129-72

INSTITUTE FOR RESEARCH ON POVERTY DISCUSSION PAPERS

SELECTION BIAS IN EVALUATING TREATMENT EFFECTS: THE CASE OF INTERACTION

Arthur S. Goldberger

UNIVERSITY OF WISCONSIN ~ MADISON

SELECTION BIAS IN EVALUATING TREATMENT EFFECTS: THE CASE OF INTERACTION

Arthur S. Goldberger

The research reported here was supported in part by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity pursuant to the provisions of the Economic Opportunity Act of 1964. The conclusions are the sole responsibility of the author. The author thanks Glen Cain for valuable suggestions.

June 1972

ABSTRACT

When individuals of varying true ability are non-randomly assigned to treatment and control groups, biased effects of treatment may be observed. In this paper, we assess the possibilities in terms of a simple test-score model which allows for an interaction between treatment and true ability. For the control group, pretest score and posttest score are assumed to be fallible measures of true ability. For the treatment group, pretest score is a fallible measure of true ability, while posttest score is a linear function of true ability, plus a random error. Posttest score is regressed on pretest score within each of the two groups.

We find that both the additive and non-additive effects of the treatment are biased when individuals are selected for treatment explicitly on the basis of true ability. However, when individuals are selected for treatment explicitly on the basis of pretest score, only the non-additive effect is biased. Furthermore, this bias is the same as that which would arise under random selection.

Our analysis is a formal one, regressions being expressed in terms of population parameters. We also evaluate the statistical efficiency of selection-on-the-basis-of-pretest relative to random selection.

SELECTION BIAS IN EVALUATING TREATMENT EFFECTS: THE CASE OF INTERACTION

Arthur S. Goldberger

1. Introduction

The argument that a spurious treatment effect will be observed when individuals of varying true ability are non-randomly assigned to treatment and control, was examined by Goldberger (1972). The examination ran in terms of a simple test-score model, with normally distributed true ability, pretest, and posttest. Regressing posttest on pretest and treatment, we found that the spurious effect arose when selection was based on true ability but not when selection was based on pretest. Although the analysis was confined to the situation where the true effect of the treatment was nil, it was clear that the findings would carry over to the situation where the true effect of the treatment was additive.

In the present paper we extend the discussion to the case of interaction -- where the true effect of the treatment may be nonadditive, varying linearly with true ability. We find that the favorable conclusions about pretest selection hold up, albeit in weakened form.

2. The Basic Model

We suppose that true ability x^* is normally distributed with expectation 0 and variance Q, and that pretest score x is an erroneous measure of true ability in the sense that

$$x = x^* + u \; , \qquad (1)$$

where u is normally distributed with expectation 0 and variance $(1 - P)Q/P$ with $0 < P < 1$. Posttest score is determined by true ability and treatment via

$$y = x^* + \alpha z + \beta zx^* + v = \alpha z + (1 + \beta z) x^* + v \tag{2}$$

where v is normally distributed with expectation 0 and variance $(1 - P)Q/P$. Here z is the binary variable indicating whether or not the individual received the treatment:

$$z = \begin{cases} 1 \text{ if received treatment} \\ 0 \text{ if did not receive treatment (i.e. control).} \end{cases}$$

We suppose that x^*, u, and v are independent and that v is independent of z.

Our selection procedures will be such that half of the population receives the treatment and half are control, so that

$$p_o = \text{Prob}\ \{z = 0\} = p_1 = \text{Prob}\ \{z = 1\} = \tfrac{1}{2},$$

$$E(z) = \tfrac{1}{2}, \qquad V(z) = 1/4 .$$

Our assumptions imply that x is normally distributed with expectation

$$E(x) = E(x^*) + E(u) = 0,$$

and variance

$$V(x) = V(x^*) + V(u) = Q/P.$$

The distribution of y will in general be non-normal with expectation

$$E(y) = E(x^*) + \alpha\ E(z) + \beta\ E(zx^*) + E(v) = \tfrac{1}{2}\ \alpha + \beta\ C(z,x^*);$$

the covariance on the right depends on the selection procedure.

The present model permits the treatment to have both an additive and a nonadditive effect on posttest score. Specifically, (2) says that

(3) $$E(y|x^*,\ z) = \alpha\ z + (1 + \beta z)\ x^*.$$

For the treatment group $(z = 1)$ the regression of posttest on true ability is

$$E(y|x^*,1) \equiv E(y|x^*,z{=}1) = \alpha + (1 + \beta)\ x^*,$$

while for the control group $(z = 0)$ the regression of posttest on true ability is

$$E(y|x^*,0) \equiv E(y|x^*,z{=}0) = x^*.$$

These two lines differ in general. If $\beta = 0$ their slopes are the same (no nonadditive effect), if $\alpha = 0$ their intercepts are the same (no additive effect). If both $\alpha = 0$ and $\beta = 0$, the two lines coincide (no effect at all) and we are reduced to the situation examined in Goldberger (1972).

Whatever the true effects may be, they would show up if y were regressed on x^*, z, and zx^* (equivalently, if y were regressed on x^* within each group separately). But in practice, with x^* unobserved, one can regress y on x, z, and zx to assess the treatment effect. Since x is an erroneous measure of x^*, this assessment may be biased. The bias presumably depends on the selection procedure -- the basis on which individuals were assigned to the treatment and control groups. We will consider three procedures: (o) random selection, (i) selection on true ability, and (ii) selection on pretest score. For each, we develop the within-group regressions of posttest on true ability and of posttest on pretest. The within-group regressions are translated into an overall regression with, and without, interaction.

When the need arises, we will presume that the true effects of the treatment are nonnegative, that is $\alpha \geq 0$ and $\beta \geq 0$.

3. Technical Digression

We first record some general results on the relation between within-group and overall regressions. Consider the joint distribution of three random variables r, s, z, where z is a binary variable taking on the value 0 with probability p_o and the value 1 with probability $p_1 = 1 - p_o$. The linear regression of s on r given that z = 0 we denote by

$$E(s|r,0) = \alpha_{oo} + \alpha_{1o}\, r \ ;$$

the linear regression of s on r given that z = 1 we denote by

$$E(s|r,1) = \alpha_{o1} + \alpha_{11}\, r \ .$$

Collectively, these within-group linear regressions will be denoted by

$$E(s|r,z) = \alpha_{oz} + \alpha_{1z}\, r \qquad (z = 0,1), \tag{4}$$

or equivalently as

$$E(s|r,z) = \alpha_{oo} + \alpha_{1o}\, r + (\alpha_{o1} - \alpha_{oo})z + (\alpha_{11} - \alpha_{1o})\, zr. \tag{5}$$

The latter form is the one which arises when a single linear regression of s on r, z, and the interaction term zr, is run over the full population. For simplicity, we use $E(.|.)$ to denote linear regressions regardless of whether the true regression function is in fact linear.

From the general theory of linear regression we know that the within-group slopes and intercepts can be expressed in terms of within-group moments as

$$\begin{aligned} \alpha_{1z} &= C(r,s|z)/V(r|z), \\ \alpha_{oz} &= E(s|z) - \alpha_{1z}\, E(r|z) \qquad (z = 0,1). \end{aligned} \tag{6}$$

Now consider the result of running a single linear-additive regression of s on r and z <u>without</u> <u>interaction</u> term over the entire population:

$$E^*(s|r,z) = \alpha_o + \alpha_1\, r + \alpha_2\, z. \tag{7}$$

These parameters can also be expressed in terms of within-group moments:

$$\begin{aligned} \alpha_o &= E(s|0) - \alpha_1\, E(r|0) \\ \alpha_1 &= \underset{z}{E}\, C(r,s|z)/\underset{z}{E}\, V(r|z), \\ \alpha_2 &= (E(s|1) - E(s|0)) - \alpha_1(E(r|1) - E(r|0)), \end{aligned}$$

where

$$\underset{z}{E}\, C(r,s|z) = p_o\, C(r,s|0) + p_1\, C(r,s|1) \tag{8}$$

and

$$\underset{z}{E}\, V(r|z) = p_o\, V(r|0) + p_1\, V(r|1) \tag{9}$$

are the average within-group covariance and variance respectively.

The parameters in the additive regression are related to those in the nonadditive regression as follows: Using (8), (9), (6) we write

$$\begin{aligned} \alpha_1 &= w\, C(r,s|0)/V(r|0) + (1\text{-}w)\, C(r,s|1)/V(r|1) \\ &= w \quad \alpha_{1o} \quad + (1\text{-}w)\alpha_{11}\ , \end{aligned}$$

where

$$w = p_o\, V(r|0)/\underset{z}{E}\, V(r|z), \quad (1\text{-}w) = p_1\, V(r|1)/\underset{z}{E}\, V(r|z).$$

Thus the overall r-slope in (7) is a weighted average of the within-group r-slopes in (4) or (5). Proceeding, we find

$$\begin{aligned} \alpha_o &= E(s|0) - (w\, \alpha_{1o} + (1 - w)\, \alpha_{11})\, E(r|0) \\ &= (E(s|0) - \alpha_{1o}\, E(r|0)) + (1 - w)(\alpha_{1o} - \alpha_{11})E(r|0) \\ &= \alpha_{oo} + (1 - w)(\alpha_{1o} - \alpha_{11})E(r|0), \end{aligned}$$

and

$$\alpha_2 = (\alpha_{o1} - \alpha_{oo}) + (\alpha_{11} - \alpha_{1o})(w\, E(r|1) + (1 - w)\, E(r|0)).$$

These formulas simplify in special cases. If the variance of r is the same within each of the two groups, $V(r|0) = V(r|1)$, then $w = p_o$ and $1 - w = 1 - p_o = p_1$, so that

$$\alpha_1 = p_o \alpha_{1o} + p_1 \alpha_{11} ,$$

$$\alpha_o = \alpha_{oo} + p_1(\alpha_{1o} - \alpha_{11})E(r|0)$$

$$\alpha_2 = (\alpha_{o1} - \alpha_{oo}) + (\alpha_{11} - \alpha_{1o})(p_o E(r|1) + p_1 E(r|0)).$$

If further $p_o = \frac{1}{2} = p_1$, then

$$\alpha_o = \alpha_{oo} + \tfrac{1}{2} (\alpha_{1o} - \alpha_{11})E(r|0), \tag{10}$$

$$\alpha_1 = \tfrac{1}{2} (\alpha_{1o} + \alpha_{11}) \tag{11}$$

$$\alpha_2 = (\alpha_{o1} - \alpha_{oo}) + (\alpha_{11} - \alpha_{1o}) E(r).$$

If also $E(r) = 0$ then this last expression simplifies to

$$\alpha_2 = (\alpha_{o1} - \alpha_{oo}). \tag{12}$$

In our applications we will have $V(r|z)'$ constant, $p_o = \frac{1}{2}$, $E(r) = 0$, so that (10)-(12) hold. The overall r-slope is just the average of the within-group r-slopes, and the overall z-coefficient is just the difference of the within-group intercepts.

4. Random Selection

In case (o), individuals are assigned to the treatment and the control groups in a manner which is random both with respect to true ability and the error component of pretest.

Regressions on true ability. Since z is independent of x*, we immediately have

$$E(x^*|z) = E(x^*) = 0, \qquad V(x^*|z) = V(x^*) = Q.$$

Then writing

$$y = \alpha z + (1 + \beta z) x^* + v,$$

we use the independence of z, x*, and v to find

$$E(y|z) = \alpha z + (1 + \beta z) E(x^*|z) = \alpha z ,$$

$$C(x^*,y|z) = (1 + \beta z) V(x^*|z) = (1 + \beta z) Q.$$

Applying (6) with y taking the role of s, and x* taking the role of r, we obtain the slopes and intercepts of the within-group regressions of posttest on true ability:

$$(13) \qquad \begin{aligned} \alpha_{1z} &= (1 + \beta z)Q/Q = (1 + \beta z) , \\ \alpha_{oz} &= \alpha z - (1 + \beta z)\, 0 = \alpha z \qquad (z = 0, 1). \end{aligned}$$

As was to be expected, these accurately capture both true effects of the treatment: $\alpha_{11} - \alpha_{1o} = \beta$, the interaction effect, and $\alpha_{o1} - \alpha_{oo} = \alpha$, the additive effect. Applying (11)-(12) with y taking the role of s, and x* taking the role of r, we obtain the r- and z- coefficients of the overall additive regression of y on x*:

$$(14) \qquad \alpha_1 = 1 + \tfrac{1}{2}\beta, \qquad \alpha_2 = \alpha .$$

The z-coefficient directly captures the additive effect of the treatment, and the interaction effect can be recovered from the r-coefficient as $\beta = 2(\alpha_1 - 1)$.

<u>Regression on pretest score</u>. With z independent of u as well as of x^*, it is independent of x, so that

$$E(x|z) = E(x) = 0, \qquad V(x|z) = V(x) = Q/P,$$

and

$$C(x,y|z) = C((x^* + u,y)|z) = C(x^*,y|z) = (1 + \beta z)\,Q.$$

Applying (6) with y taking the role of s, and x taking the role of r, we obtain the slopes and intercepts of the within-group regressions of posttest on pretest:

$$\text{(15)} \qquad \alpha_{1z} = (1 + \beta z)Q/(Q/P) = P(1 + \beta z),$$

$$\alpha_{oz} = \alpha\, z + P(1 + \beta z)\, 0 = \alpha\, z \qquad (z = 0,\ 1).$$

The interaction effect of the treatment is attenuated but the additive effect is accurately captured: $\alpha_{11} - \alpha_{1o} = P\ \beta$ but $\alpha_{o1} - \alpha_{oo} = \alpha$. Attenuation of the slopes is of course inevitable with $P < 1$ even when the selection is random. Applying (11)-(12) with y for s, and x for r, we obtain the coefficients in the overall additive regression of posttest on pretest:

$$\text{(16)} \qquad \alpha_1 = P(1 + \tfrac{1}{2}\,\beta), \qquad \alpha_2 = \alpha\,.$$

A comparison with (14) confirms that under random selection the measurement error does not bias an additive treatment effect.

5. Selection on Basis of True Ability

In case (i), individuals are assigned to the treatment group or the control group according as their true ability is below or above the mean true ability in the population:

$$z = \begin{cases} 1 \text{ if } x^* \leq 0 \\ 0 \text{ if } x^* > 0 \end{cases} .$$

Since z is determined exactly by x* it will be independent of u as well as v, but not of x.

Regressions on true ability. To start, we have the following results from Goldberger (1972, p. 12):

$$E(x^*|z) = (1 - 2z)\sqrt{2Q/\pi} , \quad V(x^*|z) = (\pi - 2)Q/\pi,$$

where $\pi = 3.14159\ldots$. Proceeding, we find

$$E(y|z) = \alpha z + (1 + \beta z) \; E(x^*|z) = \alpha \; z + (1 + \beta z)(1 - 2z)\sqrt{2Q/\pi} ,$$

$$C(x^*,y|z) = (1 + \beta z) \; V(x^*|z) = (1 + \beta z)(\pi - 2)Q/\pi.$$

Applying (6) with y taking the role of s, and x* taking the role of r, we obtain the slopes and intercepts of the within-group regressions of y on x*:

(17) $$\alpha_{1z} = 1 + \beta z , \qquad \alpha_{oz} = \alpha z \qquad (z = 0,1)$$

Once again, these accurately capture both true effects of the treatment. The non-randomness of the selection procedure does not distort effects when x* is used as the explanatory variable. Similarly, the coefficients in the overall additive regression of y on x* are

(18) $$\alpha_1 = 1 + \tfrac{1}{2}\beta , \qquad \alpha_2 = \alpha .$$

Regressions on pretest score. In view of the independence of u and z, we have immediately:

$$E(x|z) = E(x^*|z) + E(u|z) = E(x^*|z) + E(u) = (1 - 2z)\sqrt{2Q/\pi},$$

$$V(x|z) = V(x^*|z) + V(u|z) = V(x^*|z) + V(u) = (\pi - 2P)Q/(\pi P),$$

cf. Goldberger (1972, p. 12). Similarly,

$$C(x,y|z) = C(x^*,y|z) = (1 + \beta z)\ (\pi - 2)Q/\pi \quad .$$

Applying (6) with y for s, and x for r, we obtain the slopes and intercepts of the within-group regressions of y on x:

$$\alpha_{1z} = P\ \left(\frac{\pi - 2}{\pi - 2P}\right)(1 + \beta z),$$

(19) $$\alpha_{oz} = \alpha\ z + (1 + \beta z)\ E(x^*|z) - \alpha_{1z}\ E(x^*|z)$$

$$= \alpha\ z + (1 + \beta z)(1 - 2z)\ \pi(1 - P)\ \sqrt{2Q/\pi}\ /(\pi - 2P) \quad (z = 0,1).$$

Both treatment effects are biased. The difference in slopes, namely

(20) $$\alpha_{11} - \alpha_{1o} = P\ \left(\frac{\pi - 2}{\pi - 2P}\right)\ \beta,$$

is a doubly-attenuated measure of the true interaction effect β: $\alpha_{11} - \alpha_{1o} < \beta$, since $\beta > 0$. The factor $P < 1$ arises from measurement error (cf. (15)), and the factor $(\pi - 2)/(\pi - 2P) < 1$ arises from the conjunction of measurement error and non-random selection. The difference in intercepts, namely

(21) $$\alpha_{o1} - \alpha_{oo} = \alpha - (1 + \tfrac{1}{2}\beta)\ (1 - P)\sqrt{8\pi Q}/(\pi - 2P),$$

will understate the true additive treatment effect α: with $\beta > 0$, $\alpha_{o1} - \alpha_{oo} < \alpha$. As in Goldberger (1972, pp. 11-12) these spurious deleterious effects of the treatment are attributable to a selection procedure which systematically puts low-ability individuals in the treatment group and high-ability individuals in the control group.

From (11)-(12), an overall additive regression of y on x, z yields

$$\alpha_1 = P\left(\frac{\pi - 2}{\pi - 2P}\right)(1+\tfrac{1}{2}\beta), \alpha_2 = \alpha - (1 + \tfrac{1}{2}\beta)(1 - P)\sqrt{8\pi Q}/(\pi-2P). \tag{22}$$

Note that if $\alpha = 0 = \beta$, α_2 reduces to the result reported by Goldberger (1972, p. 11) for a no-true-effect situation.

6. Selection on Basis of Pretest Score

In case (ii), individuals are assigned to the treatment group or the control group according as their pretest score is below or above the mean pretest score in the population:

$$z = \begin{cases} 1 \text{ if } x \leq 0 \\ 0 \text{ if } x > 0 \end{cases}.$$

Since z is determined exactly by x it will be dependent on both x* and u, but remains independent of v.

Regressions on true ability. We start with

$$E(x^*|z) = (1 - 2z)\sqrt{2PQ/\pi}, \quad V(x^*|z) = (\pi - 2P)Q/\pi,$$

given in Goldberger (1972, pp. 15, 17). Then we find

$$E(y|z) = \alpha z + (1 + \beta z)E(x^*|z) = \alpha z + (1 + \beta z)(1 - 2z)\sqrt{2PQ/\pi},$$

$$C(x^*,y|z) = (1 + \beta z)V(x^*|z) = (1 + \beta z)(\pi - 2P)Q/\pi.$$

Applying (6) yields the slopes and intercepts of the within-group regressions of y on x*:

$$\alpha_{1z} = 1 + \beta z , \quad \alpha_{oz} = \alpha z \qquad (z = 0,\ 1). \tag{23}$$

Once again these accurately capture both treatment effects. The non-randomness of the selection procedure does not distort effects when x* is used as the explanatory variable. Similarly, the coefficients in the overall additive regression of y on x* and z are

$$\alpha_1 = 1 + \tfrac{1}{2} \beta , \quad \alpha_2 = \alpha . \tag{24}$$

Regressions on pretest. We have

$$E(x|z) = (1 - 2z) \sqrt{2Q/(\pi P)} , \quad V(x|z) = (\pi - 2)Q/(\pi P),$$

and

$$C(x,y|z) = C((x,\ \alpha z + (1 + \beta z)\ x^* + v)|z) = C((x,\ (1 + \beta z)x^*)|z)$$

$$= (1 + \beta z)C(x,\ x^*|z) = (1 + \beta z)\ (\pi - 2)Q/\pi;$$

cf. Goldberger (1972, pp. 14, 17, 18). Thus the slopes and intercepts of the within-group regressions of y on x are

$$\begin{aligned} \alpha_{1z} &= P(1 + \beta z), \\ \alpha_{oz} &= \alpha\ z + (1 + \beta z)E(x^*|z) - P(1 + \beta z)E(x|z) = \alpha\ z \qquad (z = 0,\ 1). \end{aligned} \tag{25}$$

These are identical with the random-selection results in (15). The interaction effect is attenuated: $\alpha_{11} - \alpha_{1o} = P\ \beta$, but the additive effect is correctly captured: $\alpha_{o1} - \alpha_{oo} = \alpha$. Despite the fact that selection on pretest tends to assign low-ability individuals to the treatment and high-ability individuals to the control, it generates the same measures of the treatment effect as does random selection. Because of the inevitable slope attenuation (due to measurement error) we can no longer say (as we did in the earlier paper) that pretest selection yields unbiased estimates of the treatment effect, but rather that it yields the same estimates as does random selection. In this weakened form, the results of our previous analysis hold up in the presence of interaction effects. The contrast to true-ability selection is still striking. Finally, we record the coefficients for the overall regression of y on x and z ,

$$(26) \qquad \alpha_1 = P(1 + \tfrac{1}{2}\beta) , \qquad \alpha_2 = \alpha ,$$

which coincide with the random-selection results in (16).

7. Efficiency

While the same parameters are estimated under pretest selection as under random selection, efficiency is reduced. To see this explicitly, consider the within-group regressions of posttest on pretest when the (total) sample size is 2T. The moment matrix of explanatory variables (the constant and the pretest score x) has expectation equal to T times

$$\begin{pmatrix} 1 & E(x|z) \\ E(x|z) & E(x^2|z) \end{pmatrix} .$$

The inverse of this is

$$\frac{1}{V(x|z)} \begin{pmatrix} E(x^2|z) & -E(x|z) \\ -E(x|z) & 1 \end{pmatrix}.$$

In case (o), random selection, the diagonal elements of this inverse, namely

$$m^{oo} = E(x^2|z)/V(x|z) = 1 + E^2(x|z)/V(x|z)$$

$$m^{11} = 1/V(x|z),$$

take on the values

$$m^{oo}(o) = 1 \ , \qquad m^{11}(o) = P/Q;$$

cf. p. 9 above. In case (ii), pretest selection, they take on the values

$$m^{oo}(ii) = 1 + ((1-2z)^2 \ 2Q/(\pi P))/((\pi-2)Q/(\pi P)) = 1 + 2/(\pi-2)$$

$$= \pi/(\pi-2),$$

$$m^{11}(ii) = \ \pi P \ /((\pi-2)Q) = (\pi/(\pi-2))P/Q;$$

cf. p. 13 above. Thus

$$m^{oo}(ii)/m^{oo}(o) = m^{11}(ii)/m^{11}(o) = \pi/(\pi-2)$$

which implies that the sampling variances of the within-group regression coefficients under pretest selection are

$$\pi/(\pi-2) \doteq 2.75$$

times as large as they are under random selection; this is the same efficiency loss as was found in the no-effect case, Goldberger (1972, p. 22). Our calculation presumes that the within-group disturbance variance does not change with the change in experimental design. That presumption is correct, since it can be shown that

$$V(y|x,z) = (1-P)Q((1+z)^2 + 1/P) \qquad (z = 0, 1),$$

whether the selection is random or based on pretest.

8. Comments

1. Our assumption that $E(x^*) = 0$ is not entirely innocent. Indeed it may appear that the assessment of treatment effects will be contaminated when true ability has a nonzero expectation. For example, consider the random selection case (o). If $E(x^*) = \mu$, then $E(x^*|z) = \mu$ so that $E(x|z) = \mu$, $E(y|z) = \alpha z + (1 + \beta z)\mu$, and the second line of (15) will change to

$$\alpha_{oz} = \alpha z + (1+\beta z)\mu - P(1+\beta z)\mu = \alpha z + (1-P)(1+\beta z)\mu .$$

This says that the difference between the within-group regression intercepts is

$$\alpha_{o1} - \alpha_{oo} = \alpha + \beta\ (1-P)\mu\ .$$

Does this mean that the difference in intercepts no longer captures the true additive effect α when there is a nonzero interaction effect ($\beta \neq 0$)? If so, all our earlier conclusions would be seriously misleading.

A resolution of the difficulty runs as follows. When there is an interaction effect the measure of the additive effect is essentially arbitrary.

What, after all, is <u>the</u> vertical distance between two non-parallel straight lines? The arbitrariness can be resolved by conventionally measuring the additive effect as the difference between the ordinates of the within-group lines when the abscissa is $E(x^*)$. (This lies in the middle of the relevant range of the data.) But that is just the difference in intercepts, provided that x^* is measured in terms of deviations about its expectation. Thus our assumption $E(x^*) = 0$ did not limit the domain of the analysis but rather adopted, in effect, the convention that additive effects are to be measured at $E(x^*)$. The entire problem, of course, disappears when there is no interaction.

2. Our presumption that both true effects of the treatment are non-negative may be somewhat misleading. If $\alpha > 0$ and $\beta > 0$, then the lines

$$E(y|x^*, 1) = \alpha + (1+\beta)x^* , \qquad E(y|x^*, 0) = x^*$$

will cross, at the point

$$x^*_o = -\alpha/\beta < 0.$$

Consequently, while the treatment may be beneficial to individuals whose true ability exceeds x^*_o, it would be detrimental to individuals whose true ability lies below x^*_o. This specification may be objectionable in that it fails to do justice to one's notions about a beneficial treatment. To avoid the objection, we must fall back on an assumption that the relevant range of the data is $x^* \geq x^*_o$; with $\alpha \geq 0$ and $\beta \geq 0$ the treatment is then beneficial (or at least non-detrimental) for everyone. (The entire problem disappears when there is no interaction). Alternatively, one could incorporate the presumption that the treatment has a nonnegative effect from the start by replacing (3) by

$$E(y|x^*,z) = \alpha z + x^* + \beta z x^{**}$$

where

$$x^{**} = \begin{cases} 0 & \text{if } x^* < \bar{x}^* \\ x^* & \text{if } x^* \geq \bar{x}^* \end{cases}$$

with $\bar{x}^*$ being a prespecified value of true ability (perhaps 0). Analysis of this "kinky-interaction" formulation, however, lies outside the scope of the present paper.

3. To conclude, we provide some information on the magnitude of the bias which arises when selection is based on true ability. Confining attention to the no-interaction situation (β=0), we have from (22):

$$\alpha_2 - \alpha = - (1-P)\sqrt{8\pi Q}/(\pi-2P).$$

To obtain meaningful units, we measure the bias in terms of standard deviations of pretest score (= standard deviations of control posttest score); thus

$$\text{Bias} = (\alpha_2 - \alpha)/\sqrt{Q/P} = - (1-P)\sqrt{8\pi P}/(\pi-2P).$$

The bias is tabulated below for selected values of P:

P	.50	.60	.70	.75	.80	.85	.90	.95	1.0
Bias	-.83	-.80	-.72	-.66	-.58	-.48	-.35	-.20	0 .

REFERENCE

Goldberger, A. S., "Selection Bias in Evaluating Treatment Effects: Some Formal Illustrations," University of Wisconsin, Institute for Research on Poverty: Discussion Paper 123-72, April 1972.