



Analysis of treatment response data from eligibility designs

Siddhartha Chib^{a,*}, Liana Jacobi^b

^a Olin Business School, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA

^b Department of Economics, University of Melbourne, Victoria 3010, Australia

ARTICLE INFO

Article history:

Received 15 June 2007

Received in revised form

5 March 2008

Accepted 14 April 2008

Available online 20 May 2008

Keywords:

Confounding

Eligibility designs

Heterogeneity

Instrumental variable

Marginal likelihood

Markov chain Monte Carlo

Non-compliance

Non-randomly assigned treatment

Potential outcomes

Treatment effect

Partial compliance

Principal stratification

ABSTRACT

In this paper, we develop and compare two alternative approaches for calculating the effect of the actual intake when treatments are randomized, but compliance with the assignment in the treatment arm is less than perfect for reasons that are correlated with the outcome. The approaches are based on different identification assumptions about these unobserved confounders. In the first approach, which stems from [Sommer, A., Zeger, S., 1991. On estimating efficacy in clinical trials. *Statistics in Medicine* 10, 45–52], the unobserved confounders are modeled by a discrete indicator variable that represents subject-type, defined in terms of the potential intake in the face of each possible assignment. In the second approach, confounding is modeled without reference to subject-type in the spirit of the Roy model. Because the two models are non-nested, and model comparison and assessment of the approaches in a real data setting is one of our central goals, we formulate the discussion from a Bayesian perspective, comparing the two models in terms of marginal likelihoods and Bayes factors, and in terms of inferences about the treatment effects. The latter we calculate from a predictive perspective in a way that is different from that in the literature, where typically only a point summary of that effect is calculated. Our real data analysis focuses on the JOBS II eligibility trial that was implemented to test the effectiveness of a job search seminar in decreasing the negative mental health effects commonly associated with job loss. We provide a comparative analysis of the data from the two approaches with prior distributions that are both reasonable in the context of the data and comparable across the model specifications. We show that the approaches can lead to different evaluations of the treatment.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Consider an experiment where subjects are randomized into either a treatment arm or control arm but intake in the treatment arm is not necessarily the same as the assignment. Such an experiment, which we refer to as an eligibility trial, often arises in the testing of the effectiveness of new interventions, for instance a medical treatment or a social program. In this paper, we develop and compare two alternative approaches for calculating the effect of the actual intake from such a trial when the lack of compliance in the treatment arm is potentially caused by unobserved factors that also affect the outcome. The approaches are based on different identification assumptions about these unobserved confounders. Because the two approaches are non-nested, we formulate the discussion from a Bayesian perspective, comparing the two models in terms of marginal likelihoods and Bayes factors, and in terms of inferences about the treatment effects. The latter we calculate from a predictive perspective in a way that is different from that in the

literature where typically only a point summary of that effect is calculated.

The specific problem may be described formally as follows. We have a random sample of n individuals. For the i th individual in the sample, the random treatment assignment is indicated by $z_i = l$ ($l = 0, 1$), where $l = 0$ indicates assignment to the control arm and $l = 1$ assignment to the treatment arm. The treatment intake is $x_i \in \{0, 1\}$ which is necessarily the same as z_i only if $z_i = 0$. In the treatment arm, however, the observed intake is not necessarily the same as the assignment. In other words,

$$x_i = \begin{cases} 0 & \text{if } z_i = 0 \\ 0 \text{ or } 1 & \text{if } z_i = 1. \end{cases} \quad (1.1)$$

The observed outcome is y_i which we assume is continuous. The main concern is that the intake and assignment are not the same in the treatment arm for reasons that are potentially related to the outcome. Observed confounders, that is, covariates that simultaneously affect the outcome and the intake in the treatment arm, are denoted by $\mathbf{w}_i : p \times 1$ and the objective of the analysis is to isolate the effect of the treatment intake on the outcome, given the sample data $\mathbf{y} = (y_1, \dots, y_n)$, $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{z} = (z_1, \dots, z_n)$, and the observed confounders $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$.

* Corresponding author. Tel.: +1 314 935 4657.

E-mail addresses: chib@wustl.edu (S. Chib), ljacobi@unimelb.edu.au (L. Jacobi).

It is well known that with hidden confounders the determination of the treatment effect requires auxiliary, untestable assumptions (Pearl, 2000). One set of assumptions has been extensively explored in the recent literature (for example, Sommer and Zeger (1991), Imbens and Rubin (1997), Frangakis and Rubin (1999), Hirano et al. (2000), Jo (2002), Yau and Little (2001), Ten have et al. (2003), Barnard et al. (2003), Frangakis et al. (2004), Levy et al. (2004), Mealli et al. (2004)). In econometrics, the literature on the LATE effect (for example, Imbens and Angrist (1994)) parallels these developments. In this approach, the unobserved confounders are modeled as a discrete variable that represents subject type. In particular, each subject is viewed as having two possible potential intakes, x_{il} , where x_{il} is the potential intake under the assignment $z_i = l$. Only one of these potential intakes is realized depending on the observed assignment. In other words,

$$x_i = x_{i0}(1 - z_i) + x_{i1}z_i.$$

Now a subject is one of four types depending on the values of both potential intakes: complier (if $x_{i0} = 0$ and $x_{i1} = 1$), never-taker (if $x_{i0} = 0$ and $x_{i1} = 0$), defier (if $x_{i0} = 1$ and $x_{i1} = 0$) and always-taker (if $x_{i0} = 1$ and $x_{i1} = 1$). Under certain untestable assumptions, most importantly, the absence of always takers (because these cannot be identified in the eligibility set-up where subjects in the control arm have no possibility of getting the training), the absence of defiers (the monotonicity assumption), and that assignment is a proper instrumental variable, it becomes possible to find the causal effect of the actual intake on the outcome for the strata of compliers.

In many important problems it may be helpful to contrast the subject type approach with an approach that does not rely on the notion of a discrete confounder variable. Such an approach, which we term the general confounder approach, stems from the so-called Roy model. The Roy approach has been common in the non-eligibility setting, for example, Heckman and Honore (1990), Heckman and Vytlacil (2005) and, from the Bayesian perspective, by amongst others, Chib and Hamilton (2000, 2002), Deb et al. (2006), Chib (2007) and Chib and Jacobi (2007). The confounder in this case is a continuous latent variable (see for example Heckman et al. (2001)). In this paper we show how this assumption can be used to formulate a model for analyzing eligibility data. The model is conceptually simple and easily fit.

An important point to emphasize is that our modeling and inferential procedures are based on Chib (2007) and, therefore, do not require the unknowable joint distribution of the potential outcomes. As a result of this, in the type model, our analysis is simpler in terms of the required prior inputs and computational intensity than the approaches of Hirano et al. (2000), Barnard et al. (2003) and others. For example, the latter paper assumes a zero correlation between the potential outcomes, which is equivalent to assuming a prior distribution with all its mass concentrated on zero for the correlation parameter. We need to make no such assumption. Another point relates to our approach for fitting the general confounder model. Although the overall approach is drawn from that of Chib (2007) the details are different because intake in the control arm is non-stochastic, whereas in the model of Chib (2007) this was not the case. Finally, we note that within each framework, we compute a range of treatment effects, including the average and quantile treatment effects, that are based on the marginal predictive distributions of the potential outcomes.

Because one of our objectives is to the formal comparison of the models in terms of marginal likelihoods and Bayes factors, the prior distribution on the parameters in the different models in our empirical analysis is constructed in a reasoned way from another sample of subjects that was exposed to the same experiment. In addition, the prior distributions are formulated to reflect the assumption of a neutral treatment effect and the assumption of a

positive treatment effect. We explain our method for arriving at these prior distributions below. The empirical analysis shows that the two alternative approaches can lead to different evaluations of the treatment.

The remainder of the paper is organized as follows. In Section 2 we consider the type confounder approach and develop the framework for estimation and calculation of the treatment effect, and in Section 3 we provide the corresponding analysis from the general confounder assumption. The question of model comparisons via Bayes factors and marginal likelihoods is taken up in Section 4. Section 5 presents results on the performance of the inferential methods in simulation studies while Section 6 provides a detailed analysis of data from the 1991 JOBS II Intervention Project, a job search seminar for the recently unemployed that was designed to prevent the deterioration in mental health associated with job loss. We focus on finding the effect of the job-training on subsequent depression scores under each approach. We conclude the paper with remarks in Section 7.

2. Type confounder approach

2.1. Model formulation

We begin our discussion with a formulation of the subject-type approach in fully Bayesian terms without involvement of the joint distribution of potential outcomes and from the new, simplifying perspective, that subject-type is a discrete unobserved confounder. Following Hirano et al. (2000), Jo (2002) and Frangakis et al. (2004), the distribution of the confounder is modeled with covariates.

Under the notations introduced in the previous section, we start by modeling the joint density of each potential outcome y_{ij} and intake given the treatment assignment and the covariates

$$p(y_i, x_i = 0 | \mathbf{w}_i, z_i = l) \equiv p(y_{i0}, x_i = 0 | \mathbf{w}_i, z_i = l)$$

$$p(y_i, x_i = 1 | \mathbf{w}_i, z_i = l) \equiv p(y_{i1}, x_i = 1 | \mathbf{w}_i, z_i = l)$$

subject to the restriction on intake given in (1.1). The notation emphasizes the point that y_i (the observed outcome for the i th observation) is a place holder for the appropriate potential outcome y_{ji} .

The idea of formulating the problem in terms of these joint densities is due to Chib (2007) where it is pointed out that the joint distribution of the potential outcomes is not necessary for inferences about the causal treatment effect in settings with unobserved confounders, and also not particularly useful in the Bayesian context. By proceeding without the joint distribution of the potential outcomes no prior assumptions about unidentified correlation parameter(s) are needed in the model formulation and no unobserved outcomes have to be simulated in the model fitting, simplifying the computation of such models.

Now let $s_i = k$ indicate a discrete confounder random variable that affects both the intake x_i and the potential outcome y_{ji} . Suppose that this discrete confounder takes two possible values $\{0, 1\}$ that represent subject type, namely $\{n, c\}$, for never-taker and complier, respectively. Then, conditional on $z_i = l$ and $s_i = k$, the joint density of (y_i, x_i) (equivalently (y_{ji}, x_i)) can be factored as

$$p(y_i, x_i = j | \mathbf{w}_i, z_i = l, s_i = k) \\ \equiv p(y_{ji}, x_i = j | \mathbf{w}_i, z_i = l, s_i = k) \quad (2.1)$$

$$= p_j(y_i | \mathbf{w}_i, s_i = k) \Pr(x_i = j | y_{ji}, \mathbf{w}_i, z_i = l, s_i = k) \quad (2.2)$$

where $p_j(y_i | \mathbf{w}_i, s_i = k)$ is the density of y_{ji} conditional on the latent subject type and the second term is the conditional mass function of $x_i = j$. The former density does not involve $z_i = l$ on account of the so-called exclusion restriction. Notice, too, that the second term is either zero or one, for any value of y_i or \mathbf{w}_i . This can be seen from Table 1 which gives the distribution of type by

Table 1
Distribution of types by treatment arm and intake

	$x_i = 0$	$x_i = 1$
$z_i = 0$	c, n	–
$z_i = 1$	n	c

treatment arm and intake. For example, if $z_i = 0$ and $s_i = 1$, then $x_i = 0$, so that $\Pr(x_i = 0|y_i, \mathbf{w}_i, z_i = 0, s_i = 1) = 1$. In addition, if $z_i = 1$ and $s_i = 0$, then $x_i = 0$, implying that $\Pr(x_i = 0|y_i, \mathbf{w}_i, z_i = 1, s_i = 0) = 1$. Thus, given $z_i = l$ and $s_i = k$, the intake is fully determined.

To derive the joint density of the outcome and the intake, let $I_{lj} = \{i : z_i = l \text{ and } x_i = j\}$ denote the sample indices of the subjects in each of the three non-empty cells of Table 1. Also let

$$\Pr(s_i = 1|\mathbf{v}_i, \boldsymbol{\alpha}) = q_{ci} = \phi(\mathbf{v}'_i \boldsymbol{\alpha})$$

denote the probability of subject type c , where ϕ is the cdf of the standard normal distribution. We assume that this probability is a function of the $q \times 1$ vector of pre-treatment variables \mathbf{v}_i that is a subset of \mathbf{w}_i . It is independent of z_i because of the random assignment of subjects to the treatment arms. Since we do not observe the subject type in the control arm, the joint density of y_i and $x_i = j$ conditional on $z_i = l$, is given by appropriately averaging over possible types:

$$p(y_i, x_i = j|\mathbf{w}_i, z_i = l) = \begin{cases} (1 - q_{ci}) p_0(y_i|\mathbf{w}_i, s_i = 0) + q_{ci} p_0(y_i|\mathbf{w}_i, s_i = 1) & \text{if } i \in I_{00} \\ (1 - q_{ci}) p_0(y_i|\mathbf{w}_i, s_i = 0) & \text{if } i \in I_{10} \\ q_{ci} p_1(y_i|\mathbf{w}_i, s_i = 1) & \text{if } i \in I_{11}. \end{cases} \quad (2.3)$$

This expression does not involve the mass function of the intake due to the discussion surrounding (2.1).

Thus, from (2.3) we see that the modeling of $(y_i, x_i = j)$ requires three type and treatment state specific distributions, $p_0(y_i|\mathbf{w}_i, s_i = 0)$, and $p_j(y_i|\mathbf{w}_i, s_i = 1)$, for $j = 0, 1$. For specificity, assume that these are student- t with known degrees of freedom ν and of the form

$$p_0(y_i|\mathbf{w}_i, s_i = 0) = t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0n}, \eta_{0n}^2) \\ p_j(y_i|\mathbf{w}_i, s_i = 1) = t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{jc}, \eta_{jc}^2), \quad j = 0, 1 \quad (2.4)$$

where $\boldsymbol{\beta}_{0n}$ and $\boldsymbol{\beta}_{jc}$ are type and treatment state-specific p -dimensional vectors of regression parameters, η_{0n}^2 and η_{jc}^2 are the corresponding dispersion parameters and $t_\nu(\cdot|\mu, s^2)$ is the student- t density function with ν degrees of freedom ($\nu > 2$), mean μ and variance $\nu s^2/(\nu - 2)$. The student- t assumption is a generalization of the common Gaussian assumption in the literature.

We complete the model specification with a prior distribution for the model parameters $\boldsymbol{\psi} = (\boldsymbol{\beta}, \boldsymbol{\eta}^2, \boldsymbol{\alpha})$ where $\boldsymbol{\beta} = (\boldsymbol{\beta}_{0c}, \boldsymbol{\beta}_{0n}, \boldsymbol{\beta}_{1c})$ and $\boldsymbol{\eta}^2 = (\eta_{0c}^2, \eta_{0n}^2, \eta_{1c}^2)$. We assume that the parameters are mutually independent and that the prior density is of the form

$$\pi(\boldsymbol{\psi}) = \pi(\boldsymbol{\alpha}) \prod_{j=0,1} \prod_{k \in \mathcal{K}_j} \pi(\boldsymbol{\psi}_{jk})$$

where $\pi(\boldsymbol{\alpha}) = \mathcal{N}_q(\boldsymbol{\alpha}|\boldsymbol{\alpha}_0, A_0)$, $\boldsymbol{\psi}_{jk} = (\boldsymbol{\beta}_{jk}, \eta_{jk}^2)$, $\mathcal{K}_0 = \{c, n\}$, $\mathcal{K}_1 = \{c\}$,

$$\pi(\boldsymbol{\psi}_{jk}) = \mathcal{N}_p(\boldsymbol{\beta}_{jk}|\boldsymbol{\beta}_{jk,0}, \mathbf{B}_{jk,0}) \mathcal{I}\mathcal{G}\left(\eta_{jk}^2 \left| \frac{n_{jk,0}}{2}, \frac{d_{jk,0}}{2} \right. \right),$$

and $\mathcal{I}\mathcal{G}(\cdot|n, d)$ is the inverse-gamma density.

2.2. Posterior distribution and computations

We now discuss how the model we have specified can be fit. We rely on Markov chain Monte Carlo (MCMC) methods to sample from the posterior distribution. Because our formulation is different than that in the literature, the MCMC fitting method is also different. Let $\boldsymbol{\psi} = (\boldsymbol{\beta}, \boldsymbol{\eta}^2, \boldsymbol{\alpha})$ denote the model parameters and $\pi(\boldsymbol{\psi}|\mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{W}) \propto f(\mathbf{y}, \mathbf{x}|\mathbf{z}, \mathbf{W}, \boldsymbol{\psi})\pi(\boldsymbol{\psi})$, the posterior distribution which is proportional to the product of the likelihood $f(\mathbf{y}, \mathbf{x}|\mathbf{z}, \mathbf{W})$ of the observed data and the prior distribution $\pi(\boldsymbol{\psi})$, where the likelihood function is

$$f(\mathbf{y}, \mathbf{x}|\mathbf{z}, \mathbf{W}, \boldsymbol{\psi}) = \prod_{i=1}^n p(y_i, x_i = j|\mathbf{w}_i, z_i = l) \\ = \prod_{i \in I_{00}} (1 - q_{ci}) t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0n}, \eta_{0n}^2) + q_{ci} t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0c}, \eta_{0c}^2) \\ \times \prod_{i \in I_{10}} (1 - q_{ci}) t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0n}, \eta_{0n}^2) \prod_{i \in I_{11}} q_{ci} t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{1c}, \eta_{1c}^2). \quad (2.5)$$

The posterior distribution is not tractable because of the mixture distribution in the control arm. MCMC computations can be simplified, however, by including $\mathbf{s}_{00} = \{s_i : i \in I_{00}\}$ as unknown parameters. In addition, it is helpful to rewrite the student- t density $t_\nu(\cdot|\mu, s^2)$ in familiar form as $\mathcal{N}(\cdot|\mu, \lambda_i^{-1} s^2)$, where $\lambda_i \sim \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$, and then augment the parameter space to include $\boldsymbol{\lambda} = \{\lambda_i : i \leq n\}$.

Under these assumptions and augmentations, the posterior density of interest becomes $\pi(\boldsymbol{\psi}, \boldsymbol{\lambda}, \mathbf{s}_{00}|\mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{W})$ which is proportional to

$$\pi(\boldsymbol{\psi}) \prod_{i=1}^n \mathcal{G}\left(\lambda_i \left| \frac{\nu}{2}, \frac{\nu}{2} \right. \right) \prod_{i \in I_{00}} \mathcal{N}(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0s_i}, \lambda_i^{-1} \eta_{0s_i}^2) p(s_i|\boldsymbol{\alpha}) \\ \times \prod_{i \in I_{10}} I[s_i = 0] (1 - \phi(\mathbf{v}'_i \boldsymbol{\alpha})) \mathcal{N}(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0n}, \lambda_i^{-1} \eta_{0n}^2) \\ \times \prod_{i \in I_{11}} I[s_i = 1] \phi(\mathbf{v}'_i \boldsymbol{\alpha}) \mathcal{N}(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{1c}, \lambda_i^{-1} \eta_{1c}^2) \quad (2.6)$$

where $p(s_i|\boldsymbol{\alpha}) = I[s_i = 0][1 - \phi(\mathbf{v}'_i \boldsymbol{\alpha})] + I[s_i = 1][\phi(\mathbf{v}'_i \boldsymbol{\alpha})]$ is the probability mass function of s_i . It is easy to see that the prior times the likelihood function in (2.5) emerges from this augmented distribution if one integrates out $\{s_i\}$ and $\{\lambda_i\}$ term by term.

MCMC sampling of this distribution can be implemented in a completely natural way. One possible algorithm (that we have extensively tested to be efficient) proceeds in the following fashion.

- Sample \mathbf{s}_{00} and $\boldsymbol{\lambda}$ jointly conditioned on the data $(\mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{W})$ and the remaining unknowns. This is done by sampling s_i for subjects in I_{00} (marginalized over λ_i) with probability

$$\Pr(s_i = 1|y_i, x_i, \boldsymbol{\alpha}, \boldsymbol{\beta}_{0c}, \boldsymbol{\beta}_{0n}, \eta_{0c}^2, \eta_{0n}^2) \\ = \frac{q_{ci} t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0c}, \eta_{0c}^2)}{q_{ci} t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0c}, \eta_{0c}^2) + (1 - q_{ci}) t_\nu(y_i|\mathbf{w}'_i \boldsymbol{\beta}_{0n}, \eta_{0n}^2)}. \quad (2.7)$$

Once s_i has been so generated for subjects in I_{00} , we know the type of all n subjects in the sample (since subjects in I_{10} are of type $s_i = 0$ and those in I_{11} are of type $s_i = 1$). We can now sample each λ_i from the Gamma density

$$\pi(\lambda_i|y_i, x_i, \mathbf{w}_i, s_i, \boldsymbol{\beta}, \boldsymbol{\eta}^2) \\ = \mathcal{G}\left(\lambda_i \left| \frac{\nu + 1}{2}, \frac{\nu + (y_i - \mathbf{w}'_i \boldsymbol{\beta}_{jk}) \eta_{j,k}^{-2} (y_i - \mathbf{w}'_i \boldsymbol{\beta}_{jk})}{2} \right. \right).$$

- In the second and third steps, conditional on the subject types $\mathbf{s} = \{s_i : i \leq n\}$ and the sub-samples of individuals $i \in I_{jk}$, we sample the regression parameters β_{jk} and the variance parameters η_{jk}^2 . Specifically, let

$$\mathbf{y}_{jk} = \{y_i : i \in I_{jk}\}; \quad \mathbf{x}_{jk} = \{x_i : i \in I_{jk}\}$$

$$\mathbf{W}_{jk} = \{\mathbf{w}_i : i \in I_{jk}\}; \quad \lambda_{jk} = \{\lambda_i : i \in I_{jk}\}$$

denote the sub-sample of observations, by intake state j and type k , on the outcome, intake, covariates and latent scale, respectively. Then, our sampling of β_{jk} is from

$$\pi(\beta_{jk} | \mathbf{y}_{jk}, \mathbf{x}_{jk}, \mathbf{W}_{jk}, \eta_{jk}^2, \lambda_{jk})$$

$$= \mathcal{N} \left(\beta_{jk} \mid \mathbf{B}_{jk} \left\{ \mathbf{B}_{jk,0}^{-1} \beta_{jk,0} + \sum_{i \in I_{jk}} \lambda_i \mathbf{w}_i \eta_{jk}^{-2} y_i \right\}, \right.$$

$$\left. \left\{ \mathbf{B}_{jk,0}^{-1} + \sum_{i \in I_{jk}} \lambda_i \mathbf{w}_i \eta_{jk}^{-2} \mathbf{w}_i' \right\}^{-1} \right) \quad (2.8)$$

and that of η_{jk}^2 from

$$\pi(\eta_{jk}^2 | \mathbf{y}_{jk}, \mathbf{x}_{jk}, \mathbf{W}_{jk}, \beta_{jk}, \lambda_{jk})$$

$$= \mathcal{I}\mathcal{G} \left(\eta_{jk}^2 \mid \frac{n_{jk,0} + n_{jk}}{2}, \frac{d_{jk,0} + \sum_{i \in I_{jk}} \lambda_i (y_i - \mathbf{w}_i' \beta_{jk})^2}{2} \right) \quad (2.9)$$

where n_{jk} denotes the number of individuals in the set I_{jk} .

- We finish our MCMC algorithm by sampling α from $\pi(\alpha | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{s}_{00})$ via the Metropolis Hastings algorithm (Chib and Greenberg, 1994, 1995) as the distribution is not available in closed form. The proposal value α^+ is drawn from a tailored multivariate- t distribution $t_{20}(\mu, \mathbf{V})$ where μ is the approximate mode of

$$\ln \left[\prod_{i \in I_{0c} \cup I_{1c}} \phi(s_i | \mathbf{v}_i' \alpha) \prod_{i \in I_{0n}} 1 - \phi(s_i | \mathbf{v}_i' \alpha) \right]$$

and \mathbf{V} is the inverse Hessian of the density evaluated at μ . The proposal value α^+ is then accepted with probability

$$\alpha_{MH} = \min\{1, r_1 r_2\}$$

where

$$r_1 = \frac{\pi(\alpha^+) \prod_{i \in I_{0c} \cup I_{1c}} \Phi(s_i | \mathbf{v}_i' \alpha^+) \prod_{i \in I_{0n}} (1 - \Phi(s_i | \mathbf{v}_i' \alpha^+))}{\pi(\alpha) \prod_{i \in I_{0c} \cup I_{1c}} \Phi(s_i | \mathbf{v}_i' \alpha) \prod_{i \in I_{0n}} (1 - \Phi(s_i | \mathbf{v}_i' \alpha))}$$

and

$$r_2 = \frac{t(\alpha | \mu, \mathbf{V}, 20)}{t(\alpha^+ | \mu, \mathbf{V}, 20)}.$$

2.3. Inferring treatment effects for compliers

The treatment effects analysis investigates whether the actual intake of the training program has a positive causal effect on outcomes for compliers. Chib (2007) has shown that the Bayesian predictive approach is useful in drawing inferences about causal treatment effects. In this section we show how these predictive distributions can be calculated. We use these distributions to compute various treatment effects, such as quantile treatment effects, and a predictive version of the common complier-average causal effect. We call this predictive effect the predictive average causal effect (PACE).

We begin our predictive analysis by considering a subject that is randomly drawn from the subpopulation of compliers. Let $y_{jc,n+1}$ denote the potential outcome for subject $n + 1$ who is a complier. The relevant predictive densities $p(y_{jc,n+1} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}, s_{n+1} = 1)$ are given by

$$\int p(y_{jc,n+1} | \mathbf{w}_{n+1}, \beta_{jc}, \eta_{jc}^2) I(s_{n+1} = 1) p(s_{n+1} | \mathbf{v}_{n+1}, \alpha)$$

$$\times \pi(\beta_{jc}, \eta_{jc}^2, \alpha, \mathbf{w}_{n+1}, \mathbf{v}_{n+1} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z})$$

$$\times d\beta_{jc} d\eta_{jc}^2 d\alpha d\mathbf{w}_{n+1} d\mathbf{v}_{n+1} \quad (2.10)$$

where $p(y_{jc,n+1} | \mathbf{w}_{n+1}, \beta_{jc}, \eta_{jc}^2)$ is $t_v(y_{jc,n+1} | \mathbf{w}_{n+1}' \beta_{jc}, \eta_{jc}^2)$ and the unknowns are marginalized with respect to the posterior distribution.

Since the integral cannot be calculated analytically, we use the method of composition to generate draws from $p(y_{jc,n+1} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}, s_{n+1} = 1)$. This is done by appending the following steps at the end of each MCMC iteration.

- First, we randomly sample $\mathbf{w}_{n+1}^{(g)}$ and $\mathbf{v}_{n+1}^{(g)}$ from the full set of covariates.
- Next, we sample $s_{n+1}^{(g)} = I[\mathbf{v}_{n+1}^{(g)} \alpha^{(g)} + u_{n+1}^{(g)} > 0]$, where $u_{n+1}^{(g)} \sim \mathcal{N}(0, 1)$.
- We then check compliance: If $s_{n+1}^{(g)} = 1$ we draw $y_{jc,n+1}^{(g)} (j = 0, 1)$ from

$$t_v(y_{jc,n+1} | \mathbf{w}_{n+1}^{(g)} \beta_{jc}^{(g)}, \eta_{jc}^{2(g)}).$$

Otherwise we skip and move to the next step in the chain.

The resulting draws $(y_{0c,n+1}^{(1)}, \dots, y_{0c,n+1}^{(G)})$ and $(y_{1c,n+1}^{(1)}, \dots, y_{1c,n+1}^{(G)})$ are stored and used to compute the treatment effects of interest. For example,

$$\text{PACE} = \mathbb{E}(y_{1c,n+1} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}, s_{n+1} = 1)$$

$$- \mathbb{E}(y_{0c,n+1} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}, s_{n+1} = 1) \quad (2.11)$$

is computed from the simulated draws in an obvious manner.

3. General confounder approach

3.1. Modeling

Consider now an alternative way of modeling unobserved confounding in which the hidden confounder, instead of being discrete with a two point distribution, is assumed to be non-specific with a distribution of some particular form. As in Section 2, we follow Chib (2007) and phrase the problem in terms of the joint density of $p(y_{ji}, x_i = j | \mathbf{w}_i, z_i = l)$, for $l = 0, 1$, without involving the joint distribution of the potential outcomes. For subjects in the control arm, $\Pr(x_i = 0 | \mathbf{w}_i, z_i = 0) = 1$, and it is only necessary to specify a marginal model of the potential outcome y_{i0} . In other words, joint modeling of the potential outcome and the intake is only needed for subjects in the treatment arm. For generality, we allow the marginal distributions of the potential outcomes to be different. Just as the type confounder model, we impose the exclusion restriction that the marginal distribution of y_{ji} does not depend on the assignment $z_i = l$. Specifically, we assume that

$$p(y_{ji}, x_i = j | \mathbf{w}_i, z_i = l)$$

$$= \begin{cases} p_0(y_{0i} | \mathbf{w}_i) & \text{if } i \in I_0 \\ p_j(y_{ji}, x_i = j | \mathbf{w}_i, z_i = 1), \quad j = 0, 1 & \text{if } i \in I_1 \end{cases} \quad (3.1)$$

where $I_l = \{i : z_i = l\}$ denotes the sample indices of a subject with $z_i = l$. On decomposing the joint density in terms of a marginal density of the outcome and a conditional probability of the intake,

and writing out the $j = 0$ and $j = 1$ cases fully, we can write the latter model explicitly as

$$p(y_{ji}, x_i = j | \mathbf{w}_i, z_i = l) = \begin{cases} p_0(y_{0i} | \mathbf{w}_i) & \text{if } i \in I_{00} \\ p_0(y_{0i} | \mathbf{w}_i) \Pr(x_i = 0 | \mathbf{w}_i, z_i = 1, y_{0i}) & \text{if } i \in I_{10} \\ p_1(y_{1i} | \mathbf{w}_i) \Pr(x_i = 1 | \mathbf{w}_i, z_i = 1, y_{1i}) & \text{if } i \in I_{11} \end{cases} \quad (3.2)$$

where, as before, $I_{lj} = \{i : z_i = l \text{ and } x_i = j\}$. Thus, for subjects in the control arm the model is specified in terms of the marginal density $p(y_{0i} | \mathbf{w}_i)$. Because of the maintained exclusion restriction, this marginal density is the same as the marginal density of y_{0i} of subjects in the treatment arm who forgo the treatment.

The task at hand, therefore, is to specify the form of

$$p(y_i, x_i = j | \mathbf{w}_i, z_i = 1) \equiv p(y_{ji}, x_i = j | \mathbf{w}_i, z_i = 1), \quad j = 0, 1.$$

We generate these two joint densities from the model

$$\begin{aligned} y_{ji} &= \mathbf{w}'_i \boldsymbol{\beta}_j + \varepsilon_{ji}, \quad j = 0, 1 \\ x_i^* &= \mathbf{v}'_i \boldsymbol{\gamma} + u_i \\ x_i &= I \{x_i^* > 0\} \end{aligned}$$

where \mathbf{v}_i is a vector of pretreatment variables that is a subset of \mathbf{w}_i and x_i^* is a latent variable. Note that the role of $z_i = 1$ is implicit in this formulation since it serves merely to pinpoint the fact that when $z_i = 1$ the outcome and intake are jointly determined. To model continuous unobserved confounders we assume that

$$(\varepsilon_{ji}, u_i) | \lambda_i \sim \mathcal{N}_2(\mathbf{0}, \lambda_i^{-1} \boldsymbol{\Omega}_j)$$

where

$$\boldsymbol{\Omega}_j = \begin{pmatrix} \eta_j^2 & \omega_j \\ \omega_j & 1 \end{pmatrix}$$

is up to scale the conditional covariance matrix between the y_{ji} and x_i^* . Here λ_i is a positive random-variable that is assumed to be iid gamma $(\frac{\nu}{2}, \frac{\nu}{2})$ for some known value $\nu > 2$. We could entertain other possible mixing distributions but our choice of the gamma distribution leads to the bivariate student- t distribution which is simple to work with and comparable to the one in the type confounder model.

If we let $\boldsymbol{\beta} = (\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\gamma})$, $\sigma_j^2 = \eta_j^2 - \omega_j^2$ (for reasons presented below) and $\boldsymbol{\psi}_j = (\boldsymbol{\beta}_j, \omega_j, \sigma_j^2)$, it follows that the joint density of y_{ji} and x_i^* for subjects in I_1 is

$$p^*(y_{ji}, x_i^* | \mathbf{w}_i, z_i = 1, \boldsymbol{\psi}_j, \boldsymbol{\gamma}, \lambda_i) = \mathcal{N}_2(y_{ji}, x_i^* | \mathbf{X}_{j,i} \boldsymbol{\beta}, \lambda_i^{-1} \boldsymbol{\Omega}_j) \quad (3.3)$$

where

$$\mathbf{X}_{j,i} = \begin{pmatrix} \mathbf{w}'_i \times (1-j) & \mathbf{w}'_i \times j & 0 \\ 0 & 0 & \mathbf{v}'_i \end{pmatrix}.$$

Therefore, marginal of λ_i , the joint densities

$$p^*(y_{0i}, x_i^* | \mathbf{w}_i, z_i = 1, \boldsymbol{\psi}_0, \boldsymbol{\gamma})$$

and

$$p^*(y_{1i}, x_i^* | \mathbf{w}_i, z_i = 1, \boldsymbol{\psi}_1, \boldsymbol{\gamma})$$

are bivariate student- t . From here on, integrating out the latent scale while paying particular attention to the interval of integration, we get that for $i \in I_1$,

$$\begin{aligned} p(y_i, x_i = j | \mathbf{w}_i, z_i = 1, \boldsymbol{\psi}_j, \boldsymbol{\gamma}) &\equiv p(y_{ji}, x_i = j | \mathbf{w}_i, z_i = 1, \boldsymbol{\psi}_j, \boldsymbol{\gamma}) \\ &= p_j(y_i | \mathbf{w}_i, \boldsymbol{\beta}_j, \eta_j^2) \int_{A_j} p^*(x_i^* | \mathbf{w}_i, z_i = l, y_i, \boldsymbol{\beta}_j, \eta_j^2, \omega_j) dx_i^* \\ &= t_\nu(y_i | \mathbf{w}'_i \boldsymbol{\beta}_j, \eta_j^2) T_{\nu+1} \left\{ (2j-1) \frac{\mu_{ji}}{h_{ji} \phi_j} \right\} \end{aligned} \quad (3.4)$$

where A_j is the set $(-\infty, 0)$ if $j = 0$ and $(0, \infty)$ if $j = 1$, t_ν is the student- t density with ν degrees of freedom, $T_{\nu+1}\{\cdot\}$ is the cdf of the standard t -density,

$$\begin{aligned} \mu_{ji} &= \mathbf{v}'_i \boldsymbol{\gamma} + \omega_j \eta_j^{-2} (y_i - \mathbf{w}'_i \boldsymbol{\beta}_j) \\ h_{ji}^2 &= [\nu(\nu+1)] [1 + (y_i - \mathbf{w}'_i \boldsymbol{\beta}_j)^2 \eta_j^{-2} / \nu] \\ \phi_j^2 &= 1 - \omega_j^2 / \eta_j^2. \end{aligned}$$

Setting $j = 0$ in Eq. (3.3), and noting that the intake is non-stochastic when $z_i = 0$, it follows that

$$\begin{aligned} p(y_i, x_i = 0 | \mathbf{w}_i, z_i = 0, \boldsymbol{\beta}_0, \eta_0^2) &= p(y_{0i}, x_i = 0 | \mathbf{w}_i, z_i = 0, \boldsymbol{\beta}_0, \eta_0^2) \\ &= t_\nu(y_i | \mathbf{w}'_i \boldsymbol{\beta}_0, \eta_0^2). \end{aligned} \quad (3.5)$$

3.2. Prior

To complete the model specification we now supply a prior density for the model parameters $\boldsymbol{\psi} = (\boldsymbol{\gamma}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_1)$. Specifically, we assume that σ_j^2 is distributed inverse-gamma and $\boldsymbol{\beta}_j$ and ω_j are Gaussian:

$$\pi(\boldsymbol{\psi}) = \mathcal{I} \mathcal{G} \left(\sigma_j^2 \mid \frac{n_{j,0}}{2}, \frac{d_{j,0}}{2} \right) \mathcal{N}_p(\boldsymbol{\beta}_j | \boldsymbol{\beta}_{j,0}, \mathbf{B}_{j,0}) \mathcal{N}(\omega_j | m_{j,0}, M_{j,0})$$

where the quantities indexed by zero are the prior hyperparameters, and p is the dimension of $\boldsymbol{\beta}_j$. We assume that $\boldsymbol{\psi}_0$ and $\boldsymbol{\psi}_1$ are a priori independent. Notice that σ_j^2 is the determinant of $\boldsymbol{\Omega}_j$ and our choice of an inverse-gamma distribution for it ensures that $\boldsymbol{\Omega}_j$ is positive definite. We note for future use that these assumptions imply that the prior density of the $p+1$ -dimensional vector $\tilde{\boldsymbol{\beta}}_1 = (\boldsymbol{\beta}_1, \omega_1)$ is $\mathcal{N}_{p+1}(\tilde{\boldsymbol{\beta}}_1 | \tilde{\boldsymbol{\beta}}_{1,0}, \tilde{\mathbf{B}}_{1,0})$ with mean $\tilde{\boldsymbol{\beta}}_{1,0} = (\boldsymbol{\beta}_{1,0} | m_{1,0})$ and covariance matrix

$$\tilde{\mathbf{B}}_{1,0} = \begin{pmatrix} \mathbf{B}_{1,0} & \mathbf{0} \\ \mathbf{0} & M_{1,0} \end{pmatrix}.$$

We further assume that $\boldsymbol{\gamma}$ is distributed as $\mathcal{N}_q(\boldsymbol{\gamma} | \boldsymbol{\gamma}_0, \mathbf{G}_0)$, independent of $\boldsymbol{\psi}_j$. Putting these assumptions together, the prior on $\boldsymbol{\psi} = (\boldsymbol{\beta}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_1)$ is of the form

$$\begin{aligned} \pi(\boldsymbol{\psi}) &= \mathcal{N}_q(\boldsymbol{\gamma} | \boldsymbol{\gamma}_0, \mathbf{G}_0) \prod_{j=0}^1 \mathcal{I} \mathcal{G} \left(\sigma_j^2 \mid \frac{n_{j,0}}{2}, \frac{d_{j,0}}{2} \right) \\ &\quad \times \mathcal{N}_p(\boldsymbol{\beta}_j | \boldsymbol{\beta}_{j,0}, \mathbf{B}_{j,0}) \mathcal{N}(\omega_j | m_{j,0}, M_{j,0}). \end{aligned} \quad (3.6)$$

3.3. Posterior distribution and fitting

Estimation of $\boldsymbol{\psi}$ from the posterior density $\pi(\boldsymbol{\psi} | \mathbf{y}, \mathbf{x}) \propto \pi(\boldsymbol{\psi}) f(\mathbf{y}, \mathbf{x} | \boldsymbol{\psi})$ requires MCMC methods. Following Albert and Chib (1993), let $\mathbf{x}_i^* = \{x_i^* : i \in I_1\}$, and $\boldsymbol{\lambda} = \{\lambda_i : i \leq n\}$ and consider the joint density

$$\begin{aligned} p(y_i, x_i^*, x_i = j | \mathbf{w}_i, z_i = l, \boldsymbol{\beta}, \boldsymbol{\psi}_j, \lambda_i) &\equiv p(y_{ji}, x_i^*, x_i = j | \mathbf{w}_i, z_i = l, \boldsymbol{\beta}, \boldsymbol{\psi}_j, \lambda_i) \\ &= p^*(y_{ji}, x_i^* | \mathbf{w}_i, z_i = l, \boldsymbol{\beta}, \boldsymbol{\psi}_j, \lambda_i) [I\{x_i^* < 0\}^{1-j} + I\{x_i^* > 0\}^j] \end{aligned}$$

where $p^*(y_{ji}, x_i^* | \mathbf{w}_i, z_i = l, \boldsymbol{\beta}, \boldsymbol{\psi}_j, \lambda_i)$ is the bivariate normal density given in expression (3.3). We now focus on the distribution $\pi(\boldsymbol{\psi}, \mathbf{x}_i^*, \boldsymbol{\lambda} | \mathbf{y}, \mathbf{x})$ which is proportional to:

$$\begin{aligned} \pi(\boldsymbol{\psi}) &\prod_{i=1}^n \mathcal{G} \left(\lambda_i \mid \frac{\nu}{2}, \frac{\nu}{2} \right) \prod_{i \in I_{00}} \mathcal{N}(y_i | \mathbf{w}'_i \boldsymbol{\beta}_0, \lambda_i^{-1}, \eta_0^2) \\ &\quad \times \prod_{i \in I_{10}} \mathcal{N}_2(y_i, x_i^* | \mathbf{X}_{0,i} \boldsymbol{\beta}, \lambda_i^{-1} \boldsymbol{\Omega}_0) I\{x_i^* < 0\} \\ &\quad \times \prod_{i \in I_{11}} \mathcal{N}_2(y_i, x_i^* | \mathbf{X}_{1,i} \boldsymbol{\beta}, \lambda_i^{-1} \boldsymbol{\Omega}_1) I\{x_i^* > 0\} \end{aligned} \quad (3.7)$$

involving both the marginal density of the outcome (for subjects with intake $x_i = 0$) and the joint density of the outcome and the intake (for subjects with intake $x_i = 1$).

Given the generality of MCMC sampling methods, this posterior density can be sampled in various ways. To do the sampling as efficiently and conveniently as possible, we employ a blocking scheme that is coarse, in the sense that it minimizes the number of blocks. For instance, we sample the latent variables x_i^* , for $i \in I_{10}$ and $i \in I_{11}$, and the scale parameters λ_i , for $i \leq n$ in one block. We also sample ω_1 and β_1 in one block because it can be difficult to estimate the degree of confounding when $x = 1$ because subjects in this case are only observed under one level of the assignment (note that this problem does not occur for ω_0 and β_0 since the data in the set I_{00} supplies information about β_0 , independent of ω_0). Finally, we sample $\zeta_0 = (\sigma_0^2, \omega_0)$ in one block because we found by experimentation that this leads to better mixing in the output. We mention that the target density for the latter step is not in closed form because of the fact that the density of the outcomes in the control arm implicitly contains ω_0 since, in our parameterization, $\eta_0^2 = \sigma_0^2 + \omega_0^2$. We therefore make use of the Metropolis–Hastings (M–H) algorithm to sample that density.

In particular, our MCMC sampling algorithm is composed of the following five steps.

- In the first step we jointly sample the latent treatment variables x_i^* , for $i \in I_{10}$ and $i \in I_{11}$, and the scale parameters λ_i , for $i \leq n$, sampling x_i^* marginalized over λ_i from

$$t_{v+1}(x_i^* | m_{ji}, h_{ji}^2 \phi_j^2) \{I(x_i^* < 0)^{1-j} + I(x_i^* > 0)^j\}$$

followed by the update of λ_i for these subjects ($i \in I_1$) from a gamma density

$$\mathcal{G} \left(\lambda_i | \frac{v_1}{2}, \frac{d_1}{2} \right)$$

where $v_1 = v_0 + 2$ and $d_1 = v_0 + (\mathbf{y}_i^* - \mathbf{X}_{1i}\beta)' \Omega_i^{-1} (\mathbf{y}_i^* - \mathbf{X}_{1i}\beta)$. In the final step of the first block we update λ_i for subjects in the control group ($i \in I_{00}$) from a gamma density, where the parameters v_1 and d_1 are now given by $v_1 = v_0 + 1$ and $d_1 = v_0 + (y_i - \mathbf{w}_i'\beta_0)\eta_0^{-2}(y_i - \mathbf{w}_i'\beta_0)$.

- In the second step we sample γ based on the observations in the sets I_{10} and I_{11} from

$$\pi(\gamma | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{x}_1^*, \beta_0, \beta_1, \sigma_0^2, \omega_0, \sigma_1^2, \omega_1, \lambda) = \mathcal{N}_p(\gamma | \hat{\gamma}, \mathbf{G}) \quad (3.8)$$

with mean

$$\hat{\gamma} = \mathbf{G} \left\{ \mathbf{G}_0^{-1} \gamma_0 + \sum_{i \in I_{10}} \lambda_i \mathbf{w}_i (1 - \omega_0^2 \eta_0^{-2}) \bar{x}_{i0}^* + \sum_{i \in I_{11}} \lambda_i \mathbf{w}_i (1 - \omega_1^2 \eta_1^{-2}) \bar{x}_{i1}^* \right\}$$

where $\bar{x}_{ij}^* = (x_i^* - \omega_j \eta_j^{-2} (y_i - \mathbf{w}_i'\beta_j))$, and covariance $\mathbf{G} = \{\mathbf{G}_0^{-1} + \sum_{i \in I_{10}} \lambda_i \mathbf{w}_i (1 - \omega_0^2 \eta_0^{-2}) \mathbf{w}_i + \sum_{i \in I_{11}} \lambda_i \mathbf{w}_i (1 - \omega_1^2 \eta_1^{-2}) \mathbf{w}_i\}^{-1}$.

- Then in the third step we update β_0, β_1 and ω_1 in one block, sampling β_0 from

$$\pi(\beta_0 | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{x}_1^*, \gamma, \sigma_0^2, \omega_0, \lambda) = \mathcal{N}_p(\beta_0 | \hat{\beta}_0, \mathbf{B}_0) \quad (3.9)$$

with mean

$$\hat{\beta}_0 = \mathbf{B}_0 \left\{ \mathbf{B}_{0,0}^{-1} \beta_{0,0} + \sum_{i \in I_{00}} \lambda_i \mathbf{w}_i \eta_0^{-2} y_i + \sum_{i \in I_{10}} \lambda_i \mathbf{w}_i \sigma_0^{-2} (y_i - \omega_0 \hat{x}_i^*) \right\}$$

and covariance $\mathbf{B}_0 = \{\mathbf{B}_{1,0}^{-1} + \sum_{i \in I_{00}} \lambda_i \mathbf{w}_i \eta_0^{-2} \mathbf{w}_i' + \sum_{i \in I_{10}} \lambda_i \mathbf{w}_i \sigma_0^{-2} \mathbf{w}_i'\}^{-1}$, and $\tilde{\beta}_1 = (\beta_1, \omega_1)$ from

$$\pi(\tilde{\beta}_1 | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{x}_1^*, \gamma, \sigma_1^2, \lambda) = \mathcal{N}_{p+1} \left(\tilde{\beta}_1 | \tilde{\mathbf{B}}_1 \left\{ \tilde{\mathbf{B}}_{1,0}^{-1} \tilde{\beta}_{1,0} + \sum_{i \in I_{11}} \lambda_i \tilde{\mathbf{w}}_i \sigma_1^{-2} y_i \right\}, \tilde{\mathbf{B}}_1 = \left\{ \tilde{\mathbf{B}}_{1,0}^{-1} + \sum_{i \in I_{11}} \lambda_i \tilde{\mathbf{w}}_i \sigma_1^{-2} \tilde{\mathbf{w}}_i' \right\}^{-1} \right) \quad (3.10)$$

where $\tilde{\mathbf{w}}_i = \mathbf{w}_i | u_i$.

- In the fourth step we jointly update $\zeta_0 = (\sigma_0^2, \omega_0)$ with a M–H step. This requires a M–H step due to the fact that the density of the outcomes in the control arm involves ω_0 implicitly in the parameterization $\eta_0^2 = \sigma_0^2 + \omega_0^2$. To do this sampling efficiently, we follow Chib and Greenberg (1998), and generate the proposal values σ_0^{2+} and ω_0^+ from a tailored student- t density $t_v(\boldsymbol{\mu}, \mathbf{V})$, where $\boldsymbol{\mu}$ is the (approximate) mode of

$$\ln \left(\prod_{i \in I_{00}} \mathcal{N}(y_i | \mathbf{w}_i'\beta_0, \lambda_i^{-1} \eta_0^2) \times \prod_{i \in I_{10}} \mathcal{N}_2(y_i, x_i^* | \mathbf{X}_{0,i}\beta, \lambda_i^{-1} \Omega_0) \right)$$

for the subjects in the sets I_{00} and I_{10} , and \mathbf{V} is the inverse Hessian of this density evaluated at $\boldsymbol{\mu}$. We accept the proposal values with probability of move $\alpha_{MH} = \alpha_{MH}(\zeta_0, \zeta_0^+ | \mathbf{y}, \mathbf{x}, \mathbf{W}, \beta, \sigma_1^2, \omega_1, \mathbf{x}_1^*, \lambda)$ where

$$\alpha_{MH} = \min\{1, r_1 r_2\}, \quad (3.11)$$

where

$$r_1 = \frac{\pi(\zeta_0^+) \prod_{i \in I_{00}} \mathcal{N}(y_i | \mathbf{w}_i'\beta_0, \lambda_i^{-1} \eta_0^{2+}) \prod_{i \in I_{10}} \mathcal{N}_2(y_i, x_i^* | \mathbf{X}_{0,i}\beta, \lambda_i^{-1} \Omega_0^+)}{\pi(\zeta_0) \prod_{i \in I_{00}} \mathcal{N}(y_i | \mathbf{w}_i'\beta_0, \lambda_i^{-1} \eta_0^2) \prod_{i \in I_{10}} \mathcal{N}_2(y_i, x_i^* | \mathbf{X}_{0,i}\beta, \lambda_i^{-1} \Omega_0)},$$

$$r_2 = \frac{t_v(\zeta_0 | \boldsymbol{\mu}, \mathbf{V})}{t_v(\zeta_0^+ | \boldsymbol{\mu}, \mathbf{V})}$$

and $\eta_0^{2+} = \sigma_0^{2+} + \omega_0^{2+}$.

- In the final step, we sample σ_1^2 conditional on ω_1 and the remaining parameters from

$$\pi(\sigma_1^2 | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{x}_1^*, \beta, \lambda) = \mathcal{I} \mathcal{G} \times \left(\sigma_1^2 | \frac{v_{1,0} + n_{11}}{2}, \frac{d_{1,0} + \sum_{i \in I_{11}} \lambda_i (y_i - \mathbf{w}_i'\beta_1 - \omega_1 u_i)^2}{2} \right) \quad (3.12)$$

where n_{11} is the number of subjects in the set I_{11} .

We run our MCMC sampler for M iterations after an initial burn-in period of m_0 iterations.

3.4. Inferring treatment effects

We calculate the treatment effects in a way that parallels the approach outlined in Section 2.3. The relevant potential outcome densities of interest are the predictive densities

$$p(y_{j,n+1} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}), \quad j = 0, 1$$

where the subscript $n + 1$ refers to a subject randomly drawn from the entire population. As is usual in the Bayesian context, these predictive densities are sampled by the method of composition by appending the following steps to each iteration $g \leq M$ of the algorithm.

- First, we randomly sample $\mathbf{w}_{n+1}^{(g)}$ from the full set of covariates.
- We then draw $y_{j,n+1}^{(g)}$ from $t_v(y_{j,n+1}^{(g)} | \mathbf{w}_{n+1}^{(g)}, \beta_j^{(g)}, \eta_j^{2(g)})$, where $(\beta_j^{(g)}, \eta_j^{2(g)})$ are the values of the parameters at the g th iteration of the chain.

We use the simulated samples $(y_{0,n+1}^{(1)}, \dots, y_{0,n+1}^{(M)})$ and $(y_{1,n+1}^{(1)}, \dots, y_{1,n+1}^{(M)})$ to compute various effects of interest, including the predictive average causal effect. The average predictive treatment effect can be interpreted as the expected causal treatment gain of a random subject from the population.

An interesting point to note is that the general confounder model can also be used to deliver a treatment effect for compliers. The idea, introduced in Chib (2007), is to compute the potential treatments, x_0 and x_1 , under the two possible assignments to identify if a subject would comply with the treatment. In the eligibility design assignment into the control state implies that $x_0 = 0$. A complier is then a person who would be induced to take the treatment if assigned into the treatment arm. We generate the predictive marginal distributions of the potential outcomes for compliers by calculating $x_{1,n+1}^{(g)} = I[\mathbf{v}_i^{(g)} \boldsymbol{\gamma}^{(g)} + u_{1,n+1}^{(g)} > 0]$ at each iteration, where $u_{1,n+1}^{(g)} \sim t_v(0, 1)$. If $x_{1,n+1}^{(g)} = 1$ we compute $y_{0,n+1}^{(g)}$ and $y_{1,n+1}^{(g)}$ as described above. If $x_{1,n+1}^{(g)} = 0$ we move to the next step in the MCMC chain.

4. Estimation of the marginal likelihood

The developments in the previous sections open up the possibility of fitting eligibility design data under each framework, with the aim of appraising which model and assumptions are better supported by the data. We calculate this support through the marginal likelihood of each model. As has been discussed elsewhere, the marginal likelihood can be calculated easily by the method of Chib (1995). This method exploits the fact that on the log-scale the marginal likelihood of a model (say) M can be expressed in terms of its likelihood function, the prior and the posterior as

$$\ln m(\mathbf{y}, \mathbf{x} | M) = \ln f(\mathbf{y}, \mathbf{x} | \mathbf{z}, \mathbf{W}, \boldsymbol{\psi}^*, M) + \ln \pi(\boldsymbol{\psi}^* | M) - \ln \pi(\boldsymbol{\psi}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, M) \tag{4.1}$$

each evaluated at $\boldsymbol{\psi}^*$ (the posterior mean of the parameter vector $\boldsymbol{\psi}$). The first two terms in this expression can be found directly in each of our two models. For instance, the likelihood ordinate of the type confounder model comes from (2.5) whereas that of the general confounder model from expressions (3.4) and (3.5). What remains is the calculation of the posterior ordinate which we get as follows.

Consider first the case of the type confounder model where the parameter vector is $\boldsymbol{\psi} = (\boldsymbol{\alpha}, \eta^2, \boldsymbol{\beta}, \cdot)$, with $\boldsymbol{\beta} = (\beta_{0c}, \beta_{0n}, \beta_{1c})$ and $\eta^2 = (\eta_{0c}^2, \eta_{0n}^2, \eta_{1c}^2)$. Dropping the model index M , to estimate $\pi(\boldsymbol{\psi}^* | \mathbf{y}, \mathbf{x}, \mathbf{W})$ we utilize the decomposition

$$\pi(\boldsymbol{\psi}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}) = \pi(\eta^{2*} | \mathbf{y}, \mathbf{x}, \mathbf{W}) \pi(\boldsymbol{\alpha}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}) \times \pi(\boldsymbol{\beta}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}).$$

The first of these ordinates can be obtained by Rao-Blackwellization as

$$\hat{\pi}(\eta^{2*} | \mathbf{y}, \mathbf{x}, \mathbf{W}) = M^{-1} \sum_{g=1}^M \left(\prod_{j=0,1} \prod_{k \in K_j} \pi(\eta_{jk}^{2*} | \mathbf{y}_{jk}, \mathbf{x}_{jk}, \mathbf{W}_{jk}, \boldsymbol{\beta}_{jk}^{(g)}, \boldsymbol{\lambda}_{jk}^{(g)}) \right)$$

where $K_0 = \{c, n\}$, $K_1 = \{c\}$ and $(\boldsymbol{\beta}_{jk}^{(g)}, \boldsymbol{\lambda}_{jk}^{(g)}, q^{c(g)})$ is the g th draw from the MCMC run.

The next ordinate comes from applying a result of Chib and Jeliazkov (2001) which shows that

$$\pi(\boldsymbol{\alpha}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}) = \frac{\mathbb{E}_1(\alpha(\boldsymbol{\alpha}, \boldsymbol{\alpha}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\theta}) q(\boldsymbol{\alpha}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\theta}))}{\mathbb{E}_2 \alpha(\boldsymbol{\alpha}^*, \boldsymbol{\alpha} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\theta})} \tag{4.2}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\lambda})$, \mathbb{E}_1 denotes the expectation with respect to the posterior distribution

$$\pi(\boldsymbol{\alpha}, \boldsymbol{\theta} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*})$$

and \mathbb{E}_2 the expectation with respect to

$$\pi(\boldsymbol{\theta} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\alpha}^*) q(\boldsymbol{\alpha} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\theta}).$$

These can be estimated conveniently by the method of reduced MCMC runs by averaging the product in the numerator over the draws from a reduced MCMC run where η^2 is fixed at η^{2*} , and by averaging $\alpha(\boldsymbol{\alpha}^*, \boldsymbol{\alpha} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\theta})$ in the denominator over draws of $\boldsymbol{\theta}$ drawn from $\pi(\boldsymbol{\theta} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\alpha}^*)$ and $\boldsymbol{\alpha}$ from $q(\boldsymbol{\alpha} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\theta})$.

The output from either of the latter two reduced runs can also be used to obtain the reduced ordinate of $\boldsymbol{\beta}$ as

$$\hat{\pi}(\boldsymbol{\beta}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}) = M^{-1} \sum_{g=1}^M \pi(\boldsymbol{\beta}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \eta^{2*}, \boldsymbol{\lambda}^{(g)})$$

where the summand is the density given in (2.8).

In the case of the general confounder model the parameter is $\boldsymbol{\psi} = (\boldsymbol{\gamma}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_1)$, where as before $\boldsymbol{\psi}_j = (\boldsymbol{\beta}_j, \sigma_j^2, \omega_j)$. We decompose the posterior ordinate $\pi(\boldsymbol{\psi}^* | \mathbf{y}, \mathbf{x}, \mathbf{W})$ into a product of marginal and reduced posterior ordinates as

$$\pi(\boldsymbol{\zeta}_0^* | \mathbf{y}, \mathbf{x}, \mathbf{W}) \pi(\boldsymbol{\gamma}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\zeta}_0^*) \pi(\sigma_1^{2*} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\zeta}_0^*, \boldsymbol{\gamma}^*) \times \pi(\boldsymbol{\beta}_0^*, \boldsymbol{\beta}_1^*, \omega_1^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\zeta}_0^*, \boldsymbol{\gamma}^*, \sigma_1^{2*})$$

where $\boldsymbol{\zeta}_0^* = (\sigma_0^{2*}, \omega_0^*)$.

The first of these ordinates is estimated by the Chib and Jeliazkov (2001) approach described in connection with the ordinate of $\boldsymbol{\alpha}$ in (4.2). A similar ratio of expectations is involved with $\boldsymbol{\alpha}$ replaced by $\boldsymbol{\xi}, \boldsymbol{\alpha}^*$ by $\boldsymbol{\zeta}_0^*$, and $\boldsymbol{\theta}$ by $(\boldsymbol{\beta}, \sigma_1^2, \omega_1, \mathbf{x}_1^*, \boldsymbol{\lambda})$. Then, \mathbb{E}_1 is estimated by averaging the product in the numerator over the draws from the full MCMC run, and \mathbb{E}_2 by averaging $\alpha(\boldsymbol{\zeta}_0^*, \boldsymbol{\zeta}_0 | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\beta}, \sigma_1^2, \omega_1, \mathbf{x}_1^*, \boldsymbol{\lambda})$, where $\boldsymbol{\zeta}_0$ is drawn from $q(\boldsymbol{\zeta}_0 | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\beta}, \sigma_1^2, \omega_1, \mathbf{x}_1^*, \boldsymbol{\lambda})$ and the remaining parameters from a reduced run with $\boldsymbol{\zeta}_0$ fixed at $\boldsymbol{\zeta}_0^*$.

Notice that the output from the latter reduced run can also be used to obtain the reduced ordinate of $\boldsymbol{\gamma}$ as

$$\hat{\pi}(\boldsymbol{\gamma}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\zeta}_0^*) = M^{-1} \sum_{g=1}^M \pi(\boldsymbol{\gamma}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\zeta}_0^*, \mathbf{x}_1^{*(g)}, \boldsymbol{\beta}_0^{(g)}, \boldsymbol{\beta}_1^{(g)}, \sigma_1^{2(g)}, \omega_1^{(g)}, \boldsymbol{\lambda}^{(g)})$$

where the summand is the density given in (3.8).

Next, to estimate $\pi(\sigma_1^{2*} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\zeta}_0^*, \boldsymbol{\gamma}^*)$ we fix $(\boldsymbol{\zeta}_0, \boldsymbol{\gamma})$ at $(\boldsymbol{\zeta}_0^*, \boldsymbol{\gamma}^*)$ and continue the MCMC iterations. The draws from this reduced run produce the estimate

$$M^{-1} \sum_{g=1}^M \pi(\sigma_1^{2*} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \boldsymbol{\zeta}_0^*, \boldsymbol{\gamma}^*, \boldsymbol{\beta}_0^{(g)}, \boldsymbol{\beta}_1^{(g)}, \sigma_1^{2(g)}, \omega_1^{(g)}, \mathbf{x}_1^{*(g)}, \boldsymbol{\lambda}^{(g)})$$

where the conditional density of σ_1^2 is given by expression (3.12).

Table 2
Type confounder model: Simulation study with $\beta_{0c} = (1, 2)$, $\beta_{0n} = (-0.5, 1)$, $\beta_{1c} = (2, 3)$

Posterior means and standard deviations									
q_c	n	β'_{0c}		β'_{0n}		β'_{1c}		α	$\phi(\alpha)$
.5	250	1.41(.65)	1.93(.20)	-0.58(.29)	1.01(.10)	2.05(.32)	3.01(.11)	0.00(.12)	0.50(.05)
	500	1.01(.51)	1.97(.16)	-0.46(.21)	0.98(.07)	1.97(.23)	3.00(.08)	0.01(.09)	0.50(.03)
	1000	1.10(.40)	1.99(.12)	-0.49(.16)	1.01(.05)	1.98(.16)	3.00(.06)	0.00(.06)	0.50(.02)
.8	250	1.04(.46)	1.99(.15)	-0.55(.52)	1.02(.18)	1.94(.25)	3.02(.09)	0.89(.14)	0.81(.04)
	500	1.01(.32)	2.03(.10)	-0.55(.33)	1.03(.12)	1.92(.18)	3.01(.06)	0.81(.10)	0.79(.03)
	1000	1.00(.24)	2.02(.08)	-0.47(.25)	1.01(.09)	2.01(.13)	3.01(.05)	0.83(.07)	0.80(.02)

Reported are the average of the posterior means and standard deviations from the 20 replications of each design.

Table 3
Type confounder model: Inefficiency factors from the sampled MCMC output for the case $\bar{q}_c = 0.5$

Inefficiency factors											
Sample size	β'_{0c}		β'_{0n}		β'_{1c}		η^2_{0c}	η^2_{0n}	η^2_{1c}	α	
250	5.99	3.90	2.13	1.99	1.24	1.24	5.14	1.90	1.40	11.67	
500	5.03	3.75	2.25	2.02	1.25	1.23	4.00	1.80	1.36	11.88	
1000	5.55	4.02	2.52	2.27	1.27	1.23	3.91	1.79	1.35	13.50	

The results are averaged over 20 replications.

Last, we compute the reduced ordinate of $(\beta_0^*, \beta_1^*, \omega_1^*)$ from the output of the another reduced run, now with $(\xi_0, \gamma, \sigma_1^2)$ fixed at $(\xi_0^*, \gamma^*, \sigma_1^{2*})$, to yield

$$M^{-1} \sum_{g=1}^M \pi(\beta_0^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \xi_0^*, \gamma^*, \sigma_1^{2*}, \mathbf{x}_1^{*(g)}, \lambda^{(g)}) \times \pi(\beta_1^*, \omega_1^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \xi_0^*, \gamma^*, \sigma_1^{2*}, \mathbf{x}_1^{*(g)}, \lambda^{(g)})$$

where the first conditional density is given in Eq. (3.9) and the second in Eq. (3.10).

5. Simulation studies

We conducted some simulation studies to examine the performance of the fitting approaches.

5.1. Type confounder model

We consider a general design with one continuous covariate w_i in the outcome model that we generate from a $\mathcal{N}(2, 4)$ distribution. The coefficient vectors in the outcome model are taken to be $\beta_{0c} = (1, 2)$, $\beta_{0n} = (-0.5, 1)$ and $\beta_{1c} = (2, 3)$, and the variances are fixed at $\eta^2 = (4.00, 4.00, 4.00)$. For the simulation exercise we define the covariate vector in the type probability q_{ci} in terms of a constant and set the coefficient vector α in turn to $\phi^{-1}(.5)$ and $\phi^{-1}(.8)$ to capture different average compliance rates of 50% and 80% in the data. We consider sample sizes of 250, 500 and 1000 for each design, with 20 replications for each sample size. In each experiment, our prior-posterior analysis is conducted under the prior assumption that $\beta_{jk} \sim \mathcal{N}_2(\mathbf{0}, 25 \times \mathbf{I}_2)$ and $\alpha \sim \mathcal{N}(0, 25)$ while the hyperparameters of the inverse gamma distribution for η_{jk}^2 are set to have a prior mean of 2 and standard deviation of 6.

In Table 2 we report the posterior means and standard deviations of β_{0c} , β_{0n} , β_{1c} , α and the implied compliance probability $\phi(\alpha)$ averaged over the 20 replications. The results indicate that the algorithm is able to recover the true parameters reasonably well under both compliance probabilities and different sample sizes. Naturally, the precision of the parameter estimates measured by the posterior standard deviations improves as the sample size increases. The average inefficiency factors (a measure of the mixing of the MCMC chain) when $\bar{q}_c = 0.5$ are given in Table 3. As expected, we observe the highest inefficiency factors for

the parameter α which is updated by a Metropolis Hastings algorithm with a tailored proposal density as described in Section 2.2. In these particular simulation exercises the acceptance rate is just below 80%.

While the results in Tables 2 and 3 show that the MCMC algorithm performs well, they reveal some aspects of the type approach that can impede inference. For example, since the compliance status of the subjects in the control arm is not observed, estimation of β_{0c} can be difficult, especially when the number of compliers in the control arm is small. This is reflected in the higher posterior standard errors of β_{0c} compared to β_{0n} and, of course, β_{1c} . This difficulty also shows up in the higher inefficiency factors of β_{0c} relative to those of β_{0n} and β_{1c} . A crucial step in the estimation of β_{0c} is the correct identification of the subject type variables in the control arm. In Fig. 1 we provide more information on the success of this inference by reporting the posterior means of the compliance type probability for each subject in the control arm. The results refer to one particular replication of the simulation when $q_c = 0.5$ and $n = 500$ with 140 control arm subjects. The left graph shows that, with a few exceptions, the posterior compliance probability is above 0.5 for the 66 compliers. The right graph shows that the posterior compliance probabilities is below 0.5 for the majority of the 74 never takers in the sample.

5.2. General confounder model

We suppose as above that we have one continuous covariate $w_i \sim \mathcal{N}(2, 4)$, and let $\beta_0 = (1.00, 2.00)$, $\beta_1 = (2.00, 3.00)$ and $\gamma = (-1.00, 1.00)$. In addition, we let $\eta^2 = (4.00, 4.00)$. To examine the role of the confounding parameters on inference, we set the value of ω_j between -1.6 and 1.6 . Under our choices of η_j^2 this implies that the correlation coefficient ρ_j ranges between -0.8 and 0.8 . For each setting of these parameters, we generate 20 replications of sample size 250, 500 and 1000. In our fitting we assume that $\gamma \sim \mathcal{N}_2(\mathbf{0}, 25 \times \mathbf{I}_2)$, $\beta_j \sim \mathcal{N}_2(\mathbf{0}, 25 \times \mathbf{I}_2)$ and $\omega_j \sim \mathcal{N}(0, 16)$, for $j = 0, 1$. In addition, we specify the hyperparameters $\nu_{0,j}$ and $d_{0,j}$ of the inverse gamma prior on η_j^2 in such a way that the prior mean and standard deviation are 2 and $\sqrt{20}$ respectively.

Table 4 contains results for β_0 , β_1 , ρ_0 and ρ_1 from the fitting to the different simulated data sets. The first column of the table refers to the specific combination of ρ_0 and ρ_1 and the second column to the sample size. The remaining columns provide the posterior means and standard deviations averaged over the 20 replications. We find that the parameters are well estimated for the

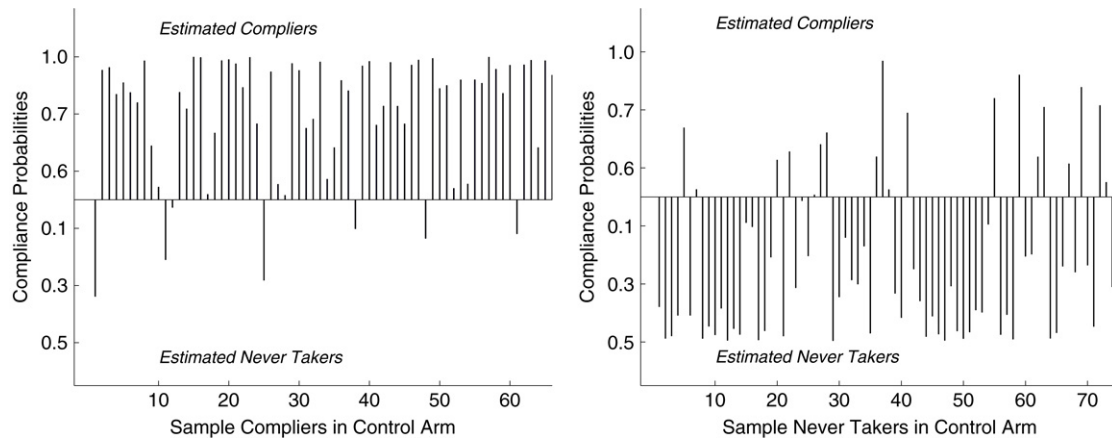


Fig. 1. Posterior means of subjects' estimated compliance probabilities by true type from one replication when $q_c = 0.5$ and $n = 500$. The left graph gives the results for the 66 compliers and the right graph for the 74 never takers in the control arm.

Table 4

General confounder model: Posterior means and standard deviations, averaged over 20 replications to data simulated for different values of ρ and n

ρ	n	β'_0	β'_1	ρ'_0	ρ'_1		
(.5,.5)	250	1.00(0.23)	2.00(0.08)	2.00(0.87)	2.99(0.23)	0.47(0.20)	0.40(0.35)
	500	0.97(0.15)	1.98(0.06)	2.16(0.61)	2.98(0.16)	0.45(0.14)	0.38(0.27)
	1000	1.00(0.11)	2.01(0.04)	1.93(0.37)	3.01(0.10)	0.49(1.00)	0.55(0.15)
(-.5,-.5)	250	0.95(0.22)	2.00(0.08)	1.67(0.88)	3.09(0.23)	-0.60(0.16)	-0.42(0.37)
	500	1.00(0.15)	1.97(0.06)	1.86(0.63)	3.03(0.17)	-0.48(0.14)	-0.44(0.27)
	1000	0.99(0.11)	2.01(0.04)	1.99(0.42)	2.99(0.11)	-0.51(0.10)	-0.46(0.19)
(.8,.8)	250	0.98(0.22)	2.00(0.08)	2.16(0.70)	2.95(0.19)	0.76(0.12)	0.68(0.25)
	500	0.96(0.15)	1.98(0.06)	2.18(0.47)	2.97(0.13)	0.75(0.09)	0.69(0.17)
	1000	1.00(0.10)	2.01(0.04)	2.04(0.29)	2.99(0.08)	0.79(0.06)	0.79(0.08)
(-.8,-.8)	250	0.99(0.21)	2.00(0.08)	1.56(0.67)	3.12(0.19)	-0.81(0.09)	-0.71(0.23)
	500	1.01(0.15)	1.97(0.06)	1.86(0.44)	3.03(0.12)	-0.78(0.08)	-0.76(0.14)
	1000	1.00(0.11)	2.01(0.04)	2.03(0.29)	2.99(0.09)	-0.79(0.06)	-0.80(0.08)

Table 5

General confounder model: Inefficiency factors for designs with $\rho = (0.5, 0.5)$ and $\rho = (0.8, 0.8)$ (second line), averaged over 20 replications

Inefficiency factors											
Sample size	β'_0		β'_1		γ'		σ_0^2	σ_1^2	ω_0	ω_1	
250	3.26	1.80	34.66	28.82	31.28	63.17	9.59	26.88	10.06	48.28	
	3.28	2.19	20.78	16.83	20.31	35.76	8.96	32.92	7.62	32.70	
500	2.49	1.42	31.12	26.09	29.51	57.23	6.90	28.31	7.89	44.68	
	2.51	1.88	27.39	23.46	20.81	36.54	10.91	24.01	9.97	34.18	
1000	2.28	1.43	20.32	15.96	25.68	54.95	7.25	28.99	8.07	31.03	
	3.03	1.82	28.54	17.73	34.53	72.83	20.12	52.46	20.20	50.59	

different degrees of confounding and sample sizes. As expected, the parameters are better estimated across all designs as the sample size is increased. Note, however, that the posterior distribution of β_1 is more dispersed than that of β_0 . Also, the posterior mean of ρ_1 , and to a smaller extent β_1 , are less well estimated than ρ_0 and β_0 , especially when the sample size is small ($n \leq 500$).

These findings reflect the point made earlier in our discussion, that the particular structure of the eligibility design makes it more difficult to separate the effect of β_1 from the effect of ω_1 . Naturally, the dependence in the joint distribution of β_1 and ω_1 manifests itself in higher inefficiency factors for these parameters, as shown in Table 5.

6. Job training and depression

We now illustrate the two approaches for dealing with unobserved confounders in eligibility designs on a real data set from the 1991 JOBS II Intervention Project at the University of Michigan (see Vinokur et al. (1995)). Under the project an eligibility experiment was implemented to test the effectiveness

of a job search seminar for the recently unemployed that was specifically designed to prevent the deterioration in mental health associated with job loss (Clark and Oswald, 1994). Trial participants were recruited from four offices of the Michigan Employment Commission, based on a pretest screening questionnaire. Using baseline information on depression symptoms, financial strain and social assertiveness, the trial identified whether a subject was at low or high risk for experiencing depression in the future. Those classified as clinically depressed were excluded.

We focus on the sample of 715 high risk respondents, of which 229 subjects were randomized into the control arm and 486 into the treatment arm. All subjects, regardless of the assignment, received a booklet on job search. The treatment arm subjects were assigned to a one week job training program that incorporated specific components to promote self-esteem and sense of control, job search skills and inoculation against setbacks. However, out of the 486 subjects randomized into the treatment arm, only 264 decided to participate in the training program. To evaluate the program, information on all sample subjects was gathered through questionnaires at different stages of the experiment. The key

Table 6
Sample summary statistics of the study data from the JOBS II Intervention Project

Variable	Explanation	All	Control arm $z = 0$		Treatment arm $z = 1$	
			$x = 0$	$x = 0$	$x = 0$	$x = 1$
<i>Depress</i>	Change in depression score	−0.44 (0.77)		−0.39		−0.51
<i>Age</i>	Age in years	36.41 (9.96)	36.05		33.25	39.29
<i>Educ</i>	School grade completed	13.35 (2.02)	13.33		12.91	13.72
<i>Married</i>	Dummy	0.62	0.59		0.62	0.64
<i>Nonwhite</i>	Dummy	0.19	0.18		0.24	0.16
<i>Econ</i>	Economic hardship	3.60 (0.87)	3.47		3.77	3.58
<i>Motivate</i>	Motivation to attend	5.32 (0.81)	5.3		5.13	5.47
<i>Assert</i>	Assertiveness	3.06 (0.91)	3.04		3.23	2.95
<i>Basedep</i>	Baseline depression score	2.45 (0.30)	2.49		2.43	2.41
<i>Baserisk</i>	Baseline risk score	1.68 (0.21)	1.69		1.68	1.67

variable for the assessment of the training program is a depression score based on an 18-item index of various stress symptoms (each on a scale from 1 to 5) that was computed for all subjects before and six months after the intervention.

In Table 6 we present the sample means and standard deviations of the variables used in the subsequent analysis. The variable *Depress* refers to the outcome variable, the change in the depression score from the baseline to 6 months after the intervention. The next four variables refer to standard demographic characteristics age, gender, marital status and race, while the remaining variables provide information on the subject's economic situation, attitude towards participation and mental health at baseline. The variable *Econ* reflects the financial strain as measured by a 3-item index on a scale from 1 to 5 while the variable *Motivate* provides information about a subject's attitude towards participation in training programs. The latter variable is based on a two-item index, each on a scale from 1 to 7, with one representing the most positive and 7 the most negative attitude. The final three variables relate to various relevant aspects of the subject's mental health at baseline. For example, *Assert* measures a subject's assertiveness based on a short 4-item index for social reticence and shyness and assertiveness, each on a scale from 1 (low) to 5 (high). In addition to the depression score at baseline, the variable *Baserisk* provides a more general measure for the risk of depression based on weighted averaged of the depression score, the financial strain score and the assertiveness score.

In the last three columns, we report the sample means of the demographic and mental health related control variables stratified by assignment and intake. The last two columns in the table show some differences in the distribution of the covariates between non-participants and participants in the treatment arm. For example, program participants were older, more educated, less assertive and more likely white. From the first row we see that the mean decline in the depression score is −0.51 for treatment participants and −0.38 for non-participants, which points to (but of course does not establish) an improvement in mental health due to the training. Corrected for confounding within the subject-type framework with normally distributed errors, Yau and Little (2001) and Skron Dahl and Rabe-Hesketh (2004), report a value of the complier average treatment of a negative 0.30 points.

We now investigate whether the different assumptions about the unobserved confounders affect the estimates of the treatment effects and the model fit (evaluated in terms of the marginal likelihood). To make the candidate approaches as similar as possible, we specify comparable marginal models for the change in depression score, for five different values of the student-*t* degrees of freedom parameter. We also employ closely related models for the complier-type probability and intake probability. In addition we formulate the prior distribution in each model to be as comparable as possible. In particular, we employ two sets of priors, the neutral and the optimistic priors. The former prior

implies no change in depression score while the latter implies a decrease in the depression score. In all, we consider thirty different model specifications (20 for the restricted and unrestricted type confounder model and 10 for the general confounder model) for a robust comparative empirical analysis.

6.1. Model specifications

We begin with the specification of the subject type model. Motivated by Yau and Little (2001) and Skron Dahl and Rabe-Hesketh (2004), we model the change in depression score ($y = \text{depress}$) for a subject in intake state j of type k as

$$p_j(y_i | \mathbf{w}_i, s_i = k) = t_v(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{jk}, \eta_{jk}^2), \quad k = \{n, c\}$$

where $j = 0$ if $k = n$ and $j = \{0, 1\}$ if $k = c$, as a function of the covariate vector

$$\mathbf{w}_i = [1, \text{basedep}, \text{baserisk}].$$

The probability of being a complier, here modeled as $q_{ci} = \phi(\mathbf{v}'_i \boldsymbol{\alpha})$, is specified in terms of the demographic and remaining mental health related variables from Table 6:

$$\mathbf{v}_i = [1, \text{age-20}, \text{motivate}, \text{educ}, \text{assert}, \text{single}, \text{econ}, \text{nonwhite}].$$

We also consider a restricted version of this type confounder model where the compliance type probability does not depend on any covariate. This specification corresponds to the type model as for example discussed in Sommer and Zeger (1991).

Next, we specify our general confounder model for the analysis of the JOBSII trial data in a way that is comparable to the subject type model. We formulate the marginal models for the change in depression score under the two possible program intake states, $j = 0, 1$, as

$$p_j(y_i | \mathbf{w}_i, \boldsymbol{\beta}_j, \eta_j^2) = t_v(y_i | \mathbf{w}'_i \boldsymbol{\beta}_j, \eta_j^2)$$

where, as above, $\mathbf{w}_i = [1, \text{basedep}, \text{baserisk}]$. Under the general confounder model the program participation decision is modeled directly through the marginal model of the binary intake variable $x_i = 0, 1$ as

$$x_i = I\{\mathbf{v}'_i \boldsymbol{\gamma} + u_i > 0\}, \quad u_i \sim t_v(0, 1),$$

that is assumed to depend on the same vector of covariates as the compliance type probability, $\mathbf{v}_i = [1, \text{age-20}, \text{motivate}, \text{educ}, \text{assert}, \text{single}, \text{econ}, \text{nonwhite}]$.

6.2. Prior specifications

We complete the model specifications by choosing prior distributions of the model parameters that imply certain beliefs about the treatment effect. Because these effects are complicated functions of the parameters, we utilize a simulation based approach to isolate the implication of the prior on the treatment effect. The general approach is as follows. We specify a prior distribution with a certain set of hyperparameters. We then

Table 7

Mean and quantiles of the empirical distribution for the change in depression score implied by the two different priors for each model

Simulated depression outcomes			Quantiles					
Model	Prior	Mean	5%	25%	50%	75%	95%	
TCM	Unrestricted	Neutral	0.00	−4.96	−3.14	−0.04	3.16	5.02
		Optimistic	−0.45	−5.47	−3.55	−0.58	2.80	4.62
TCM	Restricted	Neutral	0.02	−4.34	−2.93	−0.05	3.04	4.46
		Optimistic	−0.43	−4.82	−3.34	−0.60	2.68	4.09
GCM		Neutral	0.01	−6.35	−3.48	−0.02	3.54	6.39
		Optimistic	−0.45	−6.85	−3.91	−0.53	3.13	5.95

Results are for $\nu = 10$.

draw the model parameters from the prior distribution. Given the parameter values and the covariates and random assignment of each sample subject we then generate the subject type or the intake and the corresponding potential outcome for each subject. We repeat this process many times (say 10,000 times) and compute the distribution of the potential outcomes in terms of the means and quantiles. If these implied distributions embody the beliefs we require we accept the prior; otherwise we change the hyperparameters and re-do the simulation exercise. We utilize this approach to determine two different prior distributions. The first prior, the neutral prior, embodies the belief that the treatment effect is zero or close to zero. The second prior, the optimistic prior, embodies the belief that the treatment effect is negative. In addition, we enforce the requirement that each of these priors is similar in its implications across the two models.

Reasoning in the preceding manner, the prior in the type model

$$\pi(\psi) = \mathcal{N}_8(\alpha|\alpha_0, \mathbf{A}_0) \prod_{j=0,1} \prod_{k \in \mathcal{K}_j} \mathcal{N}_3(\beta_{jk}|\beta_{jk,0}, \mathbf{B}_{jk,0}) \times \mathcal{IG} \left(\eta_{jk}^2 \left| \frac{n_{jk,0}}{2}, \frac{d_{jk,0}}{2} \right. \right)$$

in the neutral case has hyperparameters

$$\alpha_0 = (0, 0, 0, 0, 0, 0, 0, 0), \quad \mathbf{A}_0 = 9\mathbf{I}_8$$

$$\beta_{0c,0} = (0, 0, 0), \quad \beta_{0n,0} = (0, 0, 0),$$

$$\beta_{1c,0} = (0, 0, 0), \quad \mathbf{B}_{jk,0} = 9\mathbf{I}_3$$

$$\eta_{jk,0} = 4.13, \quad d_{jk,0} = 1.06$$

whereas in the optimistic case it has hyperparameters

$$\alpha_0 = (-5.0, .03, .5, .1, 0, 0, 0, 0), \quad \mathbf{A}_0 = 9\mathbf{I}_8$$

$$\beta_{0c,0} = (2.0, -1.5, 1.0), \quad \beta_{0n,0} = (1.75, -1.5, 1.0),$$

$$\beta_{1c,0} = (1.0, -1.5, 1.0), \quad \mathbf{B}_{jk,0} = 9\mathbf{I}_3$$

$$\eta_{jk,0} = 4.13, \quad d_{jk,0} = 1.06.$$

The latter choices are loosely based on the estimation results obtained by Yau and Little (2001) from a different but related sample. We slightly adjust these hyperparameters for the restricted version of the type model. In that case, the prior mean of β_{1c} in the neutral case is set to $\beta_{1c,0} = (-0.5, 0, 0)$ and in the optimistic case to $\beta_{1c,0} = (0.5, -1.5, 1.0)$. The remaining hyperparameters are as above except for the scalar coefficient in the compliance model α_0 which is now set at 0.5. To show what these assumptions imply for the outcomes, we provide in Table 7 the mean and quantiles of the simulated distribution of the change in depression score from the neutral and the optimistic priors, averaged over 10,000 outcome distributions that were simulated from the prior. As intended, the neutral prior implies on average no change in the depression score while the optimistic prior implies a decrease in the depression score of about 0.45. Both priors imply a mean of 0.34 for the simulated binary intake variable. The table also provides the results for the simulated outcome distributions from the restricted version of the type confounder model.

We proceed in a similar fashion to set the prior

$$\pi(\psi) = \mathcal{N}_8(\gamma|\gamma_0, \mathbf{G}_0) \prod_{j=0}^1 \mathcal{IG} \left(\sigma_j^2 \left| \frac{n_{j,0}}{2}, \frac{d_{j,0}}{2} \right. \right) \times \mathcal{N}_3(\beta_j|\beta_{j,0}, \mathbf{B}_{j,0}) \mathcal{N}(\omega_j|m_{j,0}, M_{j,0})$$

in the general confounder model. The neutral prior in this case is given by the hyperparameter choices

$$\gamma_0 = \mathbf{0}, \quad \mathbf{G}_0 = 9\mathbf{I}_8$$

$$\nu_{0,j} = 4.13, \quad d_{0,j} = 1.06$$

$$\beta_{0,0} = (-.1, 0, 0), \quad \beta_{1,0} = (-.2, 0, 0), \quad \mathbf{B}_{j,0} = 9\mathbf{I}_3$$

$$m_{j,0} = 0, \quad M_{j,0} = 4.$$

Under the optimistic prior the choices are identical except that now

$$\beta_{0,0} = (1.7, -1.5, 1.0), \quad \beta_{1,0} = (.6, -1.5, 1.0).$$

As shown in Table 7 these prior assumptions imply a distribution for the change in depression score that is comparable to that from the type model. As above, the mean of the simulated binary intake variable is 0.34 under each prior specification.

6.3. Results

In Table 8 we report the log marginal likelihoods from the fitting of the type and general confounder models, for each prior and for each of five different values of the degree of freedom parameter. These and subsequent results are based on 10,000 MCMC draws beyond a burn-in of a 1000 iterations. A number of interesting patterns emerge. First, for each value of ν , the model specifications with the optimistic prior are preferred to those with the neutral prior. Second, the restricted type confounder model is dominated by the unrestricted type model. This finding is in line with the recent work on the type model (Hirano et al., 2000; Jo, 2002; Frangakis et al., 2004) and reinforces the point that modeling the compliance probability through covariates is vital. Finally, the $\nu = 5$ models with the optimistic prior have the highest marginal likelihood. Note that while the general confounder model seems to be preferred by these data, it has only a slightly larger log marginal likelihood than the type model. Given these results we focus our subsequent detailed analysis on the best supported $\nu = 5$ and optimistic prior models.

To begin, we report the parameter estimates in Table 9. Columns (2)–(4) summarize the results from the fitting of the type confounder model, while columns (5)–(6) contain the results for the general confounder model. Information in the table is separated by the outcome, type and intake models. The first eight rows of the table deal with the compliance type and intake models. We notice that, except for the intercept, the posterior means of α and γ are almost identical. This suggests that the marginal modeling of type is similar in some respects to the marginal modeling of the intake. This correspondence could not hold, of

Table 8

Log marginal likelihoods of the type and general confounder models for the neutral and optimistic priors and different values of the degree of freedom parameter

Log marginal likelihoods for model comparison						
ν	TCM				GCM	
	Unrestricted		Restricted		Neutral	Optimistic
	Neutral	Optimistic	Neutral	Optimistic		
5	-805.57	-804.15	-827.78	-827.15	-804.05	-802.92
10	-805.77	-804.32	-823.47	-822.70	-804.15	-803.12
15	-807.43	-805.94	-824.96	-824.80	-805.63	-804.65
20	-808.50	-807.00	-826.53	-825.70	-806.50	-805.51
25	-809.35	-807.86	-831.06	-830.53	-807.03	-806.17

Table 9

Estimation results from the type and the general confounder models with the optimistic prior and $\nu = 5$, based on 10,000 runs of the MCMC algorithms (1000 burn-in cycles)

	Type confounder model			General confounder model	
	α			γ	
Intercept		-4.13(0.80)		-4.01(0.88)	
Age-20		0.05(0.01)		0.04(0.01)	
Motivate		0.35(0.09)		0.37(0.09)	
Educ		0.16(0.04)		0.16(0.04)	
Assert		-0.26(0.08)		-0.25(0.09)	
Married		-0.33(0.15)		-0.36(0.15)	
Econ		-0.11(0.09)		-0.11(0.10)	
Nonwhite		-0.20(0.18)		-0.17(0.18)	
	β_{0c}	β_{0n}	β_{1c}	β_0	β_1
Intercept	0.40(0.93)	2.19(0.45)	0.55(0.42)	1.72(0.34)	0.38(0.36)
Basedep	-1.35(0.54)	-1.54(0.27)	-1.08(0.28)	-1.47(0.21)	-1.03(0.26)
Baserisk	1.53(0.73)	0.65(0.40)	0.86(0.40)	0.87(0.31)	0.73(0.38)
	η_{0c}^2	η_{0n}^2	η_{1c}^2	η_0^2	η_1^2
	0.29(0.22)	0.35(0.17)	0.27(0.18)	0.37(0.04)	0.37(0.07)
Confounding				ρ_0	ρ_1
				0.13(0.13)	0.68(0.16)

Table entries are the posterior means (standard deviation in parentheses).

Table 10

Inefficiency factors from the fitting of the type and general confounder models with the optimistic prior and $\nu = 5$

Inefficiency factors									
TCM	β'_{0c}	6.10	5.79	4.27					
	β'_{0n}	3.70	2.63	2.85					
	β'_{1c}	2.28	1.76	1.75					
	α'	2.73	2.65	3.45	2.32	2.17	2.54	4.17	2.21
GCM	β'_0	1.61	1.77	1.77					
	β'_1	4.33	2.95	3.11					
	γ'	12.69	16.97	10.70	14.73	6.35	4.45	9.08	5.95

course, in settings where the intake takes more levels, for example when the intake is an ordinal variable. Furthermore, this also suggests that it is rather vital in this context to model subject type probability with the help of covariates if the subject type model is to compete with the general confounder model. As the results from the model comparison in Table 8 show the subject type model in which the type probability is independent of covariates is least preferred.

The next five rows are concerned with the outcome models. In parallel with our findings in the simulation study, the parameter β_{0c} is estimated less precisely than β_{0n} and β_{1c} whereas in the general confounder model, the elements in β_1 are estimated less precisely than β_0 , as shown in Table 10.

In the final row of Table 9, that relate to the general confounder model, we give estimates of the posterior means and standard deviations of the correlation coefficients $\rho_j = \omega_j/\eta_j$. These estimates indicate that individuals with unobserved factors that are associated with higher depression scores are more likely to participate in the program. Just as in the simulation study, ρ_1 has a higher posterior standard deviation than ρ_0 .

Next, we present the posterior mean of the compliance probability for subjects in the control arm in Fig. 2. As above, probabilities between 0 and 0.5 (which can be taken to indicate a never taker) are plotted below the horizontal axis and those above 0.5 (and suggest a complier) are plotted above the horizontal axis. In the type approach, the treatment effect is calculated by stratifying on compliance type. We see that compliers are subjects that appear to have similar levels of the observed covariates with the exception of the marriage indicator (Fig. 3). Never takers, on the other hand, seem to be quite different from the always takers. They are younger, less motivated and less educated. We also observe that compliers and never takers have similar average values of the baseline depression and risk scores covariates that appear in the outcome model.

We now compare the estimates of the predictive treatment effects from each approach. In Figs. 4 and 5 we graph the estimated predictive marginal densities of the potential outcomes within the subject type and the general confounder approaches, respectively. In each case, the predictive density of the potential outcome from program participation (dotted line) is to the left of the density from

Table 11

Estimates of predictive average and quantile causal effects for type and general confounder models with the optimistic priors and $\nu = 5$

Estimated predictive treatment effects		Average		Quantile			
Model	Treatment effect		.05	.25	.50	.75	.95
TCM	Complier treatment effects	−0.29	−0.23	0.25	−0.29	−0.33	−0.38
GCM	Population treatment effects	−0.50	−0.50	−0.48	−0.50	−0.53	−0.55
	Complier treatment effects	−0.54	−0.48	−0.50	0.52	−0.56	−0.66

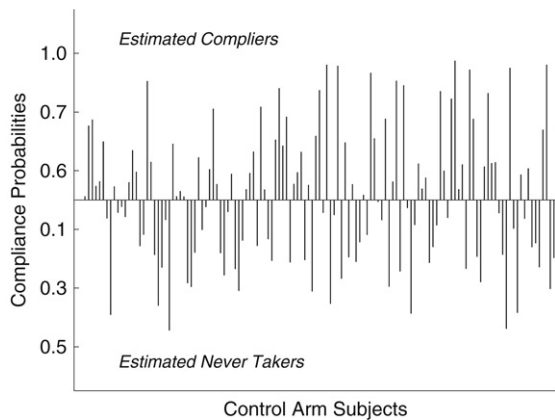


Fig. 2. Posterior compliance probabilities.

Variable	Sample Means by Intake and Type		
	$j = 0$		$j = 1$
	Complier	Never Taker	Complier
Age	40.67	32.46	39.96
Motivate	5.52	5.13	5.47
Edu	13.89	12.84	13.72
Assert	2.92	2.21	2.95
Marr	0.39	0.40	0.36
Econ	3.27	3.75	3.58
Nonw	0.12	0.25	0.16
Depress	2.52	2.45	2.41
Risk	1.69	1.67	1.67

Fig. 3. Sample means by intake and type.

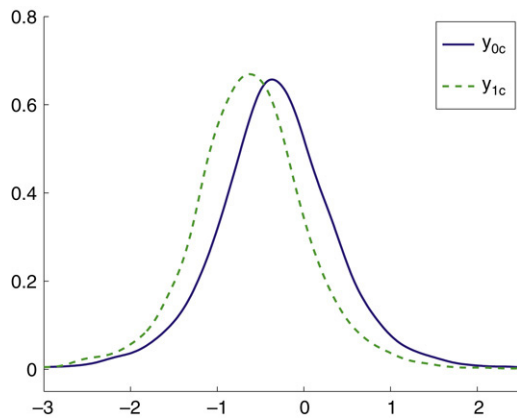


Fig. 4. Predictive marginal densities of the potential outcomes (for compliers) from the type confounder approach.

non-program participation (solid line). This implies that program participation is beneficial. The magnitude of the improvement is, however, larger from the general confounder model.

The precise numerical improvements from each model are reported in the first two rows of Table 11. We give the predictive

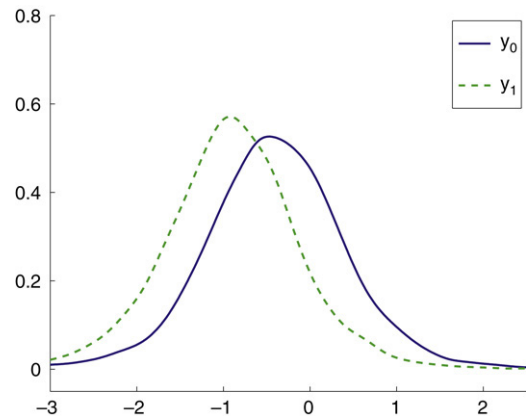


Fig. 5. Predictive marginal densities of the potential outcomes from the general confounder approach.

average causal effect and various quantile effects. As implied by the graphs, the various predictive quantile and mean treatment effects are negative, but different across the two models. Specifically, the average predictive treatment effect for compliers from the type confounder model is a negative .30 points, while the average treatment effect from the general confounder model is a negative −.50 points. We also note that the average predictive effect from the type model is similar in magnitude to the complier average causal effect found previously by Skron Dahl and Rabe-Hesketh (2004). Both average effects are substantially larger than the −0.12 points improvement suggested by the naive association (or regression) calculation. We note that in the case of the type confounder model the sign of the confounding bias is not clearly visible from the coefficient estimates of β_{0c} and β_{1c} . In addition, the quantile treatment effects vary between −0.23 and −0.38 for the complier effects in the type model, and between −0.48 and −0.55 for the general confounder model. Also noteworthy is the predicted outcome gain for the subpopulations of compliers from the general confounder model. As we see from the estimates given in the last row in Table 11, these effects are larger than those for a random subject from the population, except at the lower quantiles of the predictive distribution.

In general, these results show that competing assumptions about the unobserved confounders can produce different evaluations of the treatment. Our marginal likelihood estimates suggest that the two models are roughly equivalent for these data with a small preference for the general confounder model.

7. Conclusion

There are many problems in practice, with unobserved confounders, which share the structure of the eligibility design that we have studied in this paper. Our methods provide the means to study such problems in two competing ways, and to examine what effect, if any, those competing ways have for the estimates of the various treatment effects.

We emphasize again that our modeling and inferential techniques did not require the unknowable joint distribution of

the potential outcomes. Our fitting methods are also efficient. We noted that, in the type model, estimation of the model can be a bit difficult if the proportion of compliers is small. In the general confounder model, the main problem can be the estimation of confounding parameter in the intake state. Neither problem arose, however, in our empirical analysis.

We conclude by mentioning that the analysis described in this paper extends readily to other settings. These extensions, for example involving clustered outcomes and binary responses, are ongoing and will be reported elsewhere.

Acknowledgements

The authors thank the editor, associate editor and the referees for their constructive and helpful comments.

References

- Albert, J., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* 88, 669–779.
- Barnard, J., Frangakis, C., Hill, J., Rubin, D.B., 2003. Principal stratification approach to broken randomized experiments: A case study of school choice vouchers in New York City (with discussion). *Journal of the American Statistical Association* 98, 299–323.
- Chib, S., 1995. Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association* 90, 1313–1321.
- Chib, S., 2007. Analysis of treatment response data without the joint distribution of potential outcomes. *Journal of Econometrics* 140, 410–412.
- Chib, S., Greenberg, E., 1994. Bayes inference in regression models with ARMA (p, q) errors. *Journal of Econometrics* 64, 183–206.
- Chib, S., Greenberg, E., 1995. Understanding the Metropolis–Hastings algorithm. *American Statistician* 49, 327–335.
- Chib, S., Greenberg, E., 1998. Analysis of multivariate probit models. *Biometrika* 85, 347–361.
- Chib, S., Hamilton, B., 2000. Bayesian analysis of cross section and clustered data treatment models. *Journal of Econometrics* 97, 25–50.
- Chib, S., Hamilton, B., 2002. Semiparametric bayes analysis of longitudinal data treatment models. *Journal of Econometrics* 110, 67–89.
- Chib, S., Jacobi, L., 2007. Modeling and calculating the effect of treatment at baseline from panel outcomes. *Journal of Econometrics* 140, 781–801.
- Chib, S., Jeliazkov, I., 2001. Marginal likelihood from the Metropolis–Hastings output. *Journal of the American Statistical Association* 96, 270–281.
- Clark, A.E., Oswald, E.J., 1994. Unhappiness and unemployment. *The Economic Journal* 104, 648–659.
- Deb, P., Munkin, M.K., Trivedi, P.K., 2006. Private insurance, selection, and health care use: A bayesian analysis of a Roy-type model. *Journal of Business and Economic Statistics* 24 (1), 403–415.
- Frangakis, C.F., Rubin, D.B., 1999. Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika* 86, 365–379.
- Frangakis, C.E., Brookmeyer, R.S., Varadhan, R., Safaeian, M., Vlahov, D., Strathdee, S.A., 2004. Methodology for evaluating a partially controlled longitudinal treatment using principal stratification, with application to a needle exchange program. *Journal of the American Statistical Association* 99, 239–249.
- Heckman, J., Honore, B., 1990. The empirical content of the Roy model. *Econometrica* 58, 1121–1149.
- Heckman, J.J., Tobias, J.L., Vytlacil, E., 2001. Four parameters of interest in the evaluation of social programs. *Southern Economic Journal* 68, 210–223.
- Heckman, J., Vytlacil, E., 2005. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica* 73, 669–738.
- Hirano, K., Imbens, G.W., Rubin, D.B., Zhou, X., 2000. Assessing the effect of an influenza vaccine in an encouragement design. *Biostatistics* 1, 69–88.
- Imbens, G.W., Angrist, J., 1994. Identification and estimation of local average treatment effects. *Econometrica* 62, 467–475.
- Imbens, G.W., Rubin, D.B., 1997. Bayesian inference for causal effects in randomized trials with noncompliance. *The Annals of Statistics* 25, 305–327.
- Jo, B., 2002. Statistical power in randomized trials with noncompliance. *Psychological Methods* 7, 178–193.
- Levy, D.E., O'malley, J.A., Normand, S.T., 2004. Covariate adjustment in clinical trials with non-ignorable missing data and non-compliance. *Statistics in Medicine* 23, 2319–2339.
- Mealli, F., Imbens, G.W., Ferro, S., Biggeri, A., 2004. Analyzing a randomized trial on breast self-examination with non-compliance and missing outcomes. *Biostatistics* 5, 207–222.
- Pearl, J., 2000. *Causality*. Cambridge University Press, Cambridge.
- Sommer, A., Zeger, S., 1991. On estimating efficacy in clinical trials. *Statistics in Medicine* 10, 45–52.
- Skrondahl, A., Rabe-Hesketh, S., 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*. Chapman & Hall/CRC, New York.
- Ten have, T.R., Joffe, M., Cary, M., 2003. Causal logistic models for non-compliance under randomized treatment with univariate binary response. *Statistics in Medicine* 22, 1255–1283.
- Vinokur, A.D., Price, R.H., Schul, Y., 1995. Impact of JOBS intervention on unemployed workers varying in risk for depression. *American Journal of Community Psychology* 19, 543–562.
- Yau, L., Little, R., 2001. Inference for the complier-average causal effect from longitudinal data subject to noncompliance and missing data, with application to a job training assessment for the unemployed. *Journal of the American Statistical Association* 96, 1232–1244.