Compliance Mixture Modelling with a Zero Effect Complier Class and Missing Data.

Michael E. Sobel *
Department of Statistics
Columbia University

Bengt Muthén
Professor Emeritus
University of California, Los Angeles

June 8, 2012

1

*Corresponding author:1255 Amsterdam Avenue, Columbia University, New York, NY 10027, e-mail:michael@stat.columbia.edu

Jobs-II data

A randomized study in which unemployed subjects were assigned to either a treatment group or a control group. The treatment consisted of 6 sessions in which job search skills were taught. At several points after the implementation of the intervention, subject's level of depression was assessed. We look at depression at six months. Our analyses of these data focus on the "so-called" high risk subjects, a group of subjects deemed to be a high risk for depression. For previous analyses of these data relevant to our own, see Little and Yau (1998), Yau and LIttle (2001), Jo (2002).

Notation and Setup for Studies in which Access to Treatment is Limited to the Treatment Group.

SUTVA assumption made throughout.

For each subject $i = 1, ...n$, we observe $Z_i = 0$ if i is assigned to the control group, 1 otherwise, and $\mathbf{X}_i$, a vector of pretreatment covariates. For $z = 0, 1$, let $Y_i(z)$ denote the value of subject i's outcome under treatment assignment z, let $R_i(z) = 1$ if i reports the outcome $Y_i(z)$ under treatment assignment z, 0 otherwise, and let $D_i(z) = 1$ if subject i takes up treatment under assignment z, 0 otherwise.

OBSERVED DATA: $\{(D_i, R_i, Z_i, \mathbf{X}_i, (Y_i : R_i = 1), i = 1, ..., n)\}$

The observations are assumed to be independent and identically distributed, given treatment assignment and the covariates $\mathbf{X}$..

In this study, as subjects in the control group cannot access the treatment, $D_i(0) = 0$ for all i.

"Types" of Indiviiduals:
1) NEVER TAKERS (D(0) = D(1) = 0)
2) DEFIERS (D(0) = 1, D(1) = 0)
3. COMPLIERS (D(0) = 0, D(1) = 1)
There can be no ALWAYS TAKERS (as D(0) = 0) .
WE assume there are no DEFIERS.

We also assume compliers are of two types: those where the intervention has an effect (effect class) and those where the inttervention does not have an effect (0 effect class). Of interest are the proportion of compliers in each of the classes and the average effect of the intervention in the effect class. If we assume that exclusion holds in the 0 effect class (reasonable in a double blind experiment) and that compliers in this class are taking the treatment (for example, attending a class and getting the knowledge, not just attending), this suggests that the proportion of compliers in the 0 effect class is the proportion of subjects for whom the mediator does not affect the outcome.

Subject type is indexed by U, where $D_i(1) = 0$ if and only if subject i is a never-taker ($U_i = 1$), and $U_i = 2$ for effect class compliers, $U_i = 3$ for compliers in the 0 effect class.

In addition to SUTVA, no ALWAYS TAKERS, no DEFIERS, we shall explicitly assume throughout:
A1 (RANDOM ASSIGNMENT)

$$Y(0), Y(1), R(0), R(1), X, U \perp\!\!\!\perp Z, \tag{1}$$

where the symbol $\perp\!\!\!\perp$ denotes statistical independence.

Compliance Mixture Modelling with no Missingness

Estimands of Interest

The primary estimands are the densities $f(y(z) \mid u, \mathbf{x}, \theta_{uz}^{UZ})$, depending on parameters θ_{uz}^{UZ}, z = 0,1, u= 1,2,3, and the class proportions $\Pr(U = u \mid \mathbf{X} = \mathbf{x}, \lambda) = \pi_u(\mathbf{x}, \lambda)$, depending on parameters λ. From these, the ECACE $\mathrm{E}(Y(1) - Y(0) \mid U = 2)$, and the class proportions $\pi(u) = \Pr(U = u)$, are obtained.

We use maximum likelihood to model the distribution of Y and D, given $\mathbf{X}$ and Z. For $d = 0,1$, $r = 0,1$, $z = 0,1$, let S_{drz}^{DRZ} denote the set of observations with D=d, R=r,Z=z, Using assumption A1, the likelihood is

$$L(\xi, \lambda) = \prod_{d,z} \prod_{i \in S_{d1z}^{DRZ}} f(y_i, d_i \mid z_i, \mathbf{x}_i, \theta_{dz}^{DZ}, \lambda) = \prod_{d,z} \prod_{i \in S_{d1z}^{DRZ}} f(y_i(z_i) \mid d_i, (z_i), \mathbf{x}_i, \theta_{dz}^{DZ}, \lambda) f(d_i(z_i) \mid \mathbf{x}_i, \lambda), \tag{2}$$

where $\theta_{dz}^{DZ} = \{\theta_{uz}^{UZ} : (U = u, Z = z) \Rightarrow (D = d, Z = z)\}$, and $\theta = (\theta_{10}'^{UZ}, ..., \theta_{31}'^{UZ})' = K\xi$, where ξ is the vector of distinct elements of θ.

To estimate ξ and λ, additional assumptions are necessary.

A2 (EXCLUSION RESTRICTION FOR NEVER TAKERS)

$$f(y(1) \mid U = 1, \mathbf{X} = \mathbf{x}, \theta_{11}^{UZ}) = f(y(0) \mid U = 1, \mathbf{X} = \mathbf{x}, \theta_{10}^{UZ}). \tag{3}$$

The remaining assumptions enable us to distinguish the two complier classes from one another and identify the relative proportions of the two complier types.

A3 (MIXTURE IDENTIFIABILITY)

The mixture densities are assumed to belong to a known parametric family. For $Z = 0, D = 0$, $Z = 1, D = 0$, and $Z = 1, D = 1$, the parameters

θ_{dz}^{DZ} of the component densities of the distributions $f(y \mid d, z, \mathbf{x}, \theta_{dz}^{DZ}, \lambda)$ are assumed to be identified.

EXAMPLES

1)Mixture of Normals—A3 holds.

2) Mixture of Bernoullis—A3 does not hold.

A4 (ZERO EFFECT COMPLIER CLASS ASSUMPTION)

For U = 1,2,3, $\pi_u(\mathbf{x}, \lambda) > 0$. Further,

$$f(y(1) \mid U = 3, \mathbf{X} = \mathbf{x}, \theta_{31}^{UZ}) = f(y(0) \mid U = 3, \mathbf{X} = \mathbf{x}, \theta_{30}^{UZ}). \quad (4)$$

Assumption A4 states that all three class types have non-zero probability at each $\mathbf{x}$. For subjects in the zero effect class, $Y_i(0) = Y_i(1)$, implying assumption (A4).

For the two class model, assumption A4 is not needed, and may be replaced by the assumption $\pi_u(\mathbf{x}, \lambda) > 0$ for $U = 1, 2$

First Pass Reanalysis of the JOBS-II Data

We compare our three class model with the model Little and Yau's (1998) obtained for 502 high risk subjects with complete data from the Jobs-2 study. The outcome is the difference between the depression score at 6 months and baseline, with negative scores indicating improvement.

First, IV analysis. The mean difference in outcomes between treatment and control subjects of -.075 estimates the intent to treat estimand (ITT). 55 percent of the subjects are estimated to be compliers, which leads to an instrumental variable estimate of -.136 for the CACE, similar to the mean difference between compliers and never takers in the treatment group.

Table 1: Estimates of ECACE, class membership and parameters of compliance models for 502 high risk subjects.

	Model M_{02}		Model M_{13}			
BIC and NP	1545.89	14		1541.49	26	
ECACE	-.310	(.130)		-.389	(.187)	
$\hat{\pi}_1$	.458			.460		
$\hat{\pi}_2$	.542			.236		
$\hat{\pi}_3$				.304		
VARIABLE	λ		$\boldsymbol{\lambda_2}$		$\boldsymbol{\lambda_3}$	
Age	.079	(.015)	.080	(.018)	.077	(.018)
Education	.300	(.068)	.205	(.091)	.363	(.081)
Marital Status	.540	(.283)	.371	(.345)	.664	(.335)
Economic Hardship	.159	(.152)	.253	(.205)	-.433	(.196)
Race	-.499	(.317)	-.666	(.476)	-.423	(.360)
Motivation	.667	(.157)	.657	(.210)	.652	(.188)
Assertiveness	-376	(.143)	-.540	(.201)	-.138	(.180)

Model (M_{02})—the Little-Yau model

Model (M_{13})—our preferred 3 class model

BIC $= -2\ln L + NP \ln n$

Compliance Mixture Modelling with Missing Data

In addition to the densities $f(y(z) \mid u, \mathbf{x}, \theta_{uz}^{UZ})$ for $u = 1,2,3,\ z = 0,1$, and the class proportions $\pi_u(\mathbf{x}, \lambda)$, we now model the distribution of $R(z)$ given X and U: for $z = 0,1$, r = 0,1, $\Pr(R(z) = r \mid u, \mathbf{x}, \psi_{uz}^{UZ})$. Let $\psi = (\psi_{10}'^{UZ}, ..., \psi_{31}'^{UZ})' = L\eta$, where η is the vector of distinct elements of ψ.

Missing Data Assumptions

ASSUMPTION 5 (MISSING AT RANDOM)

$$Y \perp\!\!\!\perp R \mid Z, \mathbf{X}, U \tag{5}$$

for U observed,

$$Y, U \perp\!\!\!\perp R \mid Z, \mathbf{X} \tag{6}$$

for U unobserved.

2 CLASS MODEL— U observed when $Z = 1$, but unobserved when Z = 0.
3 CLASS MODEL—- U observed only when $Z = 1, D = 0$.
Under MAR, the response process is "ignorable" if the parameters η governing the missing data mechanism are distinct from (ξ, λ).

2 CLASS MODEL—- Frangakis and Rubin (1999)—-MAR unattractive—missingness depends on compliance and covariates in the treatment group, but only covariates in the control group.
A6 (EXTENDED LATENT IGNORABILITY)

$$Y \perp\!\!\!\perp R \mid U, Z, \mathbf{X}. \tag{7}$$

Frangakis and Rubin—NEVER TAKER RESPONSE EXCLUSION RESTRICTION: $R_i(0) = R_i(1)$ if $U_i = 1$.

NOTE: Assumption A6 + never taker response exclusion restriction is neither weaker nor stronger than the MAR assumption A5.

Mealli et al.—-A6 + COMPLIER RESPONSE EXCLUSION RESTRICTION: $R_i(0) = R_i(1)$ if $U_i = 2$. As above, this combination of assumptions is neither stronger nor weaker than the MAR assumption. Similarly to us, Mealli et al. (2004) use a parametric mixture model; however, in their analysis, the outcome Y is binary.

IMPORTANT POINTS: 1) For TWO CLASS MODEL With NORMAL MIXTURES, RESPONSE EXCLUSION RESTRICTIONS ARE NOT NEEDED.

2) For THREE CLASS MODELS, alternative identifying assumptions are needed.
If RESPONSE does not depend on the effectiveness of the intervention, the TWO COMPLIER TYPES might behave in the same fashion (assumption A7). But if RESPONSE is positively related to the effectiveness of the intervention, 0 EFFECT COMPLIERS might behave like NEVER TAKERS (assumption A8).

A7 (ZERO EFFECT/EFFECT CLASS HOMOGENEITY)
For u = 2,3 and z = 0,1,

$$\Pr(R(z) = 1 \mid U = 2, \mathbf{X} = \mathbf{x}) = \Pr(R(z) = 1 \mid U = 3, \mathbf{X} = \mathbf{x}). \qquad (8)$$

“

NOTE: A6 + A7 is weaker than MAR.

A8 (NEVER TAKER/ZERO EFFECT CLASS HOMOGENEITY)

For u = 1,3 and z = 0,1,

$$\Pr(R(z) = 1 \mid U = 1, \mathbf{X} = \mathbf{x}) = \Pr(R(z) = 1 \mid U = 3, \mathbf{X} = \mathbf{x}). \qquad (9)$$

NOTE; A6 + A8 is weaker than MAR.

Second Pass Reanalysis of the JOBS-II Data

657 high risk subjects with complete covariate data

Outcome—depression at 6 months

Table 2: Estimates of ECACE and class membership for some models with different missing data assumptions for 657 high risk subjects.

Model	BIC	NP	ECACE		Class Proportions		
M_{2MAR}: 2 class MAR	2468.69	34	-.382	(.120)	.471	.529	
M_{2LI}: 2 class LI	2474.67	35	-.404	(.139)	.471	.529	
M_{3MAR}: 3 class MAR	2461.66	45	-.240	(.107)	.462	.340	.198
M_{3ELI7}: 3 class ELI + A7	2468.11	46	-.239	(.106)	.462	.340	.198
M_{3ELI8}: 3 class ELI + A8	2467.00	46	-.653	(.393)	.472	.304	. 224

Conclude—3 CLASS MAR best, but really BIC doesn't indicate strong support for any particular model.

Using the information from the various models suggests the intervention lowers depression at six months between about .03 points (using the lower bound from the 95 % confidence interval for ECACE from M_{3MAR}) to about 1.44 points (using the upper bound from the 95 % confidence interval for ECACE from M_{3ELI8}.