ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 77, No. 3 (May, 2009), 909–931

INCENTIVES TO EXERCISE

GARY CHARNESS
University of California at Santa Barbara, Santa Barbara, CA 93106-9210, U.S.A.

URI GNEEZY
Rady School of Management, University of California at San Diego, La Jolla, CA 92093-0553, U.S.A.

Econometrica, Vol. 77, No. 3 (May, 2009), 909–931

INCENTIVES TO EXERCISE

BY GARY CHARNESS AND URI GNEEZY[1]

Can incentives be effective in encouraging the development of good habits? We investigate the post-intervention effects of paying people to attend a gym a number of times during one month. In two studies we find marked attendance increases after the intervention relative to attendance changes for the respective control groups. This is entirely driven by people who did not previously attend the gym on a regular basis. In our second study, we find improvements on health indicators such as weight, waist size, and pulse rate, suggesting the intervention led to a net increase in total physical activity rather than to a substitution away from nonincentivized ones. We argue that there is scope for financial intervention in habit formation, particularly in the area of health.

KEYWORDS: Exercise, field experiment, habit formation, incentives.

INTRODUCTION

ON SEPTEMBER 18, 2006, New York Mayor Michael Bloomberg announced a new policy he called conditional cash transfers. He said that the plan was designed to address the simple fact that the stress of poverty often causes people to make decisions—to skip a doctor's appointment or to neglect other basic tasks—that often only worsen their long-term prospects. Conditional cash transfers give them an incentive to make sound decisions instead. The intention was to provide conditional cash transfers to families of at-risk youngsters to encourage parents and young people to engage in healthy behavior, to stay in school, stay at work, and stay on track to rise out of poverty. Bloomberg also argued that the return on such investments is necessarily delayed, but that this is a clear path out of the cycle of poverty. Bloomberg's last comment is about changing peoples' habits. He believes that the cost (estimated at $42 billion) of the program is worth the benefit of this improvement in habits.

Whether or not Bloomberg is correct in his assessment, an underlying issue is whether we can construct mechanisms to induce better decision-making. As DellaVigna and Malmendier (2006) have nicely demonstrated, people make poor choices regarding membership options at a health club: people who choose to pay a flat monthly fee for membership in a gym pay more than if they would have chosen to pay a fixed cost per visit. So perhaps the incentives to exercise that are already present are ineffective or insufficiently salient. But can we improve on these existing incentives? Can we go beyond the mere

[1]We acknowledge helpful comments from Yan Chen, Martin Dufwenberg, Guillaume Fréchette, Jacob Goeree, Ulrike Malmendier, Uri Simonsohn, and Priscilla Williams, as well as seminar audiences at the Stanford Institute of Theoretical Economics, the Santa Barbara Conference on Communication and Incentives, the ESA meeting in Tucson, the ESA meeting in Lyon, CIDE in Mexico City, and Harvard University. Special thanks go to a co-editor and three anonymous referees, who suggested the second study with the biological measures reported in the paper. Charness and Gneezy each acknowledge support from the National Science Foundation.

 DOI: 10.3982/ECTA7416

identification of behavioral mistakes, and consider the issue of how a welfare maximizer would react if aware of his or her own bias?[2]

In this vein, the goal of the current paper is to test the conjecture that financial incentives can be used to develop or foster good habits. Habits are an important feature of daily life. However, people often follow a routine without much ongoing consideration about the costs and benefits of the constituent elements of this routine.

One such habit is that of regular physical exercise. The physical benefits of exercise are undeniable, as adequate exercise is associated with better health in many respects. In particular, obesity has become a prominent health issue; the National Center for Health Statistics reported that in 2003–2004 a startling 66 percent of American adults we overweight or obese. This is the highest level ever recorded; in comparison this rate was 47 percent in 1976–1980 and 55 percent in 1988–1994. Indeed, as suggested by this trend, the problem appears to be getting worse more recently: Adult obesity rates rose in 31 states from 2006 to 2007, according to the report from the Trust for America's Health (2007); rates did not decrease in any states. A new public opinion survey featured in the report finds 85 percent of Americans believe that obesity is an epidemic. This increased prevalence of obesity is paralleled by an increase in inactivity. Most jobs today are sedentary and overweight people are particularly likely to report being inactive.

Regular exercise combined with limiting calorie intake was shown to be most effective in reducing body mass (Andersen (1999)). Exercise provides health benefits even if people do not lose weight (Lee, Blair, and Jackson (1999)). There are also psychological benefits to exercising: People who exercise regularly are likely to be less depressed, have higher self-esteem, and have an improved body image (Brownell (1995)). Regular exercise may also reduce stress and anxiety (Kayman, Bruvold, and Stern (1990)).

The literature discusses four main barriers to activity (Andersen (1999)): lack of time, embarrassment at taking part in activity, inability to exercise vigorously, and lack of enjoyment of exercise. The traditional approach in economics involves providing financial incentives for people to engage in (or refrain from) various activities, but since strong (nonfinancial) incentives regarding habitual behavior are already in place without any intervention appearing to be necessary,[3] can there be much scope for intervention in the incentive structure?

[2]For a similar attempt in other economic areas, see the Thaler and Benartzi (2004) "save more tomorrow" plan. In a sense, this can be seen as complementary to the behavioral industrial-organization agenda, which considers how firms might react to consumer biases in product design, advertising, and so forth. See Ellison (2006) for a summary of the relevant literature, as well as related studies by Heidhues and Koszegi (2008) and Lauga (2008) and the recent review in DellaVigna (2009).

[3]That people are aware that exercise is beneficial is evidenced by the fact that Americans spend billions of dollars annually on diet books, exercise equipment, and weight-loss programs.

We discuss below two main hypotheses regarding the outcome of using financial incentives to shape habits. The first is the "crowding-out" hypothesis, according to which paying people for an activity (such as exercising) might destroy their intrinsic motivation to perform the task once the incentives are removed (Deci (1971), Gneezy and Rustichini (2000a, 2000b), Frey and Jegen (2001)). The alternative hypothesis is "habit-formation" behavior. The main idea here is that one's utility from consumption depends on one's past consumption (Becker and Murphy (1988)).

If it is possible to induce a beneficial habit, the policy implications are major. In this paper we undertake financial interventions, conducting two field studies in which we paid university students to attend the university's gym. In the first of our studies, we compare the behavior of three groups. All groups were given a handout regarding the benefits of exercise. One group had no further requirements; people in the other two groups were paid $25 to attend the gym once in a week, and people in one of these two groups were then paid an additional $100 to then attend the gym eight more times in the following four weeks.

We are able to observe attendance before the intervention, during the intervention, and for a period of seven weeks after the end of the intervention. The main result is that post-intervention attendance is more than twice as high for the high-incentive group as for the no-incentive group. This difference does not decline at all during the time following payment, suggesting that the effects do have some degree of persistence. There is very little difference between the behavior of the no-incentive and low-incentive groups, while there is a significant difference between the behavior of the low-incentive and high-incentive groups.

In our second study, we invited people to a first meeting in which we took biometric measures and gave them the handout regarding the benefits of exercise. They were paid $75 for this part, and were then invited to come twice more, so that we could obtain biometric information again. They were promised $50 for each of the two later meetings.

We randomly divided the participants into three groups. There were no additional requirements for the people in the control group. Participants in the second group were required to attend the gym once during the one-month intervention period, and participants in the third group were required to attend the gym eight times during the intervention period. We find a significant and persistent increase in attendance rates for people in the third group, and this increase is again entirely driven by people who had not been regular gym attendees (at least once per week). We also find improvement in the biometric measures for the third group relative to the other groups.

Our results indicate that it may be possible to encourage the formation of good habits by offering monetary compensation for a sufficient number of occurrences, as doing so appears to move some people past the "threshold" needed to engage in an activity. It may often be the case that there is initial resistance to commencing a beneficial regimen, as the startup costs loom large.

However, if people are "walked through" this process with adequate financial incentives to try the regimen regularly for a while, perhaps good habits will develop.

Note that the observation that exercising is a habitual behavior suggests that people who are interested in exercising more should try to commit themselves to exercise for a while. By doing so they affect not only their current well-being, but also their future utility, by making future exercise more beneficial.[4] This type of self-enforcing mechanism is a possible explanation of the DellaVigna and Malmendier (2006) study. As a self-control mechanism, people may choose the more expensive plan because it reduces the marginal cost of attending to zero, and people believe that this will encourage them to attend the gym in the future.

Potential applications are numerous, as much of the population seems to be aware of the benefits of some activity, but incapable of reforming without some assistance. For example, in education, Angrist, Lang, and Oreopoulos (2009) offered merit scholarships to undergraduates at a Canadian university; they have some success in improving performance, but mixed results overall.

A recent literature in economics ties habits and self-control. Laibson (1997) and O'Donoghue and Rabin (1999) discussed present-biased (hyperbolic) preferences as an explanation for persistent bad habits and addictions.[5] This relates to our study because students may over-emphasize initial setup costs for going to the gym due to hyperbolic discounting. Loewenstein and O'Donoghue (2005) developed a dual-process model in which a person's behavior is the outcome of an interaction between deliberative goal-based processes, and affective processes that encompass emotions and motivational drives. The incentives to exercise introduced in our study could help in the conflict between the two processes and increase gym visits. Bernheim and Rangel (2004) presented a model in which use among addicts may be a mistake triggered by environmental cues, which addicts may then try to avoid. In a related vein, Benabou and Tirole (2004) developed a theory of internal commitments, wherein one's self-reputation leads to self-regulation and this "willpower" enables one to maintain good behavior.

1. THE FIELD EXPERIMENTS

In our first study, we invited students (from an e-mail list of people interested in participating in experiments) at the University of Chicago to the laboratory. There was no mention of physical fitness or exercise in the recruiting materials. All participants were promised payment if, and only if, they came to the

[4]Becker and Murphy (1988) identified conditions under which past consumption of a good raises the marginal utility of present consumption; Becker (1992) applied this to habit formation. This is discussed in more detail below.

[5]See Frederick, Loewenstein, and O'Donoghue (2002) for a comprehensive review of empirical research on intertemporal choice, as well as an overview of related theoretical models.

laboratory once on a given date and again a week later. The 120 students we signed up were assigned randomly to one of the three treatments described below. The assignment to treatment was based on the arrival time to the meeting. All students at the university received a membership in the campus athletic facility as part of their fees. Each person was asked to sign a consent form allowing us to get the computerized report (based on the magnetic swipe card used to enter the gym) of his or her visits to the gym during the academic year, so we were able to obtain records concerning past attendance at this facility for all of our participants. Everyone was given a handout about the benefits of exercise; this is shown in the Supplemental Materials on the *Econometrica* website (Charness and Gneezy (2009)). Forty of these people participated in a different experiment, which was completely unrelated to exercising; this was the control group.[6]

The other 80 participants (in different sessions) were told that they would receive \$25 to visit the gym at least once during the following week and then to return to the lab to answer questions. They were told that we would be checking their computerized records. Upon returning to the lab in the following week, participants were randomly assigned to one of two treatments. For half of them this was the end of the experiment; the other half was promised an additional \$100 for attending the gym at least eight times during the next four weeks. All participants in the latter group returned after the month was over.

Our second study was conducted at University of California, San Diego, where all registered students receive a membership in the campus athletic facilities as part of their tuition. All participants were asked to sign a consent form granting us access to data on their past and future gym visits.

In all, 168 first- and second-year undergraduate students were recruited from the general campus population using e-mail lists.[7,8] All participants were paid \$175 (in installments of \$75 for the first visit and \$50 for each of the other two visits, to motivate people to show up each time) to go to a meeting room at the Rady School of Management three times (once in January, once after about one month, and once after about five months) for biometric tests. They were also asked to keep an exercise log for five weeks and to complete a questionnaire. There were no further requirements for the people in the control group; the people in the second group were additionally required to go to the gym at least once in the next month, while the people in the third group were additionally required to go to the gym at least eight times in the next month. By paying the same amount of money to all participants in this study, we control for the

[6]This was a marketing experiment studying the effect of coupons on product choice.

[7]This differs slightly from the recruiting done in Study 1, where all undergraduates were recruited. This could conceivably have affected pre-intervention attendance rates (as well as perhaps inducing the positive trend in attendance rates in time for each treatment group), as first-year students might have still been settling in.

[8]Twelve participants (see below) did not show up to all meetings and were excluded from the data.

possibility that it was the monetary payment, rather than a habit acquired by our requiring multiple gym visits, that caused the effects we observed in our first study; since the control group was paid $175 independently of gym attendance, the additional attendance for the eight-times group cannot be the result of their additional income (and possibly a corresponding additional amount of free time).

Participants who replied to the e-mail were invited for an individual meeting and were randomly assigned a treatment group according to arrival time, and given the exercise handout and a questionnaire. We then measured the individual's height, weight, body fat percentage, waist, pulse, and blood pressure.[9] We collected the exercise logs, which showed the number of days of exercise and a brief description of the kinds of exercise in both the gym and otherwise, at the second measurement appointment. The appointments for the second and third meetings were arranged by e-mail.

Hypotheses

The standard null hypothesis is that our financial intervention will not affect behavior after the end of the intervention. We formalize this as follows:

HYPOTHESIS 0: *Participants will visit the gym with the same frequency after the incentives are removed as before the incentives were introduced*.

We also test two competing hypotheses regarding the effect of this incentive. The first hypothesis is the *crowding-out* effect. Studies indicate that, in some situations, providing rewards may be counterproductive, as providing an extrinsic motivation for a task or activity may crowd out existing intrinsic motivation.[10] The formal statement of the hypothesis is next:

HYPOTHESIS 1: *Participants will visit the gym* less *frequently after the incentives are removed as compared to before the incentives were introduced*.

[9]To measure the waist circumference, the research assistant placed a tape measure around the abdomen just above the hip bone. The tape was snug and was kept parallel to the floor. Body fat percentage was measured with a conventional scale that uses the bioelectrical impedance method. A low-level electrical current is passed through the body and the impedance is measured. The result is used in conjunction with weight to determine body fat percentage. Unfortunately, the body's impedance level can be altered by many factors besides body fat, such as the amount of water in the body, skin temperature, and recent physical activity. Hence, this is a noisy measure of actual body fat. Pulse and blood pressure were measured using an automatic monitor.

[10]For early demonstrations in psychology, see Deci (1971) and Lepper and Greene (1978). See Frey and Jegen (2001) and Gneezy and Rustichini (2000a, 2000a) for demonstrations in economic settings. Benabou and Tirole (2003) presented a formal model of this issue. Fehr and Falk (2002) provided a more general framework of the psychology of incentives.

According to this hypothesis, participants are intrinsically motivated to exercise. Any extrinsic intervention, such as paying them to go to the gym, may be counterproductive in the long run by destroying the intrinsic motivation to exercise. According to this process, before the introduction of the incentives participants exercised either because it was good for them or because they simply enjoyed it. After the incentives are introduced, they may instead feel that they exercise just for the money.

Even if the incentives are large enough to motivate people to go to the gym while in force (see Gneezy and Rustichini (2000a) and Heyman and Ariely (2004) for discussions of the importance of the size of the incentive), the hypothesis is that after the incentives are removed participants will stop attending the gym because intrinsic motivation has been crowded out.

The competing hypothesis is that people who were paid to attend the gym for some period would attend the gym more frequently even after the incentives are removed.

HYPOTHESIS 2: *Participants will visit the gym* more *frequently after the incentives are removed as compared to before the incentives were introduced*.

One motivation for this hypothesis is "habit formation." Becker and Murphy (1988) identified a necessary and sufficient condition for a good to be habitual near a steady state:

$$(\sigma + 2\delta)U_{cs} > -U_{ss},$$

where δ is the depreciation rate on past consumption, σ is the rate of preference for the present, c is a consumption good, S is the stock of consumption capital, $U_{cs} = \partial^2 U/\partial c\,\partial S$ and $U_{ss} = \partial^2 U/\partial S^2$.[11] In words, an increase in one's current consumption of c increases one's future consumption of c if and only if one's behavior displays adjacent complementarity.[12,13]

Habits may be harmful or beneficial to the extent that they decrease or increase future utility. The marginal utility of today's consumption is correlated with historical consumption; changes today may have only a small effect in the short run, but increasingly large effects in the long run. In this view, "experiences influence… desires and choices partly by creating habits, addictions, and traditions" (Becker (1992, p. 335)).

[11]See Becker and Murphy (1988, pp. 679–680) for the derivation. We use the Becker (1992, p. 343) formulation.

[12]This term was first introduced Ryder and Heal (1973). An example on page 5 is "A person with adjacent complementarity [who expects to receive a heavy supper] would tend to eat a light breakfast and a substantial lunch," while this would be reversed with distant complementarity.

[13]In fact, past consumption of the good raises the marginal utility of present consumption whenever $U_{cs} > 0$.

If exercising is a form of habitual behavior, providing incentives to go to the gym for a while may increase future utility from exercising. If the marginal utility of consumption today is positively correlated with historical consumption, then this period in which people were given financial incentives to go to the gym could also induce people to go to the gym more often in the future. Hence, we call this hypothesis *habit formation*.

Note that crowding out and habit formation are not mutually exclusive. Since they work in opposite directions, the outcome could be hard to interpret. There could also be individual differences: some people may react in line with crowding out, while others react in line with habit formation. We will discuss this in light of our results.

2. RESULTS

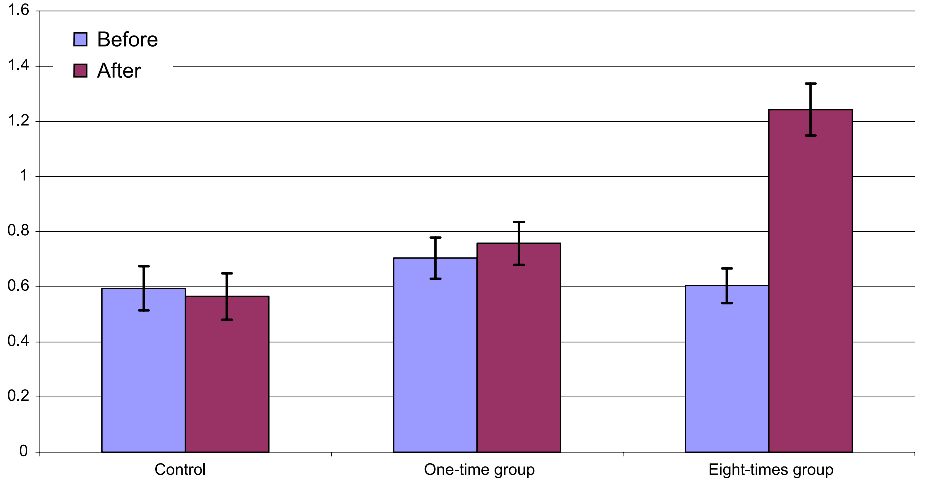
Figure 1(a) and (b) graphically presents the rate of gym attendance before and after the intervention period for Study 1 and Study 2, respectively. "Before" refers to the period before the first lab visit, while "After" refers to the period after any incentives were removed.[14,15]

In Study 1, the average attendance rate for the control group decreased slightly, from 0.59 visits per week in the eight weeks before the intervention period to 0.56 visits per week in the seven weeks after the intervention period. The corresponding change for the group required to attend only one time (henceforth the "one-time group") was from 0.70 visits per week to 0.76 visits per week. In contrast, the corresponding change for the group required to attend the gym eight additional times (henceforth the "eight-times group"[16]) was from 0.60 visits per week to 1.24 visits per week. Thus, we see an average increase of 0.64 visits, or 107 percent of the baseline, for the eight-times group, compared to a small increase for the one-time group, and a slight decline in gym visits for the control group.

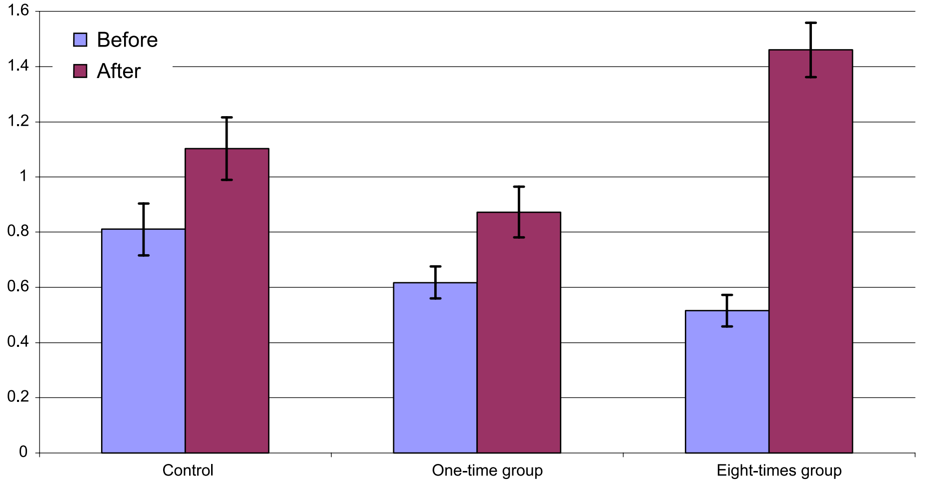
[14]We compare the same weeks in Study 1. However, since the intervention period ended earlier for the group required to attend only once than for the group required to attend eight more times, the actual post-intervention period is slightly different. Nevertheless, robustness checks show no qualitative difference for different specifications.

[15]As we only obtained the data at the end of the academic year for both Study 1 and Study 2, we did not know actual individual attendance until then, and so we paid all students who showed up for payment. It turns out that compliance was imperfect. Although all 40 people in the one-time group in Study 1 complied with the attendance rule, 3 people of 40 in the eight-times group did not fully comply (2 of these people attended six or seven times). In Study 2, 4 of the 56 people in the one-time treatment did not comply, while 5 of the 60 people in the eight-times group did not fully comply (2 of these people attended six or seven times). If we remove the people who did not fully comply with the rules from the analysis, the treatment effects reported above become stronger.

[16]We use "eight-times group" for consistency with Study 2, even though people in this group were actually required to attend nine times overall.



(a) Study 1



(b) Study 2

FIGURE 1.—Average weekly gym visits. Error bars reflect 1 standard error.

In Study 2, we observe a positive trend in attendance for all treatment groups. The average attendance rate for the control group increased from 0.81 visits per week in the 12 weeks before the intervention period to 1.10 visits per week in the 13 weeks after the intervention period; this was a 36 percent increase. The corresponding change for the one-time group was from 0.62 visits per week to 0.87 visits per week; this was a 40 percent increase. The change for the eight-times group was much greater, with an average of 0.52 weekly visits before the intervention period and 1.46 weekly visits after the intervention period; this was a 181 percent increase.[17,18]

We can also examine changes on an individual basis. In Study 1, 13 people of 40 participants (32 percent) in the eight-times group increased their average number of gym visits by more than one per week, while only 2 participants (1 participant) in the one-time group (control group) did so. The test of the equality of proportions (see Glasnapp and Poggio (1985)) finds a very significant difference between the high-incentive and no-incentive treatments, as well as between the two incentive treatments ($Z = 3.53$ and 3.15 for these comparisons, both significant at $p < 0.001$). There is no difference between the one-time group and the control group ($Z = 0.59$).

In Study 2, 40 percent of all participants in the eight-times group (24 people of 60) increased their average number of gym visits by more than one per week, while only 12 percent of the participants (7 people of 57) in the one-time group, and 13 percent of the participants (5 people of 39) in the control group did so. The test of the equality of proportions finds a very significant difference between the eight-times and control groups, as well as between the eight-times and one-time groups ($Z = 3.40$ and 2.90 for the respective comparisons, both significant at $p < 0.002$).[19] There is no difference between the control and one-time groups ($Z = 0.07$).

Regular versus Nonregular Attendees

We can view a cross section of the population by categorizing people before the intervention as regular attendees (at least one visit per week) or nonregular attendees. From the standpoint of public policy, it may well be more useful to target people who rarely (if ever) attend the gym and convert them into regular attendees than to increase the visitation rate for people who already attend the gym regularly. The effect of requiring multiple visits on the people who were not regular attendees is also particularly relevant for testing habit formation.

In Study 1, there were 27 people in the eight-times group who had not been attending the gym regularly (here and elsewhere defined as at least once per

[17]We note that the difference in pre-intervention attendance rates for the control and eight-times groups in Study 2 (0.81 versus 0.52) is not significant ($Z = 1.41$, $p = 0.159$).

[18]The observed patterns are robust to whether we use means or medians. In Study 1, the median weekly attendance for the control group and for the one-time group was 0 before the intervention and 0 after the intervention; the median weekly attendance for the eight-times group was 0 before the intervention, but 1.214 after the intervention. In Study 2, the median weekly attendance for the control group was 0.417 before the intervention and 0.615 after the intervention; the median weekly attendance for the one-time group was 0.167 before the intervention and 0.385 after the intervention; the median weekly attendance for the eight-times group was 0.167 before the intervention, but 1.000 after the intervention.

[19]A chi-squared test using these three categories shows a significant difference between the distributions in the eight-times treatment versus the other two treatments ($\chi^2_2 = 8.49$ and $\chi^2_2 = 11.66$, $p < 0.012$ and $p < 0.001$, respectively).

week); 12 of these people (44 percent) became regular attendees after being paid to go to the gym for a month; these 12 people represent 30 percent of the sample population. The average change for people who had not been regular attendees was 0.98 visits. In contrast, the average change for the 13 people who were already regular attendees was −0.07. Thus, the entire effect of the incentive for the eight-times group comes from those people who had not been regular attendees.

In Study 2, 49 of 60 people in the eight-times group had not been attending the gym regularly; 26 of these people (53 percent) became regular attendees; these 26 people represent 43 percent of the sample population. The average change for people who had not been regular attendees was 1.20 visits. In contrast, the average change for the 11 people who were already regular attendees was −0.20. Thus, as in Study 1, the entire effect in the eight-times treatment comes from those people who had not been regular attendees.

Table I illustrates the gym attendance rates before and after any intervention for the different groups in Study 1 and Study 2, by previously regular or nonregular attendees:[20]

In Study 1, we see that there is no real effect on the attendance rates of those people who were already regular attendees. In fact, there is a slight and insignificant downward drift in all treatments. Similarly, there is a slight and insignificant downward drift for ex ante nonregular attendees in the control

TABLE I

MEAN WEEKLY GYM ATTENDANCE RATES[a]

		Ex ante Regular Attendees			Ex ante Nonregular Attendees		
		Before	After	Change	Before	After	Change
Study 1	Control	1.844	1.774	−0.070	0.058	0.046	−0.012
		(0.296)	(0.376)	(0.206)	(0.036)	(0.023)	(0.020)
	One required visit	1.866	1.827	−0.040	0.077	0.181	0.104
		(0.165)	(0.211)	(0.204)	(0.040)	(0.094)	(0.106)
	Eight required visits	1.644	1.571	−0.073	0.102	1.085	0.983
		(0.127)	(0.304)	(0.264)	(0.044)	(0.234)	(0.231)
Study 2	Control group	2.433	2.677	0.244	0.250	0.560	0.310
		(0.419)	(0.465)	(0.417)	(0.047)	(0.168)	(0.160)
	One required visit	2.051	2.491	0.440	0.193	0.395	0.202
		(0.191)	(0.583)	(0.537)	(0.039)	(0.079)	(0.080)
	Eight required visits	1.901	1.706	−0.195	0.204	1.405	1.201
		(0.402)	(0.786)	(0.411)	(0.038)	(0.170)	(0.171)

[a]Standard errors are in parentheses.

[20]For the purposes of analysis, we exclude the weeks of spring and winter break, as attendance rates were, perforce, extremely low during these weeks and thus not really representative. In any case, our results are qualitatively unchanged when these weeks are included.

group. We do observe a small and insignificant increase in attendance for nonregulars in the one-time group; however, there is a large and highly significant ($t = 4.26$) effect for nonregulars in the eight-times group.

We see that in no treatment of Study 2 is there a significant effect on the attendance rates of those people who were already regular attendees. There is an upward trend for ex ante nonregular attendees in both the control group and the one-time group; this is significant for both the control group and the one-time group ($t = 1.94$ and $t = 2.52$, respectively). However, by far the largest effect (with $t = 7.02$) is observed for nonregulars in the eight-times group. This effect is significantly larger than the effect for the control group and the effect for the one-time group ($Z = 2.30$, $p = 0.011$ and $Z = 2.74$, $p = 0.003$, respectively, one-tailed test, Wilcoxon rank sum tests).

We also note that the change for regular attendees in the eight-times group is actually negative, and in contrast to the modest upward trend for regular attendees in the control and one-time groups. While this difference-in-difference is at most only marginally significant (Wilcoxon rank sum test, $Z = 1.34$; $p = 0.090$ on a one-tailed test), it does suggest the possibility that ex ante regular attendees experience some crowding out (Hypothesis 1).

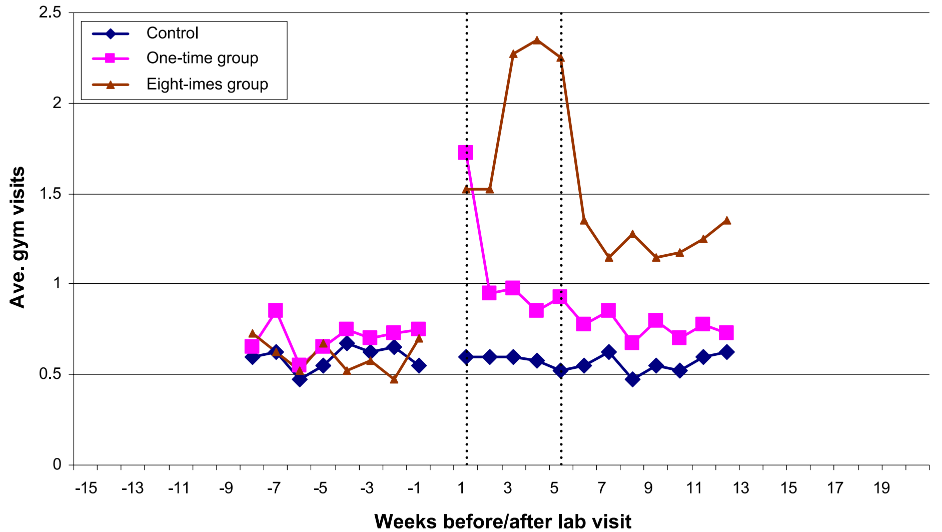
One might expect that simply requiring people to become familiar with the gym by going through the initial setup might lead to benefits. But if this were the full explanation, there should be little difference between any groups who were required to attend at least once, since they all went to the gym and incurred the setup costs. Yet we see that the increase in gym attendance, in both studies, is significantly and substantially larger for nonregular attendees in the eight-times group than in the other groups.

Thus, we see support for Hypothesis 2 over Hypothesis 0 when people who had not regularly attended the gym were required to make multiple visits (obviously we cannot test Hypothesis 2 against Hypothesis 1 for those people who had not attended the gym before the intervention, as their attendance rate cannot decrease). On the other hand, Hypothesis 0 appears to hold for the other treatments. Hypothesis 1 is generally rejected, although perhaps not for the people in Study 2 who were initially regular gym attendees.[21]

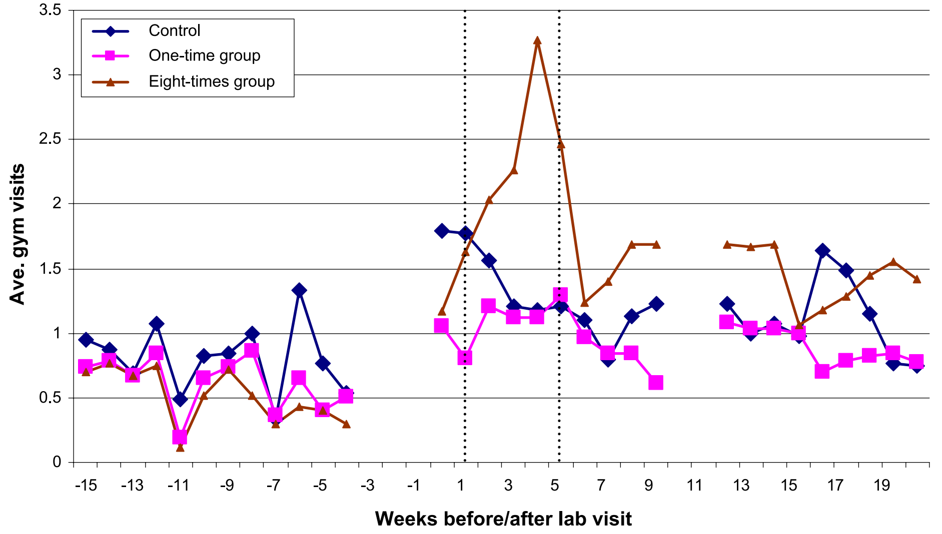
Changes Over Time

It is not surprising that the financial incentives lead to a strong effect during the incentive period. But how persistent are the post-intervention effects? Do they appear to be diminishing over time? Figures 2(a) and (b) shows that there

[21] A final concern is whether people *exceeded* the attendance requirements during the intervention period. In Study 1, 23 of 40 people in the one-time group exceeded the required attendance level during the intervention period, while 20 of 40 people in the eight-times group did so. In Study 2, 38 of 57 people in the one-time group exceeded the required attendance level during the intervention period, while 35 of 60 people in the eight-times group did so.



(a) Study 1



(b) Study 2

FIGURE 2.—Average gym visits.

is very little change in attendance rates once the intervention is over.[22] We see little if any change over the remaining time for any group.[23]

We note a tendency for people in the eight-times group to delay much of their required visits until the latter part of the required time period. In Study 1, there was little gym attendance in the early part of the intervention period. There is a pronounced peak for this group in Study 2, as by far the highest weekly average occurred in the fourth week of the intervention period. This seems consistent with self-control issues, since most standard models with uncertainty would predict some precautionary gym visits early on, to take into account later unforeseen shocks. On the other hand, we observe a sharp dip in attendance for the eight-times group in Study 2 in the week after the intervention period, suggesting some degree of intertemporal substitution.

A natural concern is the decay rate of the post-intervention attendance for the eight-times group. A regression on the gym attendance for this group in Study 1 over the time of the post-intervention period shows that the average gym attendance *increases* by the insignificant rate of 0.004 per period, so we see no signs of decay in gym attendance over time after the intervention. A similar regression for Study 2 shows that the average gym attendance drops by the insignificant rate of 0.010 per period. Given that the increase over the pre-intervention rate for this group was 1.201, the benefits would erode in 120 weeks with a linear decline (and, of course, more slowly with a constant percentage decrease from week to week).

A final issue concerns attrition. Given the payoff structure in Study 1 (people received $25 after one week and the eight-times group received $100 only after the following month), it is not surprising that there was no attrition at all, in terms of people showing up for money. In Study 2, we paid a large portion ($75) to people who came to the first session, so that some people might have felt satisfied with their earnings and dropped out. In fact, 3 people out of 42 (7 percent) dropped out of the control group, 3 people out of 60 (5 percent) dropped out of the one-time group, and 6 people out of 66 (9 percent) dropped out of the eight-times group. All of the dropouts occurred after the first session, but before the second measurement appointment; these people received the $75 that was paid at the first session. It seems natural to have a higher attrition rate in the eight-times group, since the task to earn more money is more difficult. Nevertheless, the test of proportions finds no significant difference across treatments for these attrition rates.[24]

[22]A smoothed version of Figure 2(b) (with weeks grouped together to lower the noise) can be found in the supplemental Materials on the *Econometrica* website.

[23]Both studies necessarily ended at the end of the school year, as our Institutional Review Board (IRB) approval to gather these data did not extend to the next academic year. The gap for week 0 in Figure 2(a) reflects the week of spring break, while the gaps in Figure 2(b) reflect the low-usage periods mentioned in footnote 16.

[24]We have $Z = 0.45$ for the control versus the one-time group, $Z = 0.36$ for the eight-times group versus the control group, and $Z = 0.89$ for the eight-times group versus the one-time group.

Regression Analysis

We supplement our descriptive results and nonparametric statistics with some Tobit regressions that account for the left-censoring problem.[25] These are presented in Table II.

The regressions confirm our earlier discussion. Specification 1 of Study 1 indicates that only when eight visits are required do we observe a significant increase in post-intervention gym attendance. Specification 2 shows that the effect largely vanishes for ex ante regular attendees who were required to visit the gym eight times. Specification 1 of Study 2 also indicates that only when eight visits are required do we observe a significant increase in post-intervention gym attendance. Specification 2 shows that the effect vanishes entirely (going slightly in the other direction) for ex ante regular attendees who were required to attend eight times. Note that we find no significance for gender in any regression.[26]

TABLE II

TOBIT REGRESSIONS FOR GYM ATTENDANCE RATE AFTER INTERVENTION[a]

Independent Variables	Study 1		Study 2	
	Spec. 1	Spec. 2	Spec. 1	Spec. 2
Attendance before	1.262***	1.434***	1.045***	1.195***
	[0.154]	[0.205]	[0.112]	[0.140]
One-time group	0.292	0.184	−0.022	−0.043
	[0.358]	[0.450]	[0.289]	[0.307]
Eight-times group	1.320***	1.874***	0.884***	1.234***
	[0.350]	[0.404]	[0.284]	[0.294]
Male	0.135	0.153	−0.031	−0.114
	[0.280]	[0.268]	[0.249]	[0.240]
One-time∗regular		0.198		0.230
		[0.589]		[0.480]
Eight-times∗regular		−1.527***		−1.664***
		[0.533]		[0.480]
Constant	−1.243**	−1.362***	−0.020	−0.114
	[0.366]	[0.386]	[0.250]	[0.252]
# Observations	120	120	156	156
Pseudo R^2	0.211	0.241	0.140	0.164

[a]The control-group attendee is the omitted variable in these regressions. Standard errors are in parentheses.

**indicates significance at the 5 percent level.

***indicates significance at the 1 percent level, two-tailed test.

[25]Note that these regressions address within-participant change measured on a nearly continuous variable, rather than differences in proportions with arbitrary thresholds or across individual comparisons.

[26]Consistent with the habit-formation story, we find a significant correlation between gym attendance during and after the intervention for the eight-times groups in both Study 1 and Study 2

Biometric and Questionnaire Data

As mentioned earlier, in Study 2 we gathered data on each participant's weight, body fat percentage, waist size, pulse rate, and blood pressure. The full data are presented in Appendix A. Table III summarizes the differences in these measures between the first and last measurements.

We find modest but significant differences across treatments in thechange in levels over time for body fat percentage, weight, waist size, BMI, and pulse rate.[27] Overall, with the exception of the blood-pressure measures, we see that the biometric measures of the eight-times group improved significantly relative to both the control group and (with the further exception of the pulse rate) the one-time group.[28] Thus, it appears that there are real health benefits that

TABLE III

BIOMETRIC DATA AVERAGES AND CHANGES OVER TIME—STUDY 2[a]

	Control (G1)		One-Time (G2)		Eight-Times (G3)		Difference-in-Difference (Wilcoxon–Mann–Whitney Test)		
	First	Δ	First	Δ	First	Δ	G1–G2	G2–G3	G1–G3
Body fat %	25.7	1.41	21.6	0.29	26.9	−0.78	1.12*	−1.07***	−2.19***
	(1.54)	(0.42)	(1.07)	(0.33)	(1.09)	(0.21)	[0.088]	[0.000]	[0.000]
Pulse rate	78.0	3.90	81.8	−1.75	80.2	−1.25	5.65**	−0.50	5.15**
	(1.86)	(2.08)	(1.56)	(1.45)	(1.47)	(1.78)	[0.030]	[0.974]	[0.040]
Weight (kg)	61.8	0.57	59.8	0.59	64.0	−0.34	−0.02	0.93***	0.91***
	(2.03)	(0.55)	(1.60)	(0.21)	(1.54)	(0.25)	[0.560]	[0.005]	[0.006]
BMI	22.7	0.23	21.7	0.22	23.2	−0.12	0.01	0.34***	0.35***
	(0.64)	(0.19)	(0.45)	(0.07)	(0.40)	(0.09)	[0.560]	[0.005]	[0.006]
Waist (in.)	34.3	0.07	33.0	−0.10	35.0	−0.72	0.17*	0.62**	0.79*
	(0.63)	(0.36)	(0.47)	(0.27)	(0.42)	(0.23)	[0.790]	[0.045]	[0.068]
Systolic BP	122	5.23	121	2.32	122	1.78	2.91*	0.54	3.45*
	(1.82)	(1.71)	(2.01)	(1.88)	(1.90)	(1.99)	[0.084]	[0.897]	[0.090]
Diastolic BP	74.0	2.87	75.8	1.07	74.7	2.58	1.80	−1.51	0.29
	(1.22)	(1.22)	(1.44)	(1.23)	(1.07)	(1.33)	[0.160]	[0.654]	[0.535]

[a]Standard errors are in parentheses, Two-tailed p-values are in brackets. Body mass index (BMI) is calculated using the formula: BMI = (weight in kilograms)/(height in meters)^2. "First" refers to the first measurement, which was taken in the initial week, and "Δ" indicates the change from the initial level as determined using the final measurement, taken 20 weeks later.

($r = 0.6035$ and $r = 0.6802$ in the respective studies; both correlation coefficients are significant at $p < 0.0001$).

[27]We include for convenience both weight and BMI, even though the latter is an isomorphic function of weight for a given height.

[28]Some of the significance of the improvement stems from the control group having gotten worse over time in their biometric measures. We suspect that this is not so uncommon, particularly among first-year students. There is an expression "the freshman 15," which refers to students' weight gain in the first year living away from home.

accrue from paying people to go to the gym eight times in a month.[29] There are also some differences between the control and one-time group, but these tend to have lower significance levels; perhaps there is a slight benefit from merely requiring one visit.[30]

We asked people to fill out a questionnaire at the first meeting, with questions involving the frequency of exercise, whether people wished to exercise more, whether people thought that being paid money to go to the gym would increase the amount they exercise, and, if so, whether this would have a long-term effect. Other questions pertained to the grade point average (GPA), to happiness with respect to social life and academic performance, and to the extent a change was needed in their lives. As would be expected with randomization, there were no substantial differences in the responses to these questions across treatment groups; the full summary statistics can be found in Appendix B.

On average, students at the first meeting reported exercising 2.2 times per week, while 84 percent wished to exercise more. Eighty-two percent thought that paying money to exercise would help; of these, 79 percent thought that there would be a long-term effect. On a 1–7 scale, respondents reported a mean happiness level of 5.97 with their social life and 5.40 with their academic life; an average score of 4.59 was observed for people feeling that a change was needed in their life. Overall, we found that the responses had little predictive power in terms of the change in gym attendance over time; we present regressions to this effect in Appendix B. Only the question about needing a change in one's life showed any significance, and this was not robust to different specifications. Interaction effects were not significant.[31] Thus, our results seem to be largely

[29]A reviewer suggests the following rough calibration of our health results. Suppose we assume that the only effect of the eight-times treatment was to increase the amount of exercise, and further assume that this amounted to 400 calories expended during each exercise session. The eight-times group has about 15 more exercise sessions than the control group (combining the extra sessions during the intervention period with the effect in the post-intervention periods), or spends roughly 6000 more calories, which is a little less than 1 kilogram. We find a difference-in-difference for weight of 0.91 kilograms. Again, this is only a rough calibration.

[30]We do measure the amount of gym and nongym exercise in Study 2 by asking participants to keep exercise journals during the intervention period. The people in the control group reported exercising on an average of 7.87 days, with 79 percent reporting going to the gym during this time and 26 percent reporting nongym exercise. Participants in the one-time group reported exercising on an average of 4.91 days, with 100 percent reporting going to the gym during this time and 33 percent reporting nongym exercise. People in the eight-times group reported exercising on an average of 8.58 days, with 100 percent reporting going to the gym during this time and 22 percent reporting nongym exercise. Thus, while the one-time group reports less frequent exercise than the eight-times group, there is only a small difference in frequency between the control group and the eight-times group. This is puzzling, and we suspect that self-reported data may not be the most reliable in this case.

[31]We also performed a regression (not shown) with two interaction terms for each independent variable. Only one of the coefficients for the 16 interaction terms (that for the interaction between the one-time group and one's happiness with one's social life) was significant (it was negative).

robust to the responses on our questionnaire. It appears that the mechanism by which one's exercise frequency increases is independent of one's attitudes and views.

3. CONCLUSION

Some of us have too many bad habits, such as smoking, and too few good ones, such as exercising. Could incentives be used to "improve" one's habit formation—reducing the bad ones and increasing the good ones?

This is an important public-policy question that comes to mind when discussing, for example, incentives to get an education. A major argument made by opponents of using monetary incentives in education is the risk of crowding out intrinsic incentives. Strong and robust evidence shows that the introduction of extrinsic incentives can alter the meaning of the interaction, and hence be counterproductive. In education, for example, it might result in focusing attention on test scores, instead of a more holistic approach in which test scores are only one component. A particular concern arises regarding the long-term effect of the monetary intervention. Even if incentives are effective while present, after they are removed people may revert to effort levels even lower than the initial ones. Nevertheless, some recent evidence in education (e.g., Angrist and Lavy (2009), Angrist, Lang, and Oreopoulos (2009), and Kremer, Miguel, and Thornton (2009) and the discussion therein) indicates that paying children to overcome initial resistance to engaging in a potentially beneficial activity may be quite successful.

In this paper we chose to test the effect of such extrinsic incentives on behavior that may be easier to evaluate than education, because the goal is perhaps better defined. In each of our two studies we paid one group to go to the gym for several weeks, and we observed the gym attendance for this group and all others after the incentives (if any) were removed. Two competing predictions regarding the long-term effect on exercising can arise from the existing literature. The use of incentives might weaken the intrinsic motivation to engage in exercise, such that when the incentives are removed people would exercise less than before. Alternatively, the period of time during which people were induced to exercise might be sufficient to induce the formation of a habit that will remain even after the removal of the incentives.[32]

We find a positive effect from paying people to go to the gym eight times over a period of one month, as the rate of gym visits after the intervention increased significantly in both studies. Upon closer examination, we have the encouraging result that our incentive scheme was successful in creating this positive habit of exercising more: Participants who did not attend the gym before our study began to do so during our intervention and continued to go after

[32]We reiterate that we cannot readily reject explanations (such as the ones mentioned on p. 912) other than habit formation, although we do feel that the habit-formation story fits best.

it was concluded. This result is robust to a number of factors, including gender, the expressed desire to exercise more, and satisfaction regarding one's social and academic life.

Hence, the main result of this paper is that paying people to go to the gym regularly positively reenforced this behavior. The concerns discussed above regarding a strong decline in exercising after removing the incentives were not completely rejected, as there is some slight evidence (primarily in Study 2) that imposing requirements can actually backfire with respect to people who have already been attending the gym regularly.

Finally, the evidence shows that people derive health benefits from our intervention, as the relative change in several biometric indices is significantly better for the eight-times group than the other groups in Study 2. Given the enormous sums of money spent on health care, even a modest improvement may yield large social benefits. Furthermore, if it is possible to favorably influence the habits of young people, there is at least the possibility that this improvement will last for a long time, providing social benefits for the entire period. Of course, we cannot substantiate such a strong claim; however, we do find that the gym attendance rate does not decrease significantly during the post-intervention period in either study.

Habits increase the marginal utility of engaging in an activity in the future. People seem to systematically underestimate the impact of their current actions on the utility of future action and to discount the future too much. As a result, people may underinvest in habit-forming activities (either because they fail to realize the link between current and future consumption or because they do not think that they care about the future). The implications of our findings for public policy are straightforward. Incentives to exercise work, but they should be targeted at people who currently do not exercise and must mandate enough practice hours for the habit to develop. We find that merely providing information about the benefits of exercise or even requiring one gym visit does not have much of an effect. Furthermore, paying people who currently exercise is at best a waste of money; at worst, as our findings hint, it can actually weaken post-intervention exercise habits for people who had already been exercising.

This paper clearly does not exhaust the agenda on habit formation and monetary interventions to improve behavior. There are some very important questions left unanswered, such as why do monetary interventions sometimes succeed and sometimes do not? For example, Volpp et al. (2008) found that a lottery-based financial incentive improved medication adherence for patients. However, behavior reverted back to almost exactly what it was initially once the incentives were dropped.

Another open question is the effect of incentives on bad habits. Findings in the literature on bad habits, such as smoking, are not as encouraging as our findings. For example, in cigarette smoking cessations, researchers have used punishment or rewards (Donatelle et al. (2004)), with very little success. The basic finding is that people refrain from smoking when incentives are present, but go back after the incentives are removed.

An interesting question that future research might address is why habits that we are trying to eliminate seem different from habits that we are trying to acquire. This is not a straightforward extension, since the literature on the neurobiology of addiction (see, e.g., the discussion in Bernheim and Rangel (2004)) finds that smoking and other substance addictions are qualitatively different from other "negative habits." On the other hand, a recent study by Giné, Karlan and Zinman (2008) found a positive effect for the use of incentives on smoking cessation. It is clear that more research is needed in the area of incentives and habit formation/cessation.

APPENDIX A: BIOMETRICS

TABLE A.I

BIOMETRIC DATA AVERAGES—STUDY 2[a]

	Control				One-Time				Eight-Times			
	1st	2nd	3rd	Diff	1st	2nd	3rd	Diff	1st	2nd	3rd	Diff
Body fat %	25.7	25.6	27.1	1.41	21.6	22.0	21.8	0.29	26.9	26.7	26.1	−0.78
	(1.54)	(1.55)	(1.62)	(0.42)	(1.07)	(1.11)	(1.14)	(0.33)	(1.09)	(1.09)	(1.07)	(0.21)
Pulse rate	78.0	80.3	81.9	3.90	81.8	79.9	80.1	−1.75	80.2	81.7	79.0	−1.25
	(1.86)	(1.80)	(2.12)	(2.08)	(1.69)	(1.52)	(1.59)	(1.73)	(1.47)	(1.48)	(1.81)	(1.78)
Weight (kg)	61.8	61.5	62.4	0.57	59.8	60.0	60.4	0.59	64.0	64.0	63.7	−0.34
	(2.03)	(1.95)	(1.91)	(0.55)	(1.60)	(1.56)	(1.60)	(0.21)	(1.54)	(1.51)	(1.52)	(0.25)
BMI	22.7	22.6	23.0	0.23	21.7	21.8	21.9	0.22	23.2	23.2	23.1	−0.12
	(0.64)	(0.61)	(0.61)	(0.19)	(0.45)	(0.44)	(0.45)	(0.07)	(0.40)	(0.39)	(0.39)	(0.09)
Waist (in.)	34.3	34.1	34.3	0.07	33.0	32.9	32.8	−0.10	35.0	34.7	34.3	−0.72
	(0.63)	(0.58)	(0.62)	(0.36)	(0.47)	(0.46)	(0.48)	(0.27)	(0.42)	(0.41)	(0.44)	(0.23)
Systolic BP	122	123	127	5.23	121	125	123	2.32	122	125	125	1.78
	(1.82)	(2.37)	(2.14)	(1.71)	(2.01)	(2.04)	(1.83)	(1.88)	(1.90)	(1.85)	(1.76)	(1.99)
Diastolic BP	74.0	75.9	76.8	2.87	75.8	75.6	76.8	1.07	74.7	76.3	77.3	2.58
	(1.22)	(1.14)	(1.18)	(1.22)	(1.44)	(1.11)	(1.20)	(1.23)	(1.07)	(1.08)	(1.25)	(1.33)

[a]Standard errors are in parentheses. Body mass index (BMI) is calculated using the formula BMI = (weight in kilograms)/(height in meters)^2. "1st," "2nd," and "3rd" refer to the first, second, and third times that the biometric data were taken.

APPENDIX B: QUESTIONNAIRE SUMMARY STATISTICS AND REGRESSIONS, STUDY 2

TABLE B.I

QUESTIONNAIRE SUMMARY STATISTICS[a]

	Control	One-Time	Eight-Times	Aggregate
Exercise frequency (per week)	2.391	2.202	2.058	2.194
	(0.257)	(0.264)	(0.187)	(0.136)
Wish to exercise more (yes = 1)	0.821	0.807	0.883	0.840
	(0.062)	(0.053)	(0.042)	(0.029)

(*Continues*)

TABLE B.I—*Continued*

	Control	One-Time	Eight-Times	Aggregate
Money helps (yes = 1)	0.872	0.789	0.825	0.824
	(0.054)	(0.040)	(0.043)	(0.028)
If so, long-term effect (yes = 1)	0.875	0.744	0.767	0.790
	(0.052)	(0.063)	(0.058)	(0.034)
GPA	3.324	3.077	3.267	3.211
	(0.064)	(0.065)	(0.067)	(0.039)
Happiness with social life (1–7)	5.949	5.948	6.017	5.974
	(0.107)	(0.107)	(0.087)	(0.065)
Happiness with academic affairs (1–7)	5.590	5.246	5.433	5.404
	(0.171)	(0.153)	(0.133)	(0.087)
Change needed in life (1–7)	4.500	4.456	4.783	4.594
	(0.247)	(0.194)	(0.178)	(0.116)

[a] Standard errors are in parentheses.

TABLE B.II

OLS REGRESSIONS FOR CHANGE IN ATTENDANCE RATE AND QUESTIONNAIRE RESPONSES[a]

Independent Variables	1	2	3	4
Exercise frequency (per week)	−0.056	−0.070	—	—
	[0.070]	[0.071]		
Wish to exercise more (yes = 1)	−0.435	−0.459	—	—
	[0.329]	[0.331]		
Money helps (yes = 1)	−1.034	−1.147	—	—
	[0.765]	[0.777]		
If so, long-term effect (yes = 1)	−0.057	−0.052	—	—
	[0.278]	[0.279]		
GPA	0.054	0.063	—	—
	[0.270]	[0.271]		
Happiness with social life (1–7)	0.117	0.137	—	—
	[0.138]	[0.140]		
Happiness with academic affairs (1–7)	0.025	0.020	—	—
	[0.134]	[0.136]		
Change needed in life (1–7)	0.162	0.257	0.079	0.144
	[0.080]**	[0.143]*	[0.066]	[0.123]
Change needed,* one-time		−0.196		−0.190
		[0.180]		[0.159]
Change needed,* eight-times		−0.068		0.012
		[0.176]		[0.155]
Constant	1.488	1.580	0.481***	0.471***
	[1.100]	[1.108]	[0.103]	[0.103]
No. observations	132	132	155	155
R^2	0.055	0.065	0.009	0.025

[a] The control-group attendee is the omitted variable in these regressions. *, **, *** indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively, for a two-tailed test.

REFERENCES

ANDERSEN, R. E. (1999): "Exercise, an Active Lifestyle and Obesity," *The Physician and Sport Medicine*, available at http://www.physsportsmed.com/index.php?art=psm_10.1_1999?article=1021. [910]

ANGRIST, J., D. LANG, AND P. OREOPOULOS (2009): "Incentives and Services for College Achievement: Evidence From a Randomized Trial," *American Economic Journal: Applied Economics*, 1, 136–163. [912,926]

ANGRIST, J., AND V. LAVY (2009): "The Effect of High Stakes High School Achievement Awards: Evidence From a Group Randomized Trial," *American Economic Review* (forthcoming). [926]

BECKER, G. S. (1992): "Habits, Addictions and Traditions," *Kyklos*, 45, 327–345. [912,915]

BECKER, G. S., AND K. M. MURPHY (1988): "A Theory of Rational Addiction," *Journal of Political Economy*, 96, 675–700. [911,912,915]

BENABOU, R., AND J. TIROLE (2003): "Intrinsic and Extrinsic Motivation," *Review of Economic Studies*, 70, 489–520. [914]

——— (2004): "Willpower and Personal Rules," *Journal of Political Economy*, 112, 848–886. [912]

BERNHEIM, B. D., AND A. RANGEL (2004): "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, 94, 1558–1590. [912,928]

BROWNELL, K. D. (1995): "Exercise in the treatment of obesity," in *Eating Disorders and Obesity: A Comprehensive Handbook*, ed. by K. D. Brownell and C. G. Fairburn. New York: Guilford Press, 473–478. [910]

CHARNESS, G., AND U. GNEEZY (2009): "Supplement to 'Incentives to Exercise'," *Econometrica Supplemental Material*, 77, http://www.econometricsociety.org/ecta/Supmat/7416_instructions to experimental subjects.zip. [913]

DECI, E. (1971): "Effects of Externally Mediated Rewards on Intrinsic Motivation," *Journal of Personality and Social Psychology*, 18, 105–115. [911,914]

DELLAVIGNA, S. (2009): "Psychology and Economics: Evidence From the Field," *Journal of Economic Literature* (forthcoming). [910]

DELLAVIGNA, S., AND U. MALMENDIER (2006): "Paying Not to Go to the Gym," *American Economic Review*, 96, 694–719. [909,912]

DONATELLE, R. J., D. HUDSON, S. DOBIE, A. GOODALL, M. HUNSBERGER, AND K. OSWALD (2004): "Incentives in Smoking Cessation: Status of the Field and Implications for Research and Practice With Pregnant Smokers," *Nicotine & Tobacco Research*, 6, S163–S179. [927]

ELLISON, G. (2006): "Bounded Rationality in Industrial Organization," in *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, ed. by R. Blundell, W. K. Newey, and T. Persson. Cambridge, U.K.: Cambridge University Press. [910]

FEHR, E., AND A. FALK (2002): "Psychological Foundations of Incentives," *European Economic Review*, 46, 687–724. [914]

FREDERICK, S., G. LOEWENSTEIN, AND T. O'DONOGHUE (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40, 351–401. [912]

FREY, B. S., AND R. JEGEN (2001): "Motivation Crowding Theory," *Journal of Economic Surveys*, 15, 589–611. [911,914]

GINÉ, X., D. KARLAN, AND J. ZINMAN (2008): "Put Your Money Where Your Butt Is: A Commitment Savings Account for Smoking Cessation," available at http://karlan.yale.edu/p/CARES_dec08.pdf. [928]

GLASNAPP, D., AND J. POGGIO (1985): *Essentials of Statistical Analysis for the Behavioral Sciences*. Columbus, OH: Merrill. [918]

GNEEZY, U., AND A. RUSTICHINI (2000a): "Pay Enough, or Don't Pay at All," *Quarterly Journal of Economics*, 115, 791–810. [911,914,915]

——— (2000b): "A Fine Is a Price," *Journal of Legal Studies*, 29, 1–17. [911]

HEIDHUES, P., AND B. KOSZEGI (2008): "Competition and Price Variation When Consumers Are Loss Averse," *American Economic Review*, 98, 1245–1248. [910]

HEYMAN, J., AND D. ARIELY (2004): "Effort for Payment: A Tale of Two Markets," *Psychological Science*, 15, 787–793. [915]
KAYMAN, S., W. BRUVOLD, AND J. STERN (1990): "Maintenance and Relapse After Weight Loss in Women: Behavioral Aspects," *American Journal of Clinical Nutrition*, 52, 800–807. [910]
KREMER, M., E. MIGUEL, AND R. THORNTON (2009): "Incentives to Learn," *The Review of Economics and Statistics* (forthcoming). [926]
LAIBSON, D. I. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112, 443–477. [912]
LAUGA, D. (2008): "Persuasive Advertising With Sophisticated but Impressionable Consumers," Mimeo, UCSD, available at http://management.ucsd.edu/faculty/directory/lauga/docs/persuasive_advertising.pdf. [910]
LEE, C., S. BLAIR, AND A. JACKSON (1999): "Cardiorespiratory Fitness, Body Composition, and All-Cause and Cardiovascular Disease Mortality in Men," *American Journal of Clinical Nutrition*, 69, 373–380. [910]
LEPPER, M., AND D. GREENE (1978): *The Hidden Costs of Reward: New Perspectives in the Psychology of Human Motivation*. New York: Elbaum/Wiley. [914]
LOEWENSTEIN, G., AND T. O'DONOGHUE (2005): "Animal Spirits: Affective and Deliberative Processes in Economic Behavior," Mimeo, available at http://people.cornell.edu/pages/edo1/will.pdf. [912]
NATIONAL CENTER FOR HEATH STATISTICS, http://www.cdc.gov/nchs/products/pubs/pubd/hestats/overweight/overwght_adult_03.htm.
O'DONOGHUE, T., AND M. RABIN (1999): "Doing It Now or Later," *American Economic Review*, 89, 103–124. [912]
RYDER, H., AND G. HEAL (1973): "Optimal Growth With Intertemporally Dependent Preferences," *Review of Economic Studies*, 40, 1–31. [915]
THALER, R., AND S. BENARTZI (2004): "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, 112, S164–S187. [910]
TRUST FOR AMERICA'S HEALTH (2007): "F as in Fat: How Obesity Policies Are Failing in America," available at http://healthyamericans.org/reports/obesity2007/. [910]
VOLPP, K., G. LOEWENSTEIN, A. TROXEL, J. DOSHI, M. PRICE, M. LASKIN, AND S. KIMMEL (2008): "Financial Incentive-Based Approaches for Weight Loss: A Randomized Trial," *Journal of the American Medical Association*, 300, 2631–2637. [927]

Dept. of Economics, University of California at Santa Barbara, 2127 North Hall, Santa Barbara, CA 93106-9210, U.S.A.; charness@econ.ucsb.edu

and

Rady School of Management, University of California at San Diego, Otterson Hall, 9500 Gilman Dr., #0553, La Jolla, CA 92093-0553, U.S.A.; ugneezy@ucsd.edu.

Manuscript received September, 2007; final revision received December, 2008.