Growth Mixture Modeling and Causal Inference

Booil Jo
Stanford University
booil@stanford.edu

Conference on Advances in Longitudinal Methods in the Social and Behavioral Sciences
June 17-18, 2010
Center for Integrated Latent Variable Research
Department of Measurement, Statistics & Evaluation
University of Maryland

Outline

- Causal inference and principal stratification (PS)
- Growth mixture modeling (GMM)
- Feasibility of causal inference in the GMM framework
- Some simulation studies
- Real data example – Johns Hopkins University Preventive Intervention Research Center Study
- Discussion

Accounting for Heterogeneity in Outcome Development

- Heterogeneity in outcome development or prognosis has been of great interest both in observational and experimental studies.
- Subpopulation who would develop different outcome trajectories may differently benefit from the intervention/treatment.
- This kind of inference necessitates a special kind of posttreatment intermediate variable (mediator) – latent outcome trajectory class.

Accounting for Heterogeneity in Longitudinal Intervention Studies

- The idea of discrete subpopulations is particularly appealing in the context of studying intervention or treatment effects because these subpopulations may show different treatment responses.

Accounting for Heterogeneity in Longitudinal Intervention Studies

- Once we shift our interest from simply identifying heterogeneous trajectory classes to identifying differential treatment effects for these latent classes, whether and under what conditions we can interpret the results as causal becomes a critical issue.
- Defining these conditions has not been a key component in latent variable modeling.

Why Causal Modeling

- We want to attempt causal inference taking into account heterogeneity in mediators
  - Beyond the overall intention to treat (ITT) analysis
  - Beyond exploratory moderator/mediator analyses
  - Our desire for definitive answers
- More rigorous modeling practice
  - Better utilize the data at hand
  - Better guide future trials

Potential Outcomes Approach

(e.g., Neyman, 1923; Rubin, 1978, 1980…)

- The effect of treatment is defined based on an idealized (potential) situation, in which each individual's outcome is simultaneously observed under all compared conditions.
- $Y_i(1)$: potential outcome for individual i when assigned to the treatment condition ($Z = 1$)
- $Y_i(0)$: potential outcome for individual i when assigned to the control condition ($Z = 0$)
- $Y_i(1) – Y_i(0)$: the effect of treatment assignment for individual i

Potential Outcomes Approach

- The individual level treatment effect $Y_i(1) – Y_i(0)$ is interpreted as causal given that the only cause of the difference is the treatment assignment status.
- The causal effect of treatment assignment can be defined at the average (population) level: $\mu_1 – \mu_0$
- The individual level treatment effect $Y_i(1) – Y_i(0)$ generally cannot be identified
- However, the average causal effect $\mu_1 – \mu_0$ can be identified under certain conditions

Underlying Assumptions

- In practice, we do not observe μ_1 or μ_0 either, but can identify them under the following conditions
  - Ignorable treatment assignment
  - Stable unit treatment value (SUTVA)
- Under these assumptions, observed sample means $\bar{Y}_1$ and $\bar{Y}_0$ (or more generally estimates of μ_1 and μ_0 based on the sample) can be used to identify the average causal effect $\mu_1 - \mu_0$ (as $\bar{Y}_1 - \bar{Y}_0$ or $\hat{\mu}_1 - \hat{\mu}_0$)

Ignorable Treatment Assignment

(Holland, 1986; Rubin, 1974, 1978, 1980)

- Key assumption that opens up possibilities of causal inference at the average level based on observed data.
- Treatment assignment is independent of the potential outcomes (given the observed covariates) – automatically satisfied in randomized experiments
- Individuals assigned to different treatment conditions have similar pretreatment characteristics. We can treat mean outcomes of individuals assigned to different conditions as if they were obtained from the same individuals.

Stable Unit Treatment Value (SUTVA)

(Rubin, 1978, 1980, 1990)

- Another critical assumption that makes identification of causal treatment effects possible.
  - SUTVA I: potential outcomes for each person are unaffected by the treatment assignment of other individuals
  - SUTVA II: there is only one version of each treatment

Intention to Treat Analysis

- Ignorable treatment assignment and SUTVA are identifying assumptions necessary to interpret the overall mean difference between groups (ITT effect) as causal effect.
- Further identifying assumptions are necessary when making causal inference accounting for heterogeneity in mediators.

Principal Stratification

(Frangakis & Rubin, 2002)

- Principal stratification is one way of facilitating the potential outcomes approach.
- It aims for causal inference accounting for subpopulation heterogeneity in terms of mediators.
- It means stratifying individuals based on potential values of a mediator under all treatment conditions.
- As a result, the principal stratum membership is independent of treatment assignment just like pretreatment baseline covariates (more like a moderator than mediator).

Principal Stratification

- Given the independence between the principal stratum membership (C_i) and treatment assignment status (Z_i), it is possible to draw causal inference based on subgroup (or interaction) analyses.
- Individuals with the same principal stratum membership are comparable across treatment conditions (unlike observed mediators).
- In each principal stratum, outcome of interest can be compared across treatment conditions – principal effect. Any principal effect is a causal effect.

A Widely Known Example

(Angrist, Imbens, & Rubin, 1996)

- Treatment receipt (*S*) as an intermediate variable affected by treatment assignment
- Treatment noncompliance is a very common complication in randomized experiments involving human participants
- Showed a possibility of making causal inference considering heterogeneity in intermediate variables by utilizing the potential outcomes approach.
- Principal stratification can be thought of as generalization of their strategy.

Treatment Noncompliance

(Angrist, Imbens, & Rubin, 1996)

- Random assignment to either to the treatment (Z=1) or to the control condition (Z=0).
- Individuals either receive (S=1) or do not receive (S=0) the treatment.
- Let $S_i(1)$ denote the potential treatment receipt status for individual i when assigned to the treatment condition (Z=1) and $S_i(0)$ when assigned to the control condition (Z=0).

Treatment Noncompliance

(Angrist, Imbens, & Rubin, 1996)

- Individuals are classified into 4 principal strata based on their potential treatment receipt behavior given treatment assignment

$$C_i = \begin{cases} n \text{ (never-taker)} & \text{if } S_i(1) = 0, \text{ and } S_i(0) = 0 \\ d \text{ (defier)} & \text{if } S_i(1) = 0, \text{ and } S_i(0) = 1, \\ c \text{ (complier)} & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 0, \\ a \text{ (always-taker)} & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 1. \end{cases}$$

- However, from the observed data, these four latent classes (principal strata) cannot be separated.

Four Principal Strata

(Angrist, Imbens, & Rubin, 1996)

- Mean potential outcome values and average causal effects given principal strata (γ_c is the CACE)

Proportions of Principal Strata	Mean Potential Outcome $Z=1$	$Z=0$	Average Causal Effect Given C
π_n	μ_{n1}	μ_{n0}	$\gamma_n = \mu_{n1} - \mu_{n0}$
π_d	μ_{d1}	μ_{d0}	$\gamma_d = \mu_{d1} - \mu_{d0}$
π_c	μ_{c1}	μ_{c0}	$\gamma_c = \mu_{c1} - \mu_{c0}$
π_a	μ_{a1}	μ_{a0}	$\gamma_a = \mu_{a1} - \mu_{a0}$

Underlying Assumptions

(Angrist, Imbens, & Rubin, 1996)

- To identify the average causal effect for compliers, Angrist et al. employs two identifying assumptions (in addition to ignorable treatment assignment and SUTVA).
  - Exclusion restriction
  - Monotonicity
- This effect is widely known as complier average causal effect (CACE).

Underlying Assumptions

(Angrist, Imbens, & Rubin, 1996)

- *Exclusion Restriction*: for those whose intermediate outcome (*S*) value does not change in response to treatment assignment, the distributions of the potential outcomes are independent of the treatment assignment. It applies to never-takers and always-takers. As a result,
  - $\mu_{n1} = \mu_{n0}$
  - $\mu_{a1} = \mu_{a0}$
- *Monotonicity*: there are no defiers.
  - $\pi_d = 0$

Identification of CACE

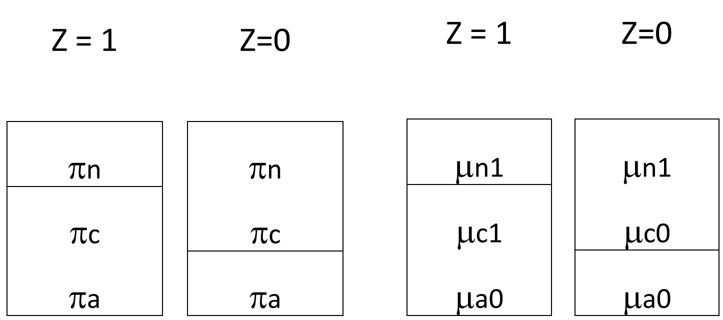
(Angrist, Imbens, & Rubin, 1996)

- Under monotonicity, $\pi_d = 0$. Then, π_c is derived as
  - $\pi_c = 1 - \pi_n - \pi_a$.
- Under exclusion restriction, $\mu_{n1} = \mu_{n0}$ and $\mu_{a1} = \mu_{a0}$
- Then, CACE is derived as

$$\gamma_c = \frac{\mu_1 - \mu_0}{\pi_c},$$

where all involved parameters are directly estimable from the observed data (e.g., $\bar{Y}_1$, $\bar{Y}_0$, and the estimated average complier probability using the sample proportions of never-takers and always-takers)

Identification of CACE



Growth Mixture Modeling (GMM)

- GMM utilizes the general latent variable modeling framework (Muthén, 2001; Muthén & Shedden, 1999)
  - Continuous latent variables capture growth trajectories (continuous heterogeneity), as in conventional mixed effects models.
  - Categorical latent variables capture subpopulation classes (discrete heterogeneity).
- GMM is an efficient way of identifying distributionally distinct latent trajectory classes on the basis of longitudinal outcome information.

Growth Mixture Modeling (GMM)

- GMM has been developed to 1) identify subpopulations that develop heterogeneous trajectory strata
- In the context of intervention/treatment studies, the method can be used to 2) identify heterogeneous intervention/treatment effects for these subpopulations.
- The second use of the method involves causal inference about the effect of the treatment.
- Very little is known about the possibility of serious causal inference in the GMM framework.

Similarities between Trajectory Strata and Principal Strata

- The two approaches share their common interest in heterogeneous subpopulations and share their discrete perspective in characterizing heterogeneity.
- Principal stratum membership is unaffected by treatment assignment, treatment effects conditioning on the principal stratum membership can be interpreted as causal effects.
- Conceptually, trajectory stratum membership is also unaffected by treatment assignment. Trajectory strata can be considered conceptual principal strata.

Differences between PS and GMM

- The two key elements in identifying causal effects in the PS framework are 1) pre-determined rules that classify individuals into principal strata and 2) partially observed stratum membership.
- In GMM, we do not have pre-determined rules (i.e., strata are decided based on empirical fitting).
- In GMM, we do not have partially observed stratum membership.
- In GMM, outcome itself is involved in the formulation of the strata. In PS, strata are formulated based on intermediate outcomes.

Causal Effect Identification in Principal Stratification

- Given partially observed stratum information and pre-determined classification rules, explicit identifying assumptions can be established to make up for missing information and to identify causal effects in principal stratification.
- Empirical model fitting is not necessary. However, the resulting causal effect estimation models may not conform with the data well.

Causal Effect Identification in GMM

- Causal effects are identified based on purely data-driven empirical model fitting.
- The resulting causal effect estimation models are likely to conform with the data well.
- Since we do not have theoretical strata, the quality of causal effect estimation relies on GMM's ability to recover the true principal trajectory strata.
- The problem is again that, we do not have theoretical strata, and therefore it is hard to evaluate the quality of causal effect estimation.

Causal Effect Estimation Based on Empirical Model Fitting

- Although GMM is conceptually a causal modeling, it is methodologically an exploratory modeling tool.
- GMM solutions and causal effects are decided mainly based on empirical model fitting.
  - Type I error?
  - Subjectivity?
  - Interpretability as in the principal stratification framework?
- Both underuse and abuse possibilities.
- Nonetheless, it can be a very useful tool if used thoughtfully.

Successful Causal Effect Identification in GMM

- True population trajectory strata are well recovered.
- True population trajectory strata are well coarsened or partitioned.
- No misleading results that are far from the truth – very large desirable or undesirable effects of the treatment while they do not really exist. In particular, we want to avoid Type I error.
- In the absence of theoretical strata, it is hard to evaluate the quality of causal effect estimation. Nonetheless, there are situations where GMM causal inference is more likely to be successful.

Simulation Study

- Random assignment to two conditions (treatment or control) shortly after T1 assessment
- Longitudinal outcome measured at 4 time points (Grades 1-4 or grades 6-9)
- Moderate distance between trajectory strata: maximum distance is **0.5-2.3** SD under both conditions
- SD of the outcome is about 1.0 SD at T1
- Total N = 500, 500 replications
- ML-EM using Mplus v6 for data generation and model estimation.

Previous Simulation Study

- Muthén & Brown (2009, Statistics in Medicine)
- Considerable distance between trajectory strata: minimum distance is **3-4** SD
- Baseline equality across randomization was utilized
- Two baseline (pre-randomization) outcome measures
- Eight post randomization outcome measures
- Total N = 100 to 500
- ML-EM using Mplus for data generation and model estimation.

Growth Mixture Model

- Assume a continuous outcome *Y* for individual *i* in the trajectory class *j* (j = 1,2,..., J) at time point *t*

$$\begin{aligned} y_{it} &= \eta_{Iij} + \eta_{Lij} W_t + \eta_{Qij} W_t^2 + \epsilon_{it}, && (1) \\ \eta_{Iij} &= \alpha_{Ij} + \lambda_I x_i + \gamma_{Ij} Z_i + \zeta_{Ii}, && (2) \\ \eta_{Lij} &= \alpha_{Lj} + \lambda_L x_i + \gamma_{Lj} Z_i + \zeta_{Li}. && (3) \end{aligned}$$

where it is assumed that $\epsilon_i \sim MN(0, \Sigma_\epsilon),\ \zeta_i \sim MN(0, \Sigma_\zeta)$. W_t is a time score at time point *t*. γ_{Ij} is the effect of treatment on the intercept and γ_{Lj} on the quadratic growth for the *j*th trajectory class.

Growth Mixture Model

- The probability π_i of belonging to a certain latent class ($C_i = j$) can be expressed in multinomial logit model as

$$logit(\pi_{0i}|x_i) = \beta_{00} + \beta_{10} x_i, \qquad (4)$$

which can vary depending on the influence of covariates *x*. There are *J* possible trajectory strata (j = 1,2,..., J). The probability π_i is a *J*-1 dimensional vector of (π_{i1}, π_{i2},..., $\pi_{i(J-1)}$), π_{ij} = Pr($C_i = j | x_i$), and logit(π_i) = (log[π_{i1}/ π_{ij}], log[π_{i2}/ π_{iJ}],..., log[$\pi_{i(J-1)}$/π_{iJ}]).

ML-EM Estimation

- To obtain the maximum likelihood estimates for the growth mixture model described in (1)-(4), we employed the EM algorithm implemented in the Mplus program (Muthén & Muthén, 1998-2010).
- The latent trajectory stratum membership is handled as missing data via the EM algorithm (for details, see Jo, Wang, Ialongo, 2009).

Basic Assumptions Revisited

- Ignorable treatment assignment: treatment assignment is independent of the potential outcomes and potential trajectory stratum membership. That is,

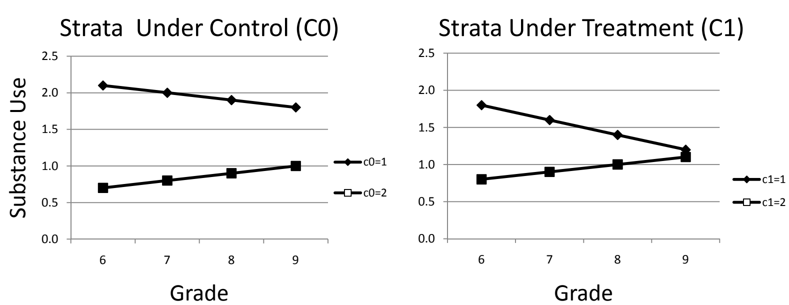
$$(Y_{it}(1), Y_{it}(0), C_i) \perp Z_i | X_i$$

  where $Y_{it}(1)$ is the potential outcome at time t when assigned to the treatment and $Y_{it}(0)$ when assigned to the control condition.
- Stable unit treatment value (SUTVA): we assume independence among subjects, but allow for the dependence across repeated measures within subjects.
- Further identifying assumptions need to be clarified!

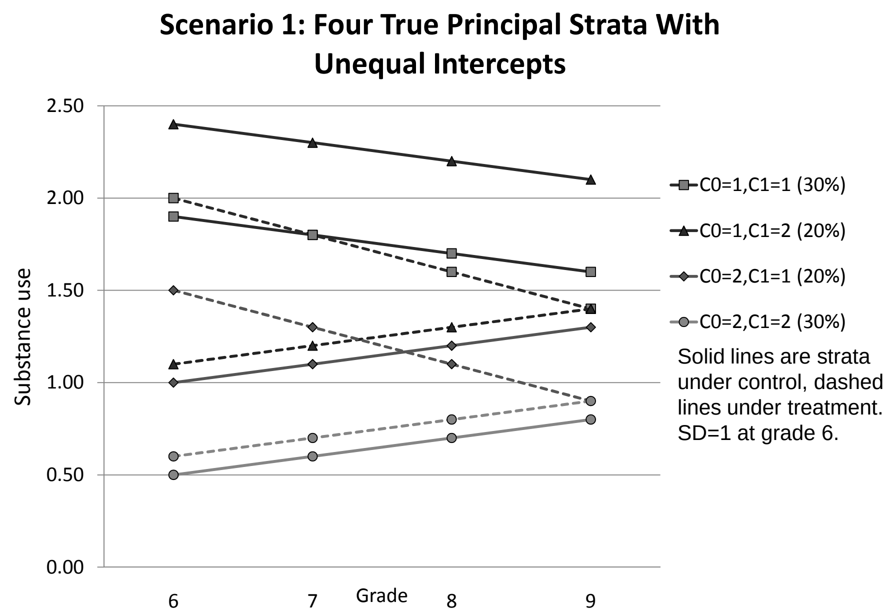
Scenario 1: Four principal strata with unequal intercepts

- Randomized intervention given during the 1st grade.
- Substance abuse is unlikely at Grade 1 (i.e., all 0)
- Substance abuse is measured at Grades 6-9
- We cannot assume the equality (even though randomized) of the outcome across treatment groups at Grade 6.

Scenario 1: Four True Principal Strata (2x2) With Unequal Intercepts



- There are 2 true trajectory strata under each condition.
- Correlation between C0 and C1 is 0.2
- 60% of C0=1 → C1=1, 40% of C0=1 → C1=2
- 40% of C0=2 → C1=1, 60% of C0=2 → C1=2



Scenario 1 simulation results: 4-class Model

(γ0, γ1 treatment effect on intercept & slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	37.5%	29.5%	37.3%
	γ0	***0.1***	0.154	0.173	0.197
	γ1	***-0.1***	-0.059	-0.038	-0.048
2	Class %	*20.0%*	11.9%	26.1%	12.4%
	γ0	***0.5***	0.236	-0.046	0.009
	γ1	***-0.3***	-0.382	-0.086	-0.022
3	Class %	*20.0%*	14.1%	20.8%	19.1%
	γ0	***-1.3***	-1.762	-0.402	-0.774
	γ1	***0.2***	0.233	-0.063	-0.065
4	Class %	*30.0%*	36.6%	23.6%	31.2%
	γ0	***0.1***	0.233	-0.226	-0.139
	γ1	***0.0***	-0.048	-0.034	-0.034

Scenario 1 simulation results: 3-class Model

(γ0, γ1 treatment effect on intercept & slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	45.9%	27.8%	33.2%
	$\gamma 0$	***0.1***	0.163	0.041	-0.179
	$\gamma 1$	***-0.1***	-0.092	-0.079	-0.031
2	Class %	*20.0%*	39.5%	27.0%	40.1%
	$\gamma 0$	***0.5***	0.247	0.228	0.244
	$\gamma 1$	***-0.3***	-0.099	-0.051	-0.076
3	Class %	*20.0%*	14.6%	45.2%	26.7%
	$\gamma 0$	***-1.3***	-1.756	-0.399	-0.433
	$\gamma 1$	***0.2***	0.178	-0.048	-0.048
4	Class %	*30.0%*			
	$\gamma 0$	***0.1***			
	$\gamma 1$	***0.0***			

Scenario 1 simulation results: 2-class Model

(γ0, γ1 treatment effect on intercept & slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*			
	$\gamma 0$	***0.1***			
	$\gamma 1$	***-0.1***			
2	Class %	*20.0%*	52.0%	34.4%	49.9%
	$\gamma 0$	***0.5***	1.027	0.352	0.272
	$\gamma 1$	***-0.3***	-0.191	-0.082	-0.073
3	Class %	*20.0%*	48.0%	65.6%	50.1%
	$\gamma 0$	***-1.3***	-1.330	-0.358	-0.437
	$\gamma 1$	***0.2***	0.093	-0.047	-0.051
4	Class %	*30.0%*			
	$\gamma 0$	***0.1***			
	$\gamma 1$	***0.0***			

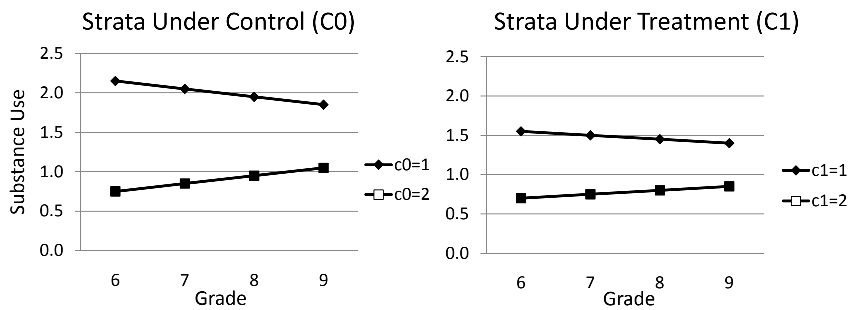
Scenario 1: Simulation Results Summary
(4 principal strata with unequal intercepts)

- The results are quite unstable across simulations using different sets of starting values.
- Both treatment effect estimates and mixing proportions can be quite off.
- In 3- and 4-class model, we may fail to detect the large treatment effect (ES=1.3 at T1) for one class.
- 2-class model may underestimate the large desirable effect on one class, or exaggerate the moderate harmful effect on another class.

Scenario 2: Two principal strata with unequal intercepts

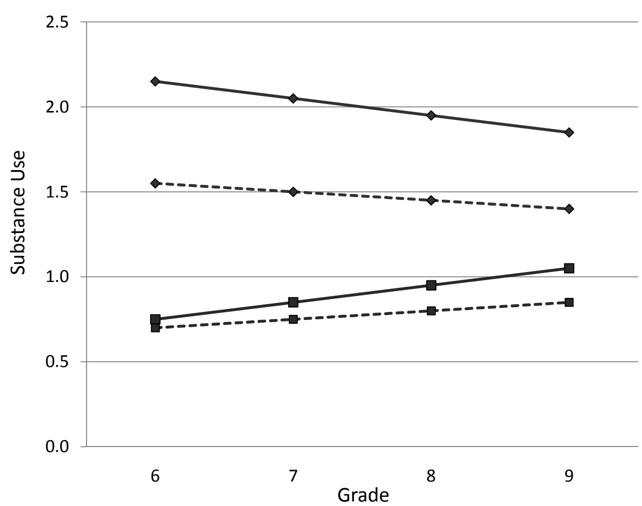
- Randomized intervention given during the 1st grade.
- Substance abuse is unlikely at Grade 1 (i.e., all 0)
- Substance abuse is measured at Grades 6-9
- We cannot assume the equality (based on randomization) of the outcome across treatment groups at Grade 6.

Scenario 2: Two True Principal Strata (2x2) With Unequal Intercepts



- There are 2 true trajectory strata under each condition.
- Correlation between C0 and C1 is **1.0**
- 100% of C0=1 → C1=1
- 100% of C0=2 → C1=2

Scenario 2: Two True Principal Strata (2x2) With Unequal Intercepts



Solid lines are strata under control, dashed lines under treatment. SD=1 at grade 6.

Scenario 2 simulation results: 3-class Model

($\gamma 0$, $\gamma 1$ treatment effect on intercept & slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	11.3%	37.9%	21.1%
	$\gamma 0$	***-0.60***	-1.912	-0.615	-0.637
	$\gamma 1$	***0.05***	0.230	0.070	0.063
2	Class %	*70.0%*	50.5%	32.0%	47.3%
	$\gamma 0$	***-0.05***	0.047	-0.048	0.037
	$\gamma 1$	***-0.05***	-0.085	-0.101	-0.077
3	Class %		38.2%	30.1%	31.6%
	$\gamma 0$		-0.008	0.120	-0.283
	$\gamma 1$		0.004	-0.070	0.020

Scenario 2 simulation results: 2-class Model

($\gamma 0$, $\gamma 1$ treatment effect on intercept & slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	38.6%	56.8%	38.3%
	$\gamma 0$	***-0.60***	-1.601	-0.601	-0.706
	$\gamma 1$	***0.05***	0.106	0.069	0.090
2	Class %	*70.0%*	61.4%	43.2%	61.7%
	$\gamma 0$	***-0.05***	0.549	0.207	0.067
	$\gamma 1$	***-0.05***	-0.109	-0.137	-0.083

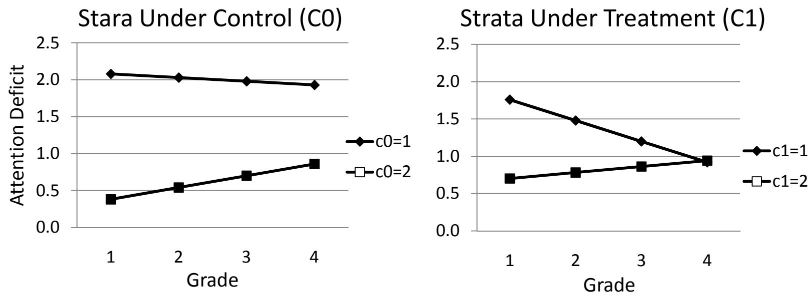
Scenario 2: Simulation Results Summary
(2 principal strata with unequal intercepts)

- The results are still quite unstable across simulations using different sets of starting values.
- Both treatment effect estimates and mixing proportions can be quite off.
- Type I error is likely (ES=1.9 or 1.6 instead of 0.6 at T1) for a desirable effect on one class.
- However, the results seem better. At least effect estimates are consistent and close to the true values in 2 out of 3 starting values.

Scenario 3: Four principal strata with equal intercepts

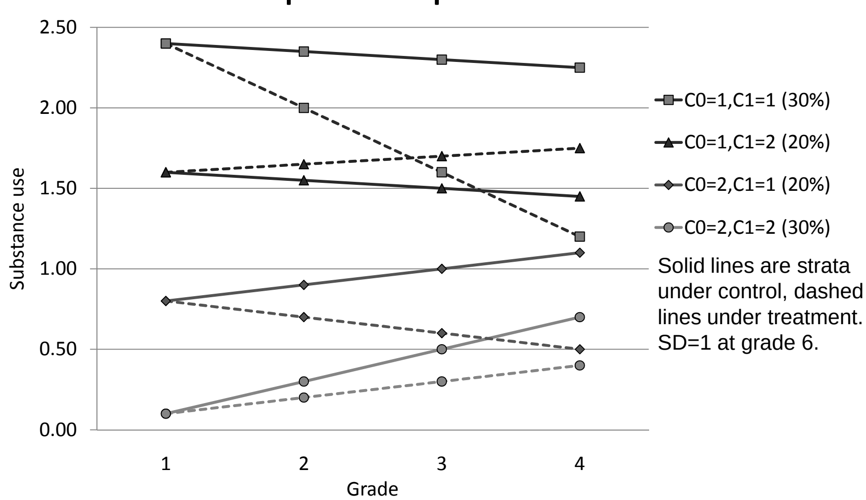
- Randomized intervention given during the 1st grade.
- Attention deficit measured at Grades 1-4.
- We can assume the baseline equality (based on randomization) of the outcome across treatment groups at Grade 1.

Scenario 3: Four True Principal Strata (2x2) With Equal Intercepts



- There are 2 true trajectory strata under each condition.
- Correlation between C0 and C1 is 0.2
- 60% of C0=1 → C1=1, 40% of C0=1 → C1=2
- 40% of C0=2 → C1=1, 60% of C0=2 → C1=2

Scenario 3: Four True Principal Strata With Equal Intercepts



Scenario 3 simulation results: 4-class Model

($\gamma 1$ treatment effect on slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	15.1%	19.0%	13.6%
	$\gamma 1$	***-0.35***	-0.314	-0.309	-0.300
2	Class %	*20.0%*	12.2%	24.4%	11.7%
	$\gamma 1$	***-0.20***	-0.403	-0.132	-0.298
3	Class %	*20.0%*	28.5%	21.6%	27.3%
	$\gamma 1$	***0.10***	-0.166	-0.177	-0.149
4	Class %	*30.0%*	44.2%	35.0%	47.4%
	$\gamma 1$	***-0.10***	-0.083	-0.120	-0.125

Scenario 3 simulation results: 3-class Model

($\gamma 1$ treatment effect on slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	15.9%	37.4%	21.1%
	$\gamma 1$	***-0.35***	-0.318	-0.294	-0.321
2	Class %	*20.0%*	31.4%	33.1%	30.3%
	$\gamma 1$	***-0.20***	-0.199	-0.116	-0.141
3	Class %	*20.0%*			
	$\gamma 1$	***0.10***			
4	Class %	*30.0%*	52.7%	29.6%	48.6%
	$\gamma 1$	***-0.10***	-0.106	-0.087	-0.103

Scenario 3 simulation results: 2-class Model

(γ1 treatment effect on slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	35.7%	50.7%	40.5%
	$\gamma 0$	***-0.35***	-0.291	-0.266	-0.291
2	Class %	*20.0%*	64.3%	49.3%	59.5%
	$\gamma 0$	***-0.20***	-0.102	-0.090	-0.091
3	Class %	*20.0%*			
	$\gamma 0$	***0.10***			
4	Class %	*30.0%*			
	$\gamma 0$	***-0.10***			

Scenario 3: Simulation Results Summary

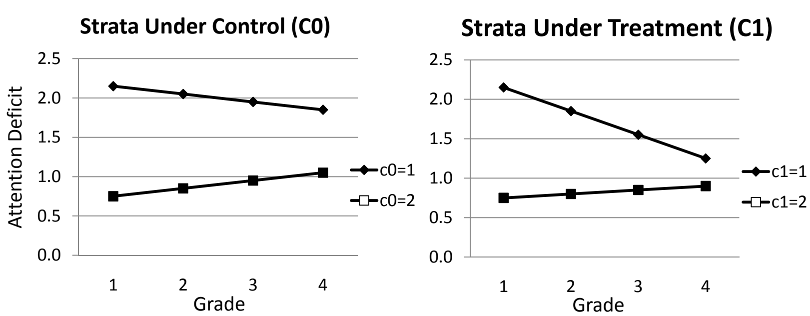
(4 principal strata with **equal** intercepts)

- Both treatment effect estimates and mixing proportions can be still somewhat off.
- The results are much more stable (esp effect estimates) when using different starting values.
- Severe Type I error is much less likely: no indication of extreme positive or negative effects.
- As the number of classes in the model decreases, treatment effects seem to get combined reasonably well without generating misleading estimates.

Scenario 4: Two principal strata with equal intercepts

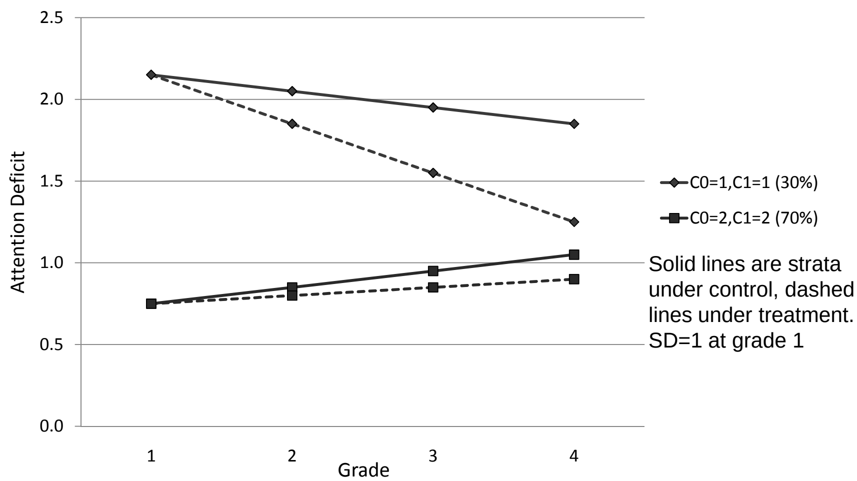
- Randomized intervention given during the 1st grade.
- Attention deficit measured at Grades 1-4.
- We can assume the equality (based on randomization) of the outcome across treatment groups at Grade 1.

Scenario 4: Two True Principal Strata (2x2) With Equal Intercepts



- There are 2 true trajectory strata under each condition.
- Correlation between C0 and C1 is **1.0**
- 100% of C0=1 → C1=1
- 100% of C0=2 → C1=2

Scenario 4: Two True Principal Strata With Equal Intercepts



Scenario 4 simulation results: 3-class Model

($\gamma 1$ treatment effect on slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	14.5%	34.6%	18.9%
	$\gamma 1$	***-0.20***	-0.227	-0.197	-0.220
2	Class %	*70.0%*	32.8%	36.4%	34.9%
	$\gamma 1$	***-0.05***	-0.151	-0.061	-0.112
3	Class %		52.8%	29.0%	46.3%
	$\gamma 1$		-0.032	-0.026	-0.041

Scenario 4 simulation results: <u>2-class</u> Model

($\gamma 1$ treatment effect on slope; SV starting value)

Strata	Parameter	True	SV1	SV2	SV3
1	Class %	*30.0%*	27.8%	60.2%	32.6%
	$\gamma 0$	***-0.20***	-0.223	-0.180	-0.225
2	Class %	*70.0%*	72.2%	39.8%	67.4%
	$\gamma 0$	***-0.05***	-0.057	0.002	-0.043

Scenario 4: Simulation Results Summary

(2 principal strata with **equal** intercepts)

- Effect estimates are close to true values and quite stable across models using different starting values.
- Mixing proportions can be still somewhat off.
- Severe Type I error is unlikely: no indication of extreme positive or negative effects.
- As the number of classes in the model increases, treatment effects seem to get partitioned reasonably well without generating misleading estimates.

All Simulation Summary

- More and intensive simulation studies are necessary to examine feasibility in more general situations
- Although GMM is conceptually a causal modeling, it is methodologically an exploratory modeling tool.
- It is important to use GMM for causal inference when it is likely to be successful.

More likely to be Successful Situations for Causal Effect Identification in GMM

- Trajectory strata are far apart from one another.
- A perfect or near perfect correlation between trajectory strata under control and under treatment.
- There is only one trajectory stratum either under the control or under the treatment condition.
- Baseline equality can be imposed (this is not always possible, e.g., drug abuse early intervention).
  - Large sample sizes. Normality holds.
  - There are pretreatment covariates that are good predictors of stratum membership.

More likely to be Successful Situations for Causal Effect Identification in GMM

- These conditions can be used as a basis for formulating plausible identifying assumptions that will support causal interpretation of GMM results.
- Once we have these assumptions, then we have latent variable causal modeling!
- The next step will be to further develop/refine identifying assumptions that are easier to work with in conducting sensitivity analysis.
- Flexible, rich modeling capability can be a big plus (Models contain assumptions, e.g., baseline equality)

An Alternative Approach

(exploratory + confirmatory)

- It is also possible to use an alternative 2-step approach where GMM and the standard causal modeling techniques are combined (Jo, Wang, Ialongo, 2009).
  - Identification of trajectory strata (step 1) and identification of causal effects (step 2) are separated.
  - Causal effect identification in step 2 is consistent with that in the PS approach.
- Conditions under which the 2-step approach is more likely to be successful need to be examined.
- These conditions are not necessarily the same for the standard GMM and the 2-step GMM approach.

JHU PIRC School Intervention Study

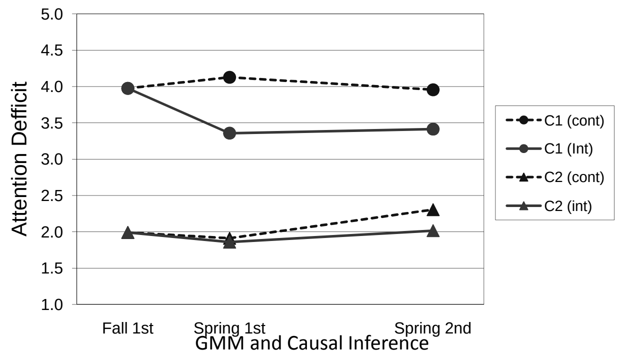
- Johns Hopkins Univ Preventive Intervention Research Center (PIRC) in 1993–1994 (Ialongo et al., 1999)
- Designed to improve academic achievement and to reduce early behavioral problems
- First-grade children were randomly assigned to the control or to intervention conditions
- Control vs. Family-School Partnership (FSP) intervention
- In the FSP condition, parents were asked to implement 66 take-home activities related to literacy and math.

JHU PIRC School Intervention Study

- Attention deficit (measured at baseline, 6, 18 months from the baseline) as the outcome. W_t = 0, 1, 3.
- A quadratic growth model.
- Based on random assignment, we assume no effect of treatment assignment on the initial status.
- MAR is assumed for missing data.
- As predictors of the growth parameters, we included baseline covariates such as parent's employment, marital status, ethnicity, education, and child's gender.

Causal Effect in JHU PIRC

- 2-class solution is preferred (BIC and experts opinion).
- The results suggest a sizable and significant effect of the intervention (SD of Y is about 1 across 3 time points) on kids at high risk (18.6%). For the majority of children, the intervention had little impact.



Causal Effect in JHU PIRC

- How seriously should we take these results?
- The two trajectory strata seem to be quite apart from each other (but not enough to guarantee a very good separation). Also, these are estimates.
- There are some covariates that may help separation.
- When GMM is conducted separately for the control and treatment groups, there are two strata under the control, but only one under the treatment, implying that two strata are enough to capture heterogeneity.
- We get very similar results using the 2-step approach used in Jo, Wang, Ialongo (2009).

Conclusions

- We can utilize latent variable modeling to improve model fit in causal models and even to identify causal effects purely based on empirical fitting.
- This development is still at a very early stage and thoughtful practice is necessary (and welcomed).
- Needs a lot more collaborative work
  - Feasibility
  - Conceptualization of causal inference in the latent variable modeling framework
  - Identification and estimation methods

References

- Jo, B. & Muthén, B. O. (2001). Modeling of intervention effects with noncompliance: A latent variable modeling approach for randomized trials. In G. A. Marcoulides and R. E. Schumacker (Eds.), Advanced Structural Equation Modeling: New Developments and Techniques (pp 57-87). Lawrence Erlbaum Associates.
- Jo, B. & Muthén, B. O. (2003). Longitudinal studies with intervention and noncompliance: Estimation of causal effects in growth mixture modeling. In N. Duan and S. Reise (Eds.), Multilevel Modeling: Methodological Advances, Issues, and Applications (pp 112-139). Lawrence Erlbaum Associates.
- Jo, B., Wang, C-P., & Ialongo, N. S. (2009). Using latent outcome trajectory classes in causal inference. Statistics and Its Interface, 2, 403-412.

References

- Muthén, B. O. (2001). & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics, 55, 463-469.
- Muthén, B. O. & Brown, C. H. (2009). Estimating drug effects in the presence of placebo response: Causal inference using growth mixture modeling. Statistics in Medicine, 23, 2937-2960.
- Muthén, B. O. et al. (2002). General growth mixture modeling for randomized preventive interventions. Biostatistics, 3, 459-475.
- Muthén, B. O. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (eds.), New Developments and Techniques in Structural Equation Modeling (pp. 1-33). Lawrence Erlbaum Associates.