EXTERNAL TREATMENT EFFECTS AND PROGRAM
IMPLEMENTATION BIAS

Tomas J. Philipson

External Treatment Effects and Program Implementation Bias
Tomas J. Philipson
NBER Technical Working Paper No. 250
January 2000
JEL No.C1, I1

ABSTRACT

This paper discusses the definition and identification of external treatment effects and experimental designs capable of detecting these effects. External effects occur when the outcome of a given individual is affected by the treatment assignments of other individuals. The paper argues that two-stage randomization schemes, which randomize allocation of treatments across communities and randomizes the treatments themselves within communities, are useful for identifying private and external treatment effects. The importance of external treatment effects are illustrated in the context of several health economics applications: the impact of R&D subsidies, smoking prevention programs for youth, and the evaluation of HIV-prevention programs currently taking place in Africa.

Tomas J. Philipson
The Irving B. Harris Graduate School of Public Policy Studies
The Department of Economics, and The Law School
The University of Chicago
1155 E.60th Street
Chicago, IL 60637
and NBER
t-philipson@uchicago.edu

1 Introduction

This paper examines the statistical inference problems associated with *external effects* in the standard econometric evaluation framework. In this standard framework, a set of treatments are randomly assigned, and the impact of these treatments on a defined outcome is studied. An implicit assumption in many studies is that only the treatments assigned to a given individual affect his outcome. This paper addresses the differences that arise when external treatment effects are allowed to operate–these effects are defined as when the outcome of a given individual is affected by the treatment assignments of other individuals. The key aspect of external treatment effects is that the *distribution* of treatments within a population, and not only the treatment of the investigated person, affects that person's outcomes – others' treatments matter to an individual conditional on a given treatment assignment. As a motivating example, consider the large HIV-prevention trials that are presently being conducted on the African continent. For HIV-prevention trials, external effects may arise when the infection rate of the control group in a trial depends on the share of treated, such as the vaccination rate in a vaccine trial. In other words, although the individuals in the control group all receive the *same* treatment, no vaccination, the assignment of treatments to others, such as the vaccination rate in the overall population, affects their outcomes. As another example, consider public R&D subsidies. In this case, external treatment effects may be present when the innovative activity of unsubsidized firms depends on the size of a subsidy program through R&D spill-overs across firms. The paper discusses a general framework to address the exact definition, identification, and estimation of such external effects in an experimental setting. This framework may be applicable to a host of economic applications, and allows one to simply assess what can be learned about large-scale program implementations from small evaluations.

External effects are of great importance because they are central to understanding how the results from a relatively small experiment can be generalized to a large scale implementation of a particular treatment. If external effects are not dealt with appropriately, evaluations showing small or no effects may turn out to have great effects when implemented, and conversely evaluations showing big effects may turn out to be insignificant when implemented. Moreover, it is shown that this is not because experiments are smaller than the programs implemented in a given study; large scale experimental evaluations may involve *larger* biases in going from evaluation to implementation depending on the nature of the external effects.

The fact that external effects are important for extrapolating from experimentation to implementation suggests that designs of evaluations or social experiments should attempt to identify them. A way to deal with the identification of external treatment effects is randomization of the allocations of treatments across communities, such as the share treated. The communities over which the share treated is randomized may be comprised of partners or couples in HIV-prevention interventions, local markets in the case of R&D subsidies, schools in learning or youth smoking interventions, or villages in broad scale public health

1

interventions. This randomization of the allocation of treatments across communities is in conjunction with randomizing the assignment of treatments themselves within these communities. Thus two stages of randomization is suggested; one of allocations of treatments to communities and another of treatments within communities as specified by those allocations. The groups across which the first step of the randomization occurs should be large enough to feasibly internalize the externality among themselves. In other words, it should include all those affected externally in their outcomes by other members of the group. For example, if the external treatment effects occur at the level of couples, then couples should be the unit of randomization. On the other hand if external effects occur at the level of the whole community, communities should be randomized as in the HIV-prevention trials in Africa. The randomization across individuals within communities identifies the standard private treatment effects while the randomization across communities identifies the external treatment effects. The external effects are picked up by comparing the outcomes of individuals who received the same treatment but were exposed to different treatment allocations, e.g. comparing the infection rate of unvaccinated people across communities with different vaccination rates or comparing unsubsidized firms in their innovative activity across markets in which the scale of the subsidy programs differ.

Given that a two step randomization process can identify external and private effects, design issues are addressed when estimating both these type of effects with finite samples. There is a tension between the optimal design for estimating the private and external effects. Private effects tend to be best estimated by having equally sized control and treatment groups but external effects tend to be best estimated by having extreme treatment shares across populations. In a vaccine trial, external effects are most efficiently estimated with very large and small vaccination rates across communities, but doing so sacrifices power to detect the private effects of vaccination within these communities.

This paper relates to many strands of previous work. The types of effects discussed here, or related phenomena, certainly have been recognized in a qualitative fashion for a long time in previous discussions. Previous work by economists on this topic dates back at least as far as Lewis' (1963) discussion regarding unions. In education, sociologists and economists have long hypothesized that peer-group effects are important in learning, see e.g. Coleman et al. (1966), Summers and Wolfe (1977), Henderson et al. (1978), and Arnott and Rowse (1987). Recent discussions relating to external effects more generally include, for example, Harris (1986), Garfinkel and Manski (1992), Manski (1995, 1997), Magden and Brookmeyer (1993), Halloran et al (1997), Philipson (1997), Heckman et al (1999), and others. The novel attempt here is to put forth a formal framework in which the common issues surrounding such effects can be explicitly defined, identified, and assessed in a quantitative fashion.[1] However, external

---

[1] Also, the paper relates to a large literature on both observational and experimental cluster sampling and randomization. However, the main rationale motivating such cluster sampling is often to reduce the costs of producing the data. These cost-advantages often justify the statistical disadvantage of cluster-sampling as it inflates standard errors beyond random sampling. Here, on the other hand, cluster randomization occurs

2

effects are different from so called endogenous effects (see e.g. Manski (1995, ch. 9)), as they concern the statistical, and often non-identifiable, relationship between group and individual outcomes. Indeed, some of the solutions offered here, such as two-stage randomization, highlight that some criticisms of evaluations with random assignments, e.g. that they do not take into account such disparate effects as 'contamination-', 'macro-', 'neighborhood-', 'peer-group-',or 'herd'- effects'., are overcome by redefining the unit of randomization to better incorporate external effects. Perhaps an overall point of the paper is that many of these separate issues, once put in a simple and common framework, may be handled quite easily by very standard methods. However, it is worth emphasizing from the start that the paper deals exclusively with an idealized experimental setting, abstracting from issues such as non-compliance, attrition, and other issues of choice that may also make experiments and implementations differ. This is not because they are not important, they are, but because the isolated effects of external effects are more transparently discussed and understood by initially abstracting from such issues.

The paper may be briefly outlined as follows. Section 2 discusses the definition of private and external treatment effects. Section 3 discusses how to identify them. Section 4 discusses estimation of external effects, and section 5 provides an empirical illustration using data from a youth smoking prevention program. Lastly, section 6 concludes.

2   Private & External Treatment Effects

I first illustrate and define external treatment effects by considering the simplest case of two treatments and a single population in which the effect operates. I then extend this to multiple treatments and populations. Let treatment assignment be represented by the treatment indicator d; an experimental treatment ($d = 1$) is allocated to a share s of the population and a control treatment ($d = 0$) is allocated to the remaining share $1 - s$. The individual outcome of interest is denoted Y and the distribution of concern is $F(y|d, s)$ which represents the outcomes of those receiving treatment d given a share of treated s. I say that there is an *external* treatment effect if $F(y|d, s)$ varies with s and a *private* treatment effect if the distribution varies with d. Thus private effects operate when one's own treatment assignment, d, affects one's own outcome as opposed to external effects that operate when the treatment assignment of others, s, affect one's own outcome. Here, external treatment effects are captured by the dependence of the conditional outcome distribution on the treatment allocation represented by the share treated.

Our discussion focuses on the mean outcome as a function of the treatment share in the community denoted by $\mu(s) = E[Y|s]$ with $\mu_d(s) = E[Y|d, s]$ being a treatment specific mean. To illustrate, consider the cases of assessing the impact of HIV-prevention methods or

for identification purposes, that is, for statistical as opposed to cost advantages.

3

public R&D subsidies[2]. For an AIDS-vaccine trial, the treatment d would be vaccination and the outcome mean μ, the infection rate. External effects arise here because the vaccination rate in the population affects infection rates of both treated and controls. In the case of R&D, the treatment d would involve a public subsidy for R&D, and the outcome mean μ represents some measure of average innovative behavior, whether in terms of inputs or output. Generally, the mean outcome in the entire population is the average of the treatment specific means weighted by the sizes of each treatment group.

$$\mu(s) = s\mu_1(s) + (1-s)\mu_0(s)$$

External effects may reveal themselves through the impact of the share treated on the overall population mean[3]. Under only private effects, when $d\mu_d/ds = 0$, the relationship is linear with the marginal impact of raising the share of treated by one percent coinciding with the treatment effect: $d\mu/ds = \mu_1 - \mu_0$. For example, if a new medical treatment was shown to increase longevity by 10 percent, then expanding the share of treated by 20 percent results in a proportional increase in the longevity of the overall population by 2 percent. The improvement in the population is proportional to the share treated. It is important to note that this would hold even when the treatment effects were heterogeneous across individuals as long as the treatment assignments were random because then the distribution of heterogeneous treatment effects would be the same in the treatment and control groups. However, under external effects, the aggregate outcome mean as a function of the treatment share may be non-linear even under random assignment of treatments. It will be convex if the share treated has a reinforcing positive external effect on the private treatment effect and it will be concave if it has a counteracting negative external effect. More precisely, the

---

[2]Aside from being useful illustrations, treatment evaluation in these contexts are also of significant substantive importance in themselves. The communities involved in HIV-trials often have over a third of their populations infected with HIV and the continued growth of the virus has therefore been argued to be one of the most important public policy issues in these regions and a major hurdle to economic development (see e.g. Over and Ainsworth (1997)). Moreover, much of economic theory stresses the implications of positive spill-overs in R&D for economic growth and other issues–making the ability to better evaluate empirically public efforts to stimulate R&D very important.

[3]The issues discussed throughout the paper generalize directly to multiple treatments and sub-populations. In the more general case, there is a population of $g = 1, .., N$ groups with treatment allocations represented by the vector $\mathbf{s} = (\mathbf{s}^1, .., \mathbf{s}^N)$ which has vector elements $\mathbf{s}^g = (\mathbf{s}^g_0, .., \mathbf{s}^g_K)$ representing the allocation across K treatments for group g. The treatment-specific outcome mean of a given group is $\mu^g_d(\mathbf{s})$ for $g = 1, .., N$ and $d = 0, .., K$. The case discussed in most of the paper concerns $N = 1$ and $K = 1$ with the other extreme case being when the groups are single individuals. Between these two extremes lies many different types of external effects, such as couple-based, school-based, or household-based external effects. A particularly common scenario would be three treatments corresponding to experimental controls and treated as well as those not participating in the experiment.

4

effect of expanding the treatment by raising the share of treated is

$$Total\,Effect \quad = \quad Private\,Effect \quad + \quad External\,Effects$$
$$\frac{d\mu}{ds} \quad = \quad [\mu_1(s) - \mu_0(s)] + [s\frac{d\mu_1}{ds} + (1-s)\frac{d\mu_0}{ds}].$$

The first term within the brackets represents the linear effect, as discussed above, when holding the private effect constant as the share of treated increases. This private effect corresponds to the composition effect that comes from changes in the treatment shares with fixed treatment specific means. However, the terms within the second bracket represent the impact of external effects on outcome means conditional on treatments (the external effects are weighted by the share of each treatment). The external effects correspond to the changes in effects *within* treatments. If there are no external effects, the second bracketed term vanishes and the marginal effect of expanding the program reduces to the standard private treatment effect. The fact that expanding the program involves both private and external effects implies that identifying only one, in particular the private effect through a conventional experiment, does not identify fully what will happen when the program is expanded.

3 External Effects and Implementation Bias

External effects are important because they affect what can be learned about implementing a large program from the results of a smaller experiment. In other words, external effects impact the infra-marginal effects that arise when going from a small experimental evaluation to a full-scale program. These effects are central to learning about the consequences of implementing a treatment by observing the treatment's experimental performance – a central goal of many evaluations. The basic question of interest is if one uses the estimated treatment effect of an experiment to extrapolate to program implementation, how close does the extrapolated result come to the actual outcomes of the larger scale program when implemented?

Given an experiment with the share of treated being s, consider implementing a program where the size of treated is expanded to $s_I >> s$. The experiment identifies the outcome means $\mu_1(s)$ and $\mu_0(s)$ of the treated and controls respectively. Using this information from the experiment, I denote the *extrapolation* mean $\varepsilon(s, s_I)$ and it is defined by

$$\varepsilon(s, s_I) \equiv s_I\mu_1(s) + (1 - s_I)\mu_0(s).$$

This extrapolated mean may be separated from the *implementation* mean $\mu(s_I)$ which occurs when actually expanding the scale of treatment during implementation.

$$\mu(s_I) \equiv s_I\mu_1(s_I) + (1 - s_I)\mu_0(s_I).$$

Thus the extrapolation mean is the one predicted when using the small scale experiment to identify the mean response under implementation as opposed to the mean which actually occurs when the program is implemented. External effects are important for the ability to extrapolate from experimental results to program implementation results because a major difference between experimentation and implementation is simply the size of the share treated or, more generally, the allocation of treatments [4]. More precisely, I define *implementation bias* $B(s, s_I)$ to be the difference between the extrapolation and implementation mean.

$$B(s, s_I) \equiv \mu(s_I) - \varepsilon(s, s_I) = s_I[\mu_1(s_I) - \mu_1(s)] + (1 - s_I)[\mu_0(s_I) - \mu_0(s)].$$

A useful illustration here is the case of no external effects versus the type of external effects that occur for R&D subsidies or vaccines as illustrated in Figures 1 and 2 below. The figures involve the (Taylor-expanded) linear specification of treatment-specific outcome means.

$$\mu_d(s) = \alpha d + \beta s, \qquad d = 0, 1$$

The private effect is thus $\alpha = \mu_1(s) - \mu_0(s)$ and the external effect is $\beta = [\mu_d(s) - \mu_d(s')]/(s - s')$. Both figures map out various outcome means as function of the program size. Figure 1 shows the standard case when there are no external effects, $\beta = 0$, so that the implementation mean coincides with the extrapolation mean and is given by the downward sloping line between the constant functions corresponding to the outcome means of the two treatment groups. Figure 2 shows the case when the outcome means fall in the share of treated, $\beta < 0$, as when higher vaccination rates lower infection rates among both vaccinated and unvaccinated individuals or when R&D subsidies lower the average cost among both subsidized and unsubsidized firms. The line between the two declining functions is the implementation mean $\mu(s_I)$ which in this case differs from the extrapolation mean $\varepsilon(s, s_I)$ by the amount of the implementation bias $B(s, s_I)$.

INSERT FIGURES 1 & 2 HERE

As the figure reveals, there will be a negative implementation bias. The case of R&D subsidies serves as a useful illustration of this negative bias. Consider when such subsidies are offered to a given share of firms after which the experimental effect of R&D related behavior is determined to be zero so that the private effect satisfies $\mu_1(s) - \mu_0(s) = 0$. This would be the extreme case of the figure when the two lines of the two treatments coincided. Then one may conclude that expanding the program is not valuable since the marginal effect of extrapolation is zero when the private effect is zero; $d\varepsilon/ds_I = 0$. Alternatively, the private effect may be small because the subsidies did affect the R&D of the subsidized firms

---

[4]I here abstract from other differences between experiments and program implementation which may also be important such as those having to do with the behavior of subjects relative to future individuals engaging in treatment, see e.g. Philipson and Desimone (1997).

but, through R&D-spillovers, the subsidies affected unsubsidized firms as well. This would correspond to a situation in which the subsidy-treatment d had no private effect on innovative activity but that the external effects within treatments nevertheless reduced production costs: $d\mu_1/ds = d\mu_0/ds < 0$. Consequently, expanding the program would affect economy wide costs although there would be no differences across subsidized and unsubsidized firms. This would make the implementation bias negative, $B(s, s_I) < 0$, because the small experimental effect of the treatment would understate the economy wide effects when implementing the R&D subsidy on a larger scale.

When external effects are present, the behavior of the bias has some interesting and non-intuitive properties. First, it is immediately obvious that a necessary condition for there to be an implementation bias is that there are external effects.

$$\frac{d\mu_1}{ds} = \frac{d\mu_0}{ds} = 0 \Rightarrow B(s, s_I) = 0$$

This occurs because with only private effects, the means outcome of the treated and controls are the same in the experiment as under program implementation so that extrapolation from experimental to implemented outcomes is feasible.

Second, it may be the case that expanding the size of the experiment raises the implementation bias.

$$\frac{dB}{ds} = -\frac{d\varepsilon}{ds} = -[s_I\frac{d\mu_1}{ds} + (1 - s_I)\frac{d\mu_0}{ds}]$$

Note that this derivative is not necessarily of the opposite sign of the bias itself, that is, *larger experiments may well involve larger implementation bias*[5]. For example, consider when the bias is **positive as in** the figure above (as in the case of R&D or vaccination). A necessary and sufficient condition for the implementation bias to rise with the size of the experiment is then:

$$\frac{dB}{ds} \geq 0 \Leftrightarrow s_I \geq \frac{\beta_0}{\beta_0 - \beta_1}.$$

Thus, as long as the program size is sufficiently large, the implementation bias will rise with the size of the experiment. Note that the effect of the size of the experiment on bias does not depend on the eventual program size when the external effects have the same effects for both controls and treated; $\beta_0 = \beta_1$ implies $\frac{d^2B}{dsds_I} = 0$.

Another non-intuitive property of the implementation bias is that the bias may fall as the size of the program to be implemented grows. This occurs when the extrapolated mean grows faster than the implementation mean with program size.

$$\frac{dB}{ds_I} \leq 0 \Leftrightarrow \frac{d\mu}{ds_I} \leq \frac{d\varepsilon}{ds_I}.$$

---

[5]This holds true even if the bias is negative so that an increase in B may involve less bias.

The effect on the extrapolation mean of expanding the size of the implemented program is simply the private treatment effect

$$\frac{d\varepsilon}{ds_I} = \mu_1(s) - \mu_0(s)$$

The effect on the implementation mean of expanding the size of the implemented program is, as discussed, comprised of both the private and external effects.

$$\frac{d\mu}{ds_I} = [\mu_1(s_I) - \mu_0(s_I)] + [s_I \frac{d\mu_1}{ds_I} + (1 - s_I)\frac{d\mu_0}{ds_I}]$$

Consequently, if the private effect falls sufficiently with program expansion, the implementation bias will rise with program size because the experimental private effect identified becomes a worse and worse approximation of the smaller private implementation effect.

4 Identification and Estimation of External Effects Through Two-Stage Randomization

Given that identifying whether private or external effects are present may have very important implications for assessing program wide implementation, it is necessary to discuss how experiments may be designed to generate the necessary independent variation to identify both external and private effects. The important question is how to perform the randomization. Once the type of randomization proposed has been done, the identification of the two types of effects is usually straightforward.[6]

To identify external effects one needs to generate variation in treatment *allocations* across groups, in addition to the variation in treatments across individuals within groups. I suggest a two-step randomization scheme in which allocations are first randomized and then treatments within the groups. Although two-stage randomization has certainly been used in practice (e.g. in community based trials), their relationship to external effects has not been stressed. The groups across which the first step of the randomization occurs should be large enough to feasibly 'internalize' the externality among themselves. For example, if the external treatment effects occur at the level of couples, then couples as opposed to individuals should be the unit of randomization using the four possible treatment combinations. On the other hand, if external effects occur at the level of the whole community, the allocation of treatments across communities should be randomized. The type of community based trials that are needed for these identification purposes can be executed within the existing format of many community or multi-center trials. An example would be the existing trials used to evaluate the effectiveness of strategies of reducing HIV infection in African countries which

---

[6]The design issues resemble those of so called *split-plot* designs, for which linear methods are well-known.

often involve many communities. For example, the HIV-prevention trial in Tanzania involves 50 communities. [7]

4.1 Identification of External Effects

It is useful to first point out that in general the existence of external effects cannot be identified from the unconditional mean $\mu(s)$. Put differently, only observing this mean function does not impose any restrictions on the treatment-specific mean functions $\mu_0(s)$ and $\mu_1(s)$. Any population outcome function $\mu(s)$ can be interpreted to include *some* external effects so that external effects can never be ruled out. To show this for two treatments, for any $\mu(s)$ if we select $\mu_0 = 0$ and $\mu_1(s) = \mu(s)/s$ then $\mu(s) = (1 - s)\mu_0 + s\mu_1(s)$. A similar argument applies for more than two treatments. Therefore, some type of functional form of on the treatment-specific mean functions is needed in order to identify the existence of external effects from the unconditional mean function $\mu(s)$.

Naturally, one can always identify them from outcomes conditional on treatments as the two stage-randomization scheme suggests. A simple two-stage randomization scheme in community based experiments would work as follows. First, one randomizes out allocations or treatment shares s in the first step. Second, after those units have been selected, the traditional randomization of the treatments d within the communities is applied according to the probabilities dictated by the s assigned to the community. For a set of communities[8] $c = 1, 2, .., C$ let $s = (s^1, ...s^C)$ be the treatment shares randomized out to each of the communities after which a share s^c of individuals in community c receive the treatment and the remaining share $1 - s^c$ the control through random assignment of the treatment d. Consequently, the two stage randomization involves randomization of treatment allocations *across* communities as well as random assignment of individual treatments *within* communities. An illustrative example of a two stage randomization scheme is the one most often employed in multi-center trials with half the community getting the treatment and half of it getting the control. Such allocations have been shown to be optimal in terms of achieving the highest level of precision in estimating the private treatment effect within the community and would here involve the vector s with each element being $s^c = (1/2, 1/2)$ for $c = 1, 2, .., C$. The problem is that this creates no variation to identify the external effects. To identify such effects, one needs variation in allocations, and not only treatments. Consequently, any allocation that is 'optimal' for identifying private treatment effects should not be replicated across the

---

[7] Naturally, such multi-center trials involve substantial costs and have therefore generally been financed by international organizations such as The World Bank or The European Commission, see e.g.Over and Ainsworth (1997).

[8] As noted previously, the analysis here would generalize in a straightfoward manner to multiple groups and treatments. The case discussed here concerns the extreme case of one big group with the other extreme case being when the groups are single individuals. Between these two extremes lie many different types of external effects, such as couple-based, school-based, or household-based external effects.

communities because no variation in allocations across communities is generated.

Consider the linear model discussed before. When there is both variation in treatments within communities and the share of treated across communities both effects are identified. The randomization of allocations *across* communities provides the necessary variation to identify the external effect β_d.

$$\beta_d = \frac{\mu_d(s) - \mu_d(s')}{s - s'}$$

Given these external effects, the randomization of treatments *within* communities provides the necessary variation to identify the private effects .

$$\mu_1(s) - \mu_0(s) = \alpha_1 - \alpha_0 + (\beta_1 - \beta_0)s$$

To illustrate, consider an example in which the two-stage scheme has forty percent, $s = 0.40$, of the population vaccinated in some communities but twenty percent, $s' = 0.20$, in others. Assume that in the more vaccinated communities the controls were infected less frequently so that the difference in infection rates among the unvaccinated controls were 10 percent; $\mu_0(s) - \mu_0(s') = 0.10$. On the other hand, consider when the difference in infection rates across communities was negligible for vaccinated individuals, possibly due to an effective vaccine; $\mu_1(s) - \mu_1(s') = 0$ The random assignment of communities then enables one to infer the external effects; $\beta_0 = \frac{\mu_0(s) - \mu_0(s')}{s - s'} = \frac{0.10}{0.20} = 0.50$ and $\beta_1 = \frac{\mu_1(s) - \mu_1(s')}{s - s'} = 0.$ and the corresponding private effects by within community treatment differences.

4.2   Estimation of External Treatment Effects

Consider when there are n observed units within each of C communities and one is to choose the shares of treated $s = (s_1, .., s_C)$ in those communities. Consider the Figure 3 that shows sample means that correspond to specific sample sizes and the shares of treated. The $2xC$ layout has different treatment allocations corresponding to columns and different treatments corresponding to rows. The first row therefore maps out the mean values among treated $\mu_1(s)$ for different shares of treated and the second row maps out the mean values of the controls $\mu_0(s)$ for the same share.

INSERT FIGURE 3 HERE

The outcome means in this figure correspond to those of Figures 1 and 2 concerning the two functions $\mu_1(s)$ and $\mu_0(s)$. Each cell contains the outcome mean for that particular allocation (column) and treatment (row). (More general external effects may always be separated in a similar manner using allocation-treatment cells.) The differences in outcome means across columns in a given row indicate the external effect within a treatment - the top

10

row being the external effects for the treated and the bottom row those of the controls. The differences in outcome means across rows given a column indicates the private effect holding the treatment allocation constant. Consider the additive (means-adjusted) specification

$$\mu_1(s) = \alpha_1 + \beta_1 s$$
$$\mu_0(s) = \alpha_0 + \beta_0 s$$

where observations on each allocation-treatment cell has variance σ^2. Running the regressions within treatment groups then gives rise to the variance of the estimator of the external effect (slope coefficient) as in[9]

$$V(\overset{\wedge}{\beta}_d) = \frac{\sigma^2}{v_s}, \qquad d = 0, 1,$$

where $v_s = \sum_{c=1}^{C}(s_c - \bar{s})^2 / C$ is the variance in the regressor made up of the treatment shares. If the private treatment effect estimator is the average private effect across allocations, its variance is

$$V(\widehat{\alpha}_1 - \widehat{\alpha}_0) = \frac{1}{C^2} \sum_{c=1}^{C} V(\widehat{\alpha}_{c1} - \widehat{\alpha}_{c0}) = \frac{1}{C^2} \sum_{c=1}^{C} \frac{\sigma}{n^2} \left[\frac{1}{s_c^2} + \frac{1}{(1 - s_c)^2} \right].$$

This displays the trade-off in efficiently estimating the private versus external effect through the variation of treatment shares across communities. In other words, there is a tension between getting efficient estimates of the private versus external effect.[10] The private effect is most efficiently estimated by having treatment shares in the middle. Indeed, splitting the sample equally between controls and treated, $s_c = 1/2$ for all $c = 1, .., C$ minimizes $V(\widehat{\alpha}_1 - \widehat{\alpha}_0)$. This is in contrast to the external treatment effect that is most efficiently estimated by having extreme values of the shares of treated for the standard reason that slope-coefficients are more efficiently estimated the greater the variation in the regressor. The variance of the private effect rises with a mean-preserving spread of the treatment shares as opposed to the variance of the external effect which falls. This trade-off in efficiency in estimating the two effects arises because of the unique aspect that the value sampled of one covariate, the share treated, determines the size of the sample by which another covariate, the private effect, can be estimated.

---

[9]The common variances across groups is a simplification, it may for example not hold for binomial outcomes.

[10]Note here that I have somewhat understated that the across community estimates of variance may be much less precise than the within community estimates due to the costs of production inherent in community designs. Whether those costs are relatively higher across than within depends on the definition of a community.

More general specifications, including interactions between treatments and controls, operate the same way. In general, I have n observations in each community with outcome y, all of which have the share of treated s, but ns observations which are treated and $n(1-s)$ observations which are controls. For a specification of the conditional mean as a function of treatments and allocations $E[Y|d,s]$, the matrix of regressor values $(1, x_1, x_2) = (1, d, s)$ would be as indicated in Figure 4 below.

INSERT FIGURE 4 HERE

The data resulting from a two stage randomization scheme will thus be comprised of n observations from each covariate value as represented by the treatment share. The unique feature is that the sample size of the other covariate is induced by the level of the first: $d = 1$ is observed for nq observations and $d = 0$ for the remaining $n(1-s)$ observations. In other words, *the treatment allocations within a community also determine the precision by which private effects can be estimated in that community.* More precisely, the variance of the private and external effects estimators is given by the usual formula for a multiple regression $V\{\hat{\beta}\} = \sigma^2 (X'X)^{-1}$. This determines the objective function for the most efficient design and the corresponding quadratic form specifying the weights the investigator wants to assign to different types of errors.

5   An Empirical Illustration from A School-Based Smoking Prevention Program

Table 1 below displays the data discussed in Donner and Klar (1994) which resulted from a school-based intervention program to prevent teenagers from smoking. For school-based smoking prevention trials for youth, a given child may react to the treatment of others through the fact that the share of children smoking imposes peer-pressure. Assessing the difference between private and external treatment effects of such prevention programs may therefore be important when the smoking of a given child depends on the share of children smoking in the school. Although this experiment was not designed and analyzed with external treatment effects in mind, the data is useful to illustrate how the above ideas could be implemented.

Table 1 shows tobacco use for each allocation-treatment combination. These sample means map out the functions $\mu_1(s)$ and $\mu_0(s)$ for the different treatment shares $(s_1, .., s_C)$. This implicitly concerns the case of two treatments ($K = 1$) and one sub-population ($N = 1$).

INSERT TABLE 1 HERE

Figure 5 below shows the estimated treatment specific means $\mu_1(s)$ and $\mu_0(s)$ when approximated by the linear specification. Using the data on the controls to estimate the

parameters of the linear specification, I obtain the estimates $(\hat{\alpha}_0, \hat{\beta}_0) = (0.050, 0.020)$ with the corresponding values for the treated $(\hat{\alpha}_1, \hat{\beta}_1) = (0.017, 0.046)$.

INSERT FIGURE 5 HERE

Figure 5 displays these two functions as well as the overall mean, $\mu(s) = s\mu_1(s) + (1 - s)\mu_0(s)$. The extrapolation bias induced by these estimates can be calculated and would represent how much a smoking-prevention program that was implemented within an entire school would affect prevalence of smokeless tobacco compared to the private effect observed within the school in an experiment. The y-axis shows this bias as a function of the share treated in the experimental evaluation measured on the x-axis. Since the figure indicates that the private treatment effect rises with the program, as $\mu_1(s)$ is less than $\mu_0(s)$ but rises at a more rapid rate, the full implementation bias falls with the share of treated.

6 Concluding Remarks

This paper discussed the definition and identification of external treatment effects as well as experimental designs capable of detecting them. Such effects were characterized by treatment allocations of some individuals affecting the outcomes of other individuals. It was argued that group based macro-randomization, as opposed to randomizing treatments across individuals, was useful for identifying external treatment effects and that this was important for differentiating implementation effects from experimental effects. The types of effects were illustrated by a youth smoking prevention program and discussed in the context of vaccination evaluations.

The paper suggests several future research questions. Since the paper only dealt with experimental designs where treatment allocations could be controlled, a corresponding analysis of external effects for observational studies, in which treatment allocations are chosen by those observed, seems warranted. In particular, it is clear that standard discussions of instrumental variables (IV) that only provides exogenous variation of treatments, as opposed to treatment allocations, do not fully deal with the identification problems introduced by external effects discussed here. Future research might, therefore, fruitfully address the exact limitations of such standard IV solutions and propose remedies to overcome them as well as the limitations of the present analysis when observational issues and issues of non-compliance are present. However, I conjecture that even under such circumstances, external effects may alter in significant ways how experimental evaluations should be interpreted as well as what they infer about the full-scale implementation of treatments as estimated through IV or other methods.

13

APPENDIX

There are many implicit assumptions that go into the practice of using private effect evaluation for assessing the implementation of an intervention. It appears useful to be more explicit about them to better understand the implicit assumptions violated when external effects are present. This Appendix therefore axiomatizes the implicit assumptions of standard private effect evaluation.

Axiomatic Characterization of Population Outcomes with Only Private Effects

Let the vector $s = (s_0, .., s_K)$ denote a probability frequency distribution (whose elements sum to unity) that represents the shares of the population that are assigned to treatments $d = 0, 1, .., K$. A asymmetric binary relation $\succ$ represents the overall population outcome ranking so that an allocation s is on a higher place on the ranking than s' if $s \succ s'$. The following result characterizes the ranking of population outcomes implicit in private effects evaluation:

The binary and asymmetric outcome ranking $\succ$ satisfies

$$s \succ s' \succ s'' \Rightarrow$$
$$\exists a, b \in (0,1) : as + (1-a)s'' \succ s' \succ bs + (1-b)s'' \quad (A1)(Allocation - Continuity)$$

$$s \succ s' \Rightarrow$$
$$as + (1-a)s'' \succ as' + (1-a)s'' \quad (A2)(Allocation - Independence)$$

if and only if there is a vector of treatment-specific means $\mu = (\mu_0, .., \mu_K)$ that do not depend on the allocations s such that

$$s \succ s' \Leftrightarrow \mu \circ s > s' \circ \mu$$

where $\circ$ represents vector products. This characterization of the outcome rankings is a direct consequence of the Mixture Space Theorem developed in decision theory as opposed to the context here of experimental program evaluation (for a proof, see Herstein and Milnor (1953)). The linear scaling of allocations that is implicit in private effect extrapolation is analogous to the linearity in probabilities of expected utility[11]. The key axiom that may be violated under external effects is Allocation-Independence axiom (A2). This axiom says that if the treated have better outcomes than the controls then a population in which some are treated always do better than a population of only controls. This assumption implies

---

[11] In the simplex representing the allocations of an expereiment with controls, treated, and non-participants, the private effects assumption implies that the iso-performance curves are linear and parallell. This obtains because as the allocation of treatment changes, the means conditional on them do not change.

14

that treatment-specific means do not depend on treatment allocations, that is, there are no external effects. To see how, let $s = s^1$ and $s' = s'' = s^0$ represent the extreme allocations where s^d is the allocation for which everyone receives the same treatment d. In this case, the mixture $as + (1 - a)s''$ represents the allocation with a share a being treated. The Allocation-Independence Assumption (A2) then says that if the treatment effect is positive, the implementation effect is positive as well. In other words, the share treated does not affect what can be learned from the experiment about implementation.

As the key qualitative condition of outcome rankings violated under external effects is Allocation Independence, relaxing this axiom will allow for external effects to be present– outcomes will depend on allocations in addition to treatments. A particular form of external effects which do not satisfy Allocation Independence is given by the following characterization.

The outcome ranking $\succ$ satisfies:

$$s \quad \succ \quad s' \succ s'' \Rightarrow$$
$$\exists a, b \quad \in \quad (0,1) : as + (1-a)s'' \succ s' \succ bs + (1-b)s'' \quad (A1)(Allocation - Continuity)$$

$$s \quad \sim \quad s' \Rightarrow$$
$$\forall a \quad \in \quad (0,1) \; \exists b \in (0,1) : as + (1-a)s'' \sim bs + (1-b)s'' \quad (A2)(Allocation - Dependence)$$

if and only if

there is a set of treatment specific means $\mu(s) = (\mu_0(s), .., \mu_K(s))$ which may depend on the allocations such that

$$s \quad \succ \quad s' \Longleftrightarrow$$
$$s \circ \mu(s) \quad > \quad s' \circ \mu(s) \,\&\, \mu_d(s) = w_d v_d [w \circ s]^{-1}, \qquad v_d \in R \quad \& \quad w \equiv (w_0, .., w_K) \in R^{K+1}$$

The proof of this proposition is omitted as it is, again, a direct reinterpretation of well-known results in decision theory, in particular the axiomatic characterization of quasi-linear orderings of Chew (1983). The Allocation Dependence assumption allows for treatment specific means being decreasing or increasing in treatment shares. It states that, as opposed to the case of Allocation Independence, there need only exist an equivalent independent allocation for which mixing preserves the outcome ranking, not that this is true for all independent allocations. Allocation Dependence implies as a special case of Allocation Independence whenever the weights w do not depend on the treatment assignments, $w_d = w$, which occurs when there are no external effects.

References

[1] Arnott, R. and J. Rowse, (1987), 'Peer Group Effects and Educational Attainment,' *Journal of Public Economics*, v32, 287-305.

[2] Brookmeyer R and Gail MH: *AIDS Epidemiology: A Quantitative Approach.* Oxford, England: Oxford University Press, 1994.

[3] Chew, S-H.,(1983), 'A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox', *Econometrica*; v51 n4, pp. 1065-92.

[4] Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McParthlan, A. M. Mood, F. D. Weinfeld, and R. L. York, *Equality of Educational Opportunity*, Washington D. C.: Government Printing Office, 1966.

[5] Donner, A., and N., Klar, (1994), 'Methods for Comparing Event Rates in Intervention Studies When the Unit of Allocation is a Cluster', *American Journal of Epidemiology*, v 140, 279-289.

[6] Garfinkel, I., and C,. Manski, (1992), 'Microexperiments and Macro Effects', in *Evaluating Training and Welfare Programs*, edited by Garfinkel, I., and C,. Manski, Cambridge: Harvard University Press.

[7] Geoffard, P-Y., and T. Philipson. "Disease Eradication: Private vs Public Vaccination." *American Economic Review* 87.1 (1997): 222-230.

[8] Halloran, E., C., Struchiner, and I., Longini, (1997), 'Study Designs for Evaluating Different Efficacy and Effectiveness Aspects of Vaccines', *American Journal of Epidemiology*, v 146, No10, 789-803.

[9] Harris, J., (1986), 'Macro-effects in Social Experiments' chapter in *Social Experimentation*, edited by J., Hausman and D., Wise, NBER, University of Chicago Press.

[10] Heckman, J., L. Lochner, and C. Taber, (1999), "General Equilibrium Cost Benefit Analysis of Education and Tax Policies", NBER Working Paper No. W 6881.

[11] Henderson, V., P. Mieszkowski and Y. Sauvageau, (1978), 'Peer Group Effects and Educational Production Functions,' *Journal of Public Economics*, v10, 97-106.

[12] Herstein,I., and J.,Milnor, (1953),'An Axiomatic Approach to Measurable Utility', *Econometrica*, v21, 291-97.

[13] Klette, T. J, Moen, J., and Z. Griliches, (1999) 'Do Subsidies to Commercial R&D Reduce Market Failures? Microeconomic Evaluation Studies', Cambridge: NBER Working Paper 6947

[14] Lewis, G., *Unionism and Relative Wages in the United States: An Empirical Enquiry.* Chicago: University of Chicago Press, 1963.

[15] Magden, L., and R., Brookmeyer, (1993),'Analysis of infectious disease data from partner studies with unknown source of infection', *Biometrics*, 49:1110-6.

[16] Manski, C., (1995), *Identification in The Social Sciences*, Cambridge and London: Harvard University Press.

[17] Manski, C., (1997), 'The Mixing Problem in Programme Evaluation', *Review of Economic Studies*; v64 n4 October 1997, pp. 537-53.

[18] Philipson, T., (1997), 'Qualitative Issues in Assessing External Treatment Effects in HIV Trials', forthcoming, *Quantitative Issues in Evaluating HIV-Prevention Programs*, edited by R. Brookmeyer, and E. Kaplan, Yale University Press.

[19] Philipson, T., and J., Desimone, (1997), 'Experiments and Subject Sampling', *Biometrika*, v 84, 221-34.

[20] Over, M., and M., Ainsworth, (1997), *Confronting AIDS: Public Priorities in a Global Epidemic,* World Bank Policy Research Report, Oxford University Press.

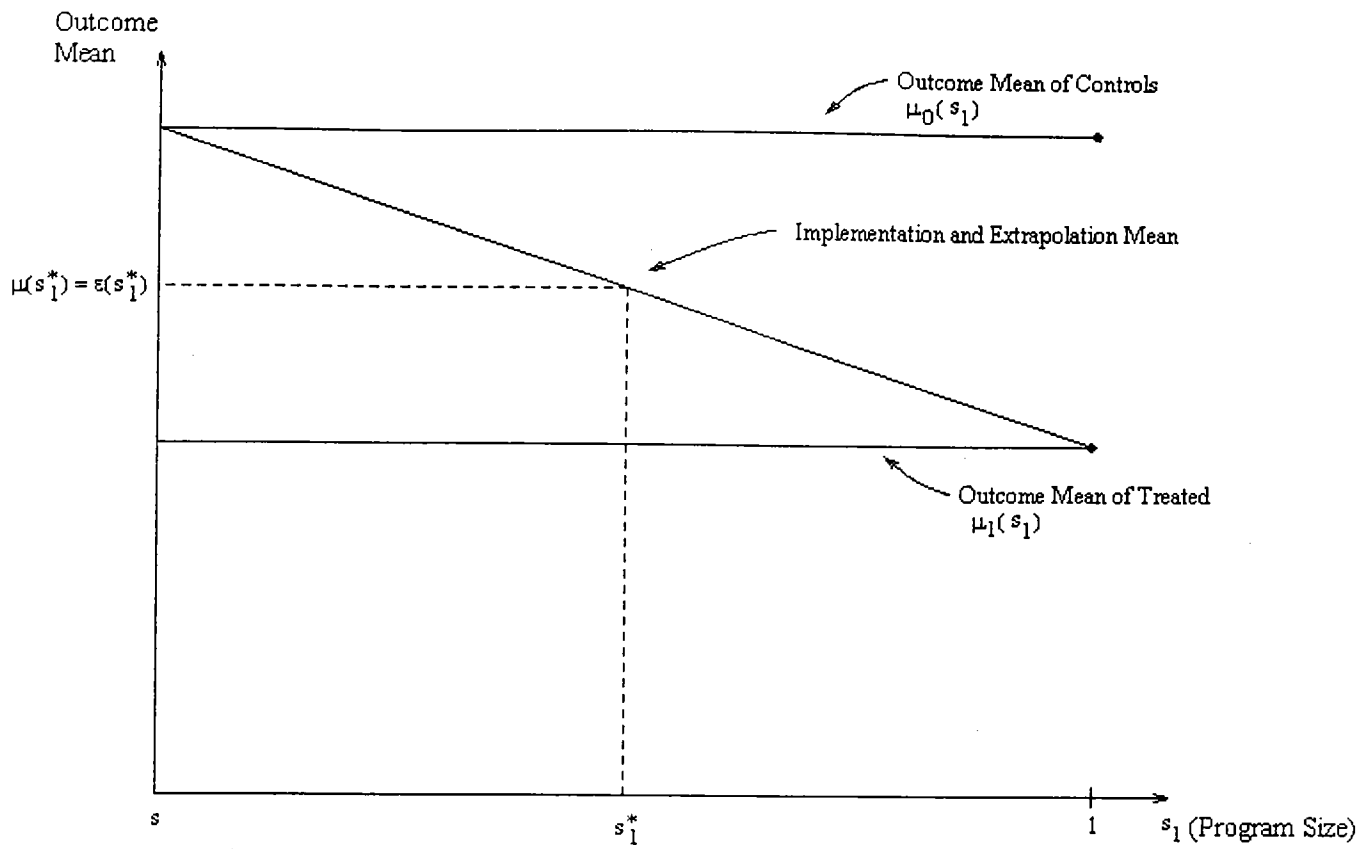[21] Summers, A., and B. Wolfe, (1977) 'Do Schools Make a Difference,' *American Economic Review*, v67, 639-652.

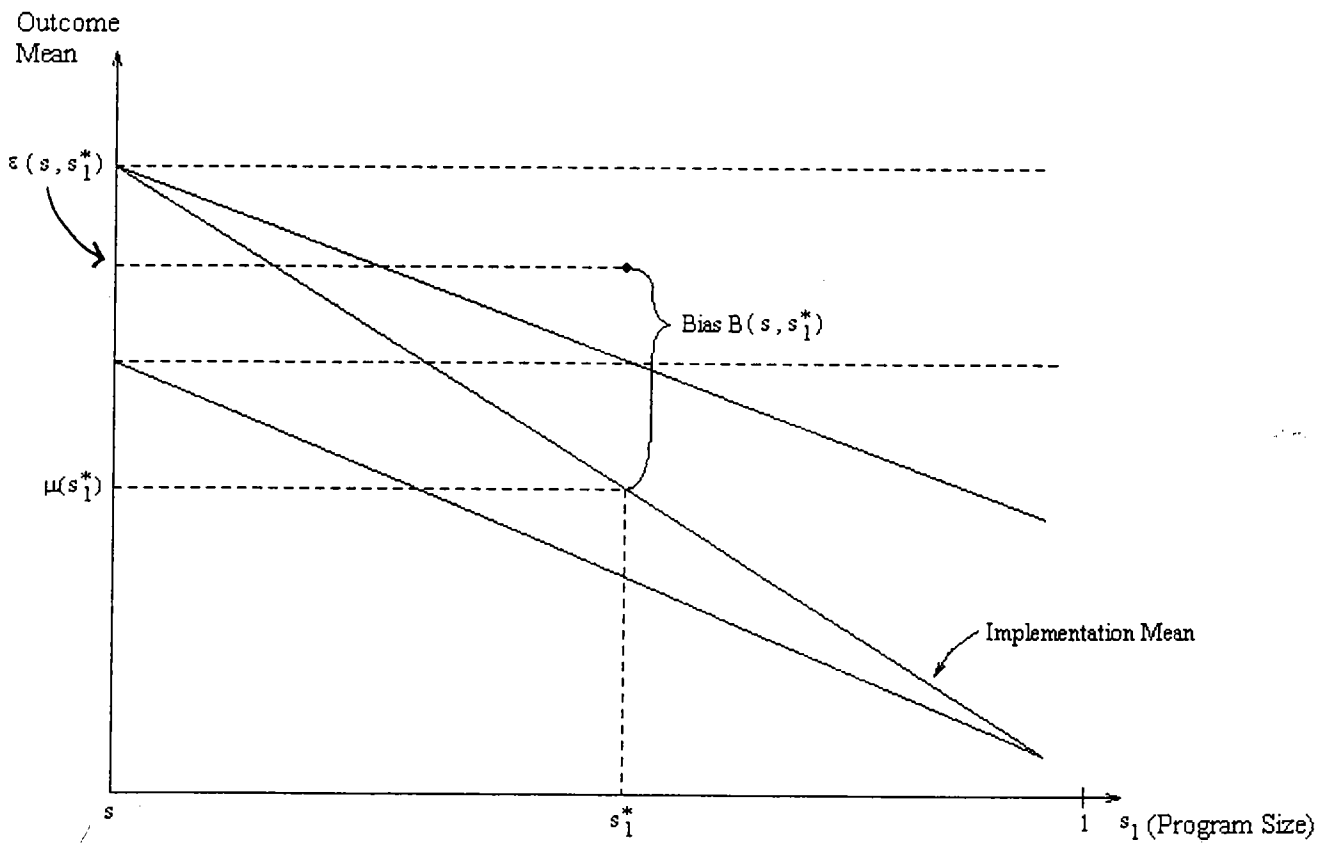**Figure 1: Extrapolation without External Treatment Effects**

Figure 2: Extrapolation with External Treatment Effects

$$
\begin{array}{c}
\textbf{Allocations (Columns)}\\
\textbf{Treatments (Rows)} \quad
\begin{array}{cccc}
\mu_1(s_1) & .. & .. & \mu_1(s_C)\\
\mu_0(s_1) & .. & .. & \mu_0(s_C)
\end{array}
\end{array}
$$

Figure 3: Outcome Means by Allocations and Treatments

$$
X = \begin{pmatrix}
1 & 1 & & s_1 \\
\vdots & \vdots & \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} ns_1 & \vdots \\
1 & 1 & & \\
1 & 0 & & \\
\vdots & \vdots & \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n(1-s_1) & \vdots \\
1 & 0 & & s_1 \\
& & & \\
\vdots & \vdots & & \vdots \\
& & & \\
1 & 1 & & s_C \\
\vdots & \vdots & \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} ns_C & \vdots \\
1 & 1 & & \\
1 & 0 & & \\
\vdots & \vdots & \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n(1-s_C) & \vdots \\
1 & 0 & & s_C
\end{pmatrix}
$$

Figure 4: Regressor Matrix for an Experiment
Estimating Private and External Effects

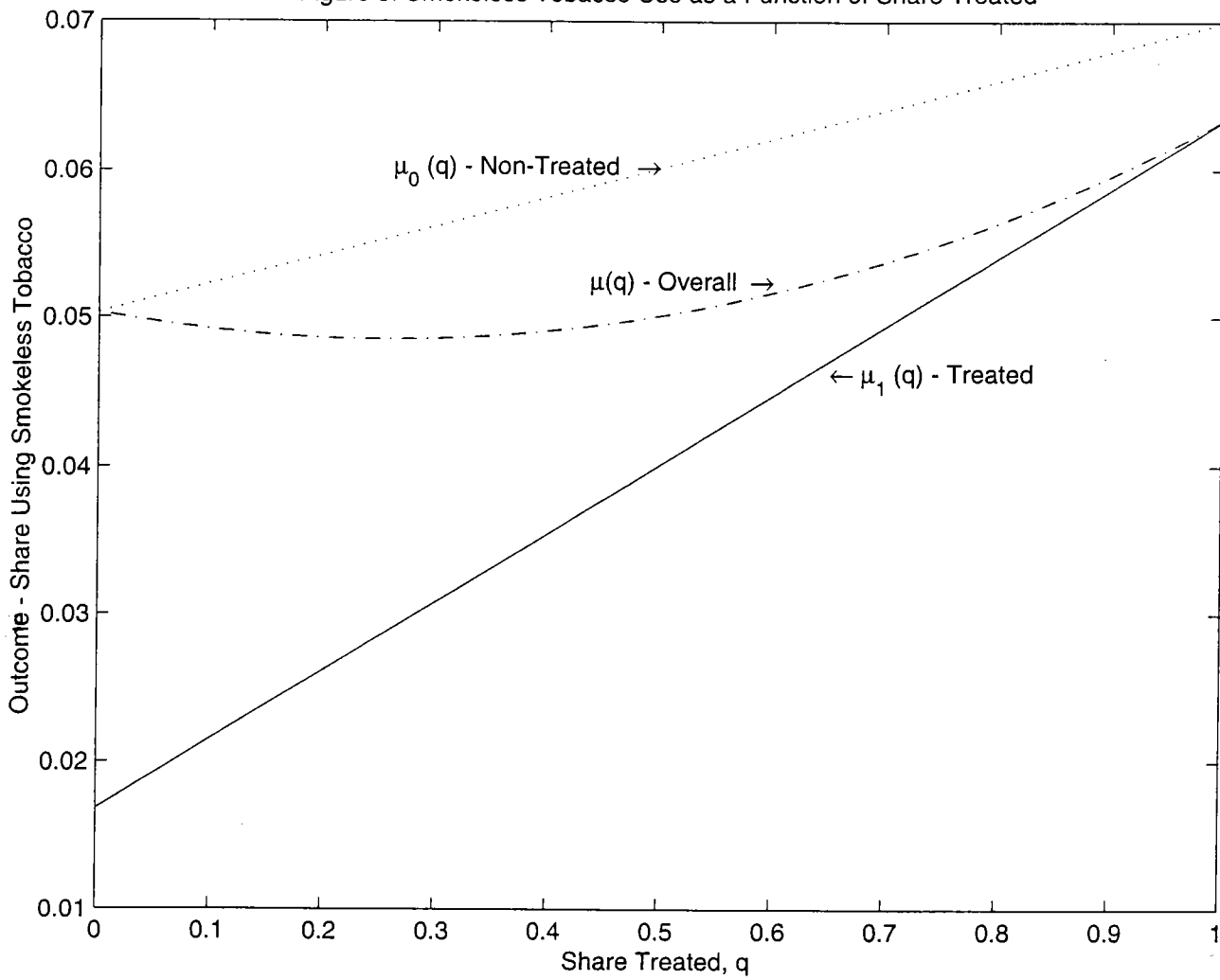Figure 5: Smokeless Tobacco Use as a Function of Share Treated

μ_0 (q) - Non-Treated $\rightarrow$

μ(q) - Overall $\rightarrow$

$\leftarrow \mu_1$ (q) - Treated

Outcome - Share Using Smokeless Tobacco

Share Treated, q

Table 1: Results of A Smokeless Tobacco Education Experiment

School	Share in Smokeless Tobacco Program	Smokeless Tobacco Use in Smokeless Program	Smokeless Tobacco Use Outside Smokeless Program	Smokeless Tobacco Use Overall
	s'	$\mu_1(s')$	$\mu_0(s')$	$\mu(s')$
1	0.27	5.9	10.2	9.0
2	0.28	6.9	1.3	2.9
3	0.29	0.0	4.9	3.4
4	0.33	1.2	1.7	1.6
5	0.35	1.8	6.9	5.1
6	0.43	1.0	12.8	7.7
7	0.48	5.2	5.8	5.5
8	0.53	3.2	1.8	2.5
9	0.64	6.0	7.2	6.5
10	0.64	8.1	8.0	8.1
11	0.68	4.6	6.7	5.3
12	0.68	2.5	4.1	3.0