Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome

Author(s): P. R. Rosenbaum and D. B. Rubin

Source: *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 45, No. 2 (1983), pp. 212-218
Published by: Wiley for the Royal Statistical Society
Stable URL: https://www.jstor.org/stable/2345524
Accessed: 25-09-2018 02:21 UTC

J. R. Statist. Soc. B (1983),
45, *No.* 2, *pp.* 212–218

Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome

By P. R. ROSENBAUM and D. B. RUBIN†

University of Wisconsin–Madison, USA

[Received May 1981. Revised May 1982]

SUMMARY

This paper proposes a simple technique for assessing the range of plausible causal conclusions from observational studies with a binary outcome and an observed categorical covariate. The technique assesses the sensitivity of conclusions to assumptions about an unobserved binary covariate relevant to both treatment assignment and response. A medical study of coronary artery disease is used to illustrate the technique.

Keywords: CATEGORICAL DATA; CAUSAL EFFECTS; INCOMPLETE DATA; LOGISTIC MODELS; NON-RANDOMIZED STUDIES; SUBCLASSIFICATION

1. INTRODUCTION AND NOTATION

Inevitably, the results of clinical studies are subject to dispute. In observational studies, one basis for dispute is obvious: since patients were not assigned to treatments at random, patients at greater risk may be over-represented in some treatment groups. This paper proposes a method for assessing the sensitivity of causal conclusions to an unmeasured patient characteristic relevant to both treatment assignment and response. Despite their limitations, observational studies will continue to be a valuable source of information, and therefore it is prudent to develop appropriate methods of analysis for them.

Our sensitivity analysis consists of the estimation of the average effect of a treatment on a binary outcome variable after adjustment for observed categorical covariates and an unobserved binary covariate u, under several sets of assumptions about u. Both Cornfield *et al.* (1959) and Bross (1966) have proposed guidelines for determining whether an unmeasured binary covariate having specified properties could explain all of the apparent effect of a treatment, that is, whether the treatment effect, after adjustment for u could be zero. Our method has two advantages: first, Cornfield *et al.* (1959) and Bross (1966) adjust only for the unmeasured binary covariate u, whereas we adjust for measured covariates in addition to the unmeasured covariate u. Second, Cornfield *et al.* (1959) and Bross (1966, 1967) only judge whether the effect of the treatment could be zero having adjusted for u, where Cornfield *et al.* (1959) employ an implicit yet extreme assumption about u. In contrast, we provide actual estimates of the treatment effect adjusted for both u and the observed categorical covariates under any assumption about u.

In principle, the ith of the N patients under study has both a binary response r_{1i} that would have resulted if he had received the new treatment, and a binary response r_{0i} that would have resulted if he had received the control treatment. In this formulation, treatment effects are comparisons of r_{1i} and r_{0i}, such as $r_{1i} - r_{0i}$. Since each patient receives only one treatment, either r_{1i} or r_{0i} is observed, but not both, and therefore comparisons of r_{1i} and r_{0i} imply some degree of speculation. Treatment effects defined as comparisons of the two potential responses, r_{1i} and r_{0i}, of individual patients are implicit in Fisher's (1953) randomization test of the sharp null hypothesis that $r_{1i} = r_{0i}$, $i = 1, \ldots, N$. Such definitions are used explicitly by Kempthorne (1952)

† *Present address*: Dr Donald B. Rubin, Departments of Statistics and Education, University of Chicago, 5734 University Avenue, Chicago, Illinois 60637, USA.

 0035–9246/83/45212 $2.00

in his discussion of randomization-based inference in experiments, and by Rubin (1977, 1978), Hamilton (1979) and Rosenbaum and Rubin (1983) in discussions of observational studies. The definition does contain some implicit assumptions, such as the assumption of non-interference between patients; see Cox (1958, Chapter 2) or Rubin (1978, Section 2.3) for discussion. Here, the N patients in the study are viewed as a simple random sample from some population, and the average treatment effect is defined as

$$E(r_1) - E(r_0) = \text{pr}(r_1 = 1) - \text{pr}(r_0 = 1) = \tau_1 - \tau_0, \tag{1.1}$$

where $E(\ )$ and $\text{pr}(\ )$ denote expectation and probability, respectively, in the population.

For the ith patient of N patients in the study ($i = 1, \ldots, N$) let z_i be the indicator for treatment assignment, and let $z_i = 1$ if patient i is assigned to the new treatment, and $z_i = 0$ if patient i is assigned to the control treatment. Suppose that patients have been stratified or subclassified into one of J subclasses on the basis of an observed categorical covariate, and that patient i falls in subclass s_i, $i = 1, 2, \ldots, N$, where s_i is an integer between 1 and J. The population model for the subclassifying variable s is assumed to be a saturated multinomial.

If the study had been a randomized block experiment with blocks defined by subclasses, treatment assignment z and response (r_1, r_0) would be conditionally independent given s, or in Dawid's (1979) notation:

$$(r_1, r_0) \perp\!\!\!\perp z \mid s. \tag{1.2}$$

Moreover, in such a study, $1 > \text{pr}(z = 0 \mid s) > 0$ for all s. These conditions are not known to hold in an observational study. Generally, we shall say that treatment assignment is *strongly ignorable given some set of covariates* $\mathbf{v}$ if $(r_1, r_0) \perp\!\!\!\perp z \mid \mathbf{v}$ and $1 > \text{pr}(z = 0 \mid \mathbf{v}) > 0$ for all possible $\mathbf{v}$. For brevity, when treatment assignment is strongly ignorable given the observed covariates, we shall say simply that *treatment assignment is strongly ignorable*. If treatment assignment is strongly ignorable, then it is ignorable in Rubin's (1978) sense, but the converse is not true.

We develop and apply a method to aid in judging the sensitivity of conclusions to certain plausible variations in assumptions about an unobserved binary covariate u. In particular, we assume that treatment assignment is not strongly ignorable given s, but is strongly ignorable given s and u; that is, we will assume that (1.2) may be false but that

$$(r_1, r_0) \perp\!\!\!\perp z \mid s, u \tag{1.3}$$

is true, where $1 > \text{pr}(z = 0 \mid s, u) > 0$ for all s, u. If conclusions are insensitive over a range of plausible assumptions about u, the number of interpretations of the data is reduced, and causal conclusions are more defensible.

An alternative approach is briefly mentioned in the review by Schlesselman (1978). In that approach, however, the parameter used to measure the effect of the treatment, namely,

$$\frac{\text{pr}(r_z = 0 \mid z = 0, u = 1, s)}{\text{pr}(r_z = 0 \mid z = 1, u = 1, s)} \tag{1.4}$$

(or r_A in Schlesselman's notation), refers only to a subpopulation of patients defined by the unobserved variable u, that is, the subpopulation with $u = 1$; as a result, as assumptions about u are varied, the meaning of the parameter changes. Unless additional assumptions are made about the absence of certain interaction effects (in particular, that expression (1.4) equals the corresponding expression with $u = 0$), this parameter is not suitable for sensitivity analyses that consider various assumptions about u. Moreover, if the problem is formulated in terms of the scalar observable response variable r_z, then it is impossible to state the crucial assumption (1.3) under which the sensitivity analysis is correct.

2. THE STRUCTURE OF THE SENSITIVITY ANALYSIS

Because, by (1.3), (r_1, r_0) and z are conditionally independent given u and s, we can write the joint distribution of (r_t, z, u, s) for $t = 0, 1$ as

$$\text{pr}\,(r_t \mid u, s)\,\text{pr}\,(z \mid u, s)\,\text{pr}\,(u \mid s)\,\text{pr}\,(s),$$

where without loss of generality

$$\text{pr}\,(s = j) = \phi_j; \quad \sum_{j=1}^{J} \phi_j = 1, \quad \text{pr}\,(u = 0 \mid s) = \pi_s, \quad s = 1, \ldots, J, \tag{2.1}$$

$$\text{pr}\,(z = 0 \mid u, s) = [1 + \exp\,(\gamma_s + u\alpha_s)]^{-1}, \quad s = 1, \ldots, J; \quad u = 0, 1; \tag{2.2}$$

and

$$\text{pr}\,(r_t = 0 \mid u, s) = [1 + \exp\,(\beta_{st} + u\delta_{st})]^{-1}, \quad s = 1, \ldots, J; \quad u = 0, 1; \quad t = 0, 1. \tag{2.3}$$

The parameters $\phi_1, \ldots, \phi_J$ give the probabilities in each of the J subclasses; π_s gives the probability that $u = 0$ in subclass s; γ_s gives the log odds of assignment to treatment 0 in subclass s when $u = 0$, and $\gamma_s + \delta_s$ gives the corresponding log odds when $u = 1$; and finally, β_{st} gives the log odds that $r_t = 0$ in subclass s when $u = 0$, while $\beta_{st} + \delta_{st}$ gives the corresponding log odds when $u = 1$. A formally similar model with an unobserved binary covariate appears in Rubin (1978, Section 4.2).

Only some of the parameters in (2.1)–(2.3) can be estimated from observed data because u is never observed and r_t is only observed when $t = z$. Obviously, the ϕ_s can be estimated from the observed proportions of patients in each subclass. As we show explicitly in Section 4, for fixed values of sensitivity parameters π_s, α_s, δ_{st} there exist unique maximum likelihood estimates of the remaining parameters, γ_s and β_{st}. Since the average treatment effect (1.1) is $\tau_1 - \tau_0$ where

$$\tau_t = \sum_{s=1}^{J} \phi_s \left[(1 - \pi_s)\,\frac{\exp\,(\beta_{st} + \delta_{st})}{1 + \exp\,(\beta_{st} + \delta_{st})} + \pi_s\,\frac{\exp\,(\beta_{st})}{1 + \exp\,(\beta_{st})} \right] \tag{2.4}$$

the maximum likelihood estimate of the treatment effect can be calculated from the maximum likelihood estimates of ϕ_s and β_{st} corresponding to fixed values of the sensitivity parameters π_s, α_s and δ_{st}.

3. AN EXAMPLE USING DATA ON CORONARY ARTERY DISEASE

Before discussing the technical details underlying the sensitivity analysis, we present the results of applying it to an example concerning symptomatic relief from coronary artery disease†. Treatment 1 is coronary artery bypass surgery; treatment 0 is medical therapy. The response (r_1, r_0) is functional improvement 6 months after cardiac catheterization, with $r_1 = 1$ or $r_1 = 0$ indicating improvement or no improvement respectively under surgical treatment, and $r_0 = 1$ or $r_0 = 0$ indicating improvement or no improvement, respectively, under medical treatment.

Here, patients are stratified by the multivariate method described in Rosenbaum and Rubin (1983) that balances 74 observed covariates. The observed proportion improved within each subclass is displayed in Table 1. In examining this table, it must be remembered that there exists evidence for a placebo effect of bypass surgery (Benson and McCalie, 1979).

When treatment assignment is strongly ignorable given the subclasses, then direct adjustment with subclass total weights yields the maximum likelihood estimate of the average treatment effect (1.1) under the saturated multinomial model for s. The directly adjusted proportions improved are

† The data are used to illustrate methodology, and do not constitute a study of coronary artery disease.

TABLE 1

Proportion of patients improved at 6 months in each subclass

Subclass†	*Treatment*	*Number of patients*	*Proportion improved*	*Standard error*
1	Surgical	26	0.54	0.10
	Medical	277	0.35	0.03
2	Surgical	68	0.70	0.06
	Medical	235	0.40	0.03
3	Surgical	98	0.70	0.05
	Medical	205	0.35	0.03
4	Surgical	164	0.71	0.04
	Medical	139	0.30	0.04
5	Surgical	234	0.70	0.03
	Medical	69	0.39	0.06

† The subclasses were constructed by the method of Rosenbaum and Rubin (1983) in which the conditional probability of surgical treatment given the observed covariates is estimated, and patients with similar estimated probabilities are placed in the same subclass. The subclasses are predictive of treatment assignments; they are not prognostic subclasses. This method balances observed covariates within each subclass.

0.36 for medicine and 0.67 for surgery, with standard errors 0.04 and 0.06 respectively (calculated following Mosteller and Tukey, 1977, Chapter 11c).

In order to study the sensitivity of estimates to the assumption of strongly ignorable assignment, we now assume that treatment assignment is not strongly ignorable, but rather that treatment assignment is strongly ignorable given s and unobserved binary covariate u. That is, we assume (1.2) is false but (1.3) is true. Table 2 displays the sensitivity of the estimate of the average treatment effect to 24 sets of assumptions about u, where α is the increase in the log odds of surgery associated with $u = 1$ rather than $u = 0$, δ_t is the increase in the log odds of improvement under treatment t associated with $u = 1$, and $\pi = \text{pr}(u = 0)$. In order to limit the size of the sensitivity analysis, $(\alpha, \delta_1, \delta_0, \pi)$ is assumed in Table 2 to be the same across all subclasses. The general method presented in Sections 2 and 4 does not require this restriction; moreover, the Appendix displays subclass specific results which do not depend on this assumption.

In Table 2, the estimates of the proportion improved vary from 0.34 to 0.38 for medicine and from 0.63 to 0.70 for surgery; these ranges of values of $\hat{\tau}_0$ and $\hat{\tau}_1$ are about the same as the standard error of the directly adjusted proportions (i.e. 0.04 for medicine, 0.06 for surgery). Consequently, we see that this hypothetical, unobserved covariate u, which has defied the cardiologists' attempt to record all variables used in assigning treatments, would have to more than triple the odds of surgery and more than triple the odds of improvement, before altering the conclusion that the proportion improved under surgery far exceeds the proportion improved under medicine. Although this difference may reflect a placebo effect of surgery, the difference does not seem to be easily explained as the result of an imbalance due to the non-randomized nature of the study.

Another way of describing this analysis is to say that we have explored the extent to which the data might be an example of Simpson's paradox. For discussion from this perspective, see Lindley and Norvick (1981).

4. CALCULATING THE MAXIMUM LIKELIHOOD ESTIMATES OF τ_1 AND τ_0

Since r_{ti} is observed only if $t = z_i$, the likelihood of the parameters given the observed data $(r_{z_i i}, z_i, s_i)$ $i = 1, \ldots, N$ is

$$\prod_{i=1}^{N} \text{pr}(r_{t_i} \mid t = z_i, s_i)\, \text{pr}(z_i \mid s_i)\, \text{pr}(s_i), \tag{4.1}$$

TABLE 2

Effects of an unobserved two-category covariate u on the probability of substantial symptomatic improvement at 6 months for medical (M) and surgical (S) patients

Effect of u = 1 vs u = 0 on treatment assignment z	*Effect of u = 1 vs u = 0 on response under M*	*Effect of u = 1 vs u = 0 on response under S*	*Fraction of patients with u = 0: π* 0.1	0.5	0.9
Doubles the odds of surgery exp (α) = 2	Halves the odds of improvement exp (δ_0) = $\frac{1}{2}$	Halves the odds of improvement exp (δ_1) = $\frac{1}{2}$	S 0.67 M 0.36	S 0.68 M 0.35	S 0.68 M 0.36
		Doubles the odds of improvement exp (δ_1) = 2	S 0.66 M 0.36	S 0.65 M 0.35	S 0.66 M 0.36
	Doubles the odds of improvement exp (δ_0) = 2	Halves the odds of improvement exp (δ_1) = $\frac{1}{2}$	S 0.67 M 0.36	S 0.68 M 0.37	S 0.68 M 0.36
		Doubles the odds of improvement exp (δ_1) = 2	S 0.66 M 0.36	S 0.65 M 0.37	S 0.66 M 0.36
Triples the odds of surgery exp (α) = 3	Reduces by $\frac{2}{3}$ the odds of improvement exp (δ_0) = $\frac{1}{3}$	Reduces by $\frac{2}{3}$ the odds of improvement exp (δ_1) = $\frac{1}{3}$	S 0.68 M 0.35	S 0.70 M 0.34	S 0.69 M 0.35
		Triples the odds of improvement exp (δ_1) = 3	S 0.66 M 0.35	S 0.63 M 0.34	S 0.65 M 0.35
	Triples the odds of improvement exp (δ_0) = 3	Reduces by $\frac{2}{3}$ the odds of improvement exp (δ_1) = $\frac{1}{3}$	S 0.68 M 0.37	S 0.70 M 0.38	S 0.69 M 0.37
		Triples the odds of improvement exp (δ_1) = 3	S 0.66 M 0.37	S 0.63 M 0.38	S 0.65 M 0.37

where, by definition,

$$\text{pr}\,(s = j) = \phi_j, \tag{4.2}$$

and, by (2.1) and (2.2)

$$\text{pr}\,(z = 0 \mid s) = \frac{\pi_s}{1 + \exp\,[\gamma_s]} + \frac{(1 - \pi_s)}{1 + \exp\,[\gamma_s + \alpha_s]}, \tag{4.3}$$

and finally, by (1.3) and (2.3),

$$\begin{aligned}\text{pr}\,(r_t = 0 \mid z = t, s) &= \text{pr}\,(r_t = 0 \mid s, u = 0)\ \text{pr}\,(u = 0 \mid z = t, s) \\ &\quad + \text{pr}\,(r_t = 0 \mid s, u = 1)\ \text{pr}\,(u = 1 \mid z = t, s) = \frac{w_{sz}}{1 + e^{\beta st}} + \frac{1 - w_{sz}}{1 + e^{\beta st + \delta st}},\end{aligned} \tag{4.4}$$

where by (2.1), (2.2) and (4.3)

$$\begin{aligned}w_{sz} &= p\,(u = 0 \mid z, s) = \frac{\pi_s \exp\,[z\gamma_s]}{1 + \exp\,[\gamma_s]} \left\{ \frac{\pi_s \exp\,[z\gamma_s]}{1 + [\exp\,[\gamma_s]} + \frac{(1 - \pi_s) \exp\,[z(\gamma_s + \alpha_s)]}{1 + \exp\,[\gamma_s + \alpha_s]} \right\}^{-1}, \\ &= \pi_s \left\{ \pi_s + (1 - \pi_s) \frac{\exp\,(z\alpha_s)\,[1 + \exp\,(\gamma_s)]}{1 + \exp\,[\gamma_s + \alpha_s]} \right\}^{-1}.\end{aligned} \tag{4.5}$$

Note that the w_{sz}'s define the conditional distribution of the unobserved covariate u given the observed treatment assignment.

Since the implied model for the $2 \times 2 \times J$ table of counts for the observed data, (r_z, z, s), is saturated, the maximum likelihood estimate of pr $(s = j) = \phi_j$ is the proportion $\hat{\text{pr}}$ $(s = j)$ of patients in subclass j, the maximum likelihood estimate of pr $(z = 0 \mid s)$ is the observed proportion $\hat{p}$ $(z = 0 \mid s)$ of patients receiving treatment 0 in subclass s, and for $z = 0$ and 1, the maximum likelihood estimate of pr $(r_t = 0 \mid t = z, s)$ is the observed proportion $\hat{\text{pr}}$ $(r_t = 0 \mid t = z, s)$ of patients with response $r_z = 0$ among patients who received treatment z in subclass s.

For each subclass s, equation (4.3) with $\hat{\text{pr}}$ $(z = 0 \mid s)$ substituted for pr $(z = 0 \mid s)$ may be solved for the maximum likelihood estimate of γ_s; having found $\hat{\gamma}_s$ and calculated $\hat{w}_{st}$, equation (4.4) with $\hat{\text{pr}}$ $(r_t = 0 \mid t = z, s)$ substituted for pr $(r_t = 0 \mid t = z, s)$ may be solved for $\hat{\beta}_{st}$. The maximum likelihood estimates of ϕ_s and β_{st} can be substituted into (2.4) to find the maximum likelihood estimate of τ_t.

Equations (4.3) and (4.4) are both of the form

$$p = \frac{a}{1 + \exp(\theta)} + \frac{1 - a}{1 + \exp(\theta + b)} . \tag{4.6}$$

To find θ for fixed a, b and $p \neq 0$, we note that the right-hand side of (4.6) is montone in $\omega = \exp(\theta)$, taking values between 0 and 1, and thus (4.6) has one positive solution in ω. To find this solution, we note that (4.6) implies a quadratic equation in ω:

$$[p \exp(b)]\ \omega^2 + [(p - a) \exp(b) + p - 1 + a]\ \omega + (p - 1) = 0. \tag{4.7}$$

To calculate $(\hat{\gamma}_s, \hat{\beta}_{s0}, \hat{\beta}_{s1})$ for fixed $(\pi_s, \alpha_s, \delta_{s1}, \delta_{s0})$:

(a) Solve (4.3) for $\hat{\gamma}_s$ using (4.7) with $p = \hat{\text{pr}}$ $(z = 0 \mid s)$, $a = \pi_s$, $b = \alpha_s$.

(b) For $t = 0, 1$, find $\hat{w}_{st}$ using (4.5).

(c) For $t = 0, 1$, solve (4.4) for $\hat{\beta}_{st}$ using (4.7) with $p = \hat{\text{pr}}$ $(r_t = 0 \mid t = z, s)$, $a = \hat{w}_{st}$, $b = \delta_{st}$.

Steps (a), (b) and (c) are repeated for $s = 1, \ldots, J$ and the results combined using (2.4). The Appendix presents some specific results for each subclass of the example from Section 2.

ACKNOWLEDGEMENTS

The authors wish to acknowledge valuable discussions with Arthur P. Dempster and Roderick J. A. Little on the subject of this paper, and thank the editor and referees for helpful comments on an earlier draft of the paper. This work was partially supported by the U.S. Army Research Office through the Mathematics Research Center at the University of Wisconsin, under Contract No. DAAG29-80-C-0041, Grant No. P30-CA-14520 from the National Cancer Institute to the Wisconsin Clinical Cancer Center, the Health Resources Administration, and the Educational Testing Service.

REFERENCES

Bensen, H. and McCallie, D. (1979) Angina pectoris and the placebo effect. *New England J. Med.*, **300**, 1424–1428.

Bross, I. D. J. (1966) Spurious effects from an extraneous variable. *J. Chronic Diseases*, **19**, 637–647.

——— (1967) Pertinency of an extraneous variable. *J. Chronic Diseases*, **20**, 487–495.

Cornfield, J. *et al.* (1959) Smoking and lung cancer: recent evidence and a discussion of some questions. *J. Nat. Cancer Inst.*, **22**, 173–200.

Cox, D. R. (1958) *The Planning of Experiments*. New York: Wiley.

Dawid, A. (1979) Conditional independence in statistical theory (with Discussion). *J. R. Statist. Soc.* B, **41**, 1–31.

Fisher, R. A. (1953) *The Design of Experiments*. London: Hafner.

Hamilton, M. A. (1979) Choosing a parameter for a 2 X 2 table or 2 X 2 X 2 table analysis. *Amer. J. Epidemiol.* **109**, 362–375.

Kempthorne, O. (1952) *The Design and Analysis of Experiments*. New York: Wiley.

Lindley, D. V. and Novick, M. R. (1981) The role of exchangeability in inference. *Ann. Statist.* **9**, 45–58.

Mosteller, C. F. and Tukey, J. W. (1977) *Data Analysis and Regression*. Reading, Mass.: Addison–Wesley.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70**, 41–55.
Rubin, D. B. (1977) Assignment to treatment group on the basis of a covariate. *J. Educ. Statist.* **2**, 1–26.
——— (1978) Bayesian inference for causal effects: the role of randomization. *Ann. Statist.*, **6** 34–58.
Schlesselman, J. J. (1978) Assessing the effects of confounding variables. *Amer. J. Epidemiol.* **108**, 3–8.

APPENDIX

Subclass Specific Results

Table 2 is based on the assumption that the sensitivity parameters $(\alpha_s, \delta_{1s}, \delta_{0s}, \pi_s)$ do not depend on the subclass s. This assumption is not crucial to the conclusions in Section 3, although it does simplify presentation. This Appendix documents the stability of the estimates of the subclass specific average response to treatment t,

$$\text{pr}(r_t = 0 \mid s) = \frac{\pi_s}{1 + \exp(\beta_{st})} + \frac{1 - \pi_s}{1 + \exp(\beta_{st} + \delta_{st})} = 1 - \tau_{st};$$

Table 3 displays values of τ_{st} when $\alpha = 3$.

As was noted in Section 4, for $s = 1, 2, \ldots, s$, $\hat{\tau}_{st}$ depends only on the observed data and $\pi_s, \alpha_s, \delta_{st}$. Therefore, we may examine the effects of allowing the sensitivity parameters to vary from subclass to subclass by simply combining subclass specific estimates $\hat{\tau}_{st}$. For example, from Table 3 when $\alpha = 3$ for $\pi_1 = 0.1$, $\delta_{10} = \frac{1}{3}$, and $\delta_{11} = 3$, we have $\hat{\tau}_{10} = 0.348$ and $\hat{\tau}_{11} = 0.524$. If, in addition, we let $\pi_s = 0.9$, $\delta_{s0} = \delta_{s1} = \frac{1}{3}$ for $s = 2, 3, 4, 5$ we have

$$\hat{\tau}_0 = (0.348 + 0.394 + 0.343 + 0.291 + 0.377)/5 = 0.351$$

and

$$\hat{\tau}_1 = (0.524 + 0.723 + 0.718 + 0.720 + 0.704)/5 = 0.678,$$

since every subclass contains the same number of patients. By similar calculations it is seen that the estimate of the average treatment effect is greater than 0.25 for all combinations of $(\alpha_s, \delta_{s1}, \delta_{s0}, \pi_s)$ in Table 3.

TABLE 3

Effects of an unobserved two-category covariate u on the probability of improvement of 6 months for medical ($t = 0$) and surgical ($t = 1$) patients within each subclass; association with treatment assignment exp (α), *equal to 3*

Association with improvement exp (δ_t)	$\pi = p(u = 0)$		*Subclass* 1	2	3	4	5
$\frac{1}{3}$	0.1	S	0.555	0.710	0.710	0.718	0.705
		M	0.348	0.396	0.343	0.287	0.366
	0.5	S	0.600	0.741	0.737	0.735	0.713
		M	0.345	0.385	0.330	0.270	0.342
	0.9	S	0.573	0.723	0.718	0.720	0.704
		M	0.348	0.394	0.343	0.291	0.377
3	0.1	S	0.524	0.685	0.686	0.699	0.693
		M	0.351	0.404	0.355	0.310	0.411
	0.5	S	0.478	0.651	0.657	0.682	0.686
		M	0.355	0.415	0.371	0.334	0.443
	0.9	S	0.508	0.682	0.686	0.703	0.697
		M	0.353	0.407	0.359	0.313	0.405