Exercising in Herds:

Treatment Size and Status Specific Peer Effects in a Randomized Exercise Intervention

Philip Babcock
John Hartman

Department of Economics
University of California, Santa Barbara
Mail Stop 9210
Santa Barbara, CA 93106-9210

August 2010

Abstract

In a field experiment using university students, we find that subjects who have been incentivized to exercise increase their recreation center usage more if they have more friends who have been incentivized, and less if they have more friends in the control group. Controls, however, are not influenced by their peers. Findings highlight subtle effects of randomization, indicating that the fraction treated in an experiment of this kind has a large influence on outcomes, and that spillovers may vary greatly by treatment status. The methodology is applicable to other settings and quantifies spillovers that previous approaches do not detect.

We thank Kelly Bedarad, Eric Bettinger, Olivier Deschenes, Peter Kuhn, Heather Royer, Bruce Sacerdote, Jon Sonstelie, and Cathy Weinberger for helpful comments. We are grateful for funding by the Hellman Family Foundation, and excellent research assistance by Jennifer Carnan, Jennifer Schulte, Allison Nuovo, Christy Helvestine, Bonnie Queen, Jessica Evans, Natalie Brechtel, and Randi Golde. We are also grateful for assistance from Chris Clontz at the Recreation Center at the University of California, Santa Barbara.

1. Introduction

Rising rates of obesity and obesity-related health problems in the U.S. have motivated a search for interventions that cause people to exercise more.[1] Economists have studied exercise interventions not only for their relevance to pressing issues of public health but for what they may tell us about behavioral theories. Pay-for-exercise programs provide a setting in which to test models of self-control and pre-commitment (O'Donoghue and Rabin, 1999 and 2001, DellaVigna and Malmendier, 2006), habit-formation (Becker, 1992), and theories of intrinsic and extrinsic motivations (Deci, 1971, Benabou and Tirole 2003, 2006, Gneezy and Rustichini 2000a, 2000b). In a seminal paper on the topic, Charness and Gneezy (2009) explore these theories and find that short-term incentives can create long-run changes in behavior. There have also been numerous studies of exercise and wellness interventions in the workplace.[2] To date, however, we have found no research on the role of social effects or spillovers in these interventions. As has become apparent in recent research on medical and agricultural interventions in developing countries, estimating externalities may be essential for understanding the generalizability of an intervention (Duflo, 2006).

Further, an analysis of exercise spillovers in networks sheds light on subtle issues rarely addressed in a growing literature on social interactions. We know of no previous work estimating treatment-size-specific and treatment-status-specific peer effects. We note, in fact, that the dominant framework used in recent studies provides no means for doing so. In many settings, including but not limited to those featuring problems of self-

[1] Centers for Disease Control and Prevention (www.cdc.gov/obesity/childhood/prevalence.html.)
[2] Anderson et al. (2009) and Conn et al (2009) are two recent meta-surveys of workplace wellness interventions. Approximately 40 percent of large employers in the U.S. offer incentive-based programs designed to improve health (Hewitt Associates, 2002).

control, these parameters are potentially quite important and worth learning about.

In this paper, we investigate the social effects of exercise incentives in a field experiment on college students. As in Charness and Gneezy (2009) and Acland and Levy (2010), we survey college students and incentivize a random subset to exercise by offering to pay them for repeated visits to the campus recreation center. A unique feature of our experiment is that prior to treatment, we elicit a detailed friendship network from the subjects, all of whom live in the same residence hall. Random assignment of treatment creates random variation in the numbers of treated and untreated peers to which the subject is exposed, and from this we identify treatment-status-specific and treatment-size-specific peer effects. This new methodology can be used in other interventions, particularly for settings in which agents have the potential to be connected to one another by measurable social ties.

2. Background

Our strategy for combating the well known problems involved in identifying social effects is based largely on Moffitt's (2001) description of a partial population intervention. Partial population designs have recently been used to estimate spillovers hypothesized to arise from information transfer in schools or departments (Duflo and Saez 2003, Miguel and Kremer 2004), from imitation or social insurance between households (Kuhn et al, forthcoming, and Angelucci and DiGiorgio, 2009), and from learning or imitation in schooling decisions (Bobonis and Finan, 2009, Lalive and Cattaneo, 2009). One could conceptualize an ideal partial population design as one which begins with a sample of social groups or villages between which there are no spillovers.

Within the treatment groups or villages, a random subset of agents or households is treated. Differences in group-level means in this setting identify several quantities of interest: a) The difference between mean outcomes for treated and control villages captures average group-level treatment effects; b) The difference between mean outcomes of treated subjects in treated villages and control subjects in untreated villages estimates the average effect of village treatment on treated subjects; c) The difference in the mean outcomes of untreated subjects in treatment and control villages captures the effect of village treatment on untreated subjects.

In practice, interventions will usually deviate from this idealization in some way. For example, in Miguel and Kremer's (2004) study of a program to fight intestinal worms in Western Kenya, treatment is randomized across villages but not within village, as the researchers cannot manipulate the decision to receive treatment. Thus, though the school-level randomization identifies an effect of school-level treatment that includes within-school spillovers, the researchers are forced to rely on non-experimental methods to isolate social effects within school.

Our paper also deviates from the idealized partial population experiment in that we use experimentally induced differences in treatment across friendship networks to quantify differences in exposure to treated peers. Thus, we replace the in-group/out-group structure of the peer group with a network. One advantage of our design is that we randomize at the individual level and do so in a way that obviates concerns about differential rates of take-up between treated subjects and controls. Because treatment in our setting is defined as *being eligible* for a cash reward for exercise, treatment and take-up are essentially identical. What our experiment will highlight is that spillovers may be present even when there is no discernible effect of group treatment on controls, i.e., *when the difference in group means described in c) of an idealized partial population experiment is zero.* This is because exposure to treated subjects may affect treated and untreated subjects in different ways.

The simple schematic in Figure 1 fleshes out these ideas by depicting an experiment in which the fraction s of a target population (or village) is treated and the fraction ($1-s$) is not.[3] If there are no spillovers and a person's outcome depends only upon his own treatment status (Rubin, 1986), the mean outcome for the population, $\mu(s)=E[Y|s]$, is the weighted average of treatment and control means

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

where treatment and control groups are indexed by subscripts 1 and 0, respectively. Figure 1A depicts this situation when a fraction s_1 of the sample is treated. The horizontal dotted lines show control and treatment means. The population mean, $\mu(s)$, is then the solid diagonal line that rises with the fraction treated, as weight is shifted from control mean to treatment mean.

If spillovers are present, control and treatment means vary with s. Vaccination for infectious diseases is a canonical example. Health outcomes of both treated and untreated subjects are expected to improve with an increase in the fraction of the population vaccinated. Figure 1B illustrates this setting by allowing $d\mu_1/ds=d\mu_0/ds>0$ (with $\mu_0(s)$ and $\mu_1(s)$ represented by dashed lines.) The projected mean for the sample, if all individual were to be treated, is then $\mu_1(1)$. The difference between $\mu_1(s_1)$ and $\mu_1(1)$ is the error associated with a projection that assumes away externalities. Philipson (2000) calls this difference "program implementation bias."

[3]Phillipson (2000) and Manski (2009) contain detailed treatments of the issues sketched here.

It need not be the case that treatment and control means vary with fraction treated in the same way. In Figure 1C, the treatment mean rises with *s* and the control mean does not. This could occur in the pay-for-exercise setting if peer effects help individuals overcome self-control problems by pre-committing to exercise with others who have been similarly incentivized. If there is little or no incentive to exercise for the untreated, their outcomes may not vary with the fraction treated.[4] We note that the comparison of means described in c) for an idealized partial population design would detect no "spillover" in this setting.[5] There are peer effects, then, that an idealized partial population experiment cannot isolate and detect.

Recent literature includes a number of concrete examples. In Kuhn et al (forthcoming), the authors estimate consumption spillovers on lottery non-particpants resulting from exposure to lottery winners. But they cannot detect whether lottery winners in a winning postal code influence each others' consumption. Angellucci and DeGiorgi (2009) estimate consumption spillovers from households eligible for transfers to wealthier, ineligible households, but cannot estimate spillovers between eligible households. Lalive and Cattaneo (2009), using similar data, summarize concerns about this kind of spillover between the treated (p. 460): "If children from poor households only interact with other children from poor households, there could be important social spillover effects that can not be detected with the PROGRESA experiment."

Figure 1 illustrates the importance of treatment-status-specific estimates of spillovers. It also highlights the difficulty of extrapolating the effects of program

[4] It is also possible that treatment and control outcomes could respond in opposite directions. This could occur in the pay-for-exercise setting if, for example, denial of the opportunity to be paid to exercise creates discouragement that is amplified by contact with treated peers.

[5] To be more precise, measured differences in group means described in a) and b) would include spillovers and private effects, but the attempt to isolate a spillover effect (in c)) would produce a zero estimate.

implementation from an experiment, absent knowledge of how outcomes vary with the size of the treatment group. We employ a method, described in the remaining sections, for identifying external effects when friendship networks are a meaningful reference group.[6]

Main findings are that treated students with more treated friends increased their usage of the recreation center more, while treated students with more untreated friends increased their recreation center usage less. Untreated subjects, however, did not alter their recreation center usage at all, and were not influenced by the treatment status of their peers. Our analysis implies that if all subjects in the target sample had been treated, the change in recreation center visits would have been 64% larger than what would have been extrapolated from an estimate that did not take into account spillovers.

3. Experimental Design

Wave I of the experiment entailed surveying and taking bio-measures of 272 incoming freshman students in Santa Catalina Residence Hall at the University of California, Santa Barbara. Students were recruited at our table outside the dormitory dining hall during move-in week, Sept.19-24, 2009, and were paid $8 for filling out a survey and giving consent to be contacted later. They were also entered in a lottery for two iPod Nanos, retail value $149.00.

In Wave II, which took place Oct.5-12, we administered a friendship survey to the

[6] Miguel and Kremer (2003), in a companion piece to their 2004 paper, also elicit friendship networks. They examine how these ties predict subjects' decisions to take up treatment. Our paper resembles this work in the use of networks, but differs in that our design is based on randomization at the individual level. Similarly, Oster and Thornton (2009) elicit friendship ties of at most 3 friends and estimate spillovers in menstrual cup usage. There, the exercise is greatly simplified in a way not usually practical. (This is because spillovers for the controls—who are not provided with the product—are deemed implausible.)

222 students who responded to our emails or returned to our table at the residence hall, and so remained in the experiment.[7] In the friendship survey, which they filled out online or at our table on a laptop, students were instructed to click on the checkboxes next to the names of the people they knew. The list of names from which they chose included only students in the experiment,[8] and they were shown at most 50 names per screen. Subjects were paid $8 for filling out the friendship survey and were entered in an iPod lottery. Treatment and control status was assigned at this time. Subjects randomized into the treatment group were informed they would be paid $80 if they visited the University Recreation Center 8 times or more between Oct. 12 and Nov. 8. We reiterate that treatment is defined here as being eligible to be paid for exercising, so the distinction between take-up and assignment essentially vanishes. By definition, all 222 students who were paid to fill out friendship surveys "took up" treatment if they were assigned to treatment, or stayed in the experiment as controls if they had been assigned to the control group.

In order to visit the Recreation Center, subjects had to show their student identification cards, which were then read by a bar-code reader. This was the source of data on recreation center visits. The treatment period ended on Nov. 8. Wave III began immediately thereafter. From Nov. 8-19, we surveyed and measured the subjects again, paid them $12 and entered them in an iPod lottery. At the end of Wave III, we paid $80 to each subject in the treatment group who had visited the facility eight times in the allotted period.

---

[7] Among those not returning for the second interview were students who dropped out of the university, switched email addresses, or left the residence hall to live elsewhere.

[8] To comply with human subjects protocols, we needed a subject's consent to include her name on the list. Of the 222 students in Wave II, 209 gave consent to have their names included on the friendship survey. Thus, there were a few subjects who could choose their friends but could not be chosen.

Several unique aspects of the research design in this pay-for-exercise experiment merit comment, as they relate directly to the identification strategy. The recreation center was about a 25-30 minute walk from the residence hall. Thus, visits to the recreation center involved nontrivial commitments of time and effort. Subjects were recruited from the same dormitory so that they would be likely to know some of the other subjects in the experiment. Approximately 20% of the students living in the dormitory participated in the experiment. The research design allowed a period for incoming freshman to form acquaintanceships with others in the residence hall, and so with other subjects in the experiment. After this period, but prior to treatment, friendship networks were elicited. Thus, the friendship networks were not influenced by treatment or treatment status. Random assignment of treatment served two purposes in this experiment. It induced random variation in exposure to treatment, which allowed identification of individual effects, and it induced random variation in exposure to treated and untreated peers, which allowed identification of external effects.

Figure 2 shows the friendship network elicited from Santa Catalina residents who were subjects in the experiment. Square nodes represent treated subjects, whereas circular nodes denote control subjects. An arrow from node A to node B implies that subject A checked the box indicating she knew subject B. Netdraw, which was used to create this graph, causes two subjects to appear farther apart on the page when the number of degrees of separation between them is larger.[9] Clustering of nodes in the diagram indicates clustering in the network. To the extent, then, that the squares are scattered randomly among the circles, the figure suggests a random distribution of treated subjects within the friendship network. Figure 3 shows the friendship group for one (circular)

[9] Borgatti, Everett, and Freeman (2002).

node, a control subject. As is apparent from the figure, the subject indicated that she knew 6 treated subjects and 11 control subjects. Random variation in numbers of treated and control subjects known by a subject, induced by random assignment to treatment, is pivotal in the analysis.[10]

4. Results

A. Individual Effects

Table 1 shows summary statistics for treatment and control groups in the Santa Catalina experiment, by age, race, gender, physical characteristics, pre-treatment recreation center usage, and treatment status of friends. The composition of control and treatment groups appears very similar in most categories. Age, race, and physical characteristics, are nearly identical for the two groups. Though treated subjects had slightly fewer treated friends than control subjects, and more men were randomized into the treatment group than the control group, the far-right column shows that none of the differences between the groups is statistically significant.

Figure 4 shows the rate of recreation center usage for treated and control subjects before and during the treatment period. Control subjects and treated subjects made 0.58 and .5 recreation center visits per week, respectively, in the pre-treatment period. During the 4-week treatment period, control subjects continued to visit the recreation center at about the same rate, while treated subjects increased their visits to an average of 1.77 per

[10] An alternative approach to the one used here would have been to stratify treatment assignment by friendship group (to insure variation in numbers of treated and untreated friends.) This would have been a greater departure from standard approaches, which focus on estimation of individual effects. Also, it would have created logistical problems in a brief 10-week quarter, as friendship data would have to be analyzed before treatment was assigned. For a discussion of trade-offs involved in estimating individual and external effects at the same time, see Philipson (2000).

week. This difference is highly significant (with p-value<0.001). We conclude that, as expected, treatment incentivized subjects to visit the recreation center.

One could worry that students simply had their cards read at the recreation center door, then turned around and left. This does not appear to have been the case. Figure 5 displays self-reported levels of exercise by treatment status, before and during the treatment period. The figure shows the number of times per week subjects exercised moderately or vigorously for 30 minutes or more. For control subjects, the number of exercise sessions dropped off during the treatment period (perhaps because the school term progressed and they grew busier), while for treated subjects, the number of sessions rose. The difference in differences is about .5 sessions per week and is statistically significant with p-value .07.

In summary, the evidence indicates that treated subjects visited the recreation center significantly more often during the treatment period than control subjects, and further, that treatment induced effort-intensive changes in exercise not limited to card-swiping visits or substitution between exercise venues.

B. Peer Effects and Recreation Visits

In the regressions summarized in Table 2, the dependent variable is the increase in recreation usage, from pre-treatment to treatment period, and the regressors are the number of treated friends, the number of untreated friends and a constant. Columns 1 and 2 summarize results by treatment group. For treated subjects, having 1 more treated friend is associated with increased recreation center usage that is .13 visits per week larger, while having 1 more untreated friend is associated with an increase that is .065

visits per week smaller. This finding is consistent with an interpretation in which treated friends help a treated subject to go to the recreation center more frequently during the treatment period (for example, by coordinating plans to create commitment devices, by reminders and encouragement, or by companionship that reduces the effort cost of exercising). Untreated subjects, in contrast, may distract, discourage or otherwise redirect a treated subject toward other activities. Column 2 indicates that unlike treated subjects, control subjects do not respond to whether their friends are treated or untreated. One possible reason for this difference is that control subjects may face no commitment problems: Having little or no incentive to visit the recreation center, they may derive no benefits from pre-commitment. We emphasize that the interpretations above are simply possibilities, and we do not claim to have established that they are the driving mechanisms. Table 2 does indicate, however, that the choices of treated subjects are sensitive to the treatment status of their friends, whereas the choices of control subjects are not.

In columns 1 and 2, a subject's response to the treatment status of her peers may be in some way confounded with the number of friends she has. An agent will tend to have more treated friends and more untreated friends if she has more friends overall. Columns 3 and 4, however, reveal that a subject's overall number of friends has virtually no influence on recreation center visits, and that this holds both for treated and control subjects. In short, the treatment composition of the friendship group matters for treated subjects, but the size of the friendship group does not.

In Table 3, the dependent variable is number of recreation visits per week in the treatment period, and we include pre-treatment recreation center usage as a regressor. In this specification, we let the data tell us how strongly pre-treatment recreation center usage influences the rate of usage during treatment. (The implicit assumption in the regressions of Table 2, which feature change in usage as the dependent variable, is that this coefficient is equal to one.) Results in Table 3 are similar to those in Table 2, and would indicate that the association between treatment composition of the friendship group and increased recreation center usage for the treated is not driven by the assumption above. Interestingly, however, for the control group, an increase of 1 visit per week in the period before treatment is associated with a change of almost exactly 1 visit per week in the treatment period. Thus pre-treatment behavior strongly predicts control-group behavior during treatment. For the treatment group, in contrast, an increase of 1 visit per week in the period before treatment is associated with a change of .57 visits per week in the treatment period. Pre-treatment behavior, for the treated, is not as strong a predictor of choices during treatment.

It may also be informative to examine the outcome that may have mattered most to the treated subjects: reaching the threshold of 8 visits and earning the $80 payment. Table 4 shows results from a linear probability model, analogous to the model in Table 3. The dependent variable is whether or not the subject visited the recreation center 8 times during the treatment period. Results are very similar in flavor to those in Table 3. Treated subjects had a .035 greater probability of exercising 8 times in the treatment period (and being paid) if they had 1 more treated friend and were .029 less likely to do so if they had 1 more untreated friend, though we note that the coefficient on treated friends is not precisely estimated (p-value=.23) Similar to the findings in Table 3, the numbers of treated and untreated friends had no clear effect on control subjects' probability of exercising 8 times.

C. Extrapolating to Implementation

How large are the effects in Tables 2 and 3, and what do they imply about extrapolation to a program that would expand treatment to the entire sample? In the experiment, treated subjects have 2.7 treated friends and 4.8 untreated friends. If treatment were to be expanded to the entire sample, treated subjects would have 4.8 more treated friends and 4.8 fewer untreated friends. Given the (strong) assumption of linearity in the external effect, the extrapolated increase in usage associated with treating the entire sample would be 4.8(.1)-4.8(-.054) =.74 visits/wk, based on the (conservative) estimates in Table 3, Column 1. This increase is statistically significant at the 5% level. The difference between treatment and control means for exercise during the treatment period was 1.15 visits per week, and was also statistically significant (p-value<0.001). The estimated external effect, then, is nearly as large as the individual effect, and program implementation bias is .74/1.15, or about 64%. Figure 5 depicts these effects graphically, using the schematic of Section 2.

Figure 5 addresses the hypothetical expansion of the fraction treated from about ⅓ of the sample to the entire sample of particpants in the experiment (the full sample being about 1/5 of the residence hall). From a policy perspective, though, this may not be the relevant counterfactual. One might want to ask what the effect would be of implementing treatment for the entire residence hall. In many interventions, it is unrealistic to imagine one could ever get 100% saturation of treatment, but here "treatment" simply means being made eligible to receive rewards for exercise. It is not difficult to imagine extending eligibility to the entire dorm. We emphasize that the field experiment included about a fifth of the residence hall, and that only individuals in the experiment were in the elicited friendship network. Thus, there may have been other relevant "untreated" friends—any friend in the residence hall who was not in the experiment. Implementation on the entire residence hall would yield more than 4.8 additional treated friends per subject, on average, and would thus be expected to produce a larger exercise effect than was extrapolated for implementation on the subpopulation in the experiment. We submit, then, that the estimate of .74 additional visits per week associated with expanding treatment to the subpopulation in the experiment, and the corresponding estimate of 64% program implementation bias, could be interpreted as lower bounds for residence hall (or broader) implementation.[11]

In the previous section, we briefly described possible mechanisms for the effects observed in the field experiment. It has long been argued, both casually and in formal models, that exercise regimens require self-control and far-sightedness to carry out. It may not be surprising, then, to have found evidence of strong peer influence in this setting. What is less expected, perhaps, is that untreated friends would seem to exert a *negative* influence on treated subjects. This highlights three implications we believe to be important. Firstly, there is evidence in this experiment that the relative size of treatment and control groups may have a large influence on behaviors. Secondly, this influence appears to differ greatly between treated subjects and controls. Thirdly, many people trying to increase their exercise levels may have difficulty doing so if they do not have

[11] We note, however, that untreated friends *in the experiment* may differ in their influence on treated subjects from untreated subjects who were never in the experiment.

one or more friends trying to reach the same goals. Randomization, itself, would seem to have many subtle effects, and these may need to be estimated, as was done here, for experiments to be interpreted meaningfully.

5. Conclusion

In a field experiment conducted on college freshmen, all of whom lived in the same residence hall, we elicited friendship networks and offered monetary incentives for using the recreation center to a treated subset. We found that treated students with more treated friends increased their usage of the recreation center more, while treated students with more untreated friends increased their recreation center usage less. Untreated subjects, however, did not alter their recreation center usage at all, and were not influenced by the treatment status of their peers.

We highlight three main contributions of this research: Firstly, the experiment provided evidence of large peer effects in a self-control setting that is of interest to researchers and policy-makers. Secondly, the findings offered empirical evidence of large variations in the magnitude of spillovers by treatment status and treatment size, highlighting potential effects of randomization in a controlled setting that are sometimes ignored. Thirdly, this research design offered a simple methodology that could be applied over a wide array of experiments and interventions that feature settings in which social networks could be expected to influence behaviors.

References

Acland, D. and M. Levy (2010), “Habit Formation and Naivete in Gym Attendance: Evidence from a Field Experiment,” mimeo, UC Berkeley.

Anderson, L., Quinn, T., Glanz, K. et al. (2009). “The Effectiveness of Worksite Nutrition and Physical Activity Interventions for Controlling Employee Overweight and Obesity: A Systematic Review” *American Journal of Preventative Medicine* 37(4): 340-357.

Angelucci, M., and G. De Giorgi (2009).”Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles’ Consumption?” *American Economic Review*, vol. 99(1), pp. 486-508.

Bobonis, G. J., and F. Finan (2009), "Neighborhood Peer Effects in Secondary School Enrollment Decisions", *Review of Economics and Statistics* 91 (4), pp. 695–716.

Becker, G. (1992) “Habits, Addictions and Traditions,” *Kyklos*, 45(3): 327-345.

Benabou, R., and Tirole, J., (2003), “Intrinsic and Extrinsic Motivation” *Review of Economic Studies*, 70(3): 489-520.

Borgatti, Steve, Martin Everett, and Linton Freeman, *Ucinet 6.87 for Windows: Software for Social Network Analysis* (Harvard, MA: Analytic Technologies, 2002).

Charness, G., and Gneezy, U. (2009), “Incentives to Exercise” *Econometrica* 77(3): 909–931.

Conn, V., Hafdahl A., and Cooper, P. (2009), “Meta-Analysis of Workplace Physical Activity Interventions” *American Journal of Preventative Medicine* 37(4): 330-339.

Deci, E. (1971), “Effects of Externally Mediated Rewards on Intrinsic Motivation” *Journal of Personality and Social Psychology*, 18: 105-115.

DellaVigna, S., and Malmendier, U. (2006), “Paying Not To Go To the Gym” *American Economic Review* 96: 694-719.

Duflo, Esther (2006), “Field Experiments in Development Economics,” Discussion paper.

Duflo, E., M. Kremer, and J. Robinson (2004). “Understanding Technology Adoption: Fertilizer in Western Kenya – Preliminary Results from Field Experiments,” Unpublished manuscript, Massachusetts Institute of Technology.

Duflo, E., and E. Saez (2003). “The Role of Information and Social Interactions in

Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, 118(3), 815–842.

Gneezy, U., and Rustichini, A. (2000). "Pay Enough or Don't Pay at All", *Quarterly Journal of Economics* 115(3): 791-810.

Hewitt Associates LLC. (2002). "Health Promotion/Managed Health Provided by Major U.S. Employers in 2001."

Kuhn, P., P Kooreman, AR Soetevent, A Kapteyn (forthcoming), "The Effects of Lottery Prizes on Winners and their Neighbors: Evidence from the Dutch Postcode Lottery," *American Economic Review.*

Lalive, R. and Cattaneo, A. (2009). "Social Interactions and Schooling Decisions," Review of Economics and Statistics, Vol. 91(3), pp. 457-477.

Manski, C.F. (2009), "Identification Of Treatment Response With Social Interactions," Mimeo, Northwestern University.

Miguel, E. and Kremer, M. (2004). "Worms: identifying impacts on education and health in the presence of treatment externalities." *Econometrica* 72(1): 159-217.

Miguel, E. and Kremer, M. (2003). "Networks, Social Learning, and Technology Adoption: The Case of Deworming Drugs in Kenya," mimeo, Harvard.

Moffitt, R. A. (2001). "Policy Interventions, Low-Level Equilibria, and Social Interactions," in *Social Dynamics*, ed. by S. N. Durlauf, and H. P. Young, pp. 45–82. MIT Press.

O'Donoghue, T. and Rabin, M. (1999), "Doing it Now or Later," *American Economic Review* 89(1): 103-124.

O'Donoghue, T. and Rabin, M. (2001), "Choice and Procrastination," *Quarterly Journal of Economics* 116(1): 121-160.

Oster, E. and Thornton, R. (2009), "Determinants of Technology Adoption: Private Value and Peer Effects in Menstrual Cup Take-Up," NBER Working Paper No.14828.

Philipson, T. (2000), "External Treatment Effects and Program Implementation Bias," NBER Working Paper, 250.

Rubin, D. (1986), "Statistics and Causal Inference: Which Ifs Have Causal Answers," *Journal of the American Statistical Association*, 81, 961-2.

Table 1 - Descriptive Statistics

	Treatment		Control		
	Mean	St. Dev.	Mean	St. Dev.	P-Value (Diff=0)
Age	18.01	0.29	18.01	0.33	0.92
Black	0.03	0.18	0.04	0.21	0.73
Hispanic	0.24	0.43	0.25	0.43	0.92
Male	0.62	0.49	0.53	0.50	0.15
Height	67.54	3.42	66.93	3.94	0.23
Weight	146.08	21.28	143.85	26.76	0.49
Visits/wk (pre)	0.50	0.76	0.58	0.82	0.45
Treated Friends	2.66	2.29	2.87	2.05	0.50
Untreated Friends	4.77	4.21	5.67	3.90	0.11
Obs	86		136		

Table 2 - Regressions of Exercise Change on Number of Treated and Untreated Peers - by Treatment Group

	1	2	3	4
	Treatment	Control	Treatment	Control
	Δvisits/wk	Δvisits/wk	Δvisits/wk	Δvisits/wk
#Treated friends	.13**	-.0027		
	(.051)	(.039)		
#Untreated friends	-.065*	.0096		
	(.033)	(.018)		
#Friends			-.0016	.0057
			(.017)	(.01)
Constant	1.2***	-.0064	1.3***	-.0083
	(.16)	(.093)	(.16)	(.09)
Observations	86	136	86	136
R-squared	0.05	0.00	0.00	0.00

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3 - Regressions of Exercise Visits on Number of Treated and Untreated Peers - by Treatment Group

	1	2	3	4
	Treatment	Control	Treatment	Control
	Visits/wk	Visits/wk	Visits/wk	Visits/wk
#Treated friends	.10**	-.0044		
	(.046)	(.039)		
#Untreated friends	-.054*	.013		
	(.032)	(.018)		
#Friends			-.0035	.007
			(.016)	(.0091)
Visits/wk(Pre)	.57***	.96***	.54***	.97***
	(.16)	(.098)	(.13)	(.061)
Constant	1.5***	.0032	1.5***	-.00039
	(.15)	(.096)	(.17)	(.09)
Observations	86	136	86	136
R-squared	0.21	0.68	0.18	0.68

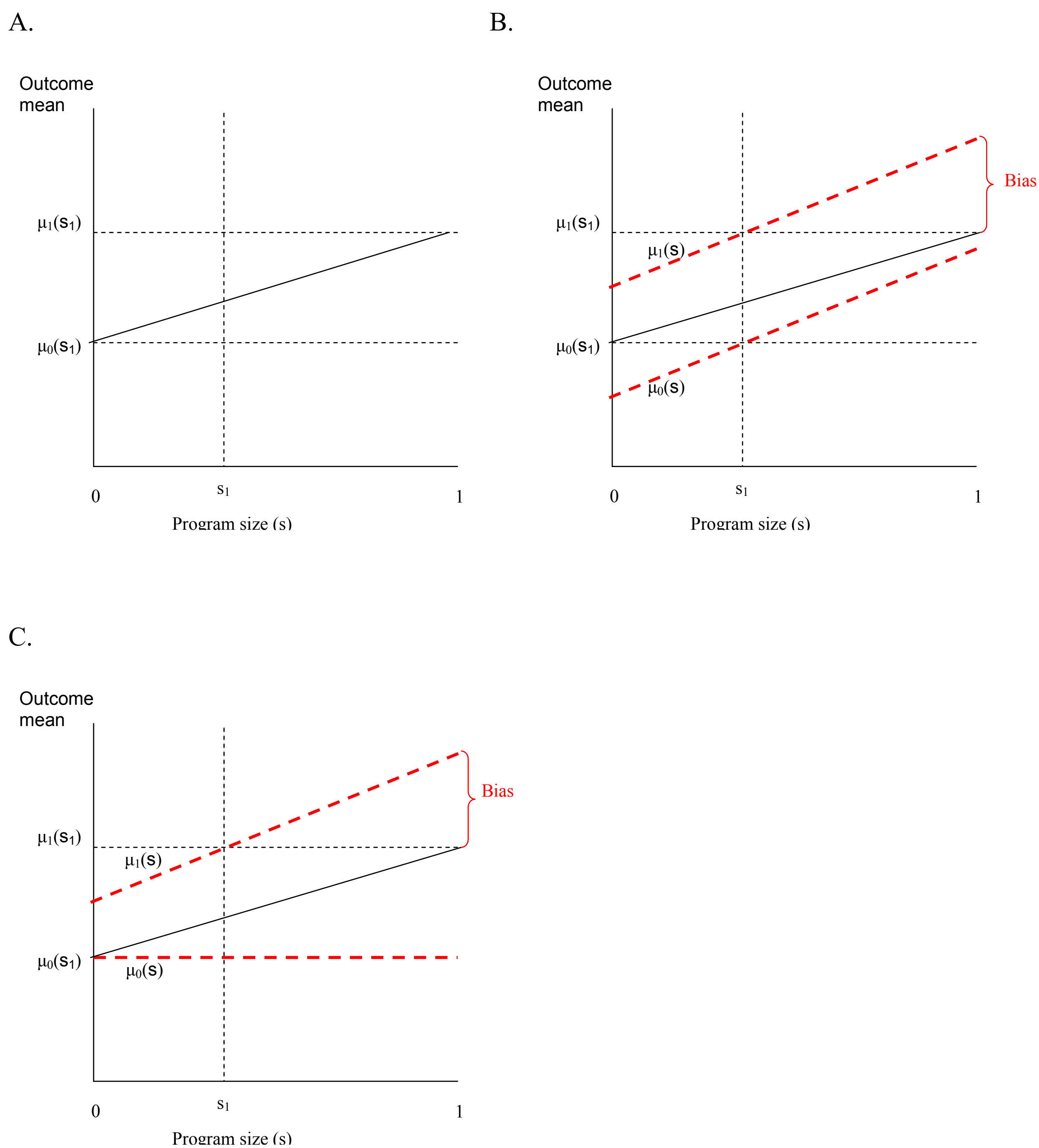
Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4- Regressions of Reaching 8-Visit Threshold on Number of Treated and Untreated Peers - by Treatment Group

	1	2	3	4
	Treatment	Control	Treatment	Control
	Visits/wk	Visits/wk	Visits/wk	Visits/wk
#Treated friends	.035	-.016		
	(.029)	(.014)		
#Untreated friends	-.027*	.01		
	(.016)	(.0078)		
#Friends			-.0065	.0016
			(.0084)	(.0041)
Visits/wk(Pre)	.17***	.26***	.16**	.26***
	(.058)	(.028)	(.067)	(.027)
Constant	.6***	-.036	.62***	-.041
	(.09)	(.041)	(.088)	(.041)
Observations	86	136	86	136
R-squared	0.09	0.44	0.07	0.43

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 1



A.
Outcome mean
$\mu_1(s_1)$
$\mu_0(s_1)$
0
s_1
1
Program size (s)
B.
Outcome mean
$\mu_1(s_1)$
$\mu_0(s_1)$
$\mu_1(s)$
$\mu_0(s)$
Bias
0
s_1
1
Program size (s)
C.
Outcome mean
$\mu_1(s_1)$
$\mu_0(s_1)$
$\mu_1(s)$
$\mu_0(s)$
Bias
0
s_1
1
Program size (s)

Figure 2

Santa Catalina Residence Hall
Friendship Network

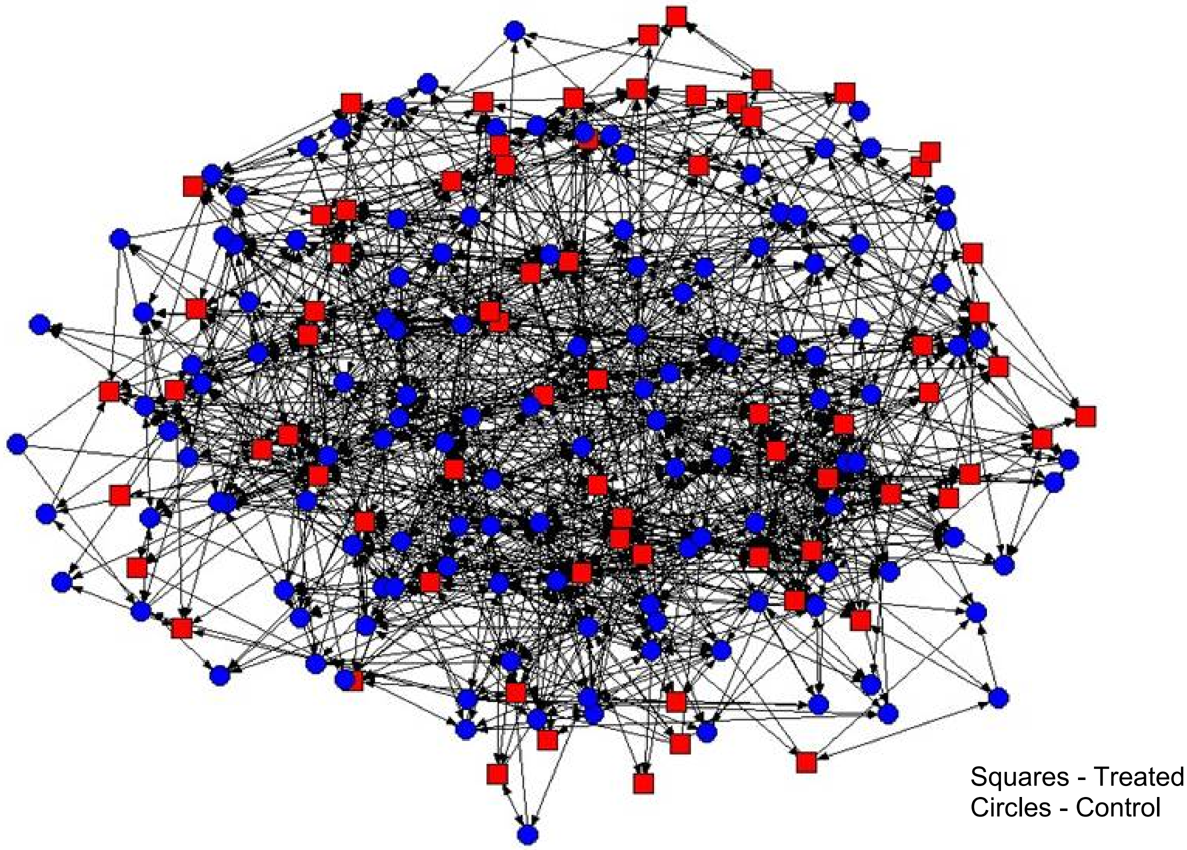
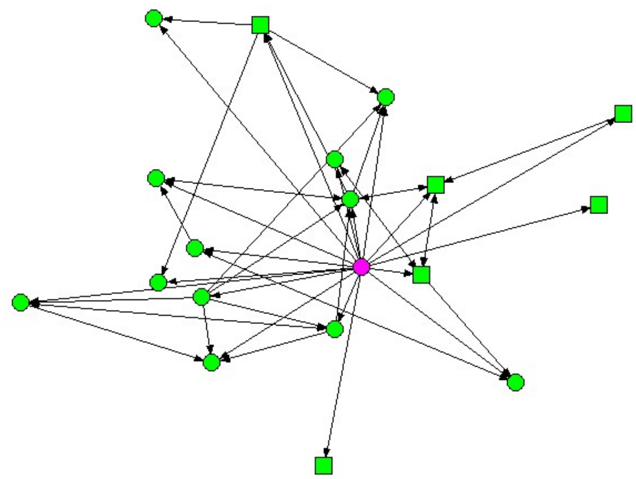


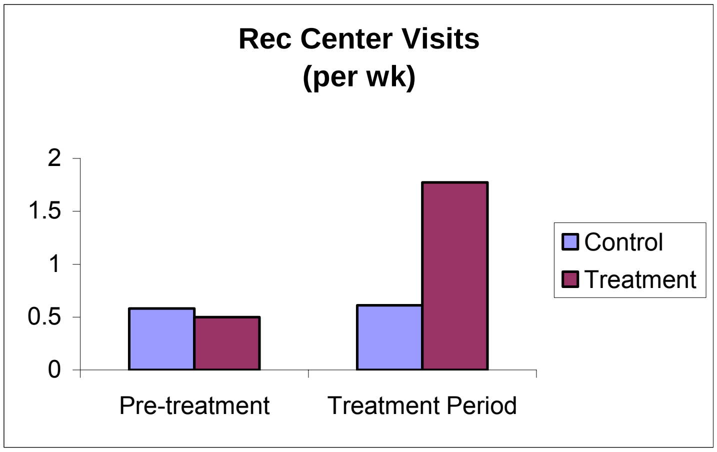
Figure 3

Friendship Links – Subject 137



Squares - Treated
Circles - Control

Figure 4



Rec Center Visits
(per wk)
2
1.5
1
0.5
0
Pre-treatment
Treatment Period
Control
Treatment

Figure 5

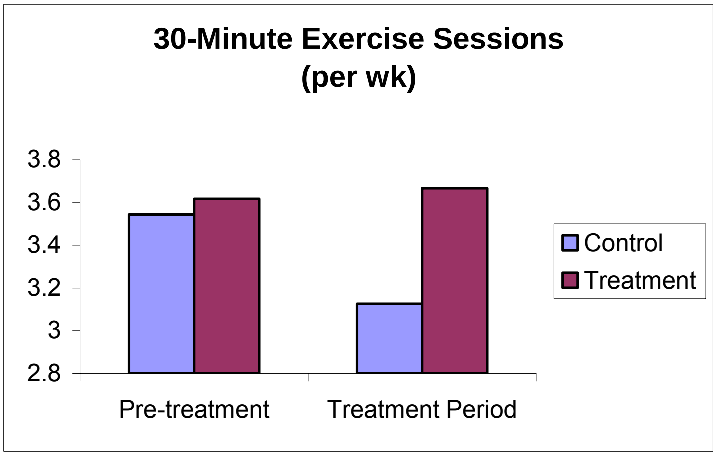

30-Minute Exercise Sessions
(per wk)
3.8
3.6
3.4
3.2
3
2.8
Pre-treatment
Treatment Period
Control
Treatment

Figure 6

Santa Catalina
Program Implementation Bias