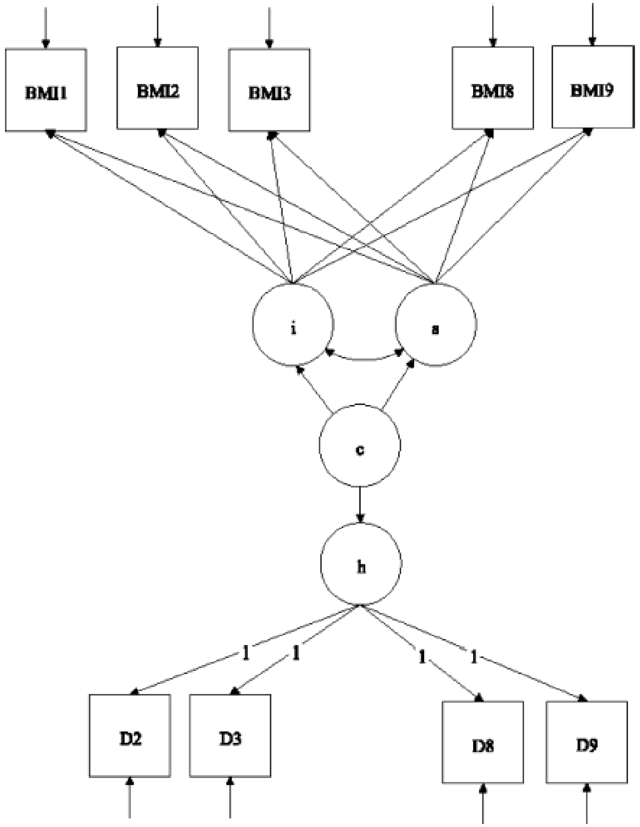


Figure 1. Schematic diagram of the joint growth mixture–discrete-time survival model. Note: BMI1–BMI9 are the observed BMI values; D2–D9 are vital status indicators, coded 0 if the respondent was alive through the end of a given time interval, 1 if the respondent died in the time interval, and missing if the respondent died previously or attrited. The means of the latent growth factors for the BMI trajectories, *i* and *s*, are allowed to vary across the trajectory classes, *c*. The mortality hazard *h* is modeled as a function of the class membership via a logistic regression model.

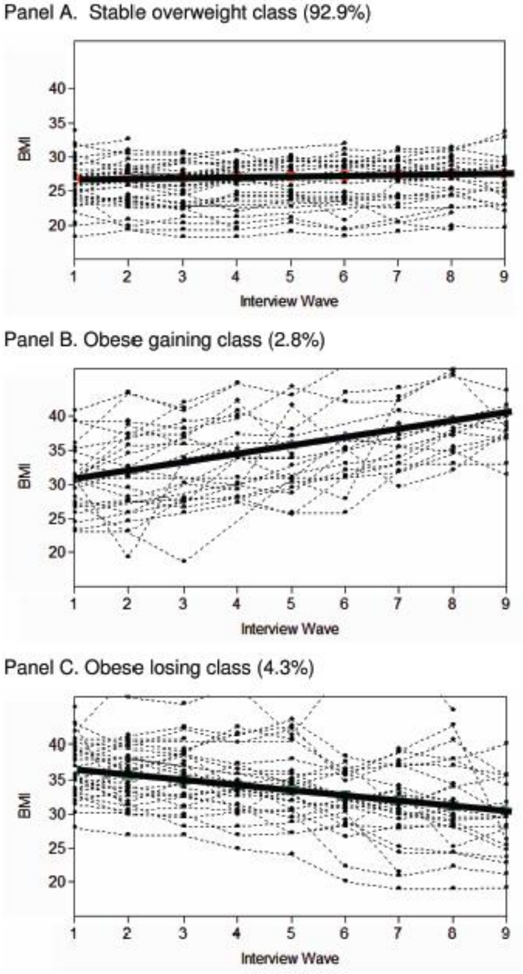
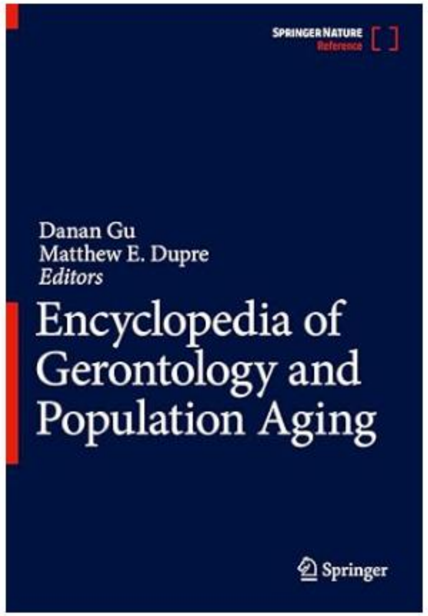


Figure 2. Three classes of BMI trajectories for men–estimated trajectory for each class and a random sample of observed individual trajectories. Panel A: Stable overweight class (92.9%); Panel B: Obese gaining class (2.8%); Panel C: Obese losing class (4.3%). Note: Results for women are visually nearly indistinguishable. The estimated sample trajectories and sample proportions for both genders are summarized in Figure 3.

Methods in Gerontology

Section Editors: *Anthony R. Bardo* and *Kenneth Carl Land*

Age-Period-Cohort Models
Big Data
Cross-Sectional and Longitudinal Studies
Dyad/Triad Studies
Ethnography
Experimental Studies and Observational Studies
False Negative/False Positive
Hierarchical Models
Item Response Theory and Modeling
Latent Class Analysis
Life Course Perspective
Likert Scale
Missing Data Concepts
Mobile Data Collection with Smartphones
Narrative Analysis
Qualitative Research/Quantitative Research
Recruitment and Retention in Aging Research
Repeated Cross-Sectional Design
Selective Bias in Longitudinal Studies
Semiparametric Methods
Structural Equation Models
Survival Analysis



- Focus on where your RQs rest
  - can't master all
    - but be broadly familiar
- Collaborate when needed
  - network outside expertise
- Published ≠ right
  - refer to solid sources and multiple examples
    - see reference lists in resources I provided
- SEM has huge scope
  - Ken Bollen and Shawn Bauldry for sociology
    - can get easily lost in psychometric literature