



RUHR

ECONOMIC PAPERS

Alfredo R. Paloyo
Arndt Rüdiger Reichert
Holger Reinermann
Harald Tauchmann

The Causal Link Between Financial Incentives and Weight Loss

An Evidence-based Survey of the Literature

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, email: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Christoph M. Schmidt
RWI, Phone: +49 (0) 201/81 49-227, e-mail: christoph.schmidt@rwi-essen.de

Editorial Office

Joachim Schmidt
RWI, Phone: +49 (0) 201/81 49-292, e-mail: joachim.schmidt@rwi-essen.de

Ruhr Economic Papers #290

Responsible Editor: Christoph M. Schmidt

All rights reserved. Bochum, Dortmund, Duisburg, Essen, Germany, 2011

ISSN 1864-4872 (online) – ISBN 978-3-86788-336-8

The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors' own opinions and do not necessarily reflect those of the editors.

Ruhr Economic Papers #290

Alfredo R. Paloyo, Arndt Rüdiger Reichert,
Holger Reinermann, and Harald Tauchmann

**The Causal Link Between
Financial Incentives and Weight Loss**
An Evidence-based Survey of the Literature

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über:
<http://dnb.d-nb.de> abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-336-8

Alfredo R. Paloyo, Arndt Rüdiger Reichert,
Holger Reinermann, and Harald Tauchmann¹

The Causal Link Between Financial Incentives and Weight Loss

An Evidence-based Survey of the Literature

Abstract

Obesity and overweight are linked to diseases that cost society a significant amount of resources. While behavior modification can reduce the problem, instigating such lifestyle changes is an uneasy task. One potential way to reduce the problem is through the use of financial incentives. In this survey, we review the available evidence with a significant emphasis on studies that yield credible estimates of the effect of financial incentives on weight loss. We find that the scientific literature on the subject has not yet satisfactorily settled whether such a mechanism is effective at eliciting the desired behavioral and health outcomes. We therefore advocate a rigorous large-scale randomized experiment to provide reliable estimates of the effect.

JEL Classification: I10, I12, I13, I18

Keywords: Financial incentives; weight loss; obesity; randomized experiment; survey

November 2011

¹ Alfredo R. Paloyo, Arndt Rüdiger Reichert, Harald Tauchmann: RWI; Holger Reinermann: Student Assistant at the Chair of Comparative Politics, University of Cologne. – We thank Viktoria Frei and Carina Mostert for research assistance. – All correspondence to Arndt Reichert, Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI), Hohenzollernstr. 1-3, 45128 Essen, Germany, E-Mail: arndt.reichert@rwi-essen.de.

1 INTRODUCTION

That obesity has reached epidemic proportions in the United States is well known. Testifying before the House of Representatives in 2003, the US Surgeon General warned that America was facing an obesity crisis. At that time, almost two thirds of Americans were either overweight or obese. More importantly, one in eight deaths in America was caused by a disease directly related to being overweight or obese.²

While figures from the US are drastic, what is extremely worrisome is the global nature of the problem. Table 1 presents comparable data on obesity³ from the OECD. Only Japan and Korea have single-digit percentages while most other countries have populations over a fifth of which is obese. Although categorized as one of the “diseases of affluence”, obesity is not exclusive to rich countries. As documented elsewhere (e.g., Prentice 2006), the obesity rate is exploding in developing countries as well, such that it is now more often referred to as the obesity pandemic, reflecting the extent of its geographic prevalence. The result is that international organizations and national policymakers have acknowledged that obesity and overweight are public health problems which require intervention.

One possible policy handle is the use of financial incentives to either reward or penalize individuals based on whether they have achieved a contractual target weight. The logic of using monetary instruments to encourage individuals to lose weight is straightforward. Unhealthy eating habits may increase a person’s utility in the short run but the subsequent decrease in utility (e.g., due to health problems associated with obesity) fails to be taken into account (Cutler et al. 2003). Most people are aware of this tradeoff and, in principle, would prefer long-run satisfaction over short-term bliss. For instance, in the US, 70 percent of obese women and around 63 percent of obese men try to lose weight (Bish et al. 2005). However, because some benefits of healthy living materialize in the distant future while food consumption yields instantaneous gratification, most people often fail to reach desired weight goals (Sassi 2010).

² “The obesity crisis in America”, Testimony Before the Subcommittee on Education Reform, Committee on Education and the Workforce, United States House of Representatives by Richard H. Carmona, Surgeon General. <http://goo.gl/KHWvY>. Accessed on 17 May 2011.

³ The World Health Organization (WHO) defines obesity as having a BMI greater than or equal to 30 kg/m². A person with a BMI greater than or equal to 25 kg/m² is considered overweight. (WHO 2011)

TABLE 1
**Percentage share of adult population with a BMI greater
than 30 kg/m²**

Country	Share in percent	Year of latest available figure
Australia	24.8	2007
Canada	24.2	2008
Chile	24.5	2003
Czech Republic	17.0	2005
Germany	14.7	2009
Ireland	23.0	2007
Japan	3.4	2008
Korea	3.8	2008
Luxembourg	20.0	2007
Mexico	30.0	2006
New Zealand	26.5	2007
Slovak Republic	16.9	2008
United Kingdom	24.5	2008
United States	33.8	2008

Source: OECD Health Data (2010).

Behavioral economists have taken this excessive discounting of future benefits relative to current costs as evidence of “time-inconsistent preferences”.⁴ The preferences are “inconsistent” because the action taken in the previous period tends to induce the economic agent to regret it in some future period. To resolve that dilemma, financial incentives may be used to bridge the gap between short-run benefits and long-run costs.

Private and social costs can also be seen as bifurcated in this case. For instance, a consumer may consider the cost of obesity for his health and future financial position but he is not cognizant of the cost to the social health insurance system. Taking into account the externalized social cost may tilt his behavior towards the healthy instead of the injurious. This disassociation of the social from the private may also be resolved by levying a so-called “fat tax”—really, a negative financial incentive—so that what is externalized may be properly internalized. The converse is also true for financial rewards for healthy behavior, which interlaces the social benefit with its private counterpart.

⁴ For an early formal description of this problem in economics, see Strotz (1955–1956). Recently, health econometric research has shown that, among other causes, obesity results from temporal decision biases (Ikeda et al. 2010).

What is yet unclear is how effective financial incentives are in reconciling the apparent inconsistency in preferences or in forcing consumers to internalize costs and benefits. While one may concede in general terms that these types of incentives nudge people towards a healthier lifestyle, the literature has not yet settled whether financial “carrots” (rewards or bonuses) or “sticks” (punishments like forfeiture of a cash deposit) are more effective. Moreover, beyond the question of weight loss, there is also the issue of weight maintenance after some reduction has already been achieved. Is the weight lost through financial incentives kept off? We aim to shed some light on this issue by offering a systematic review of the current literature and assessing the quality of the evidence presented by the various studies.

As a basis for this survey, we hold the following statements as self-evident. First, obesity and overweight are significant correlates of a large set of diseases, including—but not limited to—diabetes, high blood pressure, hyperlipidemia, numerous heart diseases, and other psychosocial problems (Mokdad et al. 2003).⁵ Second, the two conditions are principal contributors not only to total health-care costs (see, for example, Bhattacharya and Sood (2005) for the US and Konnopka et al. (2010) for Germany) but also to non-health-related costs, such as externalities in close quarters (Bhattacharya and Sood (2011) mention the example of economy-class seats in passenger aircraft). Third, the overall cost—direct and indirect, pecuniary and nonpecuniary, internal and external—of overweight and obesity is substantial, notwithstanding the savings associated with the fact that obese people tend to die younger.⁶ Finally, some individuals exhibit time-inconsistent preferences that lead them to engage in unhealthy behavior because they are unable to internalize the long-run consequences of their actions.

2 SURVEYING THE EVIDENCE

By examining the existing evidence, we wish to answer the following questions: “Can financial incentives induce weight loss and, if so, to what extent?”. Equally important, “If the weight loss was induced by financial incentives, can the subject maintain the new weight for an extended period?”

These questions have piqued the interest of many researchers across various academic disciplines. Hence, it is not surprising that numerous studies exist which

⁵ According to the WHO (2011), “44 percent of the diabetes burden, 23 percent of the ischemic heart-disease burden, and between 7 and 41 percent of certain cancer burdens are attributable to overweight and obesity.”

⁶ That obese people are likely to die younger is almost never taken into account, resulting in overestimated costs of obesity. See, e.g., McCormick et al. (2007).

attempt to provide credible answers. The variations in the incentive mechanism are almost as many as the number of studies itself. Presumably, making a direct comparison between and among studies along the lines of a meta-analysis will not be informative because of the idiosyncratic subtleties across studies. Differences in the study design, compliance rates, and indeed even the outcomes of interest certainly preclude such an approach. Instead, we conduct a narrative review of the existing literature and evaluate the strength of the evidence based on the quality of the study that generated it.

The initial source for the studies reviewed here comes from Jochelson (2007). She gives an overview of the use of financial incentives for giving up several habits that are possibly harmful for health, including those causing obesity. We supplement her literature review with material from Paul-Ebhohimhen and Avenell (2008). Finally, we searched Google Scholar and the EBSCOhost Online Research Databases for other relevant studies using the keywords “obesity”, “incentive”, “weight”, and “payment”. The full list of our extensive literature search is presented in Table A1 in the Appendix.

We distinguish our evidence-based review of the literature by emphasizing the methodological quality of the surveyed articles and by being explicit both with respect to the type of monetary incentive being analyzed (carrot/positive or stick/negative) and with respect to the outcome (weight loss or weight maintenance after weight loss). In particular, we focus on the results generated by randomized controlled trials (RCTs), with the conviction that this provides the scientific community with the most valid estimate of the effect—if any—of the incentive mechanism in place.

In order to qualify the relative strengths of the evidence, we had to appeal to a generally accepted hierarchy of evidence. In this hierarchy, the highest weight is assigned to evidence generated by RCTs. This is in accordance with the guidelines adopted by the US Preventive Services Task Force (USPSTF 1989) and the Centre for Evidence-based Medicine at the University of Oxford (CEBM 2009). Moreover, in evaluating health claims, the US Food and Drug Administration also accords the results of RCTs with the highest degree of credibility, particularly those of “double blind” RCTs (FDA 2009): “Randomized controlled trials offer the best assessment of a causal relationship between a substance and a disease because they control for known confounders of results (i.e., other factors that could affect risk of disease).”⁷

An alternative approach to RCTs is the ex-post evaluation of the outcomes of actual policies that have been implemented to encourage weight loss. These so-called “observational” studies are nonexperimental and take place in an uncontrolled

⁷ Blinding, of course, is not possible in the case of financial incentives since both the researcher and the subject cannot be made unaware of the treatment.

environment. In the present context, these can take on many forms and be implemented by various actors: among them are, for example, taxes on unhealthy foods or subsidies for healthy ones imposed by federal or local governments, health insurances offering reduced premiums to members who lose weight, or employers initiating employee health programs (e.g., Relton et al. (2011) and Cawley and Price (2011)). Some observational studies that evaluate similar reward schemes as those programmed in our selected RCTs are discussed below. These studies are informative with respect to the feasibility of financial incentive schemes outside the laboratory setting.

Although the nonrandomized design of observational studies offers some insights on the effectiveness of using monetary incentives to encourage weight loss, the credibility of the reported effect is threatened by confounding factors such as the nonrandom selection into the treatment. A well-designed and well-implemented RCT overcomes such issues, which is why it is often referred to as the “gold standard” in evaluation studies.⁸

Despite the narrow focus on RCTs, the comparisons are complicated even further because of the variability in reporting standards of the surveyed articles. Since 1996, the Consolidated Standards of Reporting Trials (CONSORT) has been in effect, which should guide researchers in reporting results of RCTs (Schulz et al. 2010).⁹ This has not always been adhered to and has thus hampered the ability of other researchers to conveniently obtain a cohesive picture painted by the scattered pieces of evidence.

A likely reason behind the deviations from CONSORT is that the issue of obesity and how to combat it has elicited an interdisciplinary company of allied researchers. Economists, in particular, are interested in the behavioral changes induced by manipulating the financial incentives faced by economic agents. Whereas one might expect epidemiologists and other public health researchers to be familiar with CONSORT (although they might not necessarily follow it), economists may not be and hence reporting conventions fall by the wayside.

⁸ To be fair, when appropriate methods are used, one can still retrieve a meaningful estimate of the effect from data obtained through observational studies. See, for example, the methodologies discussed by Imbens and Wooldridge (2009).

⁹ The latest version of CONSORT was published in 2010. In the case of nonrandomized experimental designs, the Center for Disease Control and Prevention (CDC) provides guidelines under Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) (Des Jarlais et al. 2004).

3 INCLUSION CRITERIA AND HIERARCHY OF EVIDENCE

Our principal inclusion criterion for an extensive discussion of the results of a study is its experimental design, with a special emphasis on having a genuine control group. Unfortunately, some studies compare two different treatments with each other without a third group to serve as the benchmark. The other criteria that we consider to be relevant for the credibility of the results are the effectiveness of randomization and the sample size. Rather than direct and incontrovertible evidence, we regard the results of studies that do not meet these criteria as coarse measures of the true effect. We mention these studies where appropriate but do not discuss them at length.

3.1 EXPERIMENTAL DESIGN

Ideally, the study design has to be organized into two or more experimental groups, one of which has to be an untreated (control) group to which treatment groups can be compared. The outcome for the control group serves as a proxy for the corresponding outcome for the treatment group had the treatment, in fact, not materialized. Nevertheless, we allow the absence of a control group only in the following situation: groups may receive several treatments at the same time but the financial incentive must be unique to one group and the other treatments must be common to all groups. This setup still allows us to isolate the effect of the financial incentive despite the presence of the other treatments under a more restrictive—though not necessarily unreasonable—assumption.

Consider the example in Table 2. If, as in Study Design 1, two groups receive the same Treatment A (say, a course in healthy cooking) but Experimental Group 2, in addition, receives a financial incentive tied to weight loss, the effect of the incentive can still be recovered under the assumption that a healthy-cooking course and financial incentives do not have synergistic effects.¹⁰ In this case, the effect of Treatment A is netted out. In Study Design 2, where two groups receive different treatments (say, a course in healthy cooking, Treatment A, in Experimental Group 1 and supervised exercise, Treatment B, in Experimental Group 2), and one faces a financial incentive contingent on weight loss, the impact of the incentive may not be identified.

Moreover, in practice, the differences between two treatment regimes are typically not as clear cut as our hypothetical designs in Table 2. For example, Jeffery et al. (1993a) seemingly identify the effect of financial incentives on weight loss but the experimental design does not credibly yield the effect of interest. The authors focus on providing a

¹⁰ More technically, we assume that the treatments enter the outcome equation additively and linearly.

negative incentive to program participants.¹¹ The treatment involved offering employees a weight-loss and a smoking-cessation program. Both programs included financial incentives and program participation was voluntary. The weight-loss program consisted of four rounds of 22 weeks, wherein courses were held every two weeks, covering behavioral advice for weight loss. Additionally, participants chose two parameters: (1) weight-loss goals of up to one percent of their body weight in between course meetings and (2) amounts of at least USD 5 to be deducted from their pay if they did not accomplish said weight-loss goal.

TABLE 2
Hypothetical study designs

	Study Design 1	Study Design 2
Experimental Group 1	Treatment A	Treatment A
Experimental Group 2	Treatment A + Financial Incentive	Treatment B + Financial Incentive

Source: Own representation.

While employees of the treatment firms were supported and incentivized to lose weight and to stop smoking, firms in the control group received no treatment. Such a situation does not allow us to exclusively identify the effect of financial incentives on weight loss. The purported effect on weight could be attributed to the smoking-cessation program or a combination of the programs. This is likely to be the case since smoking cessation is associated with changes in bodyweight (Austin and Grotmaker 2001).

3.2 RANDOMIZATION, COMPLIANCE, AND ATTRITION

In the studies we focus on, participants must have been randomly assigned to the experimental groups. Stratified random assignment is also allowed. Assignment mechanisms that are susceptible to selection biases (e.g., based on punctuality at health course meetings, which may be related to participant motivation) are excluded. For example, in Jeffery et al. (1978), the control group consisted of people who refused to be part of the treatment group.

Substitution and dropout biases that threaten the study design were carefully considered. A substitution bias occurs when, for instance, a person assigned to the control group is able to seek treatment on her own. Such behavior will render a simple comparison of means between the two experimental groups inadequate as a measure of the treatment effect. Nonrandom patterns in participant attrition will also bias the

¹¹ The study is uncommon and is unique among our selected articles because randomization was carried out at a more aggregated level. Instead of recruiting individual participants, 32 firms are randomized into treatment and control groups.

result. Therefore, one must be judicious in comparing studies with different substitution and dropout rates.

3.3 SAMPLE SIZE

When possible, we conduct tests to find out whether the number of observations per experimental group is sufficiently large to make assured statements on the effectiveness of financial incentives. This includes tests of whether certain assumptions (for instance, on sample distributions or variances), which are oftentimes made only implicitly, really hold.

3.4 EXCLUSIONS

Our interest is directed towards “output-oriented” incentives that are rewarded contingent on an outcome measure, such as weight. Some of the retrieved works did not actually require participants to pursue healthier living like, for example, Black and Friesen (1983) or Cameron et al. (1990), who examine financial incentives contingent on participants merely attending group meetings. Others incentivize an increase in physical activity or a change in eating habits and then measure weight loss as the outcome. We refer to this type of incentive scheme as “input-oriented” because the reward is directly linked to behavior associated with a healthy lifestyle. These studies are no less important than those focusing on output measures but a comparison of monetary incentives for a target weight with a similar incentive for increased exercise addresses a slightly different question.

3.5 FURTHER DISTINCTIONS

For the included studies, we distinguish between those that examine (a) incentives for weight loss, (b) incentives to maintain the reduced body weight, and (c) the sustainability of weight loss (i.e., whether participants are able to maintain the weight they reached under an incentive program when the incentives are no longer present).

Another characteristic along which earlier attempts differ is the type of incentives they offer. The most important distinction here is the one between negative and positive incentives. In the context of experimental research, negative incentives typically consist of initial cash deposits paid by the participants which are refundable contingent on the exhibition of the desired behavior. They therefore confront participants with the risk of losing money if they do not accomplish certain predefined goals (e.g., losing a certain amount of body weight). Positive incentives are those incentive schemes that offer a reward for the accomplishment of goals. This could be in the form of direct cash payments or a lottery.

We do not exclude either of these incentive regimes but acknowledge the necessity to carefully differentiate between the two. Given that participation in this kind of program is voluntary, participants in programs with negative incentives may need to have higher levels of motivation than participants facing positive incentives. Under a regime of negative incentives, failure to achieve the target is met with a reduction in utility; with positive incentives, participants end up—in the worst case—at the same utility level at which they started. Moreover, people tend to have an asymmetric evaluation of gains and losses. The possibility of losing money is given much more weight than an equal chance of gaining the same amount of money. This phenomenon, called “loss aversion”, has long been documented in the literature (see, e.g., Tversky and Kahnemann (1991)).

Finally, we acknowledge the setting in which the experiment takes place. Incentive schemes may work quite differently depending on the context in which they operate. For instance, participants of workplace programs may have the necessary social support to succeed in weight loss. This social network is likely to be absent in the laboratory setting.

4 RESULTS

Of the 22 studies reviewed here, only nine strictly met the inclusion criteria of which only five found that the financial incentives had any effect on the participants. Three studies which we consider to be relevant contributions are discussed below although they did not meet the inclusion criteria.

Studies that involve the use of positive incentives for weight loss include Finkelstein et al. (2007) and Jeffery et al. (1993b). The former provide evidence that positive financial incentives work while the latter is more ambivalent, suggesting that the free provision of healthy food choices is more effective at inducing weight loss.

Saccone and Israel (1978), Israel and Saccone (1979), Mahoney (1974), and Jeffery et al. (1984a) test the effectiveness of negative incentives for weight loss. Saccone and Israel (1978) provide some evidence that negative incentives could work but it is hard to interpret the effect as resulting only from the incentive mechanism in place since other factors, such as the identity of the disburser, may have had an influence. However, in general, negative incentives seem to be effective in encouraging weight loss but it does not seem to have a significant impact on maintaining the achieved weight losses for an extended period.

Volpp et al. (2008) and John et al. (2011) combine positive and negative incentives for weight loss. The results indicate that the combination of carrots and sticks encourages people to lose weight. However, the sustainability of such a weight

loss is doubtful. John et al. provide evidence that the incentives may work in the short to medium run but does not continue to help people keep the weight off.

With respect to the three studies that did not strictly meet the inclusion criteria, Kramer et al. (1986) address the question of whether financial incentives can be used to encourage people to maintain their weight after they have already lost a substantial part of it. Jeffery et al. (1983) analyze whether the amount of the bonus plays a role for the effectiveness of positive financial incentives for weight loss. Cawley and Price (2011) use observational data from an employer to investigate the effect on enrollment, attrition, and weight loss of a bonus program.

Table 3 gives an overview of the studies we reviewed and how we classified them with regard to type of incentive, incentivized behavior, and focus of the study. Volpp et al. (2008) and John et al. (2010) are not included in the table because they analyze a mixture of positive and negative incentives and could therefore not be exclusively positioned.

TABLE 3
A taxonomy for studies on the link between financial incentives and obesity and overweight

		Incentives for weight loss	Incentives for maintenance	Sustainability
Input-oriented	Negative incentive	Mahoney (1974)	Kramer et al. (1986)*	Mahoney (1974)
	Positive incentive	—	—	—
Output-oriented	Negative Incentive	Jeffery et al. (1984a) Mahoney (1974) Saccone and Israel (1978)	Kramer et al. (1986)*	Jeffery et al. (1984a) Mahoney (1974) Israel and Saccone (1979)
	Positive incentive	Finkelstein et al. (2007) Jeffery et al. (1983)* Jeffery et al. (1993b) Cawley and Price (2011)*	—	Jeffery et al. (1993b)

* These studies did not meet the inclusion criteria.
Source: Studies selected by authors.

4.1 POSITIVE INCENTIVES

Finkelstein et al. (2007) ostensibly test different levels of monetary rewards. They use a sample of 207 participants which they randomly assign to three experimental groups (see Table 4). The experiment is divided into two three-month phases. In the first phase,

one group (“front-loaded” group, $N = 64$) received USD 14 per percentage point of weight lost. The second group (“steady-payment” group, $N = 71$) received USD 7, and the third group (“back-loaded” group, $N = 72$) received no reward. In the second phase, the second group’s reward stayed the same, whereas the incentive structure of the other two rotated. While the front-loaded group received no reward, the back-loaded group received the reward of USD 14 per percentage point weight loss from baseline. As an example, if a participant from the steady-payment group lost 5 percentage points of their baseline weight, they would receive USD 35 and a further USD 35 if they maintain their weight six months after the start of the program. If this same participant were in the back-loaded group, she would receive nothing at three months and USD 70 at six months. In both cases, the participant would exit the program with USD 70.

For the first three months, Finkelstein et al. find front-loaded participants to lose more weight in absolute terms and to be more likely to lose at least five percent of their body weight than the back-loaded group. Both findings are statistically significant at the five-percent level. The same is true for steady-payment participants albeit this effect is not statistically significant. The analysis controls for age, sex, race, university education, and different levels of obesity at baseline.

TABLE 4
Study design of Finkelstein et al. (2007)

Experimental group	Reward schedule (USD per percentage point of weight loss)	
	First phase	Second phase
	Front-loaded Group	14
Steady-payment Group	7	7
Back-loaded Group	0	14

Source: Own representation.

While group assignment was randomized, we remain unconvinced that the authors are able to credibly quantify the effect of the monetary incentive on weight loss. The back-loaded group cannot serve as a control group because, at the outset, they are made aware of the monetary rewards at the end of the second phase. In other words, all groups were incentivized to lose weight from the very beginning. In order to maximize the payout at the end of the study period, the back-loaded group may have already engaged in weight-reducing activities during the first phase, thereby rendering any comparison of the first-phase results with the other experimental groups rather inappropriate.¹²

¹² A similar phenomenon was first documented by Ashenfelter (1978) when examining participation in training programs. Ashenfelter observed that the wages of participants in a

Another limitation is the high dropout rate in the study (the back-loaded group had an attrition rate of 36 and 50 percent after three and six months, respectively). Finkelstein et al. show that the odds of attending the weigh-in were, in both phases, significantly higher for those groups receiving rewards for that phase. They deal with the problem of attrition by pursuing an intent-to-treat approach—that is, by assuming that the dropouts returned to their baseline weight. Of course, there is no compelling reason to assume that this was indeed the case. An alternative would have been to exclude them from the analysis or, even better, to model participant attrition itself and use this information to correct their estimates. Unfortunately, a comparison of the results with and without the dropouts was not reported.

The results from the first phase can be seen as a weak indication of the effectiveness of positive financial incentives in encouraging weight loss. That the back-loaded group could have begun to lose weight at the start of the first phase implies that the significant impact found for the front-loaded group is likely to be underestimated. In other words, compared to a genuine treatment group, the magnitude of the effect may even be larger. We hesitate to interpret further the results from the second phase since we are unconvinced that it adequately addresses the question of the sustainability of financial incentives to induce weight loss. The reason for this is that there is no genuine control group whose weight trajectory could serve as a benchmark.

In Jeffery et al. (1993b), four different treatments are tested on a comparatively large sample of 202 participants for 18 months in total. They were randomly assigned to one control group and four treatment groups, each consisting of roughly 40 participants. The study design is presented in Table 5. All of the treatment groups received a “standard behavioral treatment” (SBT)—essentially, weight-loss advice. The meetings were held weekly for 20 weeks and once a month thereafter. In one group, the SBT was combined with providing participants specially prepared meals for free to assist with weight loss. Another group obtained the SBT plus a financial incentive. The last treatment group received a combination of all three treatments.

As a positive financial incentive, participants assigned to Treatment Groups 3 and 4 were rewarded USD 2.50 if they did not gain any weight from one week to another, USD 12.50 if they achieved 50 percent of their personal weekly weight loss goal, and USD 25 if they fully achieved it. The weekly weight-loss goals were fractions of an overall goal (one of 14, 18, or 23 kg) that participants chose for the whole study period. The analysis is based on weight at baseline and after six, 12, and 18 months. Only individuals that completed all three follow-ups were included in the analysis. Attrition, however,

training program exhibited a dip (hence, the phenomenon is called “Ashenfelter’s dip”) right before they entered the program.

was nontrivial. In the SBT and the combined incentive groups, only 65 and 85 percent, respectively, could be included.

TABLE 5
Study design of Jeffery et al. (1993b)

Experimental Group	Treatment
Control	None
Treatment Group 1	Standard behavioral treatment (SBT)
Treatment Group 2	SBT + free healthy meals
Treatment Group 3	SBT + financial incentive
Treatment Group 4	SBT + free healthy meals + financial incentive

Source: Own representation.

The authors find that financial incentives, unlike food provision, had no significant effect on weight loss. While at six months, Treatment Group 3 had lost 0.13 kg/m² (BMI units) more than Treatment Group 1, at 18 months, Treatment Group 1 had lost 0.26 kg/m² more than Treatment Groups 3 compared to baseline.¹³ Both groups regained weight after the sixth month.¹⁴ Significantly more successful at all weigh-ins were Treatment Groups 3 and 4, suggesting that the free provision of food tailored for weight loss is preferable to financial rewards or standard behavioral treatment.

There are two major limitations of Jeffery et al. (1993b). First, participants of Treatment Groups 3 and 4 were allowed to set their own weight-loss goal. It is likely that participants of the incentive groups set lower weight-loss goals than participants in the SBT group in order to have a higher likelihood to be rewarded. Thus, everything else constant, the expected weight loss at six, 12, and 18 months is lower for members of the incentive groups. Second, the attrition rates are high and different between groups. If unsuccessful participants have a higher likelihood to not show up at 18 months, the attrition rate of the SBT group should be higher because they have higher weight-loss goals. This hypothesis is supported by the data: in Treatment Group 1, the attrition rate is considerably higher than in the Treatment Group 3. Thus, it is likely that the effect of financial incentives for weight loss is biased towards zero.

¹³ Note that Treatment Group 1 is the adequate comparison group for the incentive group. The reason for this is that the "Controls" received no treatment at all while Treatment Group 3 received the SBT.

¹⁴ The authors remark that an analysis including subjects who were present at the 18-month follow-up and inputting all missing data from adjacent values yielded qualitatively similar results yet the results are neither discussed nor reported. An intention-to-treat analysis for all observations was not carried out. However, intention-to-treat results would be helpful to get an impression about the relevance of the attrition problem for the effect estimates.

4.2 NEGATIVE INCENTIVES

Saccone and Israel (1978) carry out a controlled experiment in order to find out whether input- or output-oriented incentives are more effective and whether financial refunds granted by people with which the participant has a strong personal relationship have different effects than financial rewards granted by external therapists.

The empirical analysis is based on a sample of 49 overweight adults aged 16–56 years who exhibit body weights at least 15 percent above the recommended level, to which a significant other in the participant’s family is available and who were willing to post a USD 35 deposit. Participants were randomly allocated (stratified by degree of overweight) into seven groups, one of which served as the control group. While the control group ($N = 5$) did not receive any treatment, all six treatment groups took part in a nine-week course which provided information on how to monitor food intake, how to establish a nutritionally sound diet, how to exercise, etc.

TABLE 6
Study design of Saccone and Israel (1978)

Experimental Group	Treatment
Control	None
Treatment Group 1	Program only: weight
Treatment Group 2	Program only: behavior
Treatment Group 3	Therapist: weight
Treatment Group 4	Therapist: behavior
Treatment Group 5	Significant other: weight
Treatment Group 6	Significant other: behavior

Source: Own representation.

In addition to course participation, two of these groups were requested to monitor on a daily basis body weight (Treatment Group 2, $N = 6$) or eating behavior (Treatment Group 3, $N = 8$). Yet, none of these groups were exposed to any financial incentives. During the final six weeks of the program, participants of any of the four remaining treatment groups (Treatment Groups 1, 4, 5, and 6) could achieve a refund of up to USD 5 per week. These groups differ in (i) whether the premium is refunded for successful weight loss (Treatment Groups 1 and 4) or for compliance in eating behavior (Treatment Groups 5 and 6) and (ii) whether the refund is paid directly by the therapist (Treatment Groups 1 and 5) or through a significant other (in most cases, the spouse; Treatment Groups 4 and 6) who also monitors the participant. Moreover, they attend a session where instructions for helping the participant might be given to the spouse. The sizes of these groups are between seven and eight members.

For all six treatment groups, the authors find a statistically significant decrease in weight over the treatment period of nine weeks. For the control group, an insignificant increase in body weight is found.

To address the effectiveness of financial incentives contingent on weight loss, the control group is not the appropriate reference group. While the control group indeed did not receive any treatment, the incentive groups received additional treatments other than the incentive itself. We are thus unable to isolate the exclusive impact of the financial incentive. A more appropriate comparison is between Treatment Group 1 and Treatment Group 2 (or between Treatment Group 2 and Treatment Group 4).

A comparison with all participants who could achieve a refund (irrespective of whether weight loss or changes in behavior was rewarded) yields a significant and negative effect of financial incentives on body weight. Yet, the more relevant pairwise comparisons between the course-only groups and each of the groups exposed to financial incentives yields an ambiguous picture. On one hand, a very small and insignificant effect is found for refunds granted for successful weight loss. On the other hand, participants who were refunded for compliance with nutritional advice were significantly more successful in reducing body weight than participants of the course-only groups. Another interesting result is that being refunded for compliance by someone to whom the participant has a strong personal relationship exerts a much stronger effect than being rewarded by the therapist.

Though yielding quite interesting and statistically significant results, the study still suffers from the very small number of observations that might limit opportunities for generalizing these results. Nevertheless, the finding that assigning a role to the participant's spouse or a significant other matters for the effect of financial incentives appears to be quite relevant. It suggests that changing eating habits represents a social (family) rather than an individual decision that cannot be effectively influenced on a purely individual basis. Moreover, the authors conclude that monetary reinforcement of a change in eating behavior is more effective than weight loss as the target behavior of the incentive.

In a follow-up study, Israel and Saccone (1979) required the participants of the six treatment groups to return for weigh-ins at three and 12 months in order to receive the remainder of a USD 35 deposit. During this time, they did not receive any treatment. The basic result is that, at both follow-ups, the financial incentive scheme did not result in weight-loss maintenance. This means that they are effective only on a short time horizon. This would point at financial incentives being even harmful for long-term weight loss. The authors report that a more promising long-term treatment is a monetary reinforcement for behavioral change through a family member or person in

the participant's natural environment. However, the very small number of observations and an attrition rate of 14 percent limit the informative value of the study.¹⁵

Mahoney (1974) uses negative incentives in two of three randomized experimental groups to address the twin issues of weight loss and the maintenance of weight loss. After attending group meetings at the time of project initiation, during two baseline weeks, all participants recorded their daily weight and eating habits and attended weekly weigh-ins. During the following six treatment weeks, they received standardized weight-loss and habit-improvement goals, i.e., quality, quantity, and timing as well as location of food consumption. In addition, two groups were instructed to award themselves portions of their own deposit (USD 35) depending on achieved weight-loss goals ($N = 13$) and habit-improvement goals ($N = 11$). The control group consisted of 14 participants. Nine weeks after the termination of the treatment phase, participants had a final weigh-in.

The author finds significant weight losses for all groups during the two baseline weeks. In the intervention phase, the self-reward group for habit improvement (“input-oriented”) lost 3.8 lbs. more than the control group and 3.3 percent more than the self-reward group for weight loss (“output-oriented”). The latter, however, did not lose significantly more weight than the control group. After a further nine weeks, all groups were successful in maintaining the weight losses attained at the end of treatment and did not differ in this respect.

Despite excellent weigh-in attendance rates, this analysis has substantial limitations. First, control subjects reported their pre-program weight problems as being of shorter duration than the incentive groups—implying better weight-loss chances—probably resulting in downward-biased effect estimates. Second, the power of the employed statistical Mann–Whitney test fails to reach conventional thresholds of 80 or 90 percent. This means that no conclusion regarding the ineffectiveness of output-oriented financial incentive schemes can be drawn. In other words, there remains a large risk of being wrong in rejecting the hypothesis of the effectiveness of deposits on weight-loss goals for weight reduction.

Jeffery et al. (1984a) analyze both weight loss and its sustainability. Sixty women and 55 men were recruited from a representative sample and through a newspaper advertisement. Women and men had to be overweight by at least 20 lbs. and 30 lbs., respectively. The overweight served as the weight-loss goal as well. The participants were randomly assigned to three different treatment groups, of which two required an initial

¹⁵ Although the authors mention that there appears to be no relationship between success in the program and attrition and that dropouts were unrelated to the treatment conditions.

deposit of USD 150: (1) a control group, (2) a contract group receiving a USD 30 refund for each five-pound increment of average group weight loss, (3) a second contract group receiving refunds of USD 5, USD 10, USD 20, USD 40, and USD 75 for successive five-pound individual incremental weight loss. The treatment period was 16 weeks. During that time, weekly group meetings covering nutritional, exercise, and behavioral principles were held. Sustainability of weight loss was addressed by randomly assigning participants to two different maintenance treatments which did not involve further financial incentives. No follow-up comparison between the treatment groups of the weight loss phase was carried out, i.e., the long-term effect of financial incentives were not analyzed. Thus, we exclusively concentrate on the results for the weight loss phase.

Attrition was generally low with only 11 subjects providing self-reported weight, to which the authors added 5 lbs. Employing multivariate ANOVA techniques with gender and the recruitment source as covariates, they find that percentage weight loss was significantly higher in both the constant-contract groups (10.8 percent) and the increasing-contract group (12.8 percent) compared to the control group (8.5 percent). The contract conditions did not differ significantly. Weight loss was unaffected by recruitment source although it was hypothesized that self-referrals (newspaper recruitment) would perform worse because they were more likely to have participated in prior weight-loss programs and were, on average, older.

Comparability to other studies is limited because participants were also offered to attend weekly group meetings. Moreover, attendance in the weekly group meetings differed significantly across groups, with the contract group performing better. Thus, it is unclear how much of the outcome differential is due to group meeting attendance. It is most likely that the observed treatment effect would be smaller in the absence of such group meetings.

4.3 MIXED INCENTIVES

In a relatively recent study, Volpp et al. (2008) test the effectiveness of financial incentives for weight loss among patients at a veterans' medical center. Consequently, the sample is mostly male, which limits the generalizability of the results. However, it remains interesting because of its atypical incentive scheme.

Volpp et al. use an interesting mixture of negative and positive incentives. They assigned 57 participants equally to three experimental groups, all receiving the same basic weight-monitoring program involving monthly weigh-ins, and with one serving as the control group. In the first treatment group, participants had to put in a stake from their own money, which was forfeited if their weight-loss goal of 1 lb. a week for 16 weeks was not reached. In contrast, if they reached their weight-loss goal, they could earn extra rewards up to USD 252 per month depending on their deposits. In another

treatment group, participants who reached their weight-loss goal qualified for a lottery with an expected bonus of around USD 90 per month.

The analysis is based on an intent-to-treat approach, i.e., dropouts were assumed to have returned to their baseline weight. In the control group, all participants completed the trial. In the deposit and lottery groups, 10 percent and 15 percent dropped out, respectively. Members of both treatment groups were found to have lost significantly more weight (4 lbs.) than the control participants and to have a higher likelihood of achieving the weight-loss goal. There is no significant difference in outcomes between both treatment groups. As a follow-up, Volpp et al. weighed participants again seven months after the start of the experiment. Dropout rates increased even further across the board. The results did not yield any statistical difference between the three groups.

The limitation of this study is that the observed effects may be either due to the financial incentive or the enhanced attention of the medical center for intervention-group members. Although the authors suggest that the weight-monitoring program is the same for all groups, it becomes clear later on that only the incentive participants had daily call-ins and feedback. Moreover, only members of this group had to weigh themselves each day, thus keeping better track of the weight trajectory. If monitoring per se induces participants to lose weight, it is impossible to isolate the effect of the financial incentive itself. Therefore, the effects are most likely overestimated. Additionally, assuming that the dropouts returned to their baseline weights also contributes to an overestimation of the treatment effect.

A similar experiment to Volpp et al. (2008)—in the sense that a mixture of positive and negative incentives was in play—was carried out by John et al. (2011). This is based on voluntary participation in a deposit–reward scheme. Participants were allowed to choose individual deposits with values up to USD 3 per day. However, participants could receive awards that exceed their deposits.

Sixty-six obese veterans confined at the Philadelphia Veterans Affairs Medical Center, aged between 30 and 70, took part in the experiment. All of them participated in a course where strategies for weight loss were discussed and a uniform weight-loss target of 24 lbs. in 24 weeks was set. After the treatment period, the authors initiated an eight-week maintenance phase, wherein participants were not required to lose further weight, and thereafter a 36-week follow-up phase.

Participants were randomly allocated (stratified by sex and age) to (i) a control group whose members participated in a program but was not exposed to any financial incentive (Control), (ii) a incentive treatment group for which the eight weeks of the maintenance phase were explicitly framed as a period for weight-loss maintenance

(Treatment Group 1), and (iii) an alternative incentive treatment group for which the entire period of 32 weeks was framed as a weight-loss period, i.e., the maintenance phase was not explicitly declared as such (Treatment Group 2).

The financial incentive scheme was designed as follows: for each day that participants of Treatment Groups 1 and 2 reported a weight equal to or below their daily weight-loss goal, they received twice the daily deposit (if they passed the verification of weight at the end of each month). “Thus, these participants could earn USD 84 net (USD 185 gross) per month (i.e., by making the maximum USD 3 daily deposit, and on every day of the month, truthfully reporting that they had attained their daily weight-loss goal).” In contrast, if they did not achieve the daily weight-loss goal, the deposit is forfeited. After each month, the target trajectory of weight loss was adjusted such that the final goal of 24 lbs. of weight loss could be achieved by constant daily weight loss during the rest of the treatment period.

Two outcome variables were considered: (i) weight loss after 32 weeks and (ii) weight maintenance 36 weeks after the end of the intervention period. The attrition rate was 10 percent at 32 weeks and 35 percent at the end of the intervention period. It did not significantly differ across treatment states at both weigh-ins. The analysis was carried out as an intention-to-treat analysis.

The two incentive groups did not exhibit significant deviations in weight loss after 32 weeks, most probably indicating that there is simply no framing effect. However, taking the two groups together, average weight loss for the combined group far exceeds the corresponding value for the control group. During the follow-up period, substantial weight regain was found particularly for the treatment groups although the latter still achieved an average net weight loss of 1.2 lbs. between enrollment in the study and the long-term follow-up. In contrast, the average net weight loss of the control group was 0.3 lbs. Nevertheless, for the entire period of 68 weeks (32 plus 36 weeks), average weight loss did not exhibit a significant treatment–control differential.

The experiment yields two clear-cut results: (i) financial incentives work for achieving medium-term reductions in body weight but (ii) weight loss that has been achieved by the means of financial incentives is not sustainable after incentives are terminated.

Since all participants other than those from the control group were exposed to a mixture of positive and negative financial incentives, the question of their differential efficacy cannot be adequately addressed. The assumption of dropouts returning to their baseline weight will likely lead to an overestimation of the treatment. Crucially, only the incentive participants had to weigh themselves daily and call in their weight for immediate feedback, likely contributing to the overestimation of the impact.

4-4 RELEVANT STUDIES THAT DID NOT MEET THE INCLUSION CRITERIA

There are three studies that we think deserve to be mentioned although they do not fit the criteria to be included in the detailed review. One is an observational study. The others investigate the importance of the amount of the bonus and the power of financial incentives to encourage participants to maintain their weight after having lost a significant amount of it within a randomized controlled design, respectively.

Kramer et al. (1986) randomly assigned 85 participants of a 15-week weight-loss program that had been conducted earlier into three experimental groups (two treatment groups and one control group) using stratified randomization by sex. These participants lost 10 percent or more of their body weight in the previous weight-loss program.

The authors used a negative incentive scheme. In both treatment groups, participants paid a deposit of USD 120. In one group ($N = 29$), they were refunded USD 10 each month for one year if they attended a special skills training aimed at improving knowledge on eating and exercises which support weight maintenance. In the other group ($N = 28$), they were refunded the same amount if they weighed not more than the week before and attended 12 monthly meetings. These meetings followed a nonspecific problem-solving format where participants initiated discussions about maintenance progress and problems. The control group received no contact except a six-month reminder letter and the scheduling contact for the one-year follow-up.

Twenty percent of the weight-contract participants refused to attend the final weigh-in at 12 months after treatment initiation but provided self-reported weight information. The authors added 5 lbs. to the self-reported weights to adjust for possible reporting bias. The other groups did not have such refusals and the difference in refusal of the follow-up clinical visit was statistically significant.

Kramer et al. find that all groups regained considerable amounts of weight averaging at about 40 percent of the original weight loss. The weight gains were not statistically different between each of the groups. However, the variance of weight loss was larger in the control group compared to the treatment, pointing at the latter being superior in case of risk-averse subjects. The authors further document a twice-as-big maintenance success rate in the weight-contract group than in the control and the other treatment (attendance only) group. However, due to the low number of observations, both differences are statistically insignificant.

A critical limitation of the study is that the control group did not attend monthly meetings. Since the incentive participants received not only the financial-incentive treatment but also interacted with the other incentive participants during

these meetings, the effect of the deposit cannot be singled out. Thus, it remains unclear whether the discussions about maintenance progress and problems counteracted the positive effect of the deposit contract for weight maintenance. Likewise, it may have helped participants to maintain lost weight.

The study of Jeffery et al. (1983) was not included because it does not have a “genuine” control group. They have six treatment groups that differ with respect to two dimensions: (i) group vs. individual contract and (ii) USD 30 vs. USD 150 vs. USD 300 deposit with refunds contingent on contractual target-weight achievement. The innovative design is that in the group contract interventions, participants were refunded conditional on the performance of the whole group.

The authors observe significantly greater weight losses for the group contracts during the treatment phase and at follow-up. In contrast, the amount of deposit was not significantly related to weight loss. Although in terms of percentage of participants achieving the contracted goal, groups with higher deposits performed significantly better at the end of the treatment phase. This effect did not persist at the three-month and 12-month follow-ups.

Although this study also faces serious limitations such as, for instance, high attrition rates, it confirms previous results of higher financial incentives not translating proportionally into weight loss success (we did not observe that studies with larger rewards report a higher effectiveness of their respective incentive groups) and that other control mechanisms play an important role. Here, the group contract and in Saccone and Israel (1978) the assignment of a role to the participant’s spouse or a significant other matters for the effectiveness of the financial incentives.

To investigate enrollment, attrition and weight loss, Cawley and Price (2011) analyze data from a real-world intervention. The data originate from one company that operates a weight loss program in 17 worksites. They comprise 2,407 employees who signed up for participation. The program included email and call-center support as well as quarterly weigh-ins. The worksites could choose one out of three program features that varied according to incentive schedules to lose weight. As a consequence, the employee’s enrollment options were limited to the one worksite-specific program feature.

While program Feature 1 offered quarterly rewards for weight loss, the other consisted of a deposit contract scheme (Feature 2) where participants committed themselves to pay a monthly deposit which they redeemed at the end of the year contingent on the achievement of a certain weight-loss goal, and a basic feature (Control) that did not schedule any financial incentive. Cawley and Price distinguish three treatment groups: participants of Feature 1, Feature 2, and Control. Overall, about

20 percent of the employees participated in the program.¹⁶ They find the smallest enrollment rate for Treatment Group 2 but the difference from the other groups is not significant. The highest attrition rate is observed for participants of Treatment Group 1 (76.4 percent). From Treatment Group 2, only 42.6 percent completed the program although significantly more than in Treatment Group 1. Compared to the control group, both treatment groups had a significantly higher dropout rate.

Assuming that the dropouts returned to their baseline weights, Treatment Group 2 had lost more weight than the other groups. However, only the difference from the control group in the fourth quarter is significant.

Their results with respect to the effectiveness of the financial incentive schemes have to be interpreted very cautiously because individuals were not randomly assigned to different incentive schedules. The authors make clear that the observed effects may be either due to the financial incentive or the self-selection into the weight-loss program. This means that the estimated effects may reflect the selection of worksites into specific incentive types whose attractiveness and, in turn, enrollment shares most likely depend on the employee's unobserved individual characteristics.

Nevertheless, this study is still better designed for identifying effects of financial rewards on weight loss than other studies analyzing real world interventions. As an example, in analyzing the "pounds for pounds" pilot program carried out in the NHS in southeast England, Relton et al. (2011) compare groups of participants where group membership is entirely endogenous. The endogeneity stems either from individuals being allowed to choose the characteristics of their individual weight-loss plans or from using dropouts as the comparison group.

But what can we learn from their study? That program attrition is not only a serious problem in the trials but also for their real world implementation and that the program design has a huge impact on that. Deposit contracts seem to reduce this problem, although it comes with lower enrollment rates, which were, overall, unexpectedly high.

5 CONCLUSION

The aim of this review was to gather the evidence in the literature on the effectiveness of financial incentives in promoting weight loss. To do that, we examined results from randomized controlled trials obtained from a systematic literature review. What is new in our work is that we focus on obesity treatment (in contrast to Jochelson (2007), who

¹⁶ Employees with a BMI less than 25 kg/m² might have enrolled but would not receive any payout.

surveys the effectiveness of financial incentives in manipulating health behavior in general) and that we more carefully distinguished what was incentivized and how it was incentivized (compared to Paul-Ebhohimhen and Avenell (2008)). We also adhered to a specific hierarchy of evidence to evaluate the credibility of the reported effects in the literature.

In general, the body of evidence does not point in only one direction. The issue of financial incentives and weight loss is complicated because the extant studies do not answer the same question. While some examine positive incentives, others turn their attention to negative ones—still, some combine both. Many researchers incentivize healthy behavior while others directly tie the incentive to the output measure. Table 3 therefore serves as a useful lens through which these studies can be examined.

With respect to the evaluation of output-oriented incentive schemes, the results are precocious. The majority of studies report a positive effect of monetary incentives on weight performance. Indeed, the most credible studies (i.e., Jeffery (1984a), Volpp et al. (2008), and John et al. (2011)) support this view. However, considerable doubt remains with respect to the sustainability of any weight loss achieved through financial incentives.

Incentives for the maintenance of weight are a generally under-researched field. Kramer et al. (1986) explicitly analyze weight maintenance and provide evidence that people regain up to 40 percent of the weight they have initially lost and that negative financial incentives do not seem to prevent this phenomenon. The experimental design, however, is far from ideal. Therefore, further research in that field is necessary to settle the issue.

Despite the puzzles mentioned above, it seems implausible to dismiss financial incentives as a means to reduce overweight or obesity altogether. Rather than that, the evidence suggests that there is some effect but that the literature has not fully detected how it comes about and on what it depends. Ultimately, the evidence on the effectiveness of financial incentives for weight loss has not been convincingly settled for one side or the other. This is mostly due to the lack of methodological rigor and the conservative sample sizes.

Future research on the topic should be more careful to clearly define what behavior is incentivized (input vs. output orientation, weight loss vs. maintenance) and how it is incentivized (positive vs. negative incentives). A large-scale randomized experiment is necessary to test hypotheses about these potential effects. Finally, to facilitate comparisons in the future, we advocate the use of standardized reporting guidelines.

REFERENCES

- Ashenfelter, Orly (1978). "Estimating the effect of training programs on earnings", *The Review of Economics and Statistics* 60 (1): 47-57.
- Austin, S. Bryn & Gortmaker, Steven L. (2001). "Dieting and smoking initiation in early adolescent girls and boys: a prospective study", *American Journal of Public Health* 91: 446-450.
- Bhattacharya, J. and Neeraj Sood (2005). "Health insurance and the obesity externality", *Advances in Health Economics and Health Services Research* 17(6): 279-318.
- Bhattacharya, Jay and Neeraj Sood (2011). "Who pays for obesity?", *Journal of Economic Perspectives* 25(1): 139-158.
- Bish, Connie L., Blanck, Heidi Michels, Serdula, Mary K., Marcus, Michele, Kohl, Harold W. and Kettel Khan, Laura (2005). "Diet and physical activity behaviors among Americans trying to lose weight: 2000 behavioral risk factor surveillance system", *Obesity Research* 13(3): 596-607.
- Black, David R. and James G. Friesen (1983). "Deposit money: a component in a self-directed minimal intervention program for weight control", *Behavior Therapy* 14(2): 333-340.
- Cameron, Roy, M.A. MacDonald, R.P. Schlegel, C.I. Young, S.E. Fisher, J.D. Killen, T. Rogers, L. Horlick and L.F. Shepel (1990). "Toward the development of self-help health behaviour change programs: weight loss by correspondence", *Canadian Journal of Public Health* 81(4): 275-279.
- Cawley, John and Joshua A. Price (2011). "Outcomes in a program that offers financial rewards for weight loss", in M. Grossmann and N. H. Mocan (eds.), *Economic Aspects of Obesity*, 1st ed., Chapter 4, pp. 91-126. Chicago, IL: University of Chicago Press.
- Centre for Evidence-based Medicine (2009). Levels of Evidence. URL: <http://www.cebm.net/index.aspx?o=1025>. Accessed 25 July 2011.
- Cutler, David M., Edward L. Glaeser, and Jesse M. Shapiro (2003). "Why have Americans become more obese?", *Journal of Economic Perspectives* 17 (3): 93-118.
- Des Jarlais, Don C., Cynthia Lyles, Nicole Crepaz, and the TREND Group (2004). "Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement", *American Journal of Public Health* 94 (3): 361-366.
- Finkelstein, Eric, Laura Linnan, Deborah Tate, and Ben 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): 981-989.
- Food and Drug Administration (2009). Guidance for industry: evidence-based review system for the scientific evaluation of health claims. URL: <http://goo.gl/vd66H>. Accessed July 26, 2011.

- Harris, Mary B. and Carol G. Bruner (1971). "A comparison of a self-control and a contract procedure for weight control", *Behaviour Research and Therapy*, 9 (4): 347-354.
- Imbens, Guido W. and Jeffrey M. Wooldridge (2009). "Recent developments in the econometrics of program evaluation", *Journal of Economic Literature* 47 (1): 5-86.
- Ikeda, Shinsuke, Myong-Il Kang, and Fumio Ohtake (2010). "Hyperbolic discounting, the sign effect, and the body mass index", *Journal of Health Economics* 29 (2): 268-284.
- Israel, Allen C. and Anthony J. Saccone (1979). "Follow-up of effects of choice of mediator and target of reinforcement on weight loss", *Behavior Therapy* 10 (2): 260-265.
- Jeffery Robert W., Jean L. Forster, Simone A. French, Steven H. Kelder, Harry A. Lando, Paul G. McGovern, David R. Jacobs, Judith E. Baxter (1993a). "The Healthy Worker Project: a work-site intervention for weight control and smoking cessation", *American Journal of Public Health* 83 (3): 395-401.
- Jeffery Robert W., Rena R. Wing, Carolyn Thorson, Lisa R. Burton, Cheryl Raether, Jean Harvey, Monica Mullen (1993b). "Strengthening behavioral interventions for weight loss: a randomized trial of food provision and monetary incentives", *Journal of Consulting and Clinical Psychology* 61 (6): 1038-1045.
- Jeffery, Robert W. and Simone French (1999). "Preventing weight gain in adults: the pound of prevention study", *American Journal of Public Health* 89 (5): 747-751.
- Jeffery, Robert W., Paul D. Thompson, Rena R. Wing (1978). "Effects on weight reduction of strong monetary contracts for calorie restriction or weight loss", *Behaviour Research and Therapy* 16 (5): 363-369.
- Jeffery, Robert W., Rena R. Wing, Carolyn Thorson, Lisa R. Burton (1998). "Use of personal trainers and financial incentives to increase exercise in a behavioral weight-loss program", *Journal of Consulting and Clinical Psychology* 66 (5): 777-783.
- Jeffery, Robert W., Wendy M. Bjornson-Benson, Barbara S. Rosenthal, Candace L. Kurth, Mary M. Dunn (1984a). "Effectiveness of Monetary Contracts with two repayment schedules on weight reduction in men and women from self-referred and population samples", *Behavior Therapy* 15 (3): 273-279.
- Jeffery, Robert W., Wendy M. Bjornson-Benson, Barbara S. Rosenthal, Ruth A. Lindquist, Sharon L. Johnson (1984b). "Behavioral treatment of obesity with monetary contracting: two-year follow-up", *Addictive Behaviors* 9 (3): 311-313.
- Jeffery, Robert W., Wendy M. Gerber, Barbara S. Rosenthal, Ruth 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): 242-248.
- Jochelson, Karen (2007). *Paying the patient: improving health using financial incentives*. King's Fund: London.

- John, Leslie, George Loewenstein, Andrea Troxel, Laurie Norton, Jennifer Fassbender, Kevin Volpp (2011). "Financial incentives for extended weight loss: a randomized, controlled trial", *Journal of General Internal Medicine* 26 (6): 621–626.
- Konnopka, Alexander, Melanie Bödemann, and Hans-Helmut König (2010). "Health burden and costs of obesity and overweight in Germany", *European Journal of Health Economics* 12 (4): 345–352.
- Kramer, F. Matthew, Jeffery, Robert W., Snell, Mary Kay and Forster, Jean L. (1986). "Maintenance of successful weight loss over 1 year: effects of financial contracts for weight maintenance or participation in skills training", *Behavior Therapy* 17 (3): 295–301.
- Mahoney, Michael J. (1974). "Self-reward and self-monitoring techniques for weight control", *Behavior Therapy* 5(1):48–57.
- Mavis, Brian E. and Bertram E. Stoffelmayr (1994). "Multidimensional Evaluation of Monetary Incentive Strategies for Weight Control", *The Psychological Record*, 44: 239–252.
- McCormick, Barry, I. Stone, and Corporate Analytical Team (2007). "Economic costs of obesity and the case for government intervention", *Obesity Reviews* 8 (Suppl. 1): 161–164.
- Mokdad, Ali H., Earl S. Ford, Barbara A. Bowman, William H. Dietz, Frank Vinicor, Virginia S. Bales, James S. Marks (2003). "Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001", *Journal of the American Medical Association* 289 (1): 76–79.
- OECD (2010). "OECD Health Data: Selected Data", OECD Health Statistics (Database). doi:10.1787/data-00348-en. Accessed on 17 May 2011.
- Paul-Ebhohimhen, Virginia A. and Alison Avenell (2008). "Systematic review of the use of financial incentives in treatments for obesity and overweight", *Obesity Reviews* 9 (4):355–367.
- Prentice, Andrew M. (2006). "The emerging epidemic of obesity in developing countries", *International Journal of Epidemiology* 35 (1): 93–99.
- Relton, Clare, Mark Strong and Jessica Li (2011). "The 'Pound for Pound' weight loss financial incentive scheme: an evaluation of a pilot in NHS Eastern and Coastal Kent", *Journal of Public Health*, published online 29 March 2011.
- Saccone, Anthony J. and Allen C. Israel (1978). "Effects of experimenter versus significant other-controlled reinforcement and choice of target behavior on weight loss", *Behavior Therapy* 9 (2): 271–278.
- Sassi, Franco (2010), *Obesity and the economics of prevention: fit not fat*. Paris: OECD Publishing.
- Schulz, Kenneth F., Douglas G. Altman, David Moher, and the CONSORT Group (2010). "CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials", *Journal of Clinical Epidemiology* 63 (8): 834–840.

- Strotz, Robert H. (1955-1956). "Myopia and inconsistency in dynamic utility maximization", *The Review of Economic Studies* 23 (3): 165-180.
- Tversky, Amos and Daniel Kahneman (1991). "Loss aversion in riskless choice: a reference dependent model", *Quarterly Journal of Economics* 106 (4): 1039-1061.
- US Preventive Services Task Force (1989). Guide to Clinical Preventive Services. DIANE Publishing.
- Volpp, Kevin, Leslie John, Andrea Troxel, Laurie Norton, Jennifer Fassbender, George Loewenstein (2008). "Financial incentive-based approaches for weight loss", *Journal of the American Medical Association* 300 (22): 2631-2637.
- WHO (2011). "Obesity and overweight", Fact Sheet No. 311 (March). <http://goo.gl/jgMzk>. Accessed 17 May 2011.
- Wing, Rena R., Leonard H. Epstein, Marsha Marcus, Barbara Shapira (1981). "Strong monetary contingencies for weight loss during treatment and maintenance", *Behavior Therapy* 12 (5): 702-710.

TABLE A1
Extensive overview of included and excluded studies

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
1	Black and Friesen (1983)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² not reported, mean participant above 30 percent over-weight on international standards, only women (recruited through newspaper advertisement)</p> <p>Interventions: group meetings and financial incentives for attending the meetings</p> <p>Duration: 4 month</p> <p>Groups by assignment:</p> <p>a) n=11, deposit \$10 for attendance of weigh-ins</p> <p>b) n=13, control group (meetings only).</p>	<p>Follow-up: 1 and 4 month</p> <p>Outcome:</p> <p>1. Weight</p> <p>2. Attrition</p>	<p>Type of incentive: negative</p> <p>Coupling of incentive: input-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>(i) Significantly greater weight loss at four month weigh-in in incentive group, but not at one month weigh-in</p> <p>(ii) No significant differences in attrition rates</p> <p>Limitations (excluded from survey): Incentive is not directly tied to weight loss, effect of interest cannot be identified (e.i.n.i.)</p>
2	Cameron et al. (1990)	<p>Country: Canada</p> <p>Participants: BMI (mean) kg/m² of 31.7 (males) and 41.7 (females)</p> <p>Interventions: lessons on diet, self-study exercise and behavior, homework, weekly weightings and financial interventions for different activities</p> <p>Duration: 15 weeks</p> <p>Groups by assignment:</p> <p>a) n=12, received L (mailed 15 printed lessons on diet, exercise, and behavior; \$ 21 refund for attendance at 1 year follow-up) during Follow-Up</p> <p>b) n=12, L only</p> <p>c) n=12, L+ W (weigh ins and face to face consultation as needed)</p> <p>d) n=12, L+ H (homework of one page record of changes made and weight, telephone consultation as needed)</p> <p>e) n= 12, L+ W + D (\$21 deposit refund for homework and attendance)</p> <p>f) n=10, L+H+D</p> <p>g) n=12, L+ H+ W</p> <p>h) n=9, L+ H+ W+ D.</p>	<p>Follow-up: 16 months</p> <p>Outcome:</p> <p>1. Weight (and BMI)</p>	<p>Type of incentive: negative</p> <p>Coupling of incentive: input-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>BMI reduction significantly higher than in group a) for all groups except for group b), (ii) no significant differences in BMI between treatment groups at follow-up</p> <p>Limitations (excluded from survey): e.i.n.i.</p>

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
3	Finkelstein et al. (2007)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² of 33 Employees from university and colleges with BMI >25 kg/m²</p> <p>Intervention: financial incentives per percentage point of weight lost and information on healthy weight loss at baseline</p> <p>Duration: 6 month (3 months phase 1, 3 months phase 2).</p> <p>Groups by assignment:</p> <p>a) n=72, \$ 0 after the first 3 months and 14 \$ after 6 months (back-loaded group)</p> <p>b) n=64, \$ 14 after the first 3 months and 0\$ after 6 months (front-loaded group)</p> <p>c) n=71, \$ 7 in both phases (steady-treatment group)</p>	<p>Follow-up: 3 and 6 month</p> <p>Outcome:</p> <p>1. Weight</p>	<p>Type of incentive: positive</p> <p>Coupling of the incentive: output-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>(i) Significant difference between control and incentive groups within first three months</p> <p>(ii) larger financial incentive yields higher weight-loss</p> <p>(iii) no significant difference between the groups after six months (due to equalized potential financial gains)</p> <p>Limitations:</p> <p>High drop-out rate.¹</p> <p>Back-loaded group (control group) has incentive to lose weight in first three months (underestimation of treatment effect)</p>
4	Harris and Bruner (1971)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² not reported, mean weight of 165 lbs (Experiment 1) and 144 lbs (Experiment 2)</p> <p>Intervention: financial incentives in dependence to success in weight loss and educational program</p> <p>Duration: 12 weeks (first experiment) or 16 weeks (second experiment)</p> <p>Groups by assignment:</p> <p>First experiment (n=32)</p> <p>a) n= 12, self-chosen deposit in the amount of \$ 0,5 to 1/intended weight loss</p> <p>b) n=12, educational program on behavior</p> <p>c) n=12, control group, dropped out entirely.</p> <p>Second experiment (n=18)</p> <p>a) n=8, educational program on behavior</p> <p>b) n=7, control group</p> <p>c) n=6 deposit group, declined to participate entirely and was reassigned to the other two groups</p>	<p>Follow-up: 10 month</p> <p>Outcome:</p> <p>1. Weight</p>	<p>Type of incentive: negative</p> <p>Coupling of the incentive: output-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>Experiment 1: (i) Significant weight loss in both groups; (ii) Significantly greater weight loss in the incentive group than in the one with the educational program.</p> <p>Experiment 2: No significant differences in weight loss</p> <p>Limitations (excluded from survey): Incentive participants can set own weight loss goals as well as the amount of deposit per pound lost. Incentive effect confounded by educational programs (incentive group does not obtain such a program, comparisons with control group can only be made across experiments)</p>

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
5	Jeffery et al. (1998)	<p>Country: USA</p> <p>Participants: BMI kg/m² of between 30.6-31.5</p> <p>Intervention: two strategies for improving exercise adherence and long-term weight loss in obese outpatients and financial incentives for attendance</p> <p>Duration: 24 weeks</p> <p>Groups by assignment:</p> <ul style="list-style-type: none"> a) n= 40, with diet and exercise, calorie intake diaries and weigh-ins. Progress reviews and group discussions (S) b) n= 41, S+ supervised thrice weekly walk sessions (E) c) n= 42, S+ E+ personal trainer to remind and walk participants (T) d) n= 37, S+ E+ financial incentive for attendance at exercise sessions from \$1 to \$3 per walk session (I) e) n=36, S+ E+ T+ I. 	<p>Follow-up: 18 month</p> <p>Outcome:</p> <ol style="list-style-type: none"> 1. Weight 2. Psychological status 3. Caloric intake from ft 4. Total caloric intake 5. Psychical activity levels. 	<p>Type of incentive: positive</p> <p>Coupling of the incentive: oriented temporal structure of the incentive: weight loss</p>	<p>(i) Significantly greater weight loss in group a) (S only) than in other groups, no significant differences in weight loss among other groups.</p> <p>(ii) Significantly larger increases in exercise in groups c) (S+E+T), d) (S+E+I) and e) (S+E+T+I) than in other groups.</p> <p>Limitations (excluded from survey): e.i.n.i.</p>
6	Jeffery et al. (1983)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² of 31.8</p> <p>Interventions: Behaviorally oriented program and financial incentive in dependence to success in weight loss (30 lbs.)</p> <p>Duration: 15 weeks</p> <p>Groups by assignment:</p> <p>Groups varied in amount of deposit and type of contract (refunds contingent on either individual or mean group performance)</p> <ul style="list-style-type: none"> a) n=16, individual/ \$30 deposit b) n=15, individual/ \$150 deposit c) n=14, individual/ \$300 deposit d) n=14, group/ \$30 deposit e) n=14, group/ \$150 deposit f) n=14, group/ \$300 deposit 	<p>Follow-up: 3, 12 and 24 month</p> <p>Outcome:</p> <ol style="list-style-type: none"> 1. Weight 2. % at least achieving contracted goal 	<p>Type of incentive: negative</p> <p>Coupling of the incentive: output-oriented</p> <p>temporal structure of the incentive: weight loss and weight maintenance</p>	<p>(i) Significantly greater weight losses under group contracts during treatment phase and follow-up, (ii) no significant differences in the amount of deposit, (iii) percentage of participants achieving contracted goal significantly larger for higher deposits at the end of treatment, (iv) size of deposit exhibited no significant difference at 3 and 12 months follow-up</p> <p>Limitations (excluded from survey): No comparison to control group (non-incentive), i.e., genuine effect of incentives cannot be identified.</p>

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
7	Jeffery et al. (1984a)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² of 33 (males) and 31 (females)</p> <p>Interventions:</p> <p>Financial incentive (\$ 150 deposit, refunds for weight loss or commitment), instructional program on diet, exercise and behavior; weekly group meetings; individual weight loss objectives</p> <p>Duration: 16 weeks</p> <p>Groups assignment:</p> <p>Groups varied in source: Self-referred through newspapers (SR) and existing trial pool (PR)</p> <p>a) n=19, SR; complete refund at first visit</p> <p>b) n=17, SR; fixed weekly \$30 refunds</p> <p>c) n=21, SR; increasing weekly refunds from \$5 to \$75</p> <p>d) n=21, PS; complete refund at first visit</p> <p>e) n=18, PS; fixed weekly \$30 refunds</p> <p>f) n=17, PS; increasing weekly refunds from \$5 to \$75.</p>	<p>Follow-up: 12 month</p> <p>Outcome:</p> <ol style="list-style-type: none"> 1. Weight 2. % of body weight lost 	<p>Type of incentive: negative Coupling of the incentive: output- and input-oriented temporal structure of the incentive: weight loss and weight maintenance</p>	<p>(i) Significantly greater weight loss in incentive groups in terms of percent of body weight lost; (ii) no significant difference between incentive groups; (iii) no significant difference between recruitment sources</p> <p>Follow-up: (iv) No significant differences in weight gain between contract types or recruitment sources (v) contract differentials in overall weight loss maintained better in PS groups</p> <p>Limitations:</p> <p>Follow-up results confounded by maintenance treatment</p>
8	Jeffery et al. (1984b)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² not reported</p> <p>Intervention: Behaviorally oriented program, (diet and exercise, weekly group meetings, calorie and exercise records) and financial incentive in dependence to success in weight loss (30 lbs.)</p> <p>Duration: 15 weeks</p> <p>Groups assignment:</p> <p>Groups varied in amount of deposit and type of contract (refunds contingent on either individual or mean group performance)</p> <p>g) n=16, individual/ \$30 deposit</p> <p>h) n=15, individual/ \$150 deposit</p> <p>i) n=14, individual/ \$300 deposit</p> <p>j) n=14, group/ \$30 deposit</p> <p>k) n=14, group/ \$150 deposit</p> <p>l) n=14, group/ \$300 deposit</p>	<p>Follow-up: 24, month</p> <p>Outcome:</p> <ol style="list-style-type: none"> 1. Weight 2. % at least achieving contracted goal 	<p>Type of incentive: negative Coupling of the incentive: output-oriented temporal structure of the incentive: weight loss and weight maintenance</p>	<p>Results at 24, month follow-up: (i) no significant differences under different contract sizes, (ii) net weight loss significantly larger under group contracts than under individual contracts</p> <p>Limitations:</p> <p>See Jeffery et al. (1983)</p>

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
9	Jeffery et al. (1978)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² not reported, mean weight of 232 lbs.</p> <p>Interventions: Attending courses and financial incentives (\$ 200 deposit, which were returned contingent on either attendance, calorie restriction or weight loss of 2 lbs. per week)</p> <p>Duration: 11 weeks</p> <p>Groups by assignment:</p> <p>n=31</p> <p>a) n=7, weight loss group</p> <p>b) n=10, calorie contract group</p> <p>c) n=7, attendance contract group</p> <p>d) n=7, Control group (4 drop outs).</p>	<p>Follow-up: 4 month</p> <p>Outcome:</p> <p>1. Weight.</p>	<p>Type of incentive: negative</p> <p>Coupling of the incentive: output- and input-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>(i) weight loss significantly higher in weight contract and calorie contract group than in attendance contract group, (ii) higher than in no contract group</p> <p>Limitations (excluded from survey): control group formed by self-selection and was reduced to three participants due to drop-outs.</p>
10	Jeffery et al. (1993a)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² of 25</p> <p>Employees of 32 work sites.</p> <p>Interventions: Treatment was randomly assigned on the work site level. It consisted of health education classes combined with a payroll-based incentive system and financial incentives upon reaching the self-imposed success in weight loss. The amount of paycheck deduction was also self-selected (min \$5 biweekly). Within work site employees could self-select into programs.</p> <p>Duration: 2 years</p> <p>Groups by assignment:</p> <p>a) n=16, control works sites</p> <p>b) n=16, works sites with smoking cessation and weight-loss program: financial incentives per week (self-selected, minimum \$5) for achieving self-selected weight-loss goal.</p>	<p>Follow-up: none</p> <p>Outcomes:</p> <p>1. Weight (BMI)</p> <p>2. Smoking cessation</p>	<p>Type of incentive: negative</p> <p>Coupling of the incentive: output-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>(i) low participation rates in smoking cessation program and weight loss programs</p> <p>(ii) significant decrease in smoking prevalence</p> <p>(iii) no significant difference between control and treatment sites in terms of weight loss within two years.</p> <p>Limitations (excluded from survey): Effect of financial incentive confounded by weight loss courses (control group does not receive weight loss courses)</p>

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
11	Jeffery et al. (1993b)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² between 30.8-31.1</p> <p>Interventions: Treatment consisted of health education classes combined with a payroll-based financial incentive system upon reaching a self-imposed weight loss goal.</p> <p>Duration: 18 months</p> <p>Groups by assignment:</p> <ul style="list-style-type: none"> a) n=40, no treatment control b) n=60, standard behavioral treatment (STB) of weekly meetings and weigh ins within 20 weeks, then monthly meetings and weekly weigh ins till 18 months. Self-imposed program weight loss goal (14, 18, or 23 kg) fractioned in weekly weight loss goals. c) n= 40; STB and food provision (FP) d) n=44; STB and financial incentives (FI) for weight from \$2.50 for not gaining weight within one week, \$12.50 for achieving 50 percent of the weekly self-imposed weight-loss goal, and \$25 for full achievement. e) n= 44; STB and FB and FI. 	<p>Follow-up: 6/ 12/18 month</p> <p>Outcomes:</p> <ul style="list-style-type: none"> 1. Weight (BMI) 2. Perceived barriers to abstinence 3. Caloric intake from fat 4. Total caloric intake 5. Physical activity levels 	<p>Type of incentive: negative</p> <p>Coupling of the incentive: output-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>(i) economically insignificant weight reductions in incentive group at 6 months</p> <p>(ii) no weight loss at 18 months</p> <p>(iii) statistically and economically significant effect of food provision on weight loss at all stages</p> <p>Limitations:</p> <p>Incentive participants can set own weight loss goals.</p> <p>Very large attrition rate: missing values inputted by adjacent values (effect of financial effect most likely underestimated)</p>
12	Jeffery and French (1999)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² of 26.5</p> <p>Initially recruited: 228 men and 596 women through phone solicitation, newspaper advertisements and mailings to University of Minnesota employees + 404 women through telephone solicitations and information booths in low-income neighborhoods and face-to-face recruitment in public Nutrition Programs to account for the specific obesity risks of persons with low socio-economic status.</p> <p>Interventions: Training activities and financial incentives (Raffle of \$ 200 in TN, who had sent the reply postcard)</p> <p>Duration: 3 years</p> <p>Groups by assignment:</p> <ul style="list-style-type: none"> a) n=414, no-contact control b) n=197, education through monthly newsletters c) n=198, education plus incentives for participation 	<p>Follow-up: 12/18/30 month</p> <p>Outcomes:</p> <ul style="list-style-type: none"> 1. message recognition tests of educational program 2. Reported behavior change with regard to exercise, nutrition and diet habits 3. weight 	<p>Type of incentive: positive</p> <p>Coupling of the incentive: input-oriented</p> <p>temporal structure of the incentive: weight maintenance</p>	<p>(i) Significantly better message recognition in the treatment groups compared to control group, (ii) increase in reported weighing frequency of weighing in the treatment groups (decrease in control group) and significantly smaller decrease in reported frequency of healthy weight loss practices in treatment groups, (ii) slightly better trends in body weight in treatment groups</p>

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
13	John et al. (2011)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² between 30-40, only male veterans</p> <p>Duration: 32 weeks (incl. follow-up 68 weeks)</p> <p>Interventions: (i) Weight monitoring and consultation with dietician, (ii) choice of deposit between \$0-\$3 per day (max \$504 (24 weeks)), (iii) award of twice the individual deposit for each day the individual weight loss target was met (during weeks 1-24)</p> <p>Weight loss Target: 1lb weight loss per day, i.e. 24 lb weight loss after 24 weeks (phase 1), adjusted each month such that the final target of 24 lb can be achieved by uniform daily steps (ii) for those successful in phase 1: could choose a weight loss target of 0, 0.5, or 1 lb per day for phase 2 (weeks 25 to 32)</p> <p>Groups by assignment (Equally and randomly stratified by age and sex):</p> <p>a) n=22 course only and weight loss target</p> <p>b) n=22 course; incentive scheme: final 8 weeks framed as "maintenance of weight loss period"</p> <p>c) n=22 course; incentive scheme; final 8 weeks not explicitly framed as "maintenance of weight loss period"</p>	<p>Follow-up: 36 weeks</p> <p>Outcomes: weight loss</p>	<p>Type of incentive: mixture of positive and negative incentives; output-oriented</p>	<p>(i) highly significant effect of incentive scheme on average weight loss; (ii) no significant deviation in weight loss between the two incentive groups; (iii) no significant differential in weight loss between treatment and control after follow-up</p>
14	Kramer et al. (1986)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² not reported, mean Weight of 180 lbs, successful participants of a previous 15-week weight-loss program (had lost 10 percent of body weight)</p> <p>Interventions: weight-loss program with weekly weigh-ins and courses plus financial incentives in dependence to weight maintenance and participation in skills training sessions to solidify behavioral changes</p> <p>Duration: one year</p> <p>Groups by assignment n=85;a) n=29, refund for attending monthly courses (\$ 120 deposit, \$ 10 repayment per shared session)</p> <p>b) n=28, refund for maintaining weight and attending monthly group session (\$ 120 deposit, \$ 10 repayment per shared session and weight maintenance)</p> <p>c) n=29, no treatment, only \$ 20 deposit for follow-up.</p>	<p>Follow-up: 1 year</p> <p>Outcomes: 1. weight</p>	<p>Type of incentive: negative Coupling of the incentive: mixture of output-oriented temporal structure of the incentive: weight maintenance</p>	<p>(i) forty percent regained weight in total</p> <p>(ii) no significant difference in weight regain among groups</p> <p>(iii) larger maintenance rates in the mixed incentive group</p> <p>(iv) variance of weight regain was larger in control group</p> <p>Limitations: Effect of financial incentive confounded by problem-solving group meetings (control group does not attend any meeting during intervention)</p>

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
15	Mahoney (1974)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² not reported, participants at least 20 percent overweight on international standards</p> <p>Interventions:</p> <p>Financial incentives (participants awarded themselves portions of their own deposit for attainment of either their weight loss or their habit improvement goals) \$ 35 deposit and \$5 for absences) and S= self-control program with weekly weight ins and group meetings, and given pamphlets on diary behavior control</p> <p>Duration: 8 weeks</p> <p>Groups by assignment:</p> <p>a) n=11, delayed treatment control</p> <p>b) n=13, S+ weekly self reward for weight loss</p> <p>c) n=11, S+ weekly self-reward for habit improvement</p> <p>d) n=14, S only.</p>	<p>Follow-up: 12 month</p> <p>Outcomes:</p> <p>1. % subjects maintaining weight loss</p>	<p>Type of incentive: negative</p> <p>Coupling of the incentive: output- and input-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>(i) input oriented incentive is significantly more effective than output oriented incentive scheme</p> <p>(ii) output oriented financial incentive not effective</p> <p>(iii) no group differences in weight maintenance</p> <p>Limitations:</p> <p>Control subjects have shorter pre-program duration of weight problem.</p> <p>Statistical power far below conventional levels due to small sample size (no conclusion of effectiveness of output oriented financial incentive possible)</p>
16	Mavis and Stöfelmayr (1994)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² not reported, Mean weight of 206 lbs</p> <p>Interventions:</p> <p>Weekly meetings, procurement of helpful techniques for weight loss and different financial incentives. Every participant has an \$40 deposit and the investigator provides further \$ 40.</p> <p>Duration: 14 weeks</p> <p>Groups by assignment:</p> <p>a) n=20, continuous positive reinforcement (\$8 refund per shared session)</p> <p>b) n=20, monetary response cost (\$ 8 if the weekly aim is achieved)</p> <p>c) n=20, positive reinforcement with a lottery system (deduction of \$ 8 at the 'account' for non-attainment)</p> <p>d) n=21, response cost with a lottery system (assignment of lots for lottery upon entire amount)</p> <p>e) n=20, for comparison, this group was based on an attendance-contingent monetary reward condition (Withdrawal of lots in case of non-attainment)</p>	<p>Follow-up: not mentioned.</p> <p>Outcomes:</p> <p>1. weight</p> <p>2. attendance</p> <p>3. no. of times participant achieved weight loss goal</p>	<p>Type of incentive: positive and negative</p> <p>Coupling of the incentive: output- and input-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>(i) significant weight losses in all groups but no significant differences in weight loss between groups, (ii) significantly lower achievement rates of weight loss goals in group a), (iii) significantly higher dropout rate in group d)</p> <p>Limitations (excluded from survey):</p> <p>no comparison with non-incentive control group possible, true effect of incentives cannot be identified</p>

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
17	Paul-Ebhohimhen and Avenell (2008) Meta-Analysis	Country: USA and Canada Participants: BMI (mean) kg/m ² between 29.3-31.8	Follow-up: at least 1 year Outcomes: 1. weight	Type of incentive: positive and negative Coupling of the incentive: output- and input-oriented temporal structure of the incentive: weight loss	(i) Financial incentives need to amount to at least 1.2 % of personal disposable income to induce greater weight loss than with no incentive, if at all, (ii) no differences between positive and negative incentives, (iii) input-oriented incentives induce greater weight losses than output-oriented
18	Saccocc and Israel (1978)	Country: USA Participants: BMI (mean) kg/m ² not reported, participants 16% to 100% above recommended weight (mean 46%) Interventions: Educational program on behavior and financial incentives (\$35 deposit and refunds for (a) successfully reducing weight (b) compliance with recommendations for eating behavior) Success: (a) lb/week weight reduction (awarded by \$1.5 per lb up to maximum of \$5.00 per week) (b) compliance with eating recommendations (measured at 0-63 point scale); awarded by up to \$5.00 per week Duration: 9 weeks Groups by assignment: o) n=5; control group without any treatment a) n=6, educational program only-weight (courses and self-monitoring of weight) b) n=8, program only-behavior (courses and self-monitoring of eating behavior) c) n=7, reinforcement by therapist for weight loss (courses and refunds by therapist), d) n=7, reinforcement by therapist for eating behavior change (courses and refunds by therapist) e) n=8, reinforcement by significant other spouse for weight loss" (course und refunds trough spouse) f) n=7, reinforcement by significant other spouse for change in eating behavior (course und refunds trough spouse	Follow-up: mentioned Outcomes: 1. weight	not Type of incentive: negative Coupling of the incentive: either output- oriented or input-oriented temporal structure of the incentive: weight loss	Significant effect of incentives only found for input-oriented but not for output-oriented refunds. Effect is much stronger if refund is paid by significant other.

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
19	Israel and Saccone (1979)	Follow-Up to Saccone and Israel (1978) Return for weigh-ins in order to receive USD 35 deposit of a initial deposit. No further intervention.	Follow-up: 3 and 12 months		(i) incentive scheme did not result in weight-loss maintenance at both follow-ups, (ii) more promising long-term treatment is a monetary reinforcement for behavioral change through a family member or person in the participant's natural environment. Limitations: very small number of observations and high attrition rate
20	Volpp et al. (2008)	Country: USA Participants: BMI (mean) kg/m ² of 35 Interventions: weight monitoring program with weight loss goal of 1lb per week plus financial incentives Duration: 16 weeks Groups by assignment: a) n=19, "deposit" group: self-chosen deposit between \$0.01 and \$3 per day, payout: 2*deposit only if the monthly aim was achieved b) n=19, "lottery" group: expected value of \$3 per day for reaching the daily target weight loss- Payouts only if the monthly aim was achieved. c) n=19, control group	Follow-up: 7 month Outcomes: 1. weight	Type of incentive: negative and positive Coupling of the incentive: output-oriented temporal structure of the incentive: weight loss and maintenance	(i) both incentive groups have a higher likelihood to achieve weight loss goal (ii) both incentive groups lost significantly more weight than the control group (iii) no significant difference between incentive schemes (iv) indication for effect being not sustainable Limitations: High drop-out rate. Effect of financial incentive confounded by possible enhanced attention of the medical center for incentive group members due to daily call-ins and feedback (overestimation of incentive effect)

Table A1, continued

Nr.	Study (author and year)	Study design	Outcomes	Incentives	Results and limitations
21	Wing et al. (1981)	<p>Country: USA</p> <p>Participants: BMI (mean) kg/m² not reported, Mean weight of 201 lbs</p> <p>Interventions: weekly and monthly meetings with lessons to learn techniques for weight loss and financial incentives; \$225 deposit, \$15 of which was returned at each of eight weekly treatment meetings and seven monthly maintenance meetings; where refunds were contingent on weight loss, weight loss goals equaled a cumulative loss of 2 lbs/week since beginning of the respective phase (treatment/maintenance) weight loss goals were rescheduled after if refunds were forfeited two consecutive times.</p> <p>Duration: 9 months</p> <p>Groups by assignment</p> <p>a) n=48, deposits were repaid for weight loss during treatment and attendance during maintenance</p> <p>b) n= 20, deposits were repaid for attendance during treatment and weight loss during maintenance</p>	<p>Follow-up: 7 month</p> <p>Outcomes:</p> <p>1. weight</p>	<p>Type of incentive: negative</p> <p>Coupling of the incentive: output-oriented</p> <p>temporal structure of the incentive: weight loss</p>	<p>(i) Significant weight loss in both groups, no significant differences between groups, (ii) significantly more drop outs when weight loss contingencies were in effect</p> <p>Limitations (excluded from survey): no comparison with non-incentive control group possible, true effect of incentives cannot be identified</p>

¹ If not other indicated, the study treats drop-outs as if returned to baseline weight.

Source: Studies selected and summarized by authors.