Financial Incentives for Weight Loss:
Results From a Workplace Wellness Program

John Cawley
Department of Policy Analysis and Management and
Department of Economics, Cornell University
National Bureau of Economic Research

Joshua A. Price
Department of Economics, University of Texas at Arlington

March 2, 2012

Abstract

Employers are increasingly adopting workplace wellness programs designed to improve employee health and decrease employer costs associated with health insurance and job absenteeism. This paper examines one such program that offers financial incentives for employee weight loss. We examine attrition and weight loss in response to incentives that include fixed payments and forfeitable bonds. Strengths of the study include that it is a real-world intervention implemented nationwide, a large sample (2,635 workers across 24 worksites), and a long treatment period (one year). Limitations of the study include a lack of randomization, with selection bias the likely consequence.

We document extremely high attrition and modest weight loss associated with the financial incentives in this program, which contrasts with the better outcomes associated with pilot programs in university settings. We conclude by offering suggestions, motivated by the behavioral economics literature, for increasing the effectiveness of financial incentives for weight loss.

1

Introduction

In 2009-2010, nearly three-quarters of adult men (74.1%) and two-thirds of adult women (64.5%) in the United States were either overweight or obese (Flegal et al., 2012).[1] As a result, 36% of men and 60% of women are trying to lose weight (Baradel et al., 2009). Most individuals fail in their weight loss attempts, and the majority of those who are successful in losing weight return to their baseline weight within three to five years (Crawford et al., 2000; Jones-Corneille et al., 2011).

Psychology and behavioral economics provide several explanations for why so many weight loss attempts fail.[2] First, the benefits of weight loss are not salient. Foregone health and quality of life are not visible and therefore these opportunity costs may be underestimated. Second, while the discomfort of dieting and exercise are immediate, their benefits take longer to be realized, and the effectiveness of rewards declines as they are delayed from the time of choice Ainslie (1975). Third, people may be hyperbolic discounters and thus have time inconsistent preferences (Downs and Loewenstein, 2011). In this context, people may want to do what is in their long-run interest (lose weight), but consistently succumb to the temptation to eat and be sedentary.

Financial incentives for weight loss may offer a solution to the problems of salience, immediacy, and time-inconsistency. Financial rewards have the benefit of being salient, with

---

[1] Overweight is defined as a body mass index (BMI) of greater than or equal to 25, obesity is defined as a BMI of greater than or equal to 30, and morbid obesity is defined as a BMI of greater than or equal to 40. BMI is equal to weight in kilograms divided by height in meters squared.

[2] A biological explanation for weight regain is that the homeostatic "steady state" drifts up with weight gain but does not drift down with weight loss. As a result, a dieter's metabolism slows in response to the restricted calorie intake even while they are still obese; see Catenacci et al. (2011) and Leibel et al. (1995). Behavioral economics theories can explain why people do not undertake the lifestyle changes necessary to avoid weight regain after metabolism slows.

their amount known with certainty in exchange for clearly defined objectives. They may also be paid immediately. Even small financial incentives can be effective because people tend not to compare payoffs to their income or wealth but instead "bracket" them - consider them in isolation (Read et al. 1999; Kahneman and Tversky 1979).

Financial incentives can also be structured to help people with time-inconsistent preferences stay committed to their diet and exercise regimen; e.g. one could allow people to post a bond that is forfeited if they fail to achieve their weight loss goals. Such a bond allows a person to influence their own future decision making by increasing the punishment for succumbing to short-run temptation. A forfeitable bond may be more effective than a reward of the same size because people tend to exhibit loss aversion – they dislike losing their own money more than they like winning an equal amount of someone else's money (Tversky and Kahneman 1991; Camerer 2005).

Businesses may bear some of the costs of employee obesity. In the United States, obesity raises medical care costs by $190.2 billion annually (Cawley and Meyerhoefer, 2012), increasing costs in employer-sponsored health insurance. Moreover, obesity is associated with $4.3 billion in job absenteeism costs annually (Cawley et al., 2007). These costs may ultimately be borne by workers in the form of lower wages, but employers are increasingly offering worksite health promotion programs to help employees lose weight, seeing it as a "win-win" in which the employer can save on health insurance and job absenteeism costs, and employees can become healthier and achieve their personal weight loss goals.

The purpose of this paper is to examine outcomes in one such employer-sponsored program that offers financial incentives for weight loss. Specifically, we examine attrition and

weight loss in a program that offered incentives for weight loss that varied by worksite; some employees were offered cash rewards for percentage points of baseline weight lost, some were offered the opportunity to post forfeitable bonds that would be refunded if the employees hit specific weight loss objectives, and some were offered no financial rewards for weight loss (and thus serve as a quasi control group).

This paper builds on a chapter in a conference volume that describes the intervention and reports unconditional attrition and weight loss by quarter (Cawley and Price, 2011). This paper incorporates newly-available data (including for an additional incentive scheme) and reports results from hazard models of attrition and regressions of weight loss.

Previous studies of financial rewards for weight loss were conducted outside the workplace (see the Appendix in Cawley and Price, 2011, for a comprehensive list of the papers on this topic).[3] Volpp et al. (2008) conducted a sixteen-week randomized controlled trial of 57 patients at the Philadelphia Veterans Affairs Medical Center and documented mean weight loss of 13.1 pounds in a group that entered a lottery based on their weight loss, 14.0 pounds among those who participated in a deposit contract, and 3.9 pounds in a control group. Finkelstein et al. (2007) conducted a six-month randomized trial of 207 college employees and found no statistically significant weight loss for those offered payments that were front-loaded (larger at the beginning of the program), back-loaded (larger later in the program), or steady throughout the program. Burger and Lynham (2010) examine 51 bets placed with a U.K. bookmaker regarding the bettor's weight loss, and find that despite payoffs averaging $2,332, approximately 80% of bettors fail to lose the weight and thus lose their bet.

---

[3] There is also a recent literature on financial incentives for gym attendance. Royer (2011) conducted a randomized experiment at a Fortune 500 company in which the treatment group was offered $10 per visit to the company gym, and Charness and Gneezy (2009) conducted field experiments that offered incentives for university students to visit the gym; both studies found that the incentives increased gym attendance.

There are also a substantial number of studies of financial rewards for weight loss from outside of economics, most frequently by psychologists. A recent review and meta-analysis (Paul-Ebhohimhen and Avenell 2008) identified nine published randomized controlled trials (RCTs) that used guaranteed financial incentives (i.e. certain payments, not lotteries) for weight loss, with a follow-up of at least one year. The meta-analysis was unable to reject the null hypothesis of no effect of financial incentives on weight loss; it calculated a mean weight loss of 0.4 kg at 12 months, which was not statistically significant.

Relative to previous studies, ours has both strengths and limitations. This study has a relatively large sample size (2,635); to put this in context, the combined sample size of all published RCTs of financial incentives for weight loss is much smaller: 424 as of 2007 (Paul-Ebhohimhen and Avenell, 2008). The intervention studied by this paper is relatively long: one year. Moreover, we examine data from a real-world intervention rather than a pilot program conducted by researchers, which is important because a concern about pilot programs is how well their results generalize to real-world implementation. However, an important limitation of this study is that the data are opportunistic; individuals were not randomly assigned to different incentive schedules for weight loss and thus our results likely suffer selection bias. This limitation cannot be minimized. However, so little is known about effective methods of weight loss that the Institute of Medicine states that "All types of evaluation can make an important contribution to the evidence base upon which to design policies, programs, and interventions." (IOM, 2007, p. 4). The contribution that this paper makes to the evidence base is to document the attrition and weight loss among the participants of this widespread private-sector program, which is described in the next section.

The Intervention

Our data come from a company (that we will refer to as Company X) that helps employers provide financial incentives for their employees to lose weight. After an employer contracts with Company X, Company X holds a kickoff event in the workplace that explains the program to the employees and encourages them to sign up. Participation is optional. Those who enroll select a physical activity regimen at either the foundation (easiest), intermediate, or advanced level. The program consists of several elements: 1) daily email coaching that includes information about healthy and effective methods of weight loss including decreasing calorie intake and increasing physical activity; 2) call center support; 3) weigh-ins at least once a quarter in order to monitor changes in weight; and 4) financial incentives for achieving specific weight loss targets. Only employees who are overweight (BMI of at least 25) are eligible to receive financial rewards, and once an employee is no longer overweight she cannot receive additional financial rewards for losing more weight.

The weigh-ins take place in HIPAA-compliant[4] kiosks that company X installs in the employer's workplace. Employees enter the privacy-protected kiosk and stand on a scale; their weight is recorded and sent over an internet connection to their personal webpage as well as to Company X's database. Participants can weigh themselves as often as they like, and the lowest recorded weight will count as that quarter's weight. Financial incentives are paid quarterly based on percent of baseline weight lost.[5]

---

[4] The Health Insurance Portability and Accountability Act (HIPAA) regulates the disclosure, and protects the confidentiality, of health information.

[5] We asked Company X whether people game the system by trying to weigh more at baseline (from which future weight losses are judged). They said that through the cameras installed in their kiosks they do not see people wearing heavier clothes to the baseline weigh-in than to later weigh-ins; in all cases people seem for vanity reasons to remove shoes and sweaters before weighing in. However, Company X acknowledges that they have no way to know if people (e.g.) hid weights in their pockets or shoes before the baseline weigh-in. If people engage in such deception then we would expect to see significant drops in weight at the first weigh-in after baseline but we do not find this pattern in the data.

Company X has a standard set of incentives that it proposes, but employers can modify it. In our data, there are four incentive schedules, which are listed in Table 1A (which lists the rewards for quarters 1-3) and Table 1B (which lists the rewards for quarter 4). The first incentive schedule (Standard) is Company X's standard set of incentives: the employee participants pay no fee (all costs are paid by the employer), and enrollees receive quarterly payments determined by percent of baseline weight lost to date. In the standard set of incentives, payment thresholds occur at each percentage point of weight loss up to 5% (1, 2, 3, 4, 5), then thresholds occur every 5 percentage points (5, 10, 15, 20, 25, 30) up to 30% of weight loss. The payment associated with these thresholds varies; for the first seven (1, 2, 3, 4, 5, 10, 15) the reward is a dollar per month per percentage point of weight loss (or $3, $6, $9, $12, $15, $30, and $45 per quarter). Then the per-percentage-point rewards increase: $75 for losing 20%, $105 for losing 25%, and $150 for losing 30%. Six employers (with a total of twenty worksites participating) used this standard incentives schedule.

The second incentive schedule (Modified 1), was used by one employer (with two worksites participating). It is a deposit contract in which employees who choose to participate must pay $9.95 per month (except the first month, which is free), all of which (11 * $9.95 or $109.45) is refunded if the respondent loses at least 5% of baseline weight by the end of the year-long program. If the respondent loses 10% or more of their baseline weight by year's end, they receive in addition to their refunded fees ($109.45) a $100 bonus, for a total of $229.40. Neither the refund of monthly fees nor the $100 bonus is available until quarter 4, making for a very back-loaded incentive schedule. In addition, each quarter there is a lottery in which all those who have lost weight since baseline are eligible to win gift certificates, and the "biggest loser" (as a percent of baseline weight) receives a $250 gift certificate at the end of the year.

We refer to the monthly fees as a bond because the participant posts his or her own money, which is returned contingent on achieving certain weight loss goals. However, the bond is paid in monthly installments, which may generate different behavior than if it was paid in full before beginning the program. A participant needs just a single moment of willpower to post an up-front bond, but must exercise willpower eleven times to pay all of the fees in this schedule. Before paying each of those monthly fees, the respondent may consider his likelihood of losing sufficient weight to receive a refund, and thus whether to continue participating. For this reason, attrition may be higher for refundable monthly fees than it would be for a single up-front bond.

The third schedule (Modified 2) was used in year 2 by the same employer with two worksites that used the Modified 1 schedule in year 1. Modified 2 is again a deposit contract in which participating employees pay $9.95 per month, with the first month free. However, instead of employees forfeiting all their monthly fees, or receiving a full refund, based on whether they lose 5% of their baseline weight by quarter 4, in Modified 2 the incentives are more continuous, both temporally (incentives are paid in all four quarters) and quantitatively (incentives begin with 1% of baseline weight lost and increase to 20% of baseline weight lost). For all weight loss between 1% and 20% of baseline weight, the Modified 2 rewards are greater than those in the Standard incentive schedule. At 25% weight loss the Modified 2 rewards are less than the Standard schedule; Modified 2 offers $90 for all weight loss equal to or greater than 20% whereas the Standard schedule offers in each quarter $105 for a 25% weight loss and $150 for a 50% weight loss. Those facing the Modified 2 schedule were also allowed to join a team competition, in which the each member of the winning team that quarter (defined as the team that lost the highest average percentage of weight since baseline) received $50; 86 of the 161 participants in Modified 2 opted to join a team. No other group had a team competition option.

8

The fourth schedule (Control), used by one employer (with a total of two worksites), offered no incentives for weight loss, but did include one modest incentive for participation: employees were promised $20 if they participated for the entire year (i.e. weighed in at least once in each of the four quarters). This group received all of the features of the Company X intervention (daily emails, call center access, weigh-ins at the kiosk) but were offered no incentives for weight loss, making it useful both as a quasi control group for measuring the impact of financial incentives isolated from all the other program elements, and for estimating the impact of the Company X treatment minus the financial incentives.

Hypotheses, Data, and Methods

The NIH Technology Assessment Conference Panel (1993) stated that weight loss programs should be evaluated based on two outcomes: 1) the percentage of all beginning participants who complete the program; and 2) the percentage of those completing the program who achieve various degrees of weight loss, such as 5% and 10% of baseline weight. Accordingly, we examine attrition and weight loss in this program.

We test two straightforward hypotheses: 1) individuals are less likely to quit a weight loss program when they are offered financial incentives for weight loss; and 2) individuals lose more weight when they have financial incentives for weight loss.

We face two challenges in testing these hypotheses, both of which stem from the fact that our data are opportunistic (provided to us by Company X). First, the incentive schedules were chosen by the employers. Second, the participation of employees is voluntary; there is selection by employees. We discuss each of these in turn.

Regarding selection by employers into different incentive schedules, this is a problem if employer preference for incentive structure is correlated with unobserved employee characteristics that affect attrition and weight loss. We investigated this possibility and were told by Company X that the reason one employer requested the Modified 1 and Modified 2 incentive schedules (i.e. the deposit contracts) is because that employer balked at the cost and wanted the employees to share the costs. Likewise, the employer that implemented the Control schedule (no financial rewards) wanted to save money by avoiding paying incentives. Based on this information from Company X, we assume that the selection of employers into different incentive schedules is not correlated with unobserved employee propensity for weight loss and thus does not result in selection bias, although we acknowledge that this is untestable.

A related problem is that unobserved employee characteristics may vary systematically across the four groups (Standard, Modified 1 and 2, and Control). Company X designed this intervention for office employees who spend their days in front of computers. For the most part, enrollees fit this description. The six employers (with a total of 20 worksites) in the standard incentive group include an HMO office, an HMO clinic (in which enrollees are nurses), two bank offices, a chemical company, and an insurance company. The one employer (with a total of two worksites) that instituted the modified incentive schedules is an insurance company, and the one employer (with a total of two worksites) in the control group is the administrative office of a grocery chain.

Turning to the second potential problem, selection by employees into participation, this is a limitation for generalizing results to the entire population but in another sense it is not a problem because any future intervention is likely to also be optional, and thus the outcomes among self-selected participants are relevant. Previous studies of financial incentives for weight

loss (e.g. Volpp et al., 2008; Finkelstein et al., 2007) are also based on samples of volunteers recruited to participate in a weight loss program, and are likewise not a random sample of the general population.

The total number of employees in our dataset is 2,635, with 1,580 facing the Standard incentives, 765 facing the Modified 1 incentives, 161 facing the Modified 2 incentives, and 129 in the Control group with no financial incentives. (Table 2 lists the number of employees, employers and worksites represented in each group.) The data cover 2004-2008 and include information on employer and worksite, incentive scheme offered, age, gender, baseline weight, weight at each weigh-in, the exercise regimen chosen by the participant at baseline (beginner, intermediate, or advanced), the percentage of program emails that were opened. We drop from the sample participants with baseline BMI below 25 because they were not eligible for financial incentives. Thirteen participants in the control group were dropped because they were simultaneously participating in another workplace weight-loss intervention.

For a few employers we have data on multiple years of the intervention. For these employers, we determine which employees participate in multiple years, and we keep in the sample only the initial year of participation for each individual.

Weight-loss interventions in general tend to have substantial attrition (Ware 2003; Gadbury et al. 2003). In order to examine correlates of dropping out of the program, we estimate two hazard models of attrition by quarter. The first is a basic model, with controls only for age group (40-49, 50+, with under 40 the excluded reference category), gender, indicator variables for year, baseline weight classification (obese or morbidly obese, with overweight as the excluded reference category), and the incentive scheme faced (the Control group is the excluded

reference category). The expanded model controls for all of those variables plus adds the following endogenous regressors: indicator variables for the exercise regimen chosen at baseline (foundation, intermediate or advanced), the percentage of program emails that the participant opened, an indicator for participation in the team challenge (an option in the Modified 2 group only) and percentage of baseline weight lost as of the end of the previous quarter. Based on our hypothesis that financial incentives decrease attrition, we predict that the coefficients on the indicator variables for the standard and modified incentives schemes will be negative.

We also estimate models of weight loss by quarter. We examine weight loss a variety of ways: percentage of baseline weight lost (using an OLS model), whether lost 5% of starting weight (using a probit model), whether lost 10% of starting weight (using a probit model), pounds of baseline weight lost (OLS) and units of baseline BMI lost (OLS). The first outcome (percentage of baseline weight lost) is examined because that is the outcome rewarded in every incentive scheme. The second two outcomes (whether lost 5%, and whether lost 10%, of baseline weight) are examined both because the NIH recommends evaluating weight loss programs by these outcomes, and because the Modified 1 group has large payoffs at those thresholds. The fourth and fifth outcomes (pounds and BMI units lost since baseline) are examined for the sake of comparability with previous studies, as well as to clearly express the magnitude of weight loss in the program.

We estimate both base models and expanded models of weight loss. The regressors in the base model include: age, year, gender, baseline weight classification, and indicator variables for the incentive scheme faced. The expanded model controls for those variables and adds these additional, endogenous, regressors: indicator variables for the exercise regimen chosen at

baseline, the percentage of program emails that the participant opened, and an indicator for participation in the team challenge.

There are several strategies for handling the attrition when examining weight loss outcomes. An intent-to-treat analysis includes all patients in their groups, regardless of whether they received the treatment, deviated from the protocol, or withdrew (Ware 2003). However, to implement this one must have follow-up data on the dropouts, which we do not have. Another option is to conduct a "completers" analysis, which examines data only for those who completed the study. This is likely to be biased toward showing an impact of the treatment, as those who persisted are probably those for whom the intervention was most effective (Ware 2003). Another option is baseline-carried-forward, which assumes that after dropping out the subjects return to their baseline weight. This may cause downward bias in the estimate of efficacy, as weight regain may be incomplete or slow. Based on findings reported below from models of attrition that imply that dropouts are not missing at random, we use a baseline-carried-forward approach in estimating our models of weight loss.

In all regressions we cluster standard errors by worksite because they have common food environments (e.g. cafeterias or vending machines) as well as physical activity opportunities (e.g. on-site gym, distance to parking, presence of elevators). Our data do not include any information about these environmental characteristics of worksites that would allow us to control for them directly.

Empirical Results

Summary Statistics

13

Table 3 lists summary statistics for the sample of 2,635. The vast majority of participants (78%) are female, and the average age is 44.7 years. Weight at baseline averages 203 pounds, and at baseline 39.8% of participants are overweight but not obese, 31.0% are obese but not morbidly obese, and 29.2% are morbidly obese.[6] Sixty percent of participants faced the Standard incentives, 29% were offered the Modified 1 incentives, 6% faced the Modified 2 incentives, and 5% were in the Control group that was offered no financial incentives for weight loss.

Appendix Table 1 lists summary statistics by group. In each group, women represent the overwhelming majority, average baseline BMI ranges from 31 to 33, and the average age ranges from 41 to 46. The prevalence of morbid obesity is significantly higher in the Modified 1 and Modified 2 groups than in the Control group. The percentage of program emails that were opened is significantly lower in the Control group.

Attrition

Table 4 lists unconditional cumulative attrition in each group, by quarter. In the sample as a whole, 42.9% dropped out before the end of the first quarter, and 68.0% dropped out by the end of the year. Unconditionally, attrition was significantly higher in the standard incentive group (54.9% dropped out by the end of the first quarter and 75.8% by the end of the year) than in the other three groups (in which roughly 25% dropped out by the end of the first quarter and 48-58% by the end of the year).

---

[6] The latest data for the U.S. population (Flegal et al., 2012) indicates that, among the overweight population (the right comparison group because only people who are at least overweight are eligible to participate in this program), 48.1% are overweight but not obese, 42.8% are obese but not morbidly obese and 9.1% are morbidly obese. Thus, the overweight and morbidly obese are overrepresented in this program, and the obese but not morbidly obese are underrepresented, relative to their numbers in the population of overweight individuals.

Table 5 presents marginal effects from a hazard model of attrition by quarter; column 1 is the base model with controls for gender, age, baseline weight status and incentive schedule, and column 2 adds controls for the exercise regimen chosen at baseline, the percentage of program emails that were opened, whether the individual participated in the team competition (an option only for the Modified 2 group), and the percentage of baseline weight lost as of the end of the previous quarter. We hypothesized that attrition would be lower in groups offered incentives for weight loss, but we find no evidence to support this; in no case is the marginal effect on an indicator for a reward schedule negative and statistically significant. In fact, several are significant with the opposite sign; all else equal, enrollees facing the Standard schedule were 88.1% more likely to quit the program, and those in the Modified 2 group were 54.9% more likely to quit, than those in the control group. An important caveat is that the control group was offered $20 to participate for the year, which may have decreased their attrition relative to the other groups.

One might expect attrition to be significantly lower in the Modified 1 and Modified 2 groups than in the Control group for two reasons: first, enrollees in Modified 1 and Modified 2 may be selected to be more determined or confident about weight loss because they were willing to pay monthly fees to participate; second, the refundable nature of the fees implies that, after joining, loss aversion may lead them to invest more in weight loss. However, attrition in these two groups is never significantly lower than that in the Control group.

The other correlates of attrition are: the obese are 11.3% more likely to drop out, the morbidly obese are 13.2% more likely to drop out, opening 10% more program emails is associated with a 5% lower probability of quitting the program, and those who participate in the

team competition are 46.1% less likely to drop out.[7] Given that these are endogenous variables,

these correlations may reflect selection (the more determined open more emails and sign up for

the team competition) or a causal effect of those activities; our data do not allow us to

disentangle these two explanations.

Attrition is 2.7% less likely for each additional percentage point of baseline weight that

was lost by the end of the previous quarter. This last finding confirms that those who attrite are

not missing at random: those who quit are those for whom the program wasn't working as well.

This supports the use of a baseline-carried-forward approach in the study of weight loss

outcomes.

Weight Loss

Table 6A and 6B present unconditional weight loss in pounds (6A) and percent of

baseline weight lost (6B) by group and quarter. Ignoring attrition, average weight loss among

those who persisted to the end of the year-long intervention was 7 pounds; completers lost on

average 3.3% of their baseline body weight. Average year-end weight loss was highest for those

facing the Modified 2 (11.9 lbs) and Modified 1 (8.4 lbs.) schedules. Average year-end weight

loss is much more modest for those facing the Standard incentives (5.6 lbs.) or in the Control

group (3.2 lbs.)   After quarter 2, there is little change in average weight loss from quarter-to-

quarter within groups, with the exception that in the Modified 1 group, average weight loss

doubled from quarter 3 (4.1 lbs) to quarter 4 (8.4 lbs), which is perhaps not surprising given that

the incentives in Modified 1 were almost entirely paid in quarter 4.

---

[7] We estimated a probit model of participation in the team competition in the Modified 2 group (the only group for which there was a team competition option); the only observed characteristic that significantly predicts participation in the team is that those who were aged 50 and over were 16.3 percentage points less likely to participate than those under age 40. There were no significant differences in participation by gender, baseline weight status, baseline exercise regiment, or percentage of program emails opened.

We estimate regression models to examine weight loss conditional on observed characteristics. A series of tables presents results of models of weight loss in terms of the outcomes that were directly incentivized: percent of baseline weight lost (Table 7), losing 5% of baseline weight (Table 8) and losing 10% of baseline weight (Table 9). Although the incentive schedules rewarded percentage points of baseline weight lost, it is also interesting and informative to see the results in terms of pounds lost (Table 10) and BMI units lost (Table 11). We present estimates of weight loss for each quarter, but focus on the results for quarter 4 – the end of the program. In each case we report only the baseline carried forward results, which assume that all dropouts returned to their baseline weight, based on the earlier findings from the attrition model that dropouts are not missing at random; i.e. success in the program (in the form of weight loss to date) predicts staying in the program.

Our hypothesis that individuals lose more weight when they are offered financial rewards for weight loss implies that the coefficients on the indicator variables for the Standard, Modified 1, or Modified 2 incentive schedules should be positive. However, in no case do people facing the Standard incentive schedule lose more weight than those in the control group by the end of the year-long program. In most cases, the point estimates of the coefficient are negative, indicating that those in the Standard incentive group lost *less* weight by year's end than those in the Control group who were not offered financial incentives for weight loss (however, in no case are those negative coefficients statistically significant).

However, consistent with our hypothesis, we find evidence that those facing the Modified 1 and Modified 2 incentive schedules lose more weight by year's end. Table 7 indicates that by quarter 4, those in the Modified 2 group lost between 0.78 and 1.4 additional percentage points of baseline weight. However, Tables 8 and 9 indicate that neither the Modified 1 nor Modified 2

17

group had a significantly higher probability of losing 5% or 10% of baseline weight by the end of the program. However, two caveats should be mentioned. First, although they are not statistically significant, the point estimates of the coefficients for Modified 1 and Modified 2 are substantial, implying (e.g.) a 6.2 to 8.3 percentage point higher probability of 5% weight loss by year's end for Modified 1, and a 5.5 to 9.4 percentage point higher probability of 5% weight loss by year's end for Modified 2. Second, participants in both of those groups often had a significantly higher probability than the Control group of 5% and 10% weight loss at the end of quarter 1.

We also examine other correlates of weight loss. Neither gender, age or baseline clinical weight classification is consistently associated with weight loss in Tables 7-9. However, those who open more program emails tend to lose more weight; opening 10% more program emails is associated with a one percentage point higher probability of losing 5% of baseline weight by year's end. The other consistent correlate of weight loss is participating in the team challenge (an option open only to those in the Modified 2 group); this is associated with almost an additional one percentage point of baseline weight lost by year's end (Table 7), a 3.4 percentage point higher probability of losing 5% of baseline weight by year's end (Table 8), and a 1.9 percentage point higher probability of losing 10% of baseline weight by year's end (Table 9). Obviously, because participants choose their email open rate and whether to participate in the team competition, these findings may reflect selection (more determined participants open more emails and sign up for the team competition) rather than causal effects.

Although the incentive schemes rewarded percentage points of baseline weight lost, we also estimate models for loss of pounds and BMI units because these are outcomes reported in previous studies (see, e.g. Volpp et al., 2008 and Finkelstein et al., 2007). Table 10 examines the

outcome of pounds of baseline weight lost, and Table 11 examines units of baseline BMI lost. Table 10 indicates that those who faced the Modified 1 or Modified 2 incentives lost roughly 2 more pounds by the end of the year. Results for units of baseline BMI lost (Table 11) follow the same general pattern.

Discussion:

We study outcomes in an innovative workplace wellness program that offers financial incentives for employee weight loss. We examine the two outcomes recommended by the NIH for evaluating weight loss interventions: attrition and weight loss.

More than two-thirds (68%) of all enrollees dropped out before the end of the year-long program; this is higher attrition than in virtually all previous studies of financial incentives for weight loss (see the Appendix in Cawley and Price, 2011, and Paul-Ebhohimhen and Avenell, 2008).

Importantly, attrition is higher among those unsuccessful at weight loss; each additional percentage point of baseline weight lost as of the previous quarter is associated with a 2.7 percentage point lower probability of dropping out. This confirms that those who drop out are not missing at random, and that our models of weight loss should use a baseline carried forward approach to deal with selective attrition.

The models of weight loss indicate that the Standard set of incentives does not result in any greater weight loss than is experienced by the Control group that was provided no incentives for weight loss. However, the deposit contracts used in the Modified 1 and Modified 2 groups are associated with slightly higher weight loss than that of the Control group: roughly an additional percentage point (2 pounds) of baseline weight lost by the end of the year-long

19

program. However, those in the Modified 1 and Modified 2 groups are not significantly more likely to lose 5% or 10% of their baseline weight (the outcome recommended for study by the NIH); the point estimates are substantial but imprecisely estimated.

The weight loss associated with the financial incentives we examine is generally smaller than that documented in the previous literature. For example, Volpp et al. (2008) document mean 16-week weight loss of 13.1 lbs. when rewards take the form of a lottery with a daily expected value of $3, and 14.0 lbs. when the rewards take the form of deposit contracts whose amount is chosen by the enrollee but can vary between $0 and $3 per day and is matched 1:1 if the weight loss goal is achieved. Our findings are closer to those of Finkelstein et al. (2007), which finds no significant weight loss at six months in a program that offered $7 and $14 per percentage point of baseline weight lost. Overall, our findings regarding attrition and weight loss suggest that the experience of pilot programs may be overly optimistic about what can be achieved on a larger scale, although the differences in outcomes may be due to differences in program design, a point which we return to below.

The program studied is of particular interest because it is a real-world intervention, not a pilot program designed and monitored by researchers. As a result, the data are informative about how such interventions work in the real-world. However, because it is a real-world intervention, it suffers the limitations of selection by employers of incentive schedule, and a relatively small control group (129 out of a total sample of 2,635).

A limitation of this study is that it is based on opportunistic data. Individuals were not randomly assigned to different incentive schedules for weight loss, and as a result there is likely

selection bias in our estimates. In other words, even the modest weight loss that we document may be overly optimistic and due to selection as opposed to the treatment itself.

To some extent, the high attrition and modest weight loss of this intervention may be due to a suboptimally designed incentive scheme. Research in behavioral economics offers guidance on how to improve the effectiveness of programs that offer financial incentives for weight loss. The first suggestion is based on the old adage "you get what you pay for." In Company X's program, rewards are paid based on percentage of baseline weight lost. However, what employers and insurance companies really want is for participants to become healthier – lose fat and perhaps add muscle – and thereby lower their health care costs and job absenteeism. However, the metrics of pounds, kilograms, and BMI do not distinguish fat from muscle (Burkhauser and Cawley, 2008). Rewarding loss of pounds, kilos or BMI units from baseline penalizes vigorous exercise that leads to muscle gain. Future interventions should carefully consider what outcome they want to reward; in particular, the goals of the program may be better served by rewarding the loss of fat and gain of muscle, which can be measured using methods such as Dual X-Ray Absorptiometry (DXA); see Burkhauser and Cawley (2009). One could also consider rewarding behavior change rather than weight, but it is generally more difficult to monitor behavior than to monitor weight (Paul-Ebhohimhen and Avenell, 2008). For example, Charness and Gneezy (2009) and Royer et al. (2011) observe gym attendance but not whether or how much the subjects exercised.

Given that the majority of the participants in each treatment group dropped out, a high priority is to decrease attrition. One strategy is to pay incentives more often, to make them more salient and more immediately reinforcing (Ainslie 1975; Coates et al. 1982). A meta-analysis found that vouchers were more effective at increasing drug abstinence when they were paid

immediately upon negative test results (Lussier et al. 2006). A second strategy to reduce attrition

is to pay people to participate. A third strategy is to target those most likely to drop out for extra

encouragement not to quit. This paper finds that enrollees who were obese or morbidly obese (as

opposed to merely overweight) at baseline were more likely to quit. Previous research (Texiera

et al. 2004) has found that dieting history, quality of life, and other factors predict attrition in

weight loss programs. Such information could be collected at baseline and individuals at risk of

attriting could be targeted for special encouragement.

Such programs should consider including a lottery component in their incentive

schedules, because individuals tend to overestimate the probability of unlikely events

(Kahneman and Tversky 1979) and thus a lottery can be cheaper than guaranteed payments of

the same expected value.

Bonds that are forfeited for failure to achieve goals regarding healthy behavior can be

effective (Kane et al. 2004) but take-up is low (Jeffery 1978). One option is for programs to

allow (but not require) such bonds or deposit contracts. To create an extra incentive for

adherence, any forfeited moneys can be sent to an organization that the participant abhors. The

website stickK.com allows participants to select an "anti-charity" that would receive the bond

they forfeit for failing to meet their weight loss goals; examples on their website include: the

George H.W. Bush and William Jefferson Clinton Presidential Libraries.

When facilitating the posting of a bond, it may be problematic to structure it (as

Company X did) as a monthly fee; this creates incentive for people to quit quickly if they feel

they aren't making progress. Moreover, it requires participants to repeatedly exercise willpower

to submit the next payment. Instead, a single large up-front payment should be encouraged, because that requires willpower only at one point in time.

In general, there should be greater consideration of the appropriate magnitude of the financial incentives. A review of the literature on financial incentives for weight loss concluded that "The choice of the amount, frequency or method of administration of the financial incentives was not justified in any study." (Paul-Ebhohimhen and Avenell, 2008, p. 365). Randomized experiments should be conducted to determine the elasticity of weight loss to financial incentives, by type of reward (certain payment, lottery, bond).

Offering larger incentives may increase weight loss and decrease attrition for multiple reasons. Loewenstein and Prelec (1992), summarizing empirical studies of time preference, note that large dollar amounts suffer less proportional discounting than do small dollar amounts. In other words, if the year-end payoff is small, people will succumb to temptation for only a small fraction of the year-end payoff, but if the year-end payoff is large in absolute terms, a higher percentage will be required to make them succumb to temptation today. Thus, not only does a bigger payoff represent a stronger incentive, but it is discounted less, further increasing its effectiveness.

Ainslie (1975) finds evidence that a combination of small early rewards and larger late rewards can be especially effective. This suggests that a combination of token payments at each weigh-in, combined with a year-end bond refund or other type of large bonus may especially increase adherence.

A final suggestion for the design of such programs is to beware of unintended consequences. Company X told us that their client using the Modified 1 incentives schedule

concluded that workers were using unhealthy methods to achieve weight loss just before the end of quarter 4 because the incentives were so large. The client decided to cease using such a back-loaded incentive scheme in order to avoid such unintended consequences in the future. We looked for evidence of such unhealthy behaviors in our data, which include number of weigh-ins per quarter. We hypothesized that workers engaging in purging behaviors would weigh in more often, as they engaged in an extreme behavior and then checked to see if they had crossed the threshold that qualified them for their reward. There is little evidence of such behavior in number of weigh-ins. Table 12 presents the unconditional mean number of weigh-ins by quarter and incentive scheme. Those facing the Modified 1 incentive scheme did not weigh in significantly more often during quarter 4 than they had in earlier quarters when weight loss was not rewarded; nor did they weigh in more often during quarter 4 than participants facing other incentive schemes that were not back-loaded. Table 13 presents results of regression models of number of weigh-ins during quarter 4, estimated separately by incentive scheme. We test whether there were more weigh-ins by Modified 1 participants at the threshold of the big payoffs in quarter 4 (at 5% and 10% of baseline weight lost); there is some evidence of this; those whose weight loss at end of quarter 3 was between 2.5% and 5.0% of baseline weight (i.e., those just below the payoff threshold) and those whose weight loss at end of quarter 3 was between 5.0% and 7.5% of baseline weight (i.e. those just above the payoff threshold) did weigh in significantly more often during quarter 4. However, the magnitude of the point estimates suggests a very small effect on number of weigh-ins: roughly one-third to one-half of an additional weigh-in during the entire fourth quarter.

There is a long history of financial incentives for weight loss unintentionally creating incentives for unhealthy behaviors. In the very first study to use deposit contracts to encourage

weight loss, the researcher discovered that his subjects used unhealthy weight loss strategies to get back their forfeitable bonds: "Unsolicited anecdotal reports from some of the subjects indicated that they had used extreme measures at various times to lose weight rapidly and temporarily in order to avoid aversive consequences. These measures, reportedly, included taking laxatives, diuretics, and doing vigorous exercises just before being weighed. This problem may have occurred because the contract specified that the treatment contingencies be delivered contingent upon specified weight changes rather than the behaviors that can produce these changes" (Mann, 1972, p. 108-109). This is another reminder that "you get what you pay for"; rewarding loss of fat and/or gain of muscle would not incentivize these dysfunctional responses because measured fat and muscle are not altered by diuretics, laxatives, or vomiting.

A disadvantage of pre-determined deadlines for weight loss is that participants know that they have the option to purge just before the scheduled weigh-in, and that knowledge may give them an excuse to eat more and be sedentary early on, and engage in unhealthy weight loss behaviors later. One way to avoid this is to reward participants based on randomly-timed weigh-ins.

We conclude with four directions for future research. The first is to conduct randomized controlled experiments that incorporate the lessons from behavioral economics listed above (e.g. incentivize loss of fat and gain of muscle, test whether flat amounts, lotteries, or deposit contracts are most effective).

A second direction for future research is motivated by the consistent pattern that those who chose to participate in the team competition lost more weight. Because the team competition was optional, it is impossible based on the work presented here to say whether this

25

correlation is the result of selection (more determined participants sign up for the team competition) or a causal effect. Future research could conduct field experiments to determine whether random assignment to a team competition increases weight loss.

Another important area for future research is to determine whether such financial incentives lose their effectiveness over time. Loewenstein et al. (2001) argue that individuals' response to risk is strongly related to newness; people overreact to new risks and tend to underreact to familiar ones. This suggests that an incentive program may be most effective when it is first implemented, but as it becomes routine its motivational power may deteriorate. This is an empirical question that should be a priority for future research.

The final suggestion for future research is to determine whether the introduction of external incentives leads to a depreciation of intrinsic motivation (see, e.g. Gneezy et al., 2011). Intrinsic motivation may be costly to maintain, and individuals may rationally invest less in it when an extrinsic set of incentives is introduced. Decreased intrinsic motivation may offset some or all of the benefit of the extrinsic incentives, and when the extrinsic incentives are removed the participant might engage in even more myopic unhealthy behavior than before the extrinsic incentives were introduced, until intrinsic motivation is restored to its previous level (Fishbach and Trope 2005; Lepper et al. 1973; Deci 1971). In the current context, this raises the question of whether, at the conclusion of the worksite intervention when financial incentives are ended, former participants are less motivated to control their weight than before the intervention.

Our data permit us to explore in a very limited way weight regain after termination of incentives. We have some year 2 data for the employer that implemented the Modified 1 incentive scheme in year 1. 326 employees completed the year 1 program, and 106 of them re-

26

enrolled for year 2, which began after a two-month lag. The average weight gain in this two-month period was 6.2 pounds (with a standard deviation of 5.2 pounds). To put this in perspective, these people had lost an average of 10.4 pounds in the course of year 1, so in two months they regained nearly two thirds of the weight they lost in a one-year intervention. A caveat is that we only observe weight regain among the select group that chose to re-enroll; it may be that those who regained the most after the end of incentives in year 1 were most likely to re-enroll. Moreover, we observe year 2 outcomes only for the worksite that offered the Modified 1 schedule in year 1, and this is the schedule that may have incentivized unhealthy weight loss behaviors that are particularly unlikely to lead to lasting weight gain.

Other information about weight regain after removal of financial incentives comes from a randomized controlled experiment of a 4-month program of financial incentives for weight loss (Volpp et al., 2008). It found substantial weight regain between the end of the incentives and follow-up three months later to the extent that there was no longer a significant difference in weight loss between the treatment and control groups. However, in related studies, Royer et al. (2011) and Charness and Gneezy (2009) find that incentives lead to increases in gym attendance that persist after incentives are terminated. Future studies should collect follow-up data after the removal of incentives to document weight regain, determine its predictors (in particular, whether extrinsic incentives weakens intrinsic motivation) and devise strategies to facilitate maintenance of weight loss.

Works Cited (incomplete)

Baradel, Lesley A., Cathleen Gillespie, Jana R. Kicklighter, Michele M. Doucette, Meera Penumetcha, Heidi M. Blanck. 2009. "Temporal changes in trying to lose weight and recommended weight-loss strategies among overweight and obese Americans, 1996–2003." *Preventive Medicine*, 49: 158-164.

Burger, Nicholas and John Lynham. 2010. "Betting on weight loss … and losing: personal gambles as commitment mechanisms," *Applied Economics Letters*, 17(12(: 1161-1166.

Burkhauser, Richard V., and John Cawley. 2009. "Adding Biomeasures Relating to Fatness and Obesity to the Panel Survey of Income Dynamics." *Biodemography and Social Biology*, 55(2):118–139.

Catenacci, Victoria A., Paul S. Maclean, Lorraine G. Ogden, Sarit Polsky, Holly R. Wyatt and James O. Hill. 2011. Chapter 46 in: John Cawley (editor), *The Oxford Handbook of the Social Science of Obesity*, (Oxford University Press: New York).

Cawley, John and Chad Meyerhoefer. 2012. "The Medical Care Costs of Obesity: An Instrumental Variables Approach." *Journal of Health Economics*, forthcoming.

Cawley, John, and Joshua A. Price. 2011. "Outcomes in a Program that Offers Financial Rewards for Weight Loss." Chapter 4 in: Michael Grossman and Naci Mocan (ed.), *Economic Aspects of Obesity* (Chicago IL: NBER and University of Chicago Press). pp. 91-126.

Cawley, J, Rizzo, J.A. and Haas, K. 2007. "Occupation-Specific Absenteeism Costs Associated with Obesity and Morbid Obesity." *Journal of Occupational and Environmental Medicine*, 49(12): 1317-1324.

Charness, Gary, and Uri Gneezy. 2009. "Incentives to Exercise." *Econometrica*, 77(3): 909–31.

Downs, Julie S. and George Lewenstein. 2011. "Behavioral Economics and Obesity." Chapter 9 in: John Cawley (editor), *The Oxford Handbook of the Social Science of Obesity*, (Oxford University Press: New York).

Finkelstein, E.A., Linnan, L.A., Tate, D.F., Birken, B.E. 2007. "A Pilot Study Testing the Effect of Different Levels of Financial Incentives on Weight Loss Among Overweight Employees." *Journal of Occupational and Environmental Medicine*, 49(9): 981-989.

Flegal, Katherine M., Margaret D. Carroll, Brian K. Kit, and Cynthia L. Ogden. 2012. "Prevalence of Obesity and Trends in the Distribution of Body Mass Index Among US Adults, 1999-2010." *JAMA*, 307(5): E1-E7.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4): 191-210.

Jones-Corneille, LaShanda R., Rebecca M. Stack, and Thomas A. Wadden. 2011. "Behavioral Treatment of Obesity." Chapter 44 in: John Cawley (editor), *The Oxford Handbook of the Social Science of Obesity*, (Oxford University Press: New York).

Leibel, R.L., M. Rosenbaum, and J. Hirsch. 1995. "Changes in Energy Expenditure Resulting from Altered Body Weight." *NEJM*, 332(10): 621-28.

Royer, Heather, Mark Stehr, Justin Sydnor. 2011. "Incentives, Commitments and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company." Working paper.

Volpp, K.G., John, L.K., Troxel, A.B., et al. 2008. "Financial Incentive Based Approaches for Weight Loss: A Randomized Trial." *Journal of the American Medical Association*, 300(22): 2631-2637.

Table 1: Incentive Schedules by Quarter of Intervention

Table 1A: Quarterly Rewards For Quarters 1-3

Percent Points of Baseline Weight	Standard	Modified 1	Modified 2	Control
1	3	0	15	0
2	6	0	21	0
3	9	0	30	0
4	12	0	36	0
5	15	0	45	0
6	15	0	45	0
7	15	0	51	0
10	30	0	60	0
15	45	0	75	0
20	75	0	90	0
25	105	0	90	0
30	150	0	90	0

Table 1B: Quarterly Rewards For Quarter 4

Percent Points of Baseline Weight	Standard	Modified 1	Modified 2	Control
1	3	0	15	0
2	6	0	21	0
3	9	0	30	0
4	12	0	36	0
5	15	109.45	45	0
6	15	109.45	45	0
7	15	109.45	51	0
10	30	209.45	60	0
15	45	209.45	75	0
20	75	209.45	90	0
25	105	209.45	90	0
30	150	209.45	90	0

Notes: Control was paid $20 for completing the first year, but not for weight loss. Individuals in the Modified 1 program who experienced any weight loss were entered into quarterly drawing for a gift certificate: ten $50 gift cards each quarter and ten $50 salon vouchers each quarter. Also in Modified 1, the "Biggest Loser" (defined as highest percentage of baseline weight lost) received a $250 prize. Some participants in Modified 2 elected to also participate in a team competition in which each member of the leading team (defined as highest average of percentage of baseline weight lost) received $50 that quarter. Only participants with BMI over 25 (that is, those who are overweight or obese) are eligible to receive incentives. Moreover, people can only get incentives for weight loss down to a BMI of 25—there is no financial incentive for anyone in the healthy weight (18.5 to 25) or underweight (<18.5) BMI categories to lose weight.

Table 2: Sample Information by Group

Variable	Standard Incentives	Modified 1 Incentives	Modified 2 Incentives	Control
Sample Size	1,580	765	161	129
# Employers	6	1	1	1
# Worksites	20	2	2	2

Notes: Employer and Worksites are the same for Modified 1 and Modified 2 – they were conducted in sequential years. However, all employees are unique individuals (no one in the sample participated in Modified 1 the first year and Modified 2 the second).

Table 3: Summary Statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Male	2635	0.219	0.414	0	1
Age	2635	44.740	10.084	18	77
Age < 40	2635	0.309	0.462	0	1
Age 40 - 49	2635	0.310	0.463	0	1
Age 50+	2635	0.342	0.475	0	1
Initial Weight	2635	203.409	42.670	118.2	440
Initial BMI	2635	32.790	6.104	25	70.46
Overweight	2635	0.398	0.490	0	1
Obese	2635	0.471	0.499	0	1
Morbidly Obese	2635	0.130	0.337	0	1
Foundation	2635	0.573	0.495	0	1
Intermediate	2635	0.358	0.479	0	1
Advanced	2635	0.069	0.253	0	1
Email Open Rate	1841	44.630	35.481	0	100
Standard Group	2635	0.600	0.490	0	1
Modified 1 Group	2635	0.290	0.454	0	1
Modified 2 Group	2635	0.061	0.240	0	1
Control Group	2635	0.049	0.216	0	1

**Table 4: Unconditional Cumulative Attrition
by Quarter and Incentive Schedule**

Quarter	Pooled Sample	Standard	Modified 1	Modified 2	Control
1	42.9	54.9 *	24.8	24.2	25.6
2	52.8	63.4 *	33.5	51.6 *	39.5
3	60.2	72.5 *	39.3	51.6	45.0
4	68.0	75.8 *	57.4 *	58.4	48.1

Notes: * indicates a difference between attrition in that group and attrition in the control group that is statistically significant at the 5% level.

Table 5: Hazard Model of Attrition:
Marginal Effects and [Standard Errors]

	(1)	(2)
Male	-0.013	0.017
	[0.056]	[0.057]
Age 40 - 49	-0.049	-0.079
	[0.057]	[0.058]
Age 50+	0.061	-0.003
	[0.055]	[0.056]
Obese	0.129***	0.120**
	[0.048]	[0.049]
Morbidly Obese	0.142**	0.129*
	[0.070]	[0.072]
Standard	0.747***	0.880***
	[0.120]	[0.152]
Modified 1	-0.055	0.15
	[0.127]	[0.172]
Modified 2	0.22	0.550***
	[0.151]	[0.171]
Intermediate		-0.021
		[0.050]
Advanced		-0.114
		[0.095]
Email Open Rate		-0.005***
		[0.001]
Team Competition		-0.462**
		[0.194]
Lagged Percent Weight Loss		-0.027***
		[0.004]
Observations	6091	6091

Notes: Omitted categories are age under 40, overweight, control group, foundation exercise program. Model (2) includes year fixed effects. Asterisks indicate statistical significance: * significant at 10%, ** significant at 5%, *** significant at 1%.
† indicates statistical difference with the Standard Group at the 5% level
‡ indicates statistical difference with the Standard Group at the 1% level

Table 6A: Unconditional Weight Loss (Lbs.) by Group and Quarter

Quarter	Pooled Sample	Standard	Modified 1	Modified 2	Control
1	4.48	4.26	4.23	7.72***	3.45
	(9.29)	(11.15)	(6.81)	(9.47)	(5.17)
2	5.34	5.17*	4.90	11.74***	3.08
	(10.26)	(10.29)	(9.51)	(14.38)	(7.18)
3	5.02	5.21	4.08	11.39***	3.03
	(18.01)	(24.41)	(10.20)	(16.69)	(10.17)
4	7.00	5.64	8.37***	11.93***	3.22
	(13.85)	(15.57)	(11.37)	(16.45)	(8.89)

Note: only those remaining in the program are included (i.e., attrition is ignored). Standard deviation in parentheses.

Table 6B: Unconditional Weight Loss (% of Baseline Weight) by Group and Quarter

Quarter	Pooled Sample	Standard	Modified 1	Modified 2	Control
1	2.20	2.16	2.06	3.53***	1.73
	(4.35)	(5.27)	(3.21)	(3.98)	(2.64)
2	2.60	2.56	2.38	5.26***	1.62
	(4.80)	(4.93)	(4.50)	(5.79)	(3.79)
3	2.41	2.50	2.00	5.16***	1.49
	(8.14)	(10.94)	(4.93)	(6.76)	(5.39)
4	3.29	2.47	4.15***	5.43***	1.68
	(6.05)	(6.41)	(5.44)	(6.76)	(4.93)

Note: only those remaining in the program are included (i.e., attrition is ignored). Standard deviation in parentheses

Asterisks indicate significant difference with the control group: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 7: OLS Model of Percent of Baseline Weight Lost

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Male	0.172	0.258	-0.193	0.382	0.184	0.28	-0.043	0.362
	[0.456]	[0.443]	[0.659]	[0.359]	[0.456]	[0.481]	[0.577]	[0.385]
Age 40 - 49	0.165	0.199	0.513	-0.017	0.189	0.213	0.481	0.002
	[0.126]	[0.229]	[0.328]	[0.193]	[0.135]	[0.228]	[0.305]	[0.211]
Age 50+	-0.044	0.068	0.315	-0.079	-0.02	0.071	0.278	-0.044
	[0.178]	[0.125]	[0.288]	[0.157]	[0.200]	[0.128]	[0.258]	[0.162]
Obese	-0.201	-0.016	0.013	-0.164	-0.129	0.016	0.007	-0.152
	[0.200]	[0.133]	[0.156]	[0.109]	[0.184]	[0.126]	[0.125]	[0.121]
Morbidly Obese	-0.239	-0.297	-0.183	0.163	-0.104	-0.23	-0.164	0.22
	[0.213]	[0.279]	[0.467]	[0.407]	[0.240]	[0.311]	[0.460]	[0.419]
Standard	-0.251*	0.018	-0.154	-0.182	-0.447*	-0.262	-0.159	-0.189
	[0.137]	[0.161]	[0.308]	[0.490]	[0.219]	[0.315]	[0.250]	[0.487]
Modified 1	0.328***	0.664***	0.406***	0.952*	0.519	0.434	1.341	0.756
	[0.071]	[0.104]	[0.145]	[0.477]	[0.491]	[0.399]	[1.050]	[0.514]
Modified 2	1.443***	1.604***	1.752***	1.392***	0.662***	0.732***	1.152***	0.766*
	[0.256]	[0.170]	[0.076]	[0.472]	[0.075]	[0.156]	[0.286]	[0.422]
Intermediate					0.296*	0.12	0.134	0.15
					[0.154]	[0.114]	[0.170]	[0.192]
Advanced					0.685	0.184	-0.414	-0.034
					[0.441]	[0.433]	[0.969]	[0.230]
Email Open Rate					0.010***	0.013***	0.006	0.012**
					[0.002]	[0.003]	[0.007]	[0.005]
Team Competition					1.299***	1.428***	1.125**	0.965***
					[0.433]	[0.403]	[0.502]	[0.084]
Constant	1.283***	0.828***	0.617**	0.823	1.517**	1.091*	1.633	1.014
	[0.283]	[0.250]	[0.267]	[0.516]	[0.631]	[0.602]	[1.051]	[1.223]
Observations	2635	2635	2635	2635	2635	2635	2635	2635

Notes: Omitted categories are age under 40, overweight, control group, foundation exercise program. Model (2) includes year fixed effects. Asterisks indicate statistical significance: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 8: Probit Model of 5% of Baseline Weight Lost

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Male	0.026	0	0.02	0.025	0.018	-0.003	0.013	0.018
	[0.022]	[0.034]	[0.026]	[0.041]	[0.026]	[0.039]	[0.029]	[0.044]
Age 40 - 49	0.020*	0.005	-0.001	-0.013	0.026**	0.01	0.001	-0.007
	[0.010]	[0.016]	[0.012]	[0.011]	[0.010]	[0.015]	[0.013]	[0.011]
Age 50+	-0.006	-0.008	-0.024*	-0.019	0.004	-0.001	-0.018	-0.007
	[0.009]	[0.015]	[0.015]	[0.013]	[0.011]	[0.016]	[0.014]	[0.011]
Obese	-0.011	-0.012	-0.017*	-0.028***	-0.006	-0.009	-0.017*	-0.027***
	[0.016]	[0.011]	[0.010]	[0.009]	[0.015]	[0.010]	[0.010]	[0.009]
Morbidly Obese	-0.008	-0.019	-0.013	-0.03	0.006	-0.012	-0.009	-0.027
	[0.017]	[0.015]	[0.025]	[0.026]	[0.020]	[0.019]	[0.029]	[0.029]
Standard	-0.01	0.001	-0.058***	-0.052	-0.047	-0.047	-0.100***	-0.038
	[0.010]	[0.022]	[0.020]	[0.068]	[0.029]	[0.067]	[0.027]	[0.070]
Modified 1	0.034***	0.089***	0.009	0.083	-0.006	0.036	-0.051*	0.061
	[0.008]	[0.022]	[0.018]	[0.074]	[0.040]	[0.075]	[0.029]	[0.082]
Modified 2	0.110***	0.156***	0.070***	0.093	0.044***	0.086***	0.03	0.053
	[0.025]	[0.029]	[0.023]	[0.090]	[0.008]	[0.024]	[0.022]	[0.073]
Intermediate					0.036**	0.018	0.003	0.016
					[0.015]	[0.012]	[0.011]	[0.014]
Advanced					0.074	0.044	0.04	0.029
					[0.052]	[0.050]	[0.031]	[0.026]
Email Open Rate					0.001***	0.001***	0.001***	0.001***
					[0.000]	[0.000]	[0.000]	[0.000]
Team Competition					0.072	0.071***	0.047**	0.036***
					[0.049]	[0.013]	[0.020]	[0.011]
Constant	2635	2635	2635	2635	2635	2635	2635	2635

Notes: Omitted catergories are age<40, overweight, control incentive group, and foundation exercise program. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 9: Probit Model of 10% of Baseline Weight Lost

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Male	0.019**	0.019*	0.006	0.022	0.021**	0.019	0.004	0.017
	[0.007]	[0.011]	[0.012]	[0.018]	[0.008]	[0.013]	[0.014]	[0.021]
Age 40 - 49	0.007	-0.004	-0.006	-0.015	0.007	-0.004	-0.005	-0.013
	[0.007]	[0.007]	[0.011]	[0.011]	[0.008]	[0.007]	[0.011]	[0.010]
Age 50+	-0.006	-0.009	-0.007	-0.015*	-0.006	-0.008	-0.005	-0.01
	[0.007]	[0.006]	[0.010]	[0.009]	[0.006]	[0.006]	[0.010]	[0.008]
Obese	0.006	0.006	0.008	-0.005	0.007	0.007	0.009**	-0.003
	[0.006]	[0.006]	[0.005]	[0.006]	[0.005]	[0.005]	[0.004]	[0.007]
Morbidly Obese	-0.002	-0.009	-0.005	-0.013	-0.001	-0.008	-0.003	-0.009
	[0.010]	[0.014]	[0.015]	[0.021]	[0.010]	[0.015]	[0.017]	[0.024]
Standard	0.322***	0.001	0.015	-0.001	0.300***	0.009	0.022	0.009
	[0.009]	[0.008]	[0.010]	[0.026]	[0.017]	[0.013]	[0.014]	[0.025]
Modified 1	0.595***	0.024**	0.023*	0.023	0.585***	0.037	0.027	0.021
	[0.017]	[0.010]	[0.014]	[0.030]	[0.034]	[0.024]	[0.024]	[0.033]
Modified 2	0.854***	0.063***	0.066***	0.052	0.806***	0.021	0.028	0.029
	[0.014]	[0.017]	[0.020]	[0.040]	[0.023]	[0.015]	[0.019]	[0.032]
Intermediate					0.003	0.006	0	0.009
					[0.004]	[0.006]	[0.005]	[0.008]
Advanced					0.005	0.002	0.02	0.034
					[0.009]	[0.018]	[0.022]	[0.032]
Email Open Rate					0.000***	0.000***	0.000*	0.000***
					[0.000]	[0.000]	[0.000]	[0.000]
Team Competition					0.016***	0.038	0.04	0.018
					[0.003]	[0.024]	[0.033]	[0.012]
Constant	2635	2635	2635	2635	2635	2635	2635	2635

Notes: Omitted catergories are age<40, overweight, control incentive group, and foundation exercise program. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: OLS Model of Pounds of Baseline Weight Lost

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Male	0.922	1.137	-0.049	1.363	1.003	1.234	0.306	1.341
	[1.039]	[1.045]	[1.529]	[0.848]	[1.043]	[1.122]	[1.359]	[0.899]
Age 40 - 49	0.386	0.524	1.17	0.076	0.423	0.546	1.091	0.113
	[0.260]	[0.479]	[0.685]	[0.409]	[0.277]	[0.476]	[0.647]	[0.448]
Age 50+	-0.087	0.216	0.756	-0.084	-0.063	0.204	0.671	-0.008
	[0.384]	[0.258]	[0.630]	[0.322]	[0.426]	[0.250]	[0.608]	[0.336]
Obese	0.145	0.531	0.379	0.11	0.271	0.585*	0.350*	0.127
	[0.430]	[0.320]	[0.250]	[0.140]	[0.387]	[0.295]	[0.191]	[0.169]
Morbidly Obese	0.868	0.661	0.673	1.827*	1.094*	0.765	0.685	1.941**
	[0.516]	[0.693]	[1.247]	[0.919]	[0.559]	[0.733]	[1.208]	[0.931]
Standard	-0.530*	0.15	-0.305	-0.184	-0.908**	-0.334	-0.191	0.141
	[0.296]	[0.336]	[0.631]	[0.877]	[0.434]	[0.724]	[0.639]	[0.976]
Modified 1	0.727***	1.504***	0.811***	1.999**	1.305	1.321	3.142	2.082*
	[0.131]	[0.171]	[0.268]	[0.847]	[1.061]	[0.904]	[2.228]	[1.062]
Modified 2	3.285***	3.791***	3.930***	3.208***	1.705***	1.993***	2.604***	1.944**
	[0.323]	[0.152]	[0.152]	[0.864]	[0.257]	[0.427]	[0.635]	[0.797]
Intermediate					0.494	0.172	0.234	0.332
					[0.303]	[0.231]	[0.341]	[0.404]
Advanced					1.082	0.177	-1.065	-0.183
					[0.963]	[0.925]	[2.163]	[0.494]
Email Open Rate					0.020***	0.028***	0.01	0.026**
					[0.005]	[0.006]	[0.016]	[0.010]
Team Competition					2.616***	2.917***	2.600***	1.942***
					[0.702]	[0.571]	[0.692]	[0.148]
Constant	2.009***	0.941	0.831	1.002	2.439*	1.327	2.976	1.279
	[0.639]	[0.604]	[0.640]	[0.945]	[1.299]	[1.354]	[2.396]	[2.613]
Observations	2635	2635	2635	2635	2635	2635	2635	2635

Notes: Omitted categories are age under 40, overweight, control group, foundation exercise program. Model (2) includes year fixed effects.
Asterisks indicate statistical significance: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 11: OLS Model of Baseline BMI Units Lost

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Male	0.170*	0.151	0.142	0.198**	0.156	0.143	0.116	0.178*
	[0.094]	[0.124]	[0.109]	[0.091]	[0.107]	[0.144]	[0.129]	[0.102]
Age 40 - 49	0.106**	0.108	0.072	0.041	0.119***	0.119*	0.084	0.052
	[0.040]	[0.073]	[0.094]	[0.052]	[0.042]	[0.069]	[0.093]	[0.057]
Age 50+	0.078*	0.072	0.067	0.017	0.094**	0.086*	0.087	0.04
	[0.041]	[0.046]	[0.087]	[0.049]	[0.045]	[0.044]	[0.079]	[0.051]
Obese	0.042	0.067	0.035	0.009	0.06	0.079*	0.042	0.012
	[0.051]	[0.049]	[0.031]	[0.022]	[0.049]	[0.045]	[0.028]	[0.025]
Morbidly Obese	0.139*	0.111	0.202	0.301*	0.172*	0.135	0.218	0.322**
	[0.080]	[0.113]	[0.138]	[0.148]	[0.089]	[0.122]	[0.144]	[0.151]
Standard	-0.113***	-0.034	-0.002	-0.077	-0.137**	-0.064	-0.013	0.011
	[0.039]	[0.034]	[0.031]	[0.148]	[0.066]	[0.111]	[0.087]	[0.165]
Modified 1	0.110***	0.204***	0.120***	0.295*	0.047	0.101	-0.019	0.244
	[0.014]	[0.024]	[0.033]	[0.147]	[0.095]	[0.127]	[0.116]	[0.171]
Modified 2	0.570***	0.591***	0.593***	0.492***	0.285***	0.268***	0.320***	0.264*
	[0.068]	[0.048]	[0.036]	[0.146]	[0.031]	[0.055]	[0.078]	[0.129]
Intermediate					0.063	0.042	-0.013	0.06
					[0.042]	[0.040]	[0.040]	[0.059]
Advanced					0.187	0.087	0.162	-0.011
					[0.151]	[0.161]	[0.149]	[0.072]
Email Open Rate					0.003***	0.005***	0.004**	0.004***
					[0.001]	[0.001]	[0.001]	[0.001]
Team Competition					0.484***	0.540***	0.460***	0.363***
					[0.114]	[0.105]	[0.141]	[0.026]
Constant	0.252***	0.159**	0.150*	0.156	0.107	0.052	0.191	0.066
	[0.049]	[0.076]	[0.086]	[0.156]	[0.110]	[0.156]	[0.229]	[0.334]
Observations	2635	2635	2635	2635	2635	2635	2635	2635

Notes: Omitted categories are age under 40, overweight, control group, foundation exercise program. Model (2) includes year fixed effects. Asterisks indicate statistical significance: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 12: Number of Times Weighed-in by Quarter

	Full Sample	**Standard**	**Modified 1**	**Modified 2**	**Control**
Quarter 1	1.41	1.22	1.58	1.80	1.26
	(0.997)	(0.571)	(1.213)	(1.619)	(0.637)
	Max – 13	Max – 6	Max – 12	Max – 13	Max – 5
Quarter 2	1.30	1.23	1.36	1.58	1.15
	(0.789)	(0.624)	(0.875)	(1.099)	(0.839)
	Max – 9	Max – 8	Max – 9	Max – 7	Max – 8
Quarter 3	1.35	1.37	1.34	1.54	1.10
	(0.912)	(0.921)	(0.932)	(1.053)	(0.345)
	Max - 9	Max - 8	Max - 9	Max - 8	Max - 3
Quarter 4	1.28	1.19	1.40	1.36	1.13
	(0.762)	(0.696)	(0.873)	(0.753)	(0.385)
	Max - 10	Max - 10	Max - 8	Max - 4	Max - 3

Notes: Each cell contains: mean, (standard deviation), and maximum number of weigh-ins in that quarter. In each quarter, sample excludes drop-outs, so there are no zeros included in the mean; the minimum number of weigh-ins is 1.

Table 13. OLS Model of Number of Weigh-ins during Quarter 4

	Standard	Modified 1	Modified 2	Control
Male	-0.008	-0.022	-0.028	0.051
	[0.089]	[0.111]	[0.234]	[0.096]
Age 40 - 49	0.138	-0.208*	-0.145	-0.057
	[0.091]	[0.118]	[0.246]	[0.110]
Age 50+	0.075	-0.147	-0.365	-0.223**
	[0.090]	[0.143]	[0.279]	[0.100]
Obese	0.117	0.055	-0.14	0.471***
	[0.077]	[0.109]	[0.230]	[0.107]
Morbidly Obese	0.322***	0.320*	-0.173	-0.042
	[0.116]	[0.180]	[0.320]	[0.155]
3rd Quarter Percent Weight Loss				
2.5≤x<5	0.194*	0.354***	-0.211	0.259*
	[0.106]	[0.126]	[0.313]	[0.129]
5≤x<7.5	-0.022	0.234	0.235	-0.225
	[0.156]	[0.146]	[0.303]	[0.143]
7.5≤x<10	0.21	0.184	0.456	-0.296
	[0.202]	[0.206]	[0.369]	[0.251]
10≤x<12.5	-0.022	0.308	0.087	--
	[0.269]	[0.240]	[0.513]	
≥12.5	0.139*	0.304	-0.136	-0.335**
	[0.083]	[0.213]	[0.296]	[0.127]
Intermediate	0.144*	0.103	-0.09	0.035
	[0.079]	[0.109]	[0.221]	[0.096]
Advanced	0.168	0.173	-0.4	0.031
	[0.158]	[0.186]	[0.344]	[0.166]
Email Open Rate	0.001	0	0.004	0
	[0.002]	[0.001]	[0.003]	[0.001]
Team Competition	n/a	n/a	-0.138	n/a
			[0.232]	
Constant	0.897***	1.235***	1.583***	1.109***
	[0.150]	[0.146]	[0.333]	[0.115]
Observations	383	326	67	67

Notes: Omitted categories are age under 40, overweight, foundation exercise program, and 3rd Quarter Percent Weight Loss less than 2.5%. Model (2) includes year fixed effects. Asterisks indicate statistical significance: * significant at 10%, ** significant at 5%, *** significant at 1%.

Appendix Table 1: Summary Statistics by Group

	Standard		Modified 1		Modified 2		Control	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Male	0.198	0.399	0.212	0.409	0.348	0.478	0.357	0.481
Age	46.008	10.496	42.956	8.822	41.300	9.082	44.070	10.561
Age <40	0.297	0.457	0.316	0.465	0.385	0.488	0.326	0.470
Age 40 - 49	0.296	0.457	0.335	0.472	0.311	0.464	0.341	0.476
Age 50+	0.407	0.491	0.244	0.430	0.180	0.385	0.333	0.473
Initial Weight	202.1	43.0	204.5	41.1	214.9	44.2	198.6	43.4
Initial BMI	32.8	6.2	32.8	6.0	33.4	6.2	31.3	5.7
Overweight	0.401	0.490	0.382	0.486	0.354	0.480	0.519	0.502
Obese	0.468	0.499	0.482	0.500	0.509	0.501	0.403	0.492
Morbidly Obese	0.131	0.338	0.136	0.343	0.137	0.345	0.078	0.268
Foundation	0.601	0.490	0.550	0.498	0.484	0.501	0.488	0.502
Intermediate	0.340	0.474	0.374	0.484	0.404	0.492	0.426	0.496
Advanced	0.059	0.237	0.076	0.265	0.112	0.316	0.085	0.280
Email Open Rate	41.4	35.1	51.0	35.1	42.8	35.5	28.7	32.5
Team Competition	N/A		N/A	0.472	0.501		N/A	
Start Year 2004	0.607	0.489	--		--		--	
Start Year 2005	0.128	0.334	--		--		--	
Start Year 2006	0.189	0.391	1.00		--		--	
Start Year 2007	0.077	0.266	--		1.00		1.00	
Observations	1580		765		161		129	

43