4

Outcomes in a Program that Offers Financial Rewards for Weight Loss

John Cawley and Joshua A. Price

4.1 Introduction

A variety of approaches are being used to treat obesity and encourage weight loss. One promising strategy based on psychology and behavioral economics is to offer financial incentives for weight loss. Obesity is costly to health insurance companies (Finkelstein, Fiebelkorn, and Wang 2003) and employers (Cawley, Rizzo, and Haas 2007), so for either or both of those organizations to offer monetary incentives for enrollees or employees to lose weight could be mutually beneficial.

This chapter studies data from a firm that coordinates a program of financial incentives for weight loss in various work sites in the United States. We study attrition and weight loss in three types of incentive programs: one that offers no financial rewards for weight loss, one that offers quarterly payments that rise in value with the amount of weight loss, and a third that takes deposits (bonds) that are only refunded if the employee achieves a specific weight loss goal, and also includes a quarterly lottery for those who have lost weight. Relative to previous studies of weight loss in response to financial incentives, strengths of this study include a large sample size (2,407) and a long intervention (one year).

A 2007 Institute of Medicine report on obesity prevention set the immediate next step—which it described as an essential priority action for the near future—as "learning what works and what does not work and broadly sharing that information." (Institute of Medicine 2007, 410). It also notes that "All types of evaluation can make an important contribution to the evidence base upon which to design policies, programs, and interventions." (IOM 2007, 4). This chapter makes a contribution to that effort by documenting enrollment, attrition, and weight loss in one interesting and promising intervention. This chapter presents basic patterns in the data; a subsequent chapter will estimate regression models to test specific hypotheses about attrition and weight loss.

4.2   Conceptual Framework and Previous Literature

For obese people, weight loss would likely result in substantial benefits.[1] For example, the health benefits of modest weight loss (defined as 5 to 10 percent of starting weight) include decreased blood pressure and cholesterol, and a 25 percent reduction in mortality risk for type 2 diabetics (Vidal 2002). Weight loss may also improve quality of life (Ford et al. 2001). There may also be financial benefits. Cawley (2004) finds a causal impact of weight on wages, and that obese white females earn roughly 11 percent less than healthy-weight white females. Finkelstein, Fiebelkorn, and Wang (2003) calculate that, relative to the healthy weight, the obese incur $125 higher annual out-of-pocket health care costs. With two-thirds of Americans overweight or obese (Ogden et al. 2006), and given these potential benefits of weight loss, it may not be surprising that 46 percent of all American women and 33 percent of all American men are trying to lose weight (Bish et al. 2005).

Most people fail in their attempts to lose weight,[2] and many of those who are successful in losing weight regain it in a short period of time.[3] For ex-

---

1. There are two ways researchers have sought to measure the benefits of weight loss. The first is to examine changes in outcomes associated with losing weight. The second is to compare the outcomes of individuals of different weight, and assume that the difference in outcomes is due to the difference in weight. Each has its limitations: weight loss studies often lack power, and comparisons across weight levels are confounded by differences in unobserved characteristics. Vidal (2002) assesses the evidence on the benefits of weight loss and concludes that modest weight loss (5 to 10 percent of initial body weight) improves cardiovascular risk factors and helps prevent or delay the onset of type 2 diabetes and hypertension.

2. Some obese individuals are able to lose weight by modifying their behaviors: eating less and exercising more. In the select group enrolled in the Weight Control Registry, all of whom have lost at least thirty pounds and kept it off for at least one year, 44.6 percent report losing the weight entirely on their own, that is, without the help of a commercial program, physician, or nutritionist (Wing and Phelan 2005). Clearly, such statistics do not generalize to the population; anyone who failed at initial weight loss is ineligible for this registry of people who maintained weight loss for a year.

3. Conventional wisdom is that virtually no one succeeds at maintaining weight loss. This perception has been traced back to a 1959 study of 100 obese individuals in which only 2 percent maintained loss of twenty pounds or more two years after the treatment (Stunkard and

ample, in one community-based study of weight gain prevention (Crawford, Jeffery, and French 2000), most (53.7 percent) participants *gained* weight in the first twelve months, three-quarters gained weight over three years, and only 4.6 percent lost weight and maintained the loss for three years.

Theory and evidence from psychology and behavioral economics provide several explanations for why so many weight loss attempts fail. First, the benefits of weight loss are not salient. For example, foregone quality of life and lost wages are not visible and therefore they are frequently unrecognized as opportunity costs (Bastiat 1850).

A second possible explanation for repeated failure at weight loss is that the benefits of weight loss may not be immediate. Improvements in health and labor market outcomes may not occur for some time after weight loss, and Ainslie (1975) finds consistent evidence that there is a decline in the effectiveness of rewards as the rewards are delayed from the time of choice.

A third explanation for repeated failure at weight loss is that, contrary to the standard economic model of discounted utility (Samuelson 1937), people may discount hyperbolically, which produces time-inconsistent preferences (Ainslie 1975). In this context, time-inconsistent preferences mean that people want to do what is in their long-run interest (lose weight), but they consistently succumb to the temptation to eat and be sedentary. Thaler and Shefrin (1981) describe individual decision making as a battle between a farsighted planner (who in this context wants to diet) and a myopic doer (who in this context wants to eat and be sedentary).

One intervention, financial rewards for weight loss, may offer a solution to the problems of salience, immediacy, and time-inconsistency. Financial rewards, even though they may be dwarfed in value by the other benefits of weight loss, have the benefit of being salient, with their amount and delivery date known with certainty in exchange for clearly defined objectives. Even small financial incentives can be effective because research has found that people tend not to compare payoffs to their income or wealth but instead "bracket" them—consider them in isolation (Read, Loewenstein, and Rabin 1999; Kahneman and Tversky 1979). Lotteries may be particularly cost-effective incentives for healthy behavior. People tend to overweight the probability of unlikely events and underweight the probability of likely events (Kahneman and Tversky 1979), implying that lotteries can be more attractive than certain payments, even if the two have equal expected values. Financial rewards can also be paid immediately, before other benefits of weight reduction may be realized.

Financial rewards can also be structured to help people with time-

McLaren-Hume 1959; Wing and Phelan 2005). However, the 1959 study was based on a crude diet intervention with negligible support or follow-up so its poor results may not generalize to today's much more intensive interventions.

inconsistent preferences stay committed to weight loss. In general, precommitment devices may help people with time-inconsistent preferences empower their farsighted planner (Strotz 1955–1956; Laibson 1997). In this context, one could allow people to post a bond that is automatically forfeited if they fail to achieve their weight loss goals. Such a bond allows a person to influence their own future decisions by increasing the punishment for succumbing to short-run temptation. 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), which suggests that a posted bond may be more effective than a reward of the same size. Using a bond to increase adherence to a weight loss regimen does not guarantee success. Even individuals who are aware of their time-inconsistent preferences may still be partially naive in that they overestimate their future willpower (O'Donoghue and Rabin 2001), and as a result may either post too small a bond or have too much faith in the bond as a precommitment device.

Motivated by these theories and findings, several businesses now help employers offer financial incentives for employee weight loss. In addition, several businesses help consumers post bonds that are only refunded if one achieves specific weight loss goals. The William Hill betting agency in the U.K. books wagers that the bettor cannot achieve a specified weight loss in a specific period of time and verifies the weight loss with a medical examination (Burger and Lynham 2008).[4] A company named stickK.com[5] that was founded by Yale economists Ian Ayres and Dean Karlan allows people to post bonds that are forfeited if they fail to meet their weight loss goal. However, verification is weak: success in achieving one's goal is determined (and refunds are made) based on either the honor system or through verification by a third party chosen by the bettor, and if the third party does not submit a report the self-report of the bettor is accepted.

The contribution of this chapter is to examine outcomes in a program that offers various financial rewards (including certain payments, lotteries, and refundable bonds) for weight loss. The outcomes we examine include attrition and weight loss, both in pounds and as a percentage of baseline weight.

A substantial literature confirms that financial incentives influence healthy behaviors. Kane et al. (2004) review forty-two studies of the effect of economic incentives on preventive behaviors such as immunization, smoking cessation, and exercise; they find that the economic incentives were effective at changing behavior in 73 percent of studies. Financial incentives form the basis for an innovative substance abuse treatment program known as

---

4. This market is relatively small—the annual number of applications for such bets is roughly 200 (Burger and Lynham 2008).
5. The web site's Frequently Asked Questions page states that the company's name includes two K's because "K" often symbolizes "contract" in legal writing.

contingency management. A meta-analysis found overwhelming evidence that such incentives raise compliance (drug abstinence) by an average of 30 percent (Lussier et al. 2006). Consistent with bracketing, even small financial incentives have proven effective; for example, as little as $2.50 for a single negative test result for cocaine (Higgins, Allesi, and Dantona 2002).

Specific to the current context, there is mixed empirical evidence on the extent to which weight loss is responsive to financial rewards. A recent review and meta-analysis (Paul-Ebhohimhen and Avenell 2007) 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 rewards on weight loss; it calculated a mean weight loss of 0.4 kg at twelve months, which was not statistically significant. A broader set of studies (including, e.g., those with nonrandomized designs or shorter follow-up) are listed in appendix table 4A.1.[6]

Relative to past studies, ours has several advantages. This study has a relatively large sample size (2,407); for comparison, the sample size of all published RCTs of financial incentives for weight loss combined totals 424 (treatment $N = 252$, control $N = 172$) (Paul-Ebhohimhen and Avenell 2007). The intervention studied by this chapter also covers a relatively long time period (one year). Moreover, we examine data from a real-world intervention rather than one constructed by, and overseen by researchers, which is important because a criticism of studies of weight loss programs is that it is unclear how the results of pilot programs generalize to real-world implementation. A limitation of this study, however, is that it is opportunistic data; individuals were not randomly assigned to different incentive schedules for weight loss.

4.3   Description of the Intervention

Our data come from a company (that we will call Company X) that helps employers provide financial incentives for their employees to lose weight; specifically, it monitors employee weight loss and pays the rewards. After an employer contracts with Company X, Company X has a kickoff event in the workplace that explains the program to the employees and encourages them to sign up. Participation is optional. Those who sign up select a physical activity regimen at either the foundation (easiest), intermediate, or advanced level. The program consists of several elements: (a) daily e-mail coaching that includes information about healthy and effective methods of weight loss such as decreasing calorie intake and increasing physical activity in a manner consistent with the regimen the enrollee chose at baseline;

---

6. There are other studies that offer financial rewards for exercise or for attending weight loss programs, but appendix table 4A.1 is limited to studies of financial rewards for weight loss.

(b) call center support; (c) weigh-ins at least once a quarter; and (d) financial incentives for achieving specific weight loss targets. Only employees who are overweight (body mass index [BMI] of at least 25) are eligible to receive financial rewards, and no financial rewards will be paid once an employee's BMI falls below 25 (i.e., when the employee falls into the "healthy weight" category).

The weigh-ins take place in kiosks that are compliant with the Health Insurance Portability and Accountability Act of 1996 (HIPAA)[7] and which Company X installs in the employer's workplace. Employees enter the privacy-protected kiosk and stand on a scale; their body mass index is recorded and sent over an Internet connection to their personal web page as well as to Company X's database. Participants can weigh themselves as often as they like, and the lowest recorded weight will be counted as that quarter's weight. Financial rewards are paid based on percent of baseline weight lost.[8]

Company X has a standard set of incentives that it proposes, but employers can modify it. In our data, there are three incentive schedules. The first is Company X's standard set of incentives: the employee participants pay no fee (all costs are paid by the employer), and employees receive quarterly payments determined by percent of baseline weight lost to date. Table 4.1 lists the standard set of incentives: payment thresholds occur at each percentage point of weight loss up to 5 percent (1, 2, 3, 4, 5), then thresholds occur every 5 percentage points (5, 10, 15, 20, 25, 30) up to 30 percent 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 percentage point of weight loss. Then the per-percentage-point rewards increase: $25 for losing 20 percent, $35 for losing 25 percent, and $50 for losing 30 percent. These are monthly amounts that are paid quarterly, so someone who loses 5 percent of his weight and keeps it off for three months receives a $15 check for the quarter ($5 monthly payment × 3 months). Five employers (with a total of thirteen work sites participating) used this standard incentives schedule.

The second ("modified") incentive schedule, used by one employer (with two work sites participating), is shown in table 4.2 and includes both a lottery and a deposit contract (bonds). The lottery takes place each quarter and the prizes are gift certificates (ten $50 gift cards and ten $50 salon vouchers); only

7. The Health Insurance Portability and Accountability Act (HIPAA) regulates the disclosure of health information.
8. 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, for example, 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.

Table 4.1   **Financial rewards based on weight loss, "standard incentives"**

Weight loss (as % of baseline weight)	Dollar reward per month (Paid quarterly)
1	1
2	2
3	3
4	4
5	5
10	10
15	15
20	25
25	35
30	50

Notes: 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 4.2   **Financial rewards based on weight loss, "modified incentives"**

Weight loss (as % of baseline weight)	Reward
Greater than zero	Entered into quarterly drawing for gift certificates: ten $50 gift cards each quarter and ten $50 salon vouchers each quarter.
5	Complete reimbursement of monthly fees (11 * $9.95 = $109.45), paid at end of year
10	Complete reimbursement of monthly fees (11 * $9.95 = $109.45) plus $100 bonus, paid at end of year
"Biggest loser" (as % of baseline) at work site	$250 gift certificate, awarded at end of year, plus the appropriate award listed above for the specific amount of weight loss

Notes: 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.

those who had lost some weight since baseline are eligible for the drawing. The deposit contract is that employees must pay $9.95 per month (except the first month, which is free), all of which (11 × $9.95 or $109.45) is refunded at the end of the year if the respondent loses at least 5 percent of baseline weight by year's end. If the respondent loses 10 percent or more of their baseline weight, they receive in addition to their refunded fees ($109.45) a $100 bonus, for a total of $229.40. In addition, 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.

Whether a participant would receive a higher payoff in the standard or modified group depends on both quarter and magnitude of weight loss. In quarters one through three, the standard incentives are more generous than the modified incentives at all levels of weight loss, with the exception that those losing between 0.1 percent and 0.9 percent of baseline weight receive no reward in the standard incentives group, but are eligible for the lottery for gift cards in the modified incentives group. In quarter four, the standard incentives are more generous for weight loss of between 1 percent and 4 percent, but the modified incentives are more generous for weight loss of 5 percent or more.

The third ("control") schedule, used by one employer (with a total of two work sites), offered no incentives for weight loss, but did include one modest incentive to not attrite: participants 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 e-mails, call center access, weigh-ins at the kiosk) but were offered no incentives for weight loss, making it useful both as a 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.

Figure 4.1 presents a flow diagram of attrition and analysis for all three groups (standard incentives, modified incentives, control) combined.

4.4   Hypotheses

Part of our purpose in this chapter is exploratory—to measure enrollment, attrition, and weight loss in these programs. We focus in particular on attrition and weight loss as outcomes because the National Institutes of Health (NIH) Technology Assessment Conference Panel (1993) recommends using the percentage of all beginning participants who complete the program, and the percentage of those completing the program who achieve various degrees of weight loss as measures of program success. The NIH considers a loss of 10 percent of baseline weight in six months to one year to be good progress for an obese individual (USDHHS 2000).

Another purpose of this chapter is to test the following hypotheses.
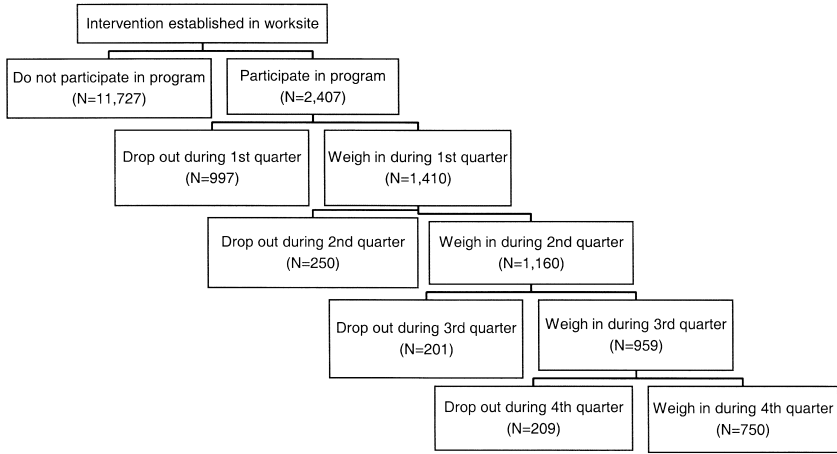
Fig. 4.1   Flow diagram of attrition and analysis

4.4.1   Hypotheses Regarding Enrollment

There will be lower enrollment in the program that required people to post forfeitable bonds. The law of demand states that the quantity demanded falls with price. The program that requires people to post a forfeitable bond raises the expected price of participation, assuming that not all possible participants expect a 100 percent probability of success (and therefore the return of their bond). The published literature confirms that, all else equal, enrollment in weight loss programs is lower if people are asked to post forfeitable bonds (e.g., Jeffery, Thompson, and Wing 1978).

Those who are willing to post a forfeitable bond will be better motivated or prepared for weight loss than those not required to post such a bond. In other words, we expect differential selection—those unwilling to post a forfeitable bond because they expect little weight loss are excluded from the modified incentives group, but are not excluded from the control group or standard incentives group. As a result, we expect that the modified incentives group will be better prepared or motivated for weight loss than the other groups.

4.4.2   Hypotheses Regarding Attrition

There will be lower attrition in the program that required employees to post bonds that are refundable based on achievement of weight loss goals. Those willing to post a bond are expected to be more motivated or determined to lose weight. Selection aside, bonds may also increase retention.

Those who attrite will have been relatively unsuccessful at weight loss. Participants enroll with incomplete information about certain costs and benefits of participating. Those that lose relatively little weight may update their prior beliefs and conclude that it is utility maximizing for them to drop out

of the intervention. This is especially true of those in the modified incentives group, who are charged a monthly fee for participation that will not be refunded if year-end weight loss is less than 5 percent of baseline weight.

4.4.3   Hypotheses Regarding Weight Loss

Weight loss will be greater for those offered financial rewards for weight loss. Both the standard incentives group and the modified incentives group were offered financial rewards for weight loss, whereas the control group was not offered any. In other words, we hypothesize that financial rewards are effective in promoting weight loss.

In quarter four, weight loss will be greater in the modified incentives group than in the standard incentives group.

This prediction is based on the magnitude of the incentives; the modified incentives group has much greater incentives for 5 percent and 10 percent weight loss by the end of quarter four. Specifically, the modified incentives group receives a refund of their $119.40 bond if at least 5 percent of weight is lost, with an additional bonus of $100 if 10 percent of weight is lost. Moreover, those achieving very high weight loss will be in competition for the $250 bonus for being the "biggest loser." In contrast, the standard incentives group is paid $5 per month for losing 5 percent of starting weight and $10 per month for losing 10 percent of starting weight (triple those amounts for the entire quarter). Relative to the standard incentives, the modified incentives create greater incentives for weight loss by the end of quarter four.

In addition, there are two reasons that the relative performance of the modified incentives group by the end of quarter four might be better than one would expect based on the magnitude of the rewards alone. First, we expect differential selection—those willing to post a bond are likely better prepared or more motivated for weight loss. Second, the research literature on loss aversion indicates that people are more motivated by a risk of losing their own money (as in the modified incentives group) than they are by the prospect of winning someone else's money (as in the standard incentives group).

In quarters one through three, weight loss will be greater in the standard incentives group than in the modified incentives group.

This prediction is also based on the magnitude of the incentives. In quarters one through three, the standard incentives group is offered $5 per month for 5 percent weight loss, and $10 per month for 10 percent weight loss (see table 4.1 for the full schedule of financial rewards). In contrast, there is no marginal reward for 5 percent or 10 percent weight loss in any of the first three quarters for the modified incentives group (those losing any weight at all are eligible for lottery prizes, but there is no additional reward for any weight loss above the trivial amount that makes one eligible for the lottery).

However, there are three reasons that the relative performance of the modified incentives group in quarters one through three might be better than one would expect based on the magnitude of the rewards alone. The first reason is differential selection. The second reason is loss aversion; the fear of losing one's money at year's end may motivate members of the modified incentives group to lose weight in the early quarters, even when there are no quarter-specific rewards for doing so. Third, it may take more than one quarter to achieve 5 percent or 10 percent weight loss, so in order to meet their year-end goals members of the modified incentives group may have to lose weight in earlier quarters, even though they have no financial incentives for meaningful weight loss in those quarters.

4.5   Methods and Data

A limitation of our data is that they are not the result of a randomized controlled trial. They are opportunistic data, provided to us by Company X. As a result, we face two challenges: (a) assignment to the three treatment groups is nonrandom: the incentive schedules were chosen by the employers; (b) the participation of employees is voluntary; there is selection by employees.

Regarding problem number one (selection by employers into different incentive schedules), we assume that this is ignorable. In other words, we assume that employer preference for incentive structure is uncorrelated with unobserved employee characteristics that affect attrition and weight loss. Company X told us that the reason that one employer requested the modified incentives schedule (with forfeitable bonds) is because the company didn't want to pay for cash rewards. In contrast, it would be problematic if the modified schedule was requested because the employer thought it would be more effective for their particular employees.

A related problem is that unobserved employee characteristics may vary systematically across the three groups. Company X designed this intervention for office employees who spend their days in front of computers; it is they, for example, who are most likely to read the daily e-mails regarding nutrition and physical activity. For the most part, enrollees fit this description. Table 4.3 lists the industries of the employers. The five employers (with a total of thirteen work sites) in the standard incentive group include a Health Maintenance Organization (HMO) office, an HMO clinic (in which enrollees are nurses), two bank offices, and an insurance company. The one employer (with a total of two work sites) that instituted the modified incentive schedule is an insurance company, and the one employer (with a total of two work sites) in the control group is the administrative office of a grocery chain. Company X tells us that the nurses (who face the standard incentive schedule) have generally been least compliant with the program; they specu-

Table 4.3   Description of employers

Employer	Description	Incentive schedule
1	HMO clinic—nurses	Standard
2	Banking office	Standard
3	HMO office	Standard
4	Banking office	Standard
5	Insurance office	Standard
6	Insurance company	Modified
7	Grocery administrative office	Control

late that it may be because they do not work in front of computers all day and thus derive less benefit from the daily e-mails and the online tracking of measured weight.

Regarding problem number two (selection by employees into participation), we consider this to be a limitation for generalizing results to the entire population, but not a problem in the sense that any similar intervention is also likely to be optional, and so the findings for a set of volunteers is most relevant. All of the studies in appendix table 4A.1 are based on volunteers recruited to participate in a weight loss program, and are likewise not a random sample of the general population.

An additional problem when studying weight loss is that there is attrition from the program. Weight loss interventions in general (even those without financial rewards) typically have substantial attrition (Ware 2003; Gadbury, Coffey, and Allison 2003). There are several strategies for handling the attrition when evaluating interventions. The definitive is the intent-to-treat analysis, which 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 all of the dropouts, which is not available in this case. 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 most likely to quit are probably those for whom the intervention was least effective (Ware 2003). Another option is last-observation-carried-forward, which assumes that the dropouts remained at their last measured weight. This also likely results in upward bias in estimates of program effectiveness, as weight regain is common (Ware 2003; Serdula et al. 1999). Another option is baseline-carried-forward, which assumes that after attriting 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. We present findings for completers analysis, last-observation-carried-forward, and baseline-carried-forward.

The total number of employees in the data set is 2,407: 1,513 facing the standard incentives, 765 facing the modified incentives, and 129 in the con-

trol group with no financial incentives. The data cover 2004 to 2008. We drop from the sample participants with baseline BMI below twenty-five because they were not eligible for financial rewards. Thirteen participants in the control group were dropped because they were simultaneously participating in another workplace weight loss intervention.

We estimate attrition rates by quarter and group. We graph the distribution of weight loss by group and quarter, both for a completers analysis (ignoring dropouts), assuming that dropouts stayed at their last measured weight (last-observation-carried-forward) and assuming that dropouts return to baseline (baseline-carried-forward). We also calculate the unconditional mean loss in pounds and percent of baseline weight lost by group and quarter, for a completers analysis, last-observation-carried-forward, and baseline-carried-forward.

4.6   Empirical Results

4.6.1   Descriptive Statistics

Table 4.4 presents the summary statistics for participants by group. Our overall sample ($N = 2,407$) consists of 1,513 participants in the standard incentives group, 765 participants in the modified incentives group, and 129 participants in the control group.

In each of these groups, men are a minority: 15.7 percent of the standard incentives group, 21.2 percent of the modified incentives group, and 35.7 percent of the control group. Average age ranges from 43.0 to 46.2 across groups, and average baseline BMI ranges between 31.3 and 32.8 across groups. In each group there is a strikingly high prevalence of morbid obesity (BMI of greater than or equal to forty). In the United States as a whole, the morbidly obese constitute 4.8 percent of the population and 7.3 percent of all overweight Americans (Ogden et al. 2006). In contrast, the morbidly obese constitute 28.7 percent of the standard incentives group, 30.5 percent of the modified incentives group, and 22.5 percent of the control group.

4.6.2   Enrollment

We hypothesized that: *There will be lower enrollment in the program that required people to post forfeitable bonds.* Table 4.5 lists the percent of the workforce that enrolled in the program, by incentive schedule. Ideally, we would know the number of employees with BMI of twenty-five or higher, because only they are eligible for financial rewards for weight loss. Instead, for the denominator we know only the total number of employees (i.e., those of all BMI). As a result, these are likely to be underestimates of the percentage of those eligible for financial rewards who enrolled in the program. Percent enrollment was 18.6 percent for the modified incentives (which required a bond), 24.8 percent for the standard incentives, and 20.3 percent

Table 4.4   Summary statistics by group

Variable	Standard incentives			Modified incentives			Control group		
	Obs.	Mean	Std. dev.	Obs.	Mean	Std. dev.	Obs.	Mean	Std. dev.
Initial BMI	1513	32.8	6.24	765	32.8	6.00	129	31.3	5.72
Male	1513	0.157	0.364	765	0.212	0.409	129	0.357	0.481
Age	1513	46.2	10.4	765	43.0	8.8	129	44.4	10.6
Height	1513	65.5	3.41	765	66.1	3.42	129	66.7	4.25
Overweight (30 > BMI > = 25)	1513	0.412	0.492	765	0.382	0.486	129	0.519	0.502
Obese (40 > BMI > = 30)	1513	0.301	0.459	765	0.314	0.464	129	0.256	0.438
Morbidly obese (BMI > = 40)	1513	0.287	0.452	765	0.305	0.461	129	0.225	0.419
Foundation exercise regimen	1513	0.601	0.490	765	0.550	0.498	129	0.488	0.502
Intermediate exercise regimen	1513	0.337	0.473	765	0.374	0.484	129	0.426	0.496
Advanced exercise regimen	1513	0.062	0.241	765	0.076	0.265	129	0.085	0.280
E-mail open rate	740	45.7	36.41	765	51.0	35.09	129	28.7	32.47

Table 4.5   **Enrollment rates**

	Control group (1)	Standard incentive group (2)	Modified incentive group (3)	p-value (1) equals (2)	p-value (1) equals (3)	p-value (2) equals (3)
Mean	0.203	0.248	0.186	0.613	0.839	0.477
(Std. dev.)	(0.100)	(0.115)	(0.024)			

Note: Enrollment rates are calculated by the fraction of those who enroll in the program by the total population of the workplace. Individuals with BMI < 25 may enroll in the program, but receive no payouts.

for the program that offered no financial rewards for weight loss but all of the other program elements (i.e., the control group). The point estimates of enrollment are consistent with our prediction that the requirement of a bond would result in lower enrollment, but the differences are not statistically significant.

We also hypothesized that: *Those who are willing to post a forfeitable bond will be better motivated or prepared for weight loss than those not required to post such a bond.* There are two variables that can give us information about the degree of such differences in selection. The first variable is the level of exercise regimen that the employee chose at the beginning of the program. If those willing to pay the monthly fees in the modified incentives group are more motivated or prepared to lose weight, one should find that they are less likely to choose the easiest exercise regimen. This is confirmed by the data. Table 4.4 indicates that the easiest exercise regimen (called Foundation) was chosen by 60.1 percent of the standard incentives group but only 55.0 percent of the modified incentives group, a difference significant at the 1 percent level. We also expected that the control group, offered $20 if they participated for the full year, would be less motivated on average and, therefore, more likely to choose the easiest exercise regimen than those in the modified incentives group, but we do not find this—an even lower percentage of the control group than the modified incentives group (48.8 percent versus 60.1 percent) chose the easiest exercise regimen, but the difference is not statistically significant.

The second variable that sheds light on difference in selectivity is the percentage of the program e-mails that enrollees read. If those willing to pay the monthly fees in the modified incentives group are more motivated or prepared to lose weight, one should find that they read a higher percentage of the program e-mails. That prediction is confirmed by the data—table 4.4 indicates that the average percentage of e-mails read was 51.0 percent for members of the modified incentives group compared to 45.7 percent for members of the standard incentives group, a difference significant at the 1 percent level. (A caveat is that this variable is missing for 51.1 percent of the

standard incentives group—it simply wasn't recorded for certain employers in certain years.)

The control group, being paid to participate, had the lowest e-mail open rate of 28.7 percent, which is significantly different from both other groups at the 1 percent level. It is interesting that the control group had the lowest percentage choosing the easiest exercise regimen (which suggests more motivation or better preparation) but the lowest e-mail open rate (which suggests lower commitment).

Overall, the patterns of both exercise regimen and e-mail opening suggest that the group required to post a bond (i.e., the modified incentives group) was selected to be better prepared and more serious about weight loss than the standard incentives group, and therefore should be less likely to attrite and more likely to lose weight.

4.6.3   Attrition

Table 4.6 lists the cumulative percentages dropping out, by quarter, for each group. In the standard incentives group, 51.2 percent of baseline participants have dropped out by the end of quarter one, and cumulative attrition rises in the three subsequent quarters to 62.1 percent, 72.0 percent and 76.4 percent. In the modified incentives group, attrition is lower: 24.8 percent after one quarter, rising in the three subsequent quarters to 33.5 percent, 39.3 percent, and 57.4 percent. Even in the control group, where participants are promised $20 if they weigh in every quarter for a year, attrition is substantial: 25.6 percent after one quarter, rising in the three subsequent quarters to 39.5 percent, 45.0 percent, and 48.1 percent. When considering the levels of attrition, one should keep in mind that enrollees were already a select sample. Participation was optional, and most employees declined to enroll.

Attrition is typically substantial in weight loss interventions of all kinds (Ware 2003; Gadbury, Coffey, and Allison 2003). However, the attrition in these groups is particularly high. For example, a recent review (Paul-Ebhohimhen and Avenell 2007) of RCTs involving financial rewards for weight loss found that the maximum attrition in any such study was 57.9 per-

Table 4.6   **Cumulative attrition, by group and quarter**

Quarter	Standard incentives (%)	Modified incentives (%)	Control group (%)
1	51.2[a,b]	24.8	25.6
2	62.1[a,b]	33.5	39.5
3	72.0[a,b]	39.3	45.0
4	76.4[a,b]	57.4[a]	48.1

[a]represents significant difference with the control group at the 5% level

[b]represents significant difference between standard and modified incentive groups at the 5% level

cent at thirteen months, far below what the standard incentives group experienced in twelve months (76.4 percent, but roughly equal to what the modified incentives group experienced at twelve months (57.4 percent). This suggests that real-world interventions may experience far higher rates of attrition than those overseen by researchers (who for the purposes of data quality undertake extensive efforts to keep enrollees from attriting), which raises questions about how well the results of pilot studies such as those in appendix table 4A.1 can be duplicated on a larger scale.

We hypothesized that: *There will be lower attrition in the program that required employees to post bonds that are refundable based on achievement of weight loss goals*. The data are consistent with this hypothesis; in every quarter, attrition is significantly lower in the modified than the standard incentives group. For example, table 4.6 shows that, by the end of quarter one, attrition in the modified incentives group is only half that in the standard incentives group (24.8 percent versus 51.2 percent). It is impossible to tell from our data whether the difference in attrition is due to selection or loss aversion. Selection was evident in the earlier finding that those in the modified incentives group were more likely to choose an advanced physical activity regimen and tend to open more program e-mails; before entering the program they may have been better prepared and more motivated to lose weight. On the other hand, those in the modified incentives group have "skin in the game" in the form of their deposits, and loss aversion may motivate them to stay in the program.

We also hypothesized that: *Those who attrite will have been relatively unsuccessful at weight loss.* Table 4.7 lists the weight loss (in pounds) by quarter, categorized by whether the participant dropped out in the following quarter or persisted in the program through the following quarter. The table is divided vertically into four panels: full sample, standard incentives group, modified incentives group, and control group. Among the full sample, those who drop out in the subsequent quarter have significantly lower average weight loss than those who persist through the next quarter, in quarters one, two, and three. For example, in the full sample, those who stay in the program through quarter two had quarter one weight loss of 4.67 pounds on average, whereas those who dropped out during quarter two had quarter one weight loss of 3.49 pounds on average. For the full sample in each quarter, the difference in mean weight loss to date is statistically significant at better than the 1 percent level. When we divide the sample by incentive schedule, the same pattern exists for those in the modified incentives group: in each of the first three quarters, weight loss to date is significantly lower among those who drop out in the following quarter than those who persist through the following quarter. Note that those in the modified incentives group have the greatest incentive to drop out if they are not making progress, because to persist requires paying monthly fees that may be forfeited. The pattern is weaker for the standard incentives group; in quarter two future dropouts

Table 4.7   **Weight loss by future attrition status**

Quarter	Persist in next quarter	Dropout next quarter	*t*-test *p*-value
	Full sample		
1	4.67	3.49	0.004
	(2.3%)	(1.8%)	
2	5.73	3.33	0.000
	(2.8%)	(1.7%)	
3	6.38	4.23	0.008
	(3.1%)	(2.0%)	
	Standard incentive group		
1	4.90	4.07	0.122
	(2.5%)	(2.1%)	
2	6.67	3.99	0.003
	(3.2%)	(2.0%)	
3	6.93	8.96	0.128
	(3.2%)	(4.3%)	
	Modified incentive group		
1	4.66	1.03	0.000
	(2.3%)	(0.6%)	
2	5.36	0.76	0.001
	(2.6%)	(0.3%)	
3	6.77	–1.91	0.000
	(3.4%)	(–1.0%)	
	Control group		
1	3.42	3.54	0.929
	(1.8%)	(1.5%)	
2	3.46	1.36	0.325
	(1.9%)	(0.6%)	
3	1.82	7.99	0.041
	(0.9%)	(4.0%)	

Note: Weight loss in pounds (percent weight loss in parentheses).

have significantly lower weight loss than those who persist through the next quarter, but the difference is not statistically significant. In quarter one and in quarter three, the sign is in the opposite direction and the difference is not statistically significant. For the control group, in no quarter do future dropouts have significantly lower weight loss to date than those who will persist in the program. On the whole, these results suggest that, for the full sample as well as for the modified incentive group in particular, those who attrite are those who have been relatively unsuccessful at weight loss. In other words, the participants who are relatively successful at losing weight are more likely to remain in the program.
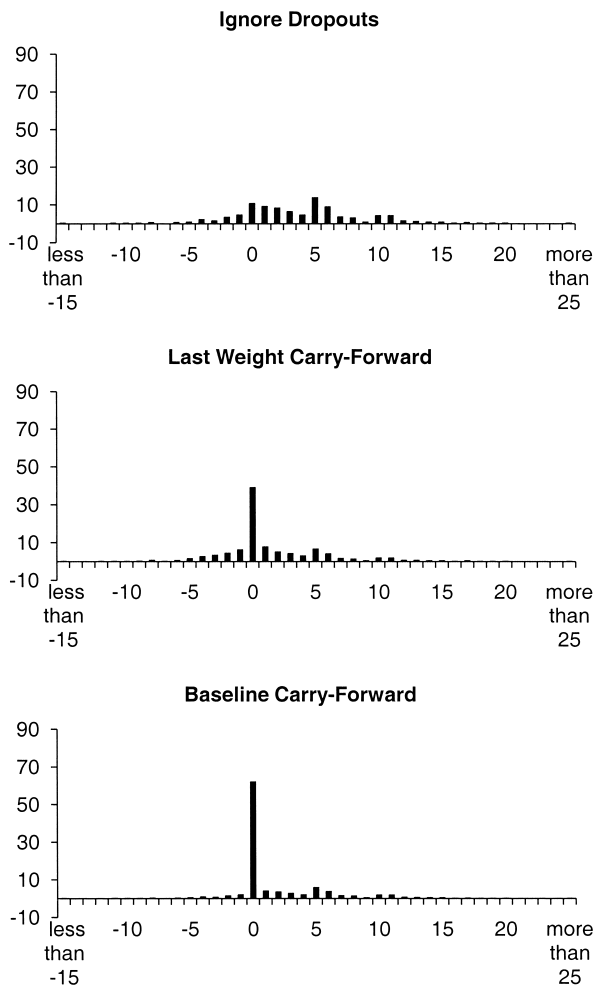
4.6.4   Weight Loss

The distribution of percent weight loss at the end of the program (end of quarter four), is shown in figure 4.2 (for the standard incentives group), figure

Ignoring Dropouts

Last Weight Carry-Forward

Baseline Carry-Forward
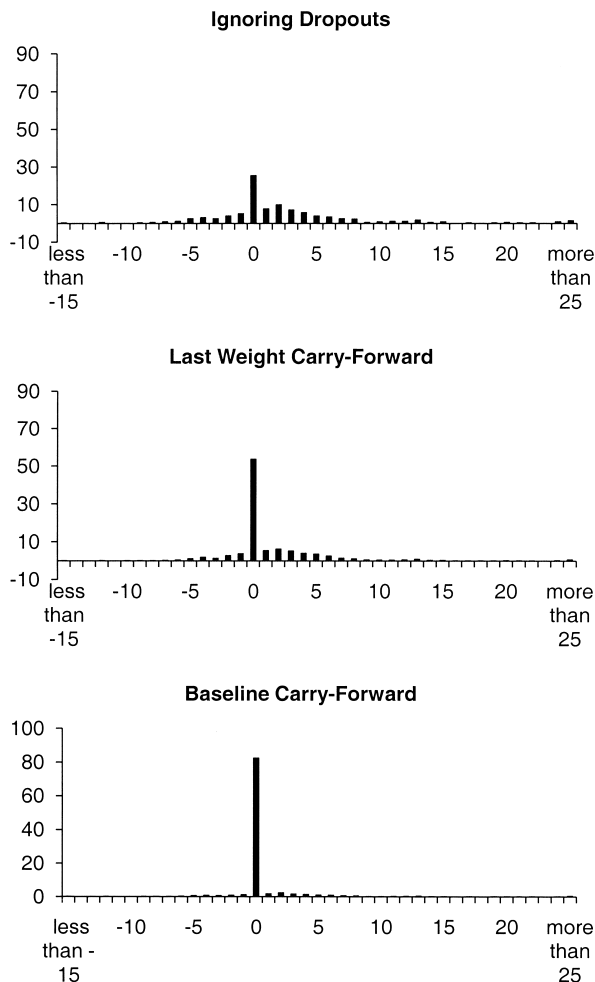
Fig. 4.2   Distribution of percent weight loss under standard incentives, quarter 4
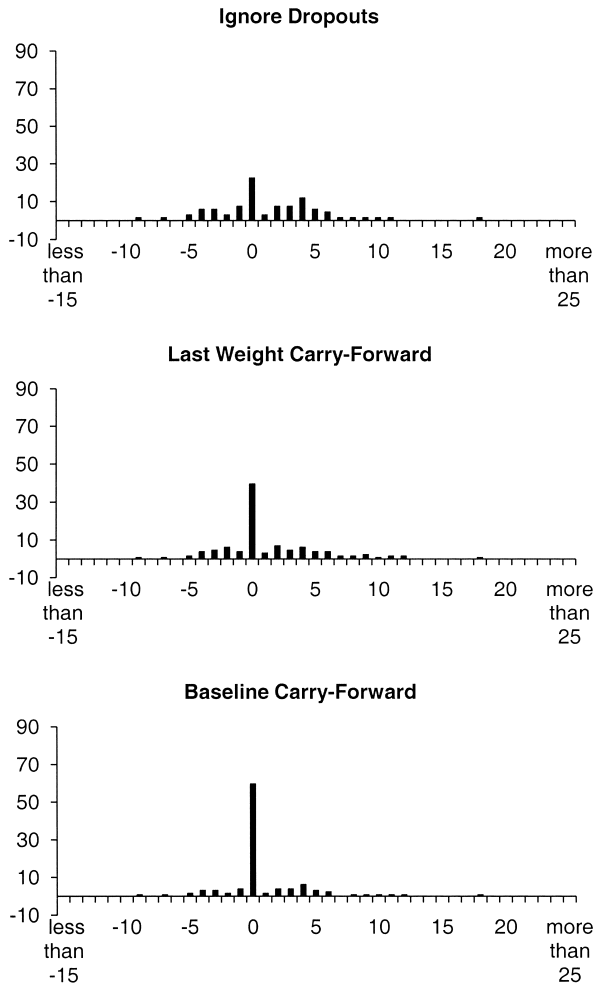
4.3 (for the modified incentives group), and figure 4.4 (for the control group). The horizontal axis shows the percent of baseline weight lost (rounded down to the nearest percentage point[9]) and the vertical axis indicates the percentage of that sample. Each figure consists of three graphs: the top graph is the distribution of weight loss in a completers analysis that ignores dropouts, the

9. We round down so that everyone indicated as having a specific percent weight loss received exactly the reward associated with that percent weight loss. If we rounded to the nearest percentage point, a participant who lost 4.6 percent of her starting weight would be rounded to 5 percent even though she would not have qualified for the financial reward associated with achieving 5 percent weight loss.

Ignore Dropouts

Last Weight Carry-Forward

Baseline Carry-Forward

Fig. 4.3   Distribution of percent weight loss under modified incentives, quarter 4

middle graph is from a last-observation-carried-forward analysis in which dropouts are assumed to have stayed at their last measured weight, and the bottom graph is the distribution of weight loss in a baseline-carried-forward analysis that assumes that every dropout returned to their baseline weight. A comparison of the top, middle, and bottom graphs confirms that how attrition is handled has a substantial impact on estimated weight loss. In the top graphs (the completers analysis), the distribution of outcomes seems more favorable (although the modal outcome is usually zero weight loss), but in the middle and bottom graphs that include information on dropouts, by far the most common outcome is that respondents lost zero weight (largely driven by the assumption of setting dropouts at baseline weight).

Ignore Dropouts

Last Weight Carry-Forward

Baseline Carry-Forward

Fig. 4.4   Distribution of percent weight loss in the control group, quarter 4

Each of the graphs in figure 4.3 indicates that more people in the modi-fied incentives group are just over the thresholds of 5 percent weight loss (at which participants are refunded their year's worth of fees, or $109.46) and 10 percent weight loss (at which they also receive a $100 bonus), than just under the thresholds. Moreover, such heaping is not apparent in the dis-tribution associated with the standard incentive schedule, which has more continuous reward thresholds. This suggests that people may be pushing to achieve the substantial payoffs associated with losing 5 percent or 10 percent of baseline weight.

We next discuss the evidence regarding our hypotheses regarding weight loss.

Weight loss will be greater for those offered financial rewards for weight loss.
We test for differences in unconditional means of weight loss in pounds and
percent of baseline weight by quarter and group. We then test for differences
in unconditional probability of losing 5 percent and 10 percent of baseline
body weight. Note that the differences between the treatment groups and
the control group can be interpreted as the effect of the financial incentives,
distinct from all of the other program elements (e.g., daily e-mails and call
center support) shared by the control group, with the caveat that there may
be differential selection by employers to incentive schedules, and that there
may be differential selection by workers into participation that differs by
incentive schedule.

Table 4.8 lists weight loss in pounds and percent of baseline weight, by
group and quarter. The cells also list the minimum and maximum weight
loss (a negative minimum weight loss indicates weight gain) for that group in
that quarter (the minimum and maximum are not affected by how dropouts
are treated, so they are entered in only the leftmost column for each group).
Because so many participants drop out, and attrition is correlated with
weight loss success, estimates of average weight loss are extremely sensitive
to how attrition is handled. We focus here on the baseline-carried-forward
analysis, which assumes that everyone who dropped out went back to their
baseline weight.

In the baseline-carried-forward analysis, average weight loss in the con-
trol group totals 2.6 pounds (1.29 percent) by the end of the first quarter,
1.9 pounds (0.98 percent) by the end of the second quarter, 1.7 pounds (0.82
percent) by the end of the third quarter, and 1.7 pounds (0.87 percent) by
the end of the fourth quarter. These can be interpreted as the unconditional
average effect of the program elements other than financial rewards (e.g.,
e-mails, call center access, and weigh-ins), because in a previous random-
ized experiment, a control group that received no treatment of any kind
experienced virtually no change in average weight after six or twelve months
(Jeffery et al. 1993). This suggests that changes in weight observed in the
control group measure the effect of all elements of Company X treatment
except financial rewards.

In the standard incentives group, average weight loss totaled 2.2 pounds
(1.13 percent) by the end of the first quarter, 2.1 pounds (1.04 percent) by
the end of the second quarter, 2.2 pounds (1.03 percent) by the end of the
third quarter, and 1.4 pounds (0.64 percent) by the end of the fourth quarter.
We fail to reject the null hypothesis of no difference between the control and
standard incentives group; in fact, average weight loss is consistently lower
in the standard incentives group than in the control group.

Despite the small average weight loss in the standard incentives group,
there are some substantial success stories; the maximum weight lost since
baseline is 58.8 lbs. in quarter one, 89.4 lbs. in quarter two, 109.2 lbs. in
quarter three, and 116.8 lbs. in quarter four. For any given mean, success

Table 4.8   Weight loss in pounds and percent of baseline weight, by group and quarter

Quarter	Standard incentives			Modified incentives			Control group		
	Ignoring dropouts	Last weight carry-forward	Baseline carry-forward	Ignoring dropouts	Last weight carry-forward	Baseline carry-forward	Ignoring dropouts	Last weight carry-forward	Baseline carry-forward
1	4.6 (2.31%) Min = −12.6 Max = 58.8	2.2[b] (1.13%)	2.2[b] (1.13)	4.2 (2.06%) Min = −11.6 Max = 34.6	3.2 (1.55%)	3.2 (1.55%)	3.4 (1.73%) Min = −6.0 Max = 19.6	2.6 (1.29%)	2.6 (1.29%)
2	5.5[a,b] (2.73%) Min = −19.8 Max = 89.4	2.7 (1.34%)	2.1[b] (1.04%)	4.9 (2.38%) Min = −18.4 Max = 52.8	3.3 (1.64%)	3.3 (1.58%)	3.1 (1.62%) Min = −8.8 Max = 30.2	2.4 (1.21%)	1.9 (0.98%)
3	7.77[a,b] (3.68%) Min = −25.4 Max = 109.2	3.2 (1.54%)	2.2 (1.03%)	4.1 (2.00%) Min = −29.8 Max = 53.6	2.6 (1.27%)	2.5 (1.21%)	3.0 (1.49%) Min = −22.2 Max = 32	2.2 (1.06%)	1.7 (0.82%)
4	6.1[b] (2.75%) Min = −25.6 Max = 116.8	3.2 (1.52%)	1.4[b] (0.64%)	8.4[a] (4.15%) Min = −30.6 Max = 61.2	3.3 (1.61%)	3.6[a] (1.77%)	3.2 (1.68%) Min = −13.8 Max = 30.8	2.9 (1.47%)	1.7 (0.87%)

Notes: A positive number indicates weight lost. A negative number (e.g., for the minimum weight loss) indicates weight gain.

[a]Represents significant difference with the control group at the 5% level.

[b]Represents significant difference between standard and modified incentive groups at the 5% level.

stories are balanced by failures; for example, the maximum weight *gain* is 12.6 lbs. in quarter one, 19.8 lbs. in quarter two, 25 lbs. in quarter three, and 25.6 lbs. in quarter four.

In the modified incentives group, average weight loss totaled 3.2 pounds (1.55 percent) by the end of the first quarter, 3.3 pounds (1.58 percent) by the end of the second quarter, 2.5 pounds (1.21 percent) by the end of the third quarter, and 3.6 pounds (1.77 percent) by the end of the fourth quarter. In quarter four (but not earlier quarters) the difference between the modified incentives group and the control group in average weight loss is statistically significant.

We also measure weight loss by success in reaching certain benchmarks. Table 4.9 lists the percent of participants losing 5 percent of baseline weight, by group and quarter, for completers, last-observation-carried-forward, and baseline-carried-forward analyses. In the baseline-carried-forward analysis, the percentage of the control group that lost 5 percent of their baseline weight, by quarter, was: 9.3 percent, 7.8 percent, 13.2 percent, and 10.1 percent.

Relative to the control group, it is generally the case that smaller percentages of the standard incentives group achieved 5 percent weight loss in each quarter (8.3 percent, 8.2 percent, 7.9 percent, and 5.4 percent); the difference is statistically significant in quarters three and four.

Relative to the control group, higher percentages of the modified incentives group achieved 5 percent weight loss in each quarter (12.6 percent, 16.5 percent, 14.0 percent, 19.5 percent); the differences are statistically significant in quarters two and four.

We also examine the probabilities of losing 10 percent of baseline weight, the outcome that the USDHHS (2000) recommends for evaluating weight loss programs. Table 4.10 lists the unconditional probabilities of losing 10 percent of baseline weight by group and quarter, for completers, last-observation-carried-forward, and baseline-carried-forward analyses. Assuming that dropouts returned to their baseline weight, the percentage of the control group that lost 10 percent of baseline weight, by quarter, was 0.0 percent, 2.3 percent, 2.3 percent, and 3.1 percent. These are comparable to the corresponding percentages for the standard incentives group (1.2 percent, 2.0 percent, 2.9 percent, and 2.4 percent; the differences are not statistically significant. Relative to the control group, higher percentages of the modified incentives group achieved 10 percent weight loss in each quarter (2.1 percent, 4.3 percent, 3.8 percent, and 6.5 percent), but the differences are not statistically significant.

We hypothesized that: *In quarter four, weight loss will be greater in the modified incentives group than in the standard incentives group.* This is true for the unconditional means in table 4.8. Assuming dropouts return to their baseline weight (baseline-carried-forward), average year-end weight loss is 3.6 pounds (1.77 percent) in the modified incentives group compared to

Table 4.9   **Percent of respondents losing 5% of baseline weight, by group and quarter**

Quarter	Standard incentives			Modified incentives			Control group		
	Ignoring dropouts (%)	Last weight carry-forward (%)	Baseline carry-forward (%)	Ignoring dropouts (%)	Last weight carry-forward (%)	Baseline carry-forward (%)	Ignoring dropouts (%)	Last weight carry-forward (%)	Baseline carry-forward (%)
1	17.1	8.3[b]	8.3[b]	16.7	12.6	12.6	12.5	9.3	9.3
2	21.6	10.8[b]	8.2[b]	24.8[a]	16.9	16.5[a]	12.8	10.1	7.8
3	28.1	13.0	7.9[a,b]	23.1	15.0	14.0	23.9	15.5	13.2
4	22.7[b]	13.3[b]	5.4[a,b]	45.7[a]	20.9	19.5[a]	19.4	17.8	10.1

[a]Represents significant difference with the control group at the 5% level.
[b]Represents significant difference between standard and modified incentive groups at the 5% level.

Table 4.10   **Percent of respondents losing 10% of baseline weight, by group and quarter**

Quarter	Standard incentives			Modified incentives			Control group		
	Ignoring dropouts (%)	Last weight carry-forward (%)	Baseline carry-forward (%)	Ignoring dropouts (%)	Last weight carry-forward (%)	Baseline carry-forward (%)	Ignoring dropouts (%)	Last weight carry-forward (%)	Baseline carry-forward (%)
1	2.4	1.2	1.2	2.8	2.1	2.1	0.0	0.0	0.0
2	5.2	2.2[b]	2.0[b]	6.5	4.3	4.3	3.8	2.3	2.3
3	10.4[b]	3.8	2.9	6.3	4.1	3.8	4.2	2.3	2.3
4	10.1[b]	4.4[b]	2.4[b]	15.3[a]	6.8	6.5	6.0	4.7	3.1

[a]Represents significant difference with the control group at the 5% level.
[b]Represents significant difference between standard and modified incentive groups at the 5% level.

1.4 pounds (0.64 percent) in the standard incentives group, a difference significant at the 1 percent level. Table 4.9 indicates that at the end of quarter four, 19.5 percent of the modified incentives group had lost 5 percent or more of their baseline weight, compared to only 5.4 percent of the standard incentives group, a difference significant at the 1 percent level. Table 4.10 shows that the percent losing 10 percent or more of baseline weight was 6.5 percent in the modified incentives group and only 2.4 percent in the standard incentives group, a difference significant at the 1 percent level.

In quarters one through three, weight loss will be greater in the standard incentives group than in the modified incentives group.

Contrary to our prediction, weight loss is greater in the modified incentives group than in the standard incentives group in quarters one through three. Table 4.8 shows that those in the modified incentives group lost an average of 3.2, 3.3, and 2.5 pounds in the first three quarters, compared to the standard incentives group average losses of 2.2, 2.1, and 2.2 pounds. This difference is statistically significant at the 1 percent level in quarters one and two. Table 4.9 shows that in each case a higher proportion of the modified incentives group than the standard incentives group achieved 5 percent weight loss: 12.6 percent versus 8.3 percent in quarter one, 16.5 percent versus 8.2 percent in quarter two, and 14.0 percent versus 7.9 percent in quarter three; in each case these differences are statistically significant at the 1 percent level. Table 4.10 shows that the probability of losing 10 percent or more of baseline weight is consistently higher in the modified incentives group than the standard incentives group, and the difference is statistically significant in quarter two.

These results suggest that the effect of greater financial incentives for the standard incentives group is swamped by some combination of more favorable selection into the modified incentives group, loss aversion, and the necessity of starting early to achieve 5 percent or 10 percent weight loss by the end of quarter four.

4.7   Discussion

A 2007 Institute of Medicine report on preventing obesity set the immediate next step—which it described as an essential priority action for the near future—as "learning what works and what does not work and broadly sharing that information." (IOM 2007, 410). It also notes that "All types of evaluation can make an important contribution to the evidence base upon which to design policies, programs, and interventions." (IOM 2007, 4). This chapter makes a contribution to that effort by documenting attrition and weight loss in a large program that offers financial incentives for weight loss.

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 2,407).

We study the two outcomes recommended by the NIH for evaluating weight loss interventions: attrition and weight loss. We find higher attrition (up to 76.4 percent after one year) than virtually all previous studies (see appendix table 4.1 and Paul-Ebhohimhen and Avenell 2007). Another recent study of real-world wagers on own weight loss also found 80 percent failure (Burger and Lynham 2008).

We find that the financial rewards in this program are associated with modest weight loss. After one year, it averages 1.4 pounds for those in the standard incentives group, 1.7 pounds for those in the control group, and 3.6 pounds for those in the modified incentives group, under the assumption that dropouts experienced no weight loss. (The additional 1.9 pounds lost by the modified incentives group relative to the control group is statistically significant, but the weight loss of the standard incentives group is not significantly different from that of the control group.) The NIH considers a loss of 10 percent of baseline weight in six months to one year to be good progress for an obese individual (USDHHS 2000). By this standard, very few participants in this program achieve good progress toward weight loss: just 2.4 percent of the standard incentives group, 3.1 percent of those in the control group, and 6.5 percent of the modified incentives group lost 10 percent of their starting weight in twelve months (neither the standard incentives nor modified incentives group is significantly different from the control group on this measure). By most measures, participants in the modified incentives group had twelve-month weight loss that was greater than those in the standard incentives group, but it is not clear how much of this is due to selection and how much is due to the difference in incentives (e.g., bonds) controlling for selection.

The weight loss associated with the program we examine is generally smaller than that documented in the previous literature. (This is especially true when one considers that selection of firms into specific incentive schedules and selection of employees into participation that differs by incentive schedule may imply that even the modest effects found here may be optimistic.) For example, Volpp et al. (2008) estimate mean sixteen-week weight loss to be 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 or bonds, 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), who find modest weight loss (between 2.0 and 4.7 lbs.) at three months, but no significant weight loss at six months, associated with financial rewards that varied between $7 and $14 per percentage point of weight lost after six months.

Likewise, Butsch et al. (2007) find no significant difference in twelve-week weight loss between a treatment group offered a $150 refund of their enrollment fee if they lost 6 percent of their initial weight, and a control group which was not eligible for such a refund.

Overall, our findings regarding attrition and weight loss suggest that the experience of pilot programs (such as those described in appendix table 4A.1) may be overly optimistic about what can be achieved on a larger scale.

To put our findings in a broader context of what works to promote weight loss, a literature review (Douketis et al. 2005) found that dietary and lifestyle therapy tends to result in less than 5 kg weight loss after two to four years, pharmacologic therapy results in 5 to 10 kg weight loss after one to two years, and surgical therapy results in 25 to 75 kg weight loss after two to four years. At this point, financial rewards remain an intriguing approach to weight loss but it remains to be seen whether they can be as effective as traditional medical approaches.

This chapter presents the basic patterns in the data. Our follow-up work will estimate hazard models of attrition and estimate regression models of weight loss to measure the change in weight associated with the incentive schedules, controlling for the observable characteristics of participants. Future research in this area should also focus on the optimal design of financial incentives for maximizing loss of excess weight, finding ways to decrease attrition, whether offering extrinsic rewards decreases intrinsic motivation, and whether weight loss is maintained after financial incentives for weight loss are removed.

Appendix

Table 4A.1   Previous literature on financial incentives for weight loss

Study	Study design	Intervention and incentives	Sample size and population	Duration	Weight loss	Attrition
Volpp et al. (2008)	Randomized controlled trial	3 groups: Deposits contract of $0–$3/day matched 1:1 Lottery for daily prize with E[V] = $3 Self-reported daily weight $20 for monthly weigh-in, unconditionally	N = 57 (19 in each of 3 groups) Patients at Philadelphia Veterans Affairs Medical Center with BMI 30–40	16 weeks	Mean weight loss: Lottery: 13.1 lbs Deposit contracts: 14.0 Control: 3.9 lbs	8.8%
Burger and Lynham (2008) working paper	Opportunistic data from William Hill betting agency for 1993–2006	Maximum bet of $65. William Hill offered odds ranging from 5:1 to 50:1; potential payoff averaged $1,926 Average duration of bet is 8 months, weight to be lost ranges from 28–168 lbs Each bettor weighed at start and end of bet by physician No control group	N = 51 Self-selected members of British population	Average of 8 months	Approximately 80% of people betting on their weight loss lose the bet	Approximately 80% of people betting on their weight loss lose the bet
Finkelstein et al. (2007)	Randomized trial, no control group	Three groups: Back loaded: $0 at 3 months, $14 per % point lost at 6 months Front loaded: $14 per % point lost at 3 months, $0 at 6 months Steady payment: $7 per % point lost at both 3 months and 6 months Weigh-ins at 3 months and 6 months Incentives only up to 10% weight loss ($140)	N = 207 (72 in back loaded, 64 in front loaded, 71 in steady payment) Overweight and obese employees at one university and 3 community colleges in NC	6 months	Mean weight loss 3 months: 2 lbs for back loaded, 4.71 lbs for front loaded, 3 lbs for steady payment Mean weight loss at 6 months not significantly different from zero	54% in back loaded, 45% in front loaded, 31% in steady payment
Butsch et al. (2007)	Sequential control-intervention, not randomized	Treatment group eligible for 50% reimbursement of enrollment fee ($150 of $300) if lose 6% of initial weight and attend 10 of 12 group sessions Control group was not eligible for reimbursement	N = 401 (241 intervention of which 59 enrolled, 160 control of which 40 enrolled) Participants in Univ. of Alabama at Birmingham EatRight lifestyle program BMI 30 and over	12 weeks	Mean weight loss: 2.25% in control group, 3.27% in intervention group; difference not statistically significant	Not stated

Table 4A.1   (continued)

Study	Study design	Intervention and incentives	Sample size and population	Duration	Weight loss	Attrition
Hubbert et al. (2003)	Propensity score matching of 4 controls to each member of intervention group	Treatment group eligible for 50% of cost of program fees ($150 of $300) if lose 6% of initial weight and attend 10 of 12 group sessions Control group was not eligible for reimbursement	N = 125; 25 in intervention group, 100 in control group Participants in Univ. of Alabama at Birmingham EatRight lifestyle program and members of UAB-owned HMO BMI 30 and over	12 weeks	Mean weight loss: 7.3 kg (6.1%) in intervention group, 4.0 kg (3.9%) in control group; both differences are statistically significant	Not stated
Jeffery, Forster, et al. (1993)	Block-randomized controlled experiment (work sites randomized)	Work sites divided evenly between treatment and control groups Treatment (Healthy Worker Project) consisted of health education classes and payroll deductions that served as bonds—refunded if achieve weight loss goals or donated to charity otherwise Goals chosen by employee and ranged from minimum of 0 lb and maximum of 1% body weight loss each week Participants chose amount of payroll deduction (minimum of $5 biweekly) 200 employees surveyed at baseline and again after 2 years (cohort) Another 200 employees surveyed after 2 years (cross-section) Weight self-reported but corrected for reporting error	32 work sites in Minneapolis / St. Paul metropolitan area Of 10,000 employees in treatment work sites, 2,041 employees participated in weight control program	2 years	No treatment effect was found for weight In cohort survey, average change in BMI was 0.08 units for control group, −0.02 units for treatment group; not statistically significant In cross-sectional survey, average change in BMI was −0.05 in both the treatment and control groups	No attrition of worksites
Jeffery, Wing, et al. (1993)	Randomized controlled experiment	Five groups: 1) control; 2) standard behavioral therapy (SBT); 3) SBT plus food provision; 4) SBT plus incentives; 5) SBT plus food provision plus incentives Weekly incentives: $0 if gained weight, $2.50 if did not gain weight; $12.50 if weight loss was 50% of goal, $25 if weight loss reached goal Weight-loss goals could be either 14, 18, or 23 kg Weight measured at baseline, 6, 12, and 18 months. There were also optional weekly weigh-ins	N = 202 men and women from Pittsburgh and Minneapolis-St. Paul, of which 40 to 41 were in each of the 5 groups Had to be 14-32 kg overweight	18 months	No effect of financial incentives or the interaction of financial incentives with food provision	11% attrition at 6 months, 13% at 12 months, 15% at 18 months

| Jeffery, Hellerstedt, and Schmid (1990) | Randomized experiment | Two groups: 1) offered a weight control newsletter program for price of $5; 2) offered the same program for free but requiring a $60 deposit that would be refunded based on (proportional to) success in weight loss. Individuals chose weight loss goals of not more than 4 lb a month. Weight self-reported (questionnaire, telephone survey). For subset of respondents, validation of self-report through measurement of weight | 6 months | Weight loss averaged about 4 lbs for $5 program and 8 lbs for incentive program | 3.8% did not return survey |
| Kramer et al. (1986) | Randomized controlled experiment | N = 1,304 residents of Bloomington, Minnesota: 1,190 in the $5 newsletter program group and 114 in newsletter plus incentive program group

Three groups: 1) monthly financial contingencies for weight maintenance; 2) monthly financial contingencies for participation in training sessions to solidify behavioral changes; 3) no treatment. $120 deposit. For each of 12 sessions not attended, participant forfeited $10. Refund also withheld if weighed more than "baseline" (post-first-treatment) weight. Withheld refunds (forfeited moneys) were distributed among those who were at or below "baseline" weight at final session. Weight measured at "baseline" and at one year | 1 year | Incentives had no impact on weight maintenance/ amount of weight regained. Average weight regain: 10.3 lbs in control group, 11.9 lbs. in group with incentives | 6 of 28 (21%) of the incentives group refused to attend final weigh-in. They self-reported weight, and 5 lbs was added to account for underreporting |
| Jeffery et al. (1984) | Randomized controlled experiment | N = 85 individuals who had already lost 10% or more of their body weight through a 15-week weight loss program

Three groups: 1) regular contract, 2) difficulty-grade contract; 3) no contract (control). All deposited $150. Immediately refunded to control group. Regular contract group received $30 for each 5-lb increment of weight loss. Difficult-grade contract group received $5 for first 5 lbs lost, $10 for second, $20 for third, $40 for fourth, and $75 for fifth | 15 weeks | Average weight loss: 26.2 lbs (12.8%) in difficulty-grade contract 21.7 lbs (10.8%) in regular contract 17.7 lbs (8.5%) in control group | 11 subjects (10%) refused to attend final weigh-in. They self-reported weight, and 5 lbs was added to account for underreporting. |
| | | N = 113 Roughly half recruited from population sample and the other half from newspaper advertisements | | | |

Table 4A.1 (continued)

Study	Study design	Intervention and incentives	Sample size and population	Duration	Weight loss	Attrition
Jeffery et al. (1983)	Randomized experiment	Six treatment groups: 3 levels of deposit ($30, $150, $300) times two types of payoff criteria: individual weight loss or mean group weight loss. All received 15-week behaviorally oriented program. Goal was 30 pounds lost. Cash refunds per week at rate of $1, $5, or $10 per pound up to 2 pounds per week. Monies not refunded for weight loss by end of program were distributed equally among those who achieved the 30-pound weight loss goal. Participants were weighed weekly	N = 89. Men in the Minneapolis area with self-reported weight at least 30 pounds above the ideal	15 weeks	Individuals rewarded for group performance lost on average 5 lbs more weight. This difference was maintained over 1 year follow-up. No significant effects of contract size	None
Coates et al. (1982)	Randomized experiment	Four treatment groups: 2 incentivized behaviors (weight loss or decrease in calorie consumption) by 2 frequencies of therapeutic contact (5 times or 1 time per week). Deposits were equal to 15 weeks' allowance or 50% of earnings from part-time work; amounts varied from $15–$240 (mean = $67.75). Source of payment: parents (51.5%), subjects (39.4%), shared (9.1%). Weight loss goal was 1 lb per week, or caloric reduction necessary to lose 1 lb per week. Monetary reward was delivered either once per week or once per week at treatment center. Weighed at each clinic visit. Food records checked	N = 36. Adolescents at least 10% above average weight-for-height	15 weeks	The treatment group receiving rewards for weight loss and coming to the clinic 5 times per week was the only group to significantly reduce the percent over-weight. Treatment effects maintained over a 6-month follow-up period. Significant correlation between initial monetary deposit and percent overweight lost. No significant difference based on whether parents or subject paid the deposit	None

Study	Design	Procedure	Subjects	Duration	Results	Dropouts
Jeffery, Thompson, and Wing (1978)	Randomized controlled experiment	Three treatment groups: deposits were returned contingent on either attendance, calorie restriction, or weight loss. Also a control group Each of the three treatment groups deposited $200. One group paid $20 for losing 2 lbs per week. Another paid $20 for calorie restrictions calculated to cause loss of 2 lbs per week. Third group paid $20 for weekly attendance	N = 31 Respondents to newspaper advertisement for people who need to lose 50 lbs or more	10 weeks	Groups rewarded for weight loss or calorie reductions lost an average of 20 lbs, significantly more weight loss than either the group rewarded for attendance (8.6 lbs) or the control group (12.4 lbs)	4 of 7 in control group quit
Mann (1972)	Single-subject reversal design	Subjects deposited a large number of valuables (e.g., money, jewelry, medals) with the researcher and signed a Contingency Contract allowing the researcher to switch them from treatment to control conditions, with the treatment being valuables being either returned or forfeited based on weight loss One valuable was returned for each 2 lb weight loss over a 2-week period. Subjects weighed every Monday, Wednesday, and Friday.	N = 8 Respondents to newspaper advertisement. All agreed to lose 25 pounds or more and had physician approval	Durations of treatments varied; total study ran at least 400 days	Average weekly weight loss of 1.6 to 1.7 pounds during treatment, regain of 1.4 pounds per week when incentives removed	None

References

Ainslie, G. 1975. Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin* 82 (4): 463–96.

Bastiat, M. F. 1850. *That which is seen, and that which is not seen: The unintended consequences of government spending.* West Valley City, Utah: Waking Lion Press, 2006.

Bish, C. L., H. M. Blanck, M. K. Serdula, M. Marcus, H. W. Kohl III, and L. K. Khan. 2005. Diet and physical activity behaviors among Americans trying to lose weight: 2000 behavioral risk factor surveillance system. *Obesity Research.* 13 (3): 596–607.

Burger, N. and J. Lynham. 2008. Betting on weight loss . . . and losing: Personal gambles as commitment mechanisms. University of California Santa Barbara. Working Paper.

Butsch, W. S., J. D. Ard, D. B. Allison, A. Patki, C. S. Henson, M. M. Rueger, K. A. Hubbert, G. L. Glandon, and D.C. Heimberger. 2007. Effects of a reimbursement incentive on enrollment in a weight control program. *Obesity* 15 (11): 2733–38.

Camerer, C. 2005. Three cheers-psychological, theoretical, empirical—For loss aversion. *Journal of Marketing Research* 42 (5): 129–33.

Cawley, J. 2004. The impact of obesity on wages. *Journal of Human Resources* 39 (2): 451–74.

Cawley, J., J. A. Rizzo, and K. Haas. 2007. Occupation-specific absenteeism costs associated with obesity and morbid obesity. *Journal of Occupational and Environmental Medicine* 49 (12): 1317–24.

Coates, T. J., R. W. Jeffery, L. E. Slinkard, J. D. Killen, and B. G. Danaher. 1982. Frequency of contact and monetary reward in weight loss, lipid change, and blood pressure reduction with adolescents. *Behavior Therapy* 13:175–85.

Crawford, D., R. W. Jeffery, and S. A. French. 2000. Can anyone successfully control their weight? Findings of a three year community-based study of men and women. *International Journal of Obesity* 24:1107–10.

Douketis, J. D., C. Macie, L. Thabane, and D. F. Williamson. 2005. Systematic review of Long-term weight loss studies in obese adults: Clinical significance and applicability to clinical practice. *International Journal of Obesity* 29:1153–67.

Finkelstein, E. A., I. C. Fiebelkorn, and G. Wang. 2003. National medical spending attributable to overweight and obesity: How much, and who's paying? *Health Affairs* Web Exclusive, W3-219.

Finkelstein, E. A., L. A. Linnan, D. F. Tate, and B. E. Birken. 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): 981–89.

Ford, E. S., D. G. Moirarty, M. M. Zack, A. H. Mokdad, and D. P. Chapman. 2001. Self-reported body mass index and health-related quality of life: Findings from the behavioral risk factor surveillance system. *Obesity Research* 9 (1): 21–31.

Gadbury, G. L., C. S. Coffey, and D. B. Allison. 2003. Modern statistical methods for handling missing repeated measurements in obesity trial data: Beyond LOCF. *Obesity* 4:175–84.

Higgins, S. T., S. M. Alessi, and R. L. Dantona. 2002. Voucher-based incentives: A substance abuse treatment innovation. *Addictive Behaviors* 27:887–910.

Hubbert, K. A., B. F. Bussey, D. B. Allison, T. M. Beasley, C. S. Henson, and D.C. Heimburger. 2003. Effects of outcome-driven insurance reimbursement on short-term weight control. *International Journal of Obesity* 27:1423–29.

Institute of Medicine. 2007. *Progress in preventing childhood obesity: How do we measure up?* Washington, DC: National Academies Press.

Jeffery, R. W., W. M. Bjornson-Benson, B. S. Rosenthal, C. L. Kurth, and M. M. Dunn. 1984. Effectiveness of monetary contracts with two repayment schedules on weight reduction in men and women from self-referred and population samples. *Behavior Therapy* 15:273–79.

Jeffery, R. W., J. L. Forster, S. A. French, S. H. Kelder, H. A. Lando, P. G. McGovern, D. R. Jacobs, and J. E. Baxter. 1993. The healthy worker project: A work-site intervention for weight control and smoking cessation. *Journal of Public Health* 83 (3): 395–401.

Jeffery, R. W., W. M. Gerber, B. S. Rosenthal, and R. A. Lindquist. 1983. Monetary contracts in weight control: Effectiveness of group and individual contracts of varying size. *Journal of Consulting and Clinical Psychology* 51 (2): 242–8.

Jeffery, R. W., W. L. Hellerstedt, and T. L. Schmid. 1990. Correspondence programs for smoking cessation and weight control: A comparison of two strategies in the Minnesota Heart Health Program. *Health Psychology* 9 (5): 585–98.

Jeffery, R. W., P. D. Thompson, and R. R. Wing. 1978. Effects on weight reduction of strong monetary contracts for calorie restriction or weight loss. *Behavior Research and Therapy* 16:363–9.

Jeffery, R. W., R. R. Wing, C. Thorson, L. R. Burton, C. Raether, J. Harvey, and M. Mullen. 1993. Strengthening behavioral interventions for weight loss: A randomized trial of food provision and monetary incentives. *Journal of Consulting and Clinical Psychology* 61:1038–45.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2): 263–92.

Kane, R. L., P. E. Johnson, R. J. Town, and M. Butler. 2004. A structured review of the effect of economic incentives on consumers' preventive behavior. *American Journal of Preventive Medicine* 27 (4): 327–52.

Kramer, F. M., R. W. Jeffery, M. K. Snell, and J. L. Forster. 1986. Maintenance of successful weight loss over one year: Effects of financial contracts for weight maintenance or participation in skills training. *Behavioral Therapy* 17 (3): 295–301.

Laibson, D. 1997. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics* 112 (5): 443–77.

Lussier, J. P., S. H. Heil, J. A. Mongeon, G. J. Badger, and S. T. Higgins. 2006. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction* 101:192–203.

Mann, R. A. 1972. The behavior-therapeutic use of contingency contracting to control an Adult behavior problem: Weight control. *Journal of Applied Behavior Analysis* 5 (2): 99–109.

National Institutes of Health (NIH) Technology Assessment Conference Panel. 1993. Methods for voluntary weight loss and control. *Annals of Internal Medicine* 119 (7): 764–70.

O'Donoghue, T., and M. Rabin. 2001. Choice and procrastination. *Quarterly Journal of Economics* 116 (1): 121–60.

Ogden, C. L., M. D. Carroll, L. R. Curtin, M. A. McDowell, C. J. Tabak, and K. M. Flegal. 2006. Prevalence of overweight and obesity in the United States, 1999–2004. *Journal of the American Medical Association* 295:1549–55.

Paul-Ebhohimhen, V., and A. Avenell. 2007. Systematic review of the use of financial incentives in treatments for obesity and overweight. *Obesity Reviews* 9:355–67.

Read, D., G. Loewenstein, and M. Rabin. 1999. Choice bracketing. *Journal of Risk and Uncertainty* 19 (1-3): 171–97.

Samuelson, P. A. 1937. A note on measurement of utility. *Review of Economic Studies* 4:155–61.

Serdula, M. K., A. H. Mokdad, D. F. Williamson, D. A. Galuska, J. M. Mendlein, and G. W. Heath. 1999. Prevalence of attempting weight loss and strategies for controlling weight. *Journal of the American Medical Association* 282 (14): 1353–58.

Strotz, R. H. (1955–1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies* 23 (3): 165–80.

Stunkard, A., and M. McLaren-Hume. 1959. The results of treatment for obesity. *Archives of Internal Medicine* 103:79–85.

Thaler, R. H., and H. M. Shefrin. 1981. An economic theory of self-control. *Journal of Political Economy* 89 (2): 392–406.

Tversky, A., and D. Kahneman. 1991. Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics* 106:1039–61.

U.S. Department of Health and Human Services, National Institutes of Health. 2000. *The practical guide: Identification, evaluation, and treatment of overweight and obesity in adults.* NHLBI Obesity Education Initiative, NIH no. #00-4084.

Vidal, J. 2002. Updated review on the benefits of weight loss. *International Journal of Obesity* 26 (Suppl 4): S25–S28.

Volpp, K. G., L. K. John, A. B. Troxel, L. Norton, J. Fassbender, and G. Loewenstein. 2008. Financial incentive based approaches for weight loss: A randomized trial. *Journal of the American Medical Association* 300 (22): 2631–37.

Ware, J. H. 2003. Interpreting incomplete data in studies of diet and weight loss. *New England Journal of Medicine* 348 (21): 2136–37.

Wing, R. R., and S. Phelan. 2005. Long-term weight loss maintenance. *American Journal of Clinical Nutrition* 82 (supplement): 222S–225S.