RUHR
ECONOMIC PAPERS

Monika Reuss-Borst
Alfredo R. Paloyo
Arndt R. Reichert
Harald Tauchmann

Heterogeneous Causal Effects of Financial Incentives on Weight Loss

Results of a Large Prospective Randomized Trial

RWI #383

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 #383

Responsible Editor: Christoph M. Schmidt

ISSN 1864-4872 (online) – ISBN 978-3-86788-438-9

Ruhr Economic Papers #383

Monika Reuss-Borst, Alfredo R. Paloyo, Arndt R. Reichert,
and Harald Tauchmann

Heterogeneous Causal Effects of Financial Incentives on Weight Loss

Results of a Large Prospective Randomized Trial

RWI

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.

http://dx.doi.org/10.4419/86788438
ISSN 1864-4872 (online)
ISBN 978-3-86788-438-9

Monika Reuss-Borst, Alfredo R. Paloyo, Arndt R. Reichert, and Harald Tauchmann[1]

Heterogeneous Causal Effects of Financial Incentives on Weight Loss – Results of a Large Prospective Randomized Trial

Abstract

In a randomized controlled trial involving 700 obese persons assigned to three experimental groups, we test whether financial incentives have heterogeneous effects on weight reduction. While two treatment groups obtain EUR 150 and EUR 300, respectively, for achieving an individually-assigned target weight within four months, a control group receives no such premium. The objective is to identify subgroups of patients (male/female, natives/migrants, etc.) who respond more than others to financial incentives. Our conclusions are the following: (1) monetary rewards effectively induce obese individuals to reduce weight across all subgroups; (2) the magnitude of the reward is relevant only for certain subgroups; (3) for people who do not lose weight via conventional weight-loss intervention programs, financial incentives are an effective supplement to induce weight loss.

JEL Classification: I10, I18, H23, C93

Keywords: Randomized experiment; financial incentives for weight loss; obesity; subgroups; effect heterogeneity

November 2012

---

1 Monika Reuss-Borst, Rehaklinik Am Kurpark, Bad Kissingen; Alfredo R. Paloyo, RWI; Arndt R. Reichert, RWI; Harald Tauchmann, RWI and CINCH. *– The authors are grateful to the Pakt für Forschung und Innovation for funding, and to Rüdiger Budde, Viktoria Frei, Karl-Heinz Herlitschke, Klaus Höhner, Julia Jochem, Mark Kerßenfischer, Lionita Krepstakies, Claudia Lohkamp, Thomas Michael, Carina Mostert, Stephanie Nobis, Adam Pilny, Margarita Pivovarova, Gisela Schubert, and Marlies Tepaß for research assistance. We thank the medical rehabilitation clinics of the German Pension Insurance of the federal state of Baden-Württemberg and the association of pharmacists of Baden-Württemberg for their support of and dedication to this experiment. In particular, we acknowledge the support of Susan Eube, Michael Falentin, Ina Hofferberth, Marina Humburg, Silke Kohlenberg, Thomas Krohm, Max Lux, Tatjass Meier, Constanze Schaal, and Wolfgang Stiels. Thanks are also due to Thomas K. Bauer, Christoph M. Schmidt, and Boris Augurzky, whose support proved invaluable in the acquisition and execution of the experiment.* *– All correspondence to Monika Reuss-Borst, RehaZentren Baden-Württemberg, Rehaklinik Am Kurpark, Kurhausstr. 9, 97688 Bad Kissingen, Germany, E-Mail: Monika.Reuss-Borst@rehaklinik-am-kurpark.de.*

1 INTRODUCTION

The global increase of obesity, together with its comorbidities, is considered to be one of the toughest challenges confronting health-care systems worldwide. Between 1980 and 2008, the worldwide prevalence of obesity more than doubled. By 2008, an estimated 1.4 billion adults globally were overweight (BMI ≥ 25 and BMI < 30 kg/m^2) and of these, 502 million adults were obese (BMI ≥ 30kg/m^2) (Finucane et al. 2001).[1] It is predicted that, by 2015, 2.3 billion adults will be overweight, with more than 700 million of them obese.

Obesity follows a strong socioeconomic gradient (Baum II and Ruhm 2009). For example, obesity rates are increasing with age and are inversely correlated with high educational level and income. Moreover, single individuals more often have normal weight than married, divorced, or widowed persons, and adolescents with a migration background have a greater risk of becoming obese than natives (Kurth and Schaffrath Rosario 2007).

The causes for the upsurge of obesity are well-known (Rosin 2008). The increase of energy-dense convenience and/or fast foods, decrease in physical activity due to more sedentary jobs, motorized transport, increasing urbanization, labor-saving devices, but also changed leisure activities (television, computer games, etc.) result in an imbalance of calorie intake and expenditure. As a result of these mainly environmental, social, and partly political changes, obesity is not only a problem in well-developed high-income countries but is also increasing in developing countries.

Being overweight or obese increases the risks of coronary heart disease, ischemic stroke, diabetes mellitus type 2, some cancers such as postmenopausal breast cancer, as well as a variety of musculoskeletal disorders (such as osteoarthritis, low back pain, and gout). Obesity is, apart from smoking, the most important preventable cause of premature death in the United States and Europe. Each year, 2.8 million people die as a result of being obese (WHO 2012).

The economic burden of obesity is tremendous. Incremental medical expenditures and the value of lost productivity, including absenteeism and presenteeism (health-related limitations at work), as a result of overweight and obesity are estimated by Finkelstein et al. (2010) to range from $322 for overweight to $6,087 for grade III obese men. For women, estimates range from $797 for overweight to $6,694 for grade III. The total annual cost attributable to obesity among full-time employees is $73.1 billion. Individuals with a BMI above 35 represent 37 percent of the obese population, but are responsible for 62 percent of excess costs. Gates et al. (2008) report that workers with a BMI equal or above 35 experience the greatest health-related work limitations, being significantly less productive than mildly obese workers.

The effective treatment of obesity is of crucial importance not only for the obese individuals themselves, but also for the social system and employers. Treatment programs mainly consist of advice on diet and physical exercise tailored to the individual patient's needs and may be best

[1] In Germany, a high-income country with highly-developed social systems, 66 percent of the male and about 50 percent of the female adult population are overweight according to a large epidemiological study with almost 20,000 participants aged between 14 and 80 years. The percentage of German adults being obese amounts to more than 20 percent (20.5 percent males; 21.2 percent females) (Max-Rubner Institut 2008).

characterized as behavioral weight-loss intervention programs. The fundamental requirement for success is to achieve an energy balance between calories spent, on the one hand, and calorie input on the other. Although behavioral interventions have been proven to be effective in various studies (Appel et al. 2011), most programs are, at best, only modestly successful and do not address all patients adequately (non-responders).

Therefore, there is increasing interest in the evaluation of the impact of potential modifiers of behavior such as financial incentives. Negative financial incentives are widely applied in the form of taxes on products such as alcohol or tobacco. The effectiveness of financial incentives in treatment for obesity and overweight is still a matter of debate. Most studies (e.g., Paul-Ebhohimhen and Avenell (2008), and Relton et al. (2011)) have issues with respect to sample size and methodology (see Paloyo et al. (2011)). Using a large sample and a randomized design, Augurzky et al. (2012) found that monetary rewards double the weight loss of medical rehabilitation patients that are advised to reduce body weight by their physicians. It remains unknown whether special subgroups of patients (natives/migrants, etc.) may respond more than others to financial incentives. It may turn out that money nudges only people with existing financial problems or special social or cultural background toward a healthier life style. If this is the case, financial incentives could be allocated more appropriately to probable responders to make disbursement more cost-effective.

Based on experimental data of 700 obese medical rehabilitation patients, the present analysis was conducted to test the primary hypothesis that randomly assigned positive, output-oriented financial incentives exhibit a heterogeneous treatment effect on weight—measured in kilograms lost—as a function of baseline characteristics. Our objective is to examine whether monetary rewards for weight loss are more influential on patients with certain baseline characteristics. If this turns out to be true, we are able to identify responder patients and non-responders. In determining which characteristics are relevant, we take the point of view of public administrators (e.g., social health insurers) who have data that are often limited to basic socio-economic variables such as gender and employment status, and without direct information on income and wealth.

The reason for this approach is that a large-scale monetary incentive scheme will naturally be managed by public entities. However, assuming that there is substantial effect heterogeneity with respect to individual characteristics, it is a priori unclear whether the "publicly observable" characteristics are actually relevant. It is entirely possible that other characteristics play a much stronger role for the effectiveness of monetary rewards. For instance, health-related behaviors may be better predictors of weight loss, and, hence, may influence the effectiveness of the monetary rewards relatively more than socio-economic characteristics. We therefore additionally classify patients according their health-related behavior exhibited prior to the experiment.

Comparisons are made between (1) Treatment Arm 1 (EUR 150) and Control Arm and (2) Treatment Arm 2 (EUR 300) and Control Arm for the following nine subgroup categories: male/female, single/in a relationship, high educational attainment/low educational attainment, native/migrant, employed/not employed, urban/rural (socio-economic characteristics), good health/bad health, frequently cooks at home/seldom cooks at home, and at least one diet in the last four years/no diet in the last four years (healthy behavior). Since the members of the control group were also asked to lose weight by their physicians during their medical rehabilitation stay, a

secondary objective is to determine which subgroup would not have lost weight without the addition of the financial incentive.

The following describe our expectations with respect to the variation in treatment effects across baseline covariates. (1) Females may respond more to the financial incentive because it typically constitutes a larger share of their labor income. They may also be rewarded more in the labor market for their weight loss, since the cost of obesity for females has been shown to be higher (e.g., Morris (2007), Hebl and Turchin (2005), and Puhl and Heuer (2009)). (2) Single individuals may be more flexible with the use of their leisure time, which enables them to better regulate both their physical activity and their diets. However, singles also lack the emotional support of a partner (Saccone and Israel 1978). (3) A high cognitive ability may be necessary to understand the benefits of losing weight, so we expect those with a high educational attainment to lose more weight (Halkjær et al. 2003). Moreover, better educated individuals may have easier access to information that is relevant for successfully losing body weight. (4) Language barriers may hinder the benefits of rehabilitation for obese migrants. However, since migrants earn, on average, less than natives (e.g., Lehmer and Ludsteck 2011) the relatively higher share of the reward in their income may induce them to lose more weight than natives. (5) The employed may be better equipped to lose weight, since they are more likely to be better educated and can afford to enroll in fitness classes. However, their leisure time constitutes a smaller part of their overall time budget compared to those who are not employed, which limits the opportunities for them to expend calories outside work. Moreover, the extra financial reward may not be seen as a significant amount to earn by losing weight. (6) Nutritional habits exhibit a substantial rural–urban differential. Though inhabitants of urban areas may be exposed to unhealthier food choices (e.g., fast food), the diet in rural areas is still more meat-oriented in Germany. In addition, urban areas afford better educational opportunities, which may mitigate the effect of the prevalence of junk food.

Among the health-related variables, (7) those in poor health are at a disadvantage because they probably face less choices with respect to the type of food they are able to eat, and may also be physically limited from exercise. Moreover, obese individuals in poor health are likely to benefit more from losing weight than their healthier counterpart. Thus, staying obese despite severe health problems may indicate a lack of self-control. (8) Those who cook at home are in better control of the ingredients that go into their food, which could mean that weight control may be easier for them. Rashad and Grossmann (2004), for example, report that eating out at fast-food and full-service restaurants are seemingly the most important factors in explaining the rise in obesity. (9) Those who have undergone a diet at least once during the last four years but nevertheless remain obese may indicate that these are people with self-control issues. We therefore expect them to respond less to financial incentives.

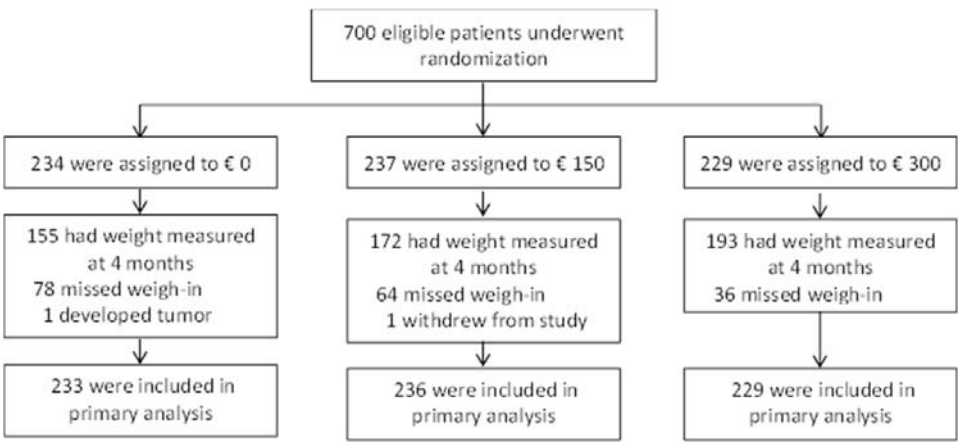
2 METHODS

The randomized controlled trial is funded by the *Pakt für Forschung und Innovation* and administered by the *Rheinisch-Westfälisches Institut für Wirtschaftsforschung* (RWI). Participants were invited to enroll in the experiment during the final week of their stay in a rehabilitation clinic in the state of Baden-Württemberg (BW) in Germany. The study protocol was approved by the

ethics commission of the Chamber of Medical Doctors of BW. A detailed description of the trial is available elsewhere (Augurzky et al. 2012) but we briefly describe it here.

The eligibility criteria involved having a body-mass index of 30 or above at admission, aged between 18 and 75 years, and a resident of BW. The exclusion criteria included, among others, pregnancy, psychological and eating disorders, tumor disease within the last five years, and alcohol or drug abuse. Between March 2010 and August 2011, a total of 700 participants were recruited, though two were later excluded because of pregnancy and cancer. The flow of the study is depicted in Figure 1.

Figure 1: Flow of the Study



Outcomes were ascertained by trained medical personnel, who measured the height and weight of participants. A detailed questionnaire was administered to collect information on socioeconomic background and other health-related variables. At clinic discharge, the physician in charge set an individual weight-loss target that had to be realized within the subsequent four months. Participants were stratified by clinic, and randomization was carried out within each stratum. The three experimental arms include a control group and two treatment groups, with the latter consisting of a group facing a EUR 150 reward and another group with a corresponding EUR 300 reward for achieving the individualized weight-loss target. Patients assigned to the treatment group received no reward if their weight loss was less than 50 percent of their target weight loss. Once the threshold 50 percent has been reached, the financial reward was disbursed proportionally, with the full amount paid out only if the weight-loss target was achieved or exceeded. Once the experiment was underway, weigh-ins were conducted in designated pharmacies using a weighing scale operated by a pharmacist to avoid self-reported weight measures.

Outcomes are examined based on the intention-to-treat principle with the assumption that dropouts return to baseline. Mean weight loss (in kg) was calculated for the two treatment arms and the control arm. The result for the control group is interpreted as baseline weight loss that cannot be attributed to financial rewards. The difference between average weight loss in the treatment arms

and baseline weight loss is regarded as incentive-induced weight loss. Treatment effects are thus estimated as the difference in weight loss between the treatment arms and the control arm of the experiment. Consequently, in assessing the effect heterogeneity of the financial incentives on weight loss according to baseline characteristics, the estimation of baseline and incentive-induced weight loss was conducted over the subgroup categories motivated above. All estimates are presented with their associated 95-percent confidence intervals.

All authors contributed in drafting the manuscript. Data preparation and management, as well as statistical analyses, were conducted at the RWI using Stata (Release 12). The authors claim responsibility for the completeness and veracity of the data and statistical analyses.

3 RESULTS

3.1 Baseline characteristics

Of the 700 participants that were randomized, 234 were assigned to the control group, 237 to the EUR 150 group, and 229 to the EUR 300 group; two were eventually excluded from the sample, while a total of 178 missed the follow-up weigh-in (i.e., 25 percent dropped out); see Figure 1. The average target weight loss is 6.5 percent from a starting average weight of 113 kg. Variables measured at baseline were shown to be balanced between groups (Table 1). About a third of the sample are women and about 80 percent were born in Germany (natives). The mean age is 48 years and most of the participants are employed. All mean comparisons between groups except the percentage of German natives and the percentage of the employed for the EUR 300 group indicate covariate balance at baseline.

We regard participants as employed if they are either full- or part-time employed, marginally employed, or have not provided information on the type of their employment. High educated participants either finished the 10th grade of secondary school or hold a university-entrance diploma (*Abitur*). Individuals with a good self-reported health status answered the question on perceived health with good or very good (the other response options were satisfactory, poor, and bad). Considering the incidence of an acute disease, participants were asked whether any of the following acute diseases occurred during the last four weeks: diarrhea, gastrointestinal infections, influenza, other viral infections, heavy cold, and any other acute disease. The city variable indicates individuals living in municipalities that cover more than one ZIP code. Furthermore, participants were asked whether they frequently cook at home by themselves and whether they made any effort (e.g., a diet) to lose weight during the last four years.

Table 1: Descriptive Statistics (Mean Values and Standard Deviations at Baseline)

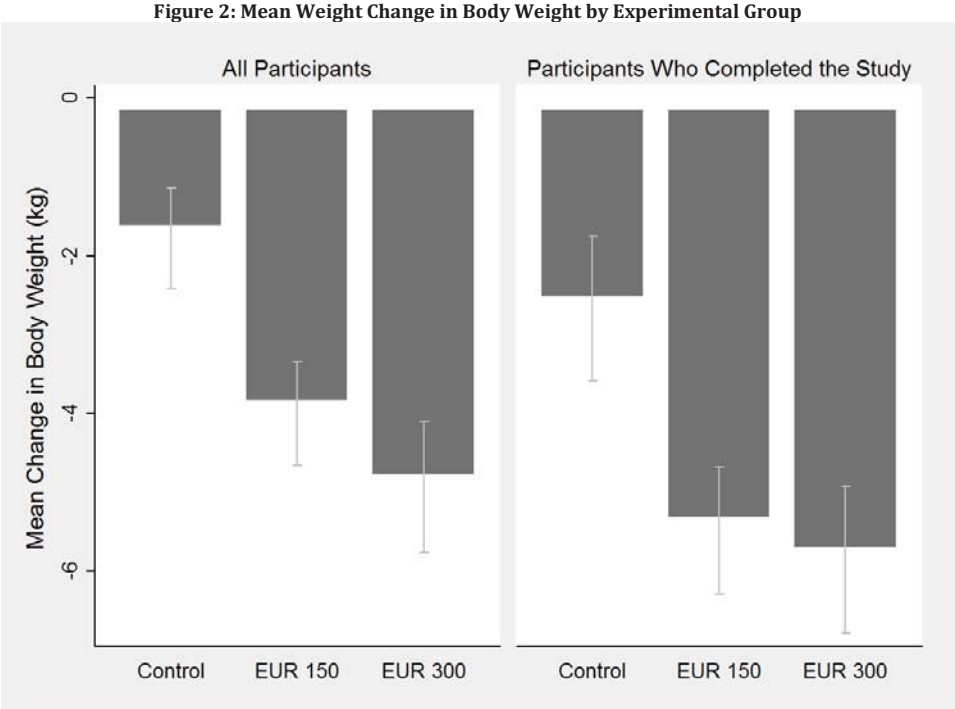
	All	Control	EUR 150	EUR 300
Initial weight (before rehab, kg)	117 ± 23	115 ± 22	118 ± 24	118 ± 24
Starting weight (after rehab, kg)	113 ± 22	111 ± 20	114 ± 23	114 ± 23
Target weight loss (%)	6.5 ± 1.2	6.4 ± 0.9	6.5 ± 1.3	6.5 ± 1.4
Bad Kissingen (%)	33.4	32.2	35.6	32.3
Bad Mergentheim (%)	41.7	41.6	40.7	42.8
Glottertal (%)	6.60	7.30	5.90	6.60
Isny (%)	18.3	18.9	17.8	18.3
Female (%)	32.2	29.1	31.3	36.3
Age (years)	48 ± 9	48 ± 10	48 ± 9	47 ± 9
Single (%)	20.4	18.5	20.9	21.8
Married (%)	61.1	63.1	61.3	59.0
Native (%)	78.9	79.5	75.1	82.2^{+}
Employed (%)	82.4	85.8	81.7	79.5*
High educational attainment (%)	34.5	35.2	36.0	32.3
Population density (1000/km^2)	7.5 ± 7.3	6.9 ± 7.1	7.0 ± 7.8	7.3 ± 7.1
City (%)	23.4	22.1	23.0	25.1
Good self-reported health (%)	35.1	30.9	37.0	36.8
Acute disease within last 4 weeks (%)	73.0	73.8	73.9	71.3
Blood Sugar Level (mmol/l)	188 ± 37	188 ± 37	186 ± 37	200 ± 38
Cholesterol Level (mmol/l)	101 ± 24	100 ± 23	102 ± 25	101 ± 24
Cooking at home frequently (%)	78.1	78.6	80.4	75.2
At least one diet during last 4 years (%)	57.2	56.9	54.7	60.0

* Deviation from control group significant at 5%. $^{+}$ Deviation from EUR 150 group significant at 5%. ± Standard deviations of binary variables omitted. Bad Kissingen, Bad Mergentheim, Isny, and Glottertal refer to the locations of the four rehabilitation clinics. Covariate balance is warranted by respective tests (Hansen and Bowers 2008).

3.2 Average treatment effect

The results in Figure 2 show the mean weight loss for each experimental arm in the trial. Replacing the missing values for the dropouts with their initial values, it is shown that the control group lost almost 2 kg, while the two treatment arms lost between about 4 to 5 kg (left panel of the figure). Weight loss in both treatment arms was significantly higher than in the control arms, revealing that financial incentives can motivate people to lose weight significantly. Using information only on those who completed the study (completed-case or per-protocol analysis) shows a similar weight-loss pattern, though the mean weight loss is slightly higher for all groups (right panel of the figure).

Figure 2: Mean Weight Change in Body Weight by Experimental Group



The two panels of the graph illustrate the well-known selection bias one expects to observe in per-protocol analyses. This can occur since it is probable that those who foresee themselves as benefitting the most from the treatment will have a higher likelihood of completing the study. This endogeneity in sample attrition makes the analysis based on the intention-to-treat principle more reliable.

3.3 Treatment heterogeneity

The estimation results for mean baseline weight change (α) and for mean incentive-induced weight change (δ), i.e. additional weight loss attributable to the monetary rewards, are presented in Table 2. The table further displays the statistical significance of these effects and the equality of baseline as well as incentive effects between the subgroups of interest.

Overall, the estimates for the baseline weight change range from −0.179 kg (95% CI [−1.843, +1.485]) for those who are not employed to −2.427 kg (95% CI [−3.496, −1.358]) for those who have not undergone a diet in the last four years. The estimates for the incentive-induced weight change range from −0.962 kg (95% CI [−2.669, +0.745]) for urban citizen (EUR 150) to −4.272 kg (95% CI [−6.385, −2.159]) for singles (EUR 300). Only in three instances (migrants, those living in urban areas, and those who have not undergone a diet in the last four years) do we not observe the EUR 150 to induce a statistically significant weight loss (indicated by *). Increasing the financial incentive to EUR 300, however, results in a significant weight loss for all these groups. EUR 150 induced the highest weight loss for those who seldom cook at home (−3.374 kg, 95% CI [−5.042, −1.707]), while EUR 300 translated to maximum incentive-induced weight loss for singles (−4.272 kg, 95% CI [−6.385, −2.159]).

We do not observe any significant general treatment effect heterogeneity (indicated by #) for the financial incentives across subgroups but note that higher financial rewards only correspond to higher observed weight loss (indicated by +) for specific subgroups (for women, migrants, those in good health, and those who frequently cook at home). The largest difference is for those who are in otherwise good health (apart from being obese): a 100-percent increase in financial incentives corresponds to a 232-percent increase in weight loss (1.807 kg, 95% CI [−3.493, −0.121] to 4.204 kg, 95% CI [−5.890, −2.518]).

Moreover, there is some heterogeneity in the baseline weight change. Not being motivated by financial rewards, men lost more weight than women, participants in a relationship lost more weight than singles (at 10-percent significance level), the employed lost more weight than those who are not employed, and those who frequently cook at home lost more weight than those who seldom cook at home. In fact, in each of these categories, there is one subgroup for which no significant baseline weight reduction is found. However, when faced with a financial reward (irrespective of the amount), these groups of people are able to significantly lose weight. This means that financial incentives also work for people who do not benefit from conventional intervention programs. Notably, these groups are exactly those which exhibit the most substantial weight loss when given a bonus for losing weight.

Table 2 – Baseline Weight Change and Incentive-induced Weight Change by Selected Subgroups

	Baseline Weight Change		Incentive-induced Weight Change				
	α	CI	δ_1 (EUR 150)	CI	δ_2 (EUR 300)	CI	Obs
Women	−0.592[#]	[−1.737, 0.553]	−2.290[*,+]	[−3.684,−0.897]	−4.007[*,+]	[−5.401,−2.613]	222
Men	−2.241[*,#]	[−2.993,−1.489]	−2.213[*]	[−3.385,−1.041]	−2.935[*]	[−4.107,−1.763]	467
Single	−0.560	[−2.040, 0.919]	−3.115[*]	[−5.228,−1.002]	−4.272[*]	[−6.385,−2.159]	142
In relationship	−2.054[*]	[−2.747,−1.361]	−2.024[*]	[−3.041,−1.007]	−2.918[*]	[−3.935,−1.900]	555
High educ.	−1.685[*]	[−2.647,−0.724]	−1.643[*]	[−3.054,−0.232]	−2.974[*]	[−4.385,−1.563]	241
Low educ.	−1.864[*]	[−2.717,−1.012]	−2.353[*]	[−3.524,−1.182]	−3.324[*]	[−4.495,−2.152]	435
Native	−1.909[*]	[−2.662,−1.155]	−2.445[*]	[−3.507,−1.383]	−3.029[*]	[−4.091,−1.967]	542
Migrant	−1.404[*]	[−2.446,−0.363]	−1.141[+]	[−2.690, 0.408]	−3.261[*,+]	[−4.810,−1.712]	145
Employed	−2.042[*,#]	[−2.718,−1.367]	−2.061[*]	[−3.048,−1.074]	−3.034[*]	[−4.021,−2.047]	574
Not employed	−0.179[#]	[−1.843, 1.485]	−3.326[*]	[−5.733,−0.919]	−4.240[*]	[−6.647,−1.834]	123
Urban	−1.457[*]	[−2.727,−0.187]	−0.962	[−2.669, 0.745]	−2.832[*]	[−4.539,−1.126]	160
Rural	−1.889[*]	[−2.622,−1.156]	−2.578[*]	[−3.662,−1.494]	−3.492[*]	[−4.576,−2.408]	524
Good health	−1.818[*]	[−3.131,−0.505]	−1.807[*,+]	[−3.493,−0.121]	−4.204[*,+]	[−5.890,−2.518]	192
Bad health	−1.783[*]	[−2.496,−1.070]	−2.165[*]	[−3.233,−1.098]	−2.857[*]	[−3.925,−1.790]	493
Cook at home	−2.046[*,#]	[−2.764,−1.328]	−2.032[*,+]	[−3.088,−0.975]	−3.342[*,+]	[−4.398,−2.285]	539
Seldom	−0.276[#]	[−1.411, 0.860]	−3.374[*]	[−5.042,−1.707]	−3.478[*]	[−5.145,−1.811]	151
Diet in past	−1.302[*]	[−2.058,−0.546]	−2.961[*]	[−4.181,−1.742]	−3.938[*]	[−5.158,−2.719]	395
No diet in past	−2.427[*]	[−3.496,−1.358]	−1.280	[−2.692, 0.132]	−2.161[*]	[−3.573,−0.749]	296

[*] Weight change significant at 5%, [+] $(\delta_2 - \delta_1)$ significant at 5%, [#] equality of weight change between the two subgroups significant at 5%. Two-sided *t*-tests used. The estimates and standard errors are obtained from subgroup-specific ordinary least-square regressions of weight change on the two treatment group indicators and a constant term.

4 DISCUSSION

The study is a large-scale randomized controlled trial designed to test the effectiveness of financial incentives in inducing weight loss among obese patients. The trial demonstrated that a positive output-based financial incentive is an effective instrument to achieve a significant weight loss among the obese. However, a general heterogeneity in the treatment effect of financial incentives is

neither observed for socio-economic characteristics nor for exhibited health-related behaviors in the study. Nevertheless, there is some evidence for particular subgroups (for women, migrants, those in good health, and those who frequently cook at home) that increasing the amount of financial incentive has a directly proportional effect on the amount of weight loss. Moreover, for migrants, urban citizens, and those who have not undergone a diet in the last four years, EUR 150 is not enough to induce further weight loss. Only the higher monetary reward motivated them to lose significantly more weight as compared to baseline. This means that the effect of promising more money for weight loss generally depends on specific individual characteristics.

Another important result is the capacity of financial incentives to induce people to lose weight when they would otherwise have failed to do so. This is observed for women, singles, those who are not employed, and those who seldom cook at home. For these groups, it is not sufficient simply to give them a personalized target weight loss, medical advice, and weight-loss counseling. The pattern seems to follow an income gradient, as these groups of people are more likely to have an economically disadvantaged position relative to their counterparts (e.g., male vs. women or employed vs. not working). In fact, if we take the observations from people who belong to the first quartile of the income distribution, we find a similar pattern—that is, being prompted to realize a certain weight loss target alone does not generate a statistically significant change in weight, but financial incentives (whether EUR 150 or EUR 300) do.

It is for these people where a financial incentive could have a relevant impact, since the other groups exhibited weight loss anyway (though the magnitude is smaller) even without the marginal incentive. Therefore, considering a fixed budget constraint, if the objective of the policy based on financial incentives is to ensure that everyone loses weight, a larger part of the monetary amount might be better allocated to these groups. The publicly observable personal characteristics are informative enough for an efficient allocation of resources.

It is not obvious, however, that such a scenario should be the goal of public policy. For instance, while inducing everyone to lose weight is desirable, ensuring that the weight loss is medically significant nevertheless remains a relevant factor. In this case, the obese facing higher financial incentives are most likely to lose the most weight and therefore reap the benefits of such a weight loss. In other words, although ensuring that everyone loses at least 1 kg is an attractive proposition, it is not necessarily better than increasing the likelihood of a particular group of people to lose at least 5 percent of their current weight, which is often regarded as a threshold to improve the health status of the obese (Vidal 2002).

Although the size of the estimated effect of EUR 300 is higher for all subgroups as compared to the effect of the EUR 150 reward, the difference between the two effects is only statistically significant for women, those who were not born in Germany, those already in otherwise good health, and those who frequently cook at home. Therefore, while there is some evidence of a positive relationship between the size of the financial incentive and the resulting amount of weight loss, the results, as a whole, indicate that the presence of a financial incentive itself is more important than its size.

Finally, although initial hypotheses about treatment-effect heterogeneity were stated at the outset, statistical tests do not indicate any evidence of its existence. In fact, heterogeneity as a function of baseline characteristics was primarily observed for the control arm of the experiment, where

participants were exclusively given a personalized target weight loss, and for the difference between the two treatment arms. Overall, this provides evidence against the hypothesis that the effect of financial incentives operates by activating a person's inherent ability to lose weight. If this were the case, one would observe stronger effects among those who have greater ability irrespective of the size of the reward (e.g., because of less time constraints). It rather appears that financial incentives represent an instrument for weight reduction on their own and, hence, may aid individuals who otherwise lack the means or ability to control his or her weight.

We want to emphasize that before the experiment took place, participants had undergone a medical rehabilitation treatment where they had received an individual weight-loss goal, advice, and counseling. For this reason, it is possible that financial incentives are only effective conditional on the assignment of a specific target weight and the knowledge of effective measures to lose weight. Therefore, we are inclined to conclude that monetary rewards in combination with advice and counseling are the key to success. Moreover, a large-scale monetary incentive scheme seems to be feasible even if general practitioners are purposefully excluded in order to avoid conflicts of interests. The present study serves as an example for its feasibility since the participants have been spread out across the whole German federal state of Baden-Württemberg which is in size and population comparable to Switzerland.

With regard to statistical multiple testing, it is important to remember at the outset that the evidence indicates that there is a statistically significant overall effect for both amounts of the financial incentives. This result is robust for all subgroup categories analyzed. So, while there is the risk for researchers to perform ex-post subgroup analyses to report significant effects that are likely the result of mere chance, this risk is mitigated in this case because the overall effect is evident. Thus, the subgroup analyses performed in this study are not susceptible to inadvertently finding significant impacts when there are none. Even if there was much of a concern, our conclusion is that in fact there is no general treatment effect heterogeneity that can be observed from the data. Adjusting the *p*-value (e.g., Bonferroni-type adjustments) to account for multiple testing will only raise the threshold for which to find a significant effect. Increasing that threshold in this case will not change the general conclusion although the finding of heterogeneity in the effect of raising the reward from EUR 150 to EUR 300 may be overturned.

On the other hand, since subsample analyses lower the power of the statistical test, it raises the possibility that the finding of no weight-loss effect for certain subgroups is due to chance. In other words, we increase the likelihood of failing to reject the null hypotheses when in fact it is false. This particular trial involved a larger number of participants (e.g., Volpp et al. (2008) had 57 and Finkelstein et al. (2007) had 207 participants), which is a distinct advantage of the current study. Since the power of the statistical test is a positive function of the number of observations, the effect on the power by conducting subsample analyses may be mitigated by the unusually high number of observations relative to other studies of this kind.

In summary, the present paper shows that financial incentives effectively help obese individuals to reduce weight across all subgroups, whereas the magnitude of the reward seems to be relevant only for certain subgroups. Heterogeneity is only found with respect to the effect of increasing the promised reward. Individual characteristics seem to matter more in the absence of monetary

rewards since the paper reveals that control-group members of some subgroups did not significantly lose weight. We provide evidence that monetary rewards have the capacity to induce people to lose significant weight even if their individual characteristics imply unfavorable weight-loss prospects, i.e., people for whom the conventional weight-loss intervention programs are unsuccessful. If the objective of a policy based on financial incentives is to ensure that everyone loses weight, a larger part of the monetary amount might be better allocated to these groups. The question of whether financial incentives are also useful to achieve weight maintenance is currently the subject of a subsequent investigation.

LITERATURE

Appel, L. J, Clark, J. M., Yeh, H. C., Wand, N. Y., Coughlin, J. W., Daumit, G., Miller, E. R., Dalcin, A., Jerome, G. J., Geller, S., Noronha, G., Pozefsky, T., Charleston, J., Reynolds, J. B., Durkin, N., Rubin, R. R., Louis, T. A., and Brancati, F. L. (2011). "Comparative effectiveness of weight-loss interventions in clinical practice", *New England Journal of Medicine* **365**(21):1959–1968.

Augurzky, B., Bauer, T. K., Reichert, A. R., Schmidt, C. M., and Tauchmann, H. (2012). "Does Money Burn Fat? Evidence from a Randomized Experiment", Ruhr Economic Paper No. 368.

Baum II, C. L. and Ruhm, C. J. (2009). "Age, socioeconomic status and obesity growth", *Journal of Health Economics* **28**(3):635–648.

Finkelstein, E. A., DiBonaventura, M., Burgess, S. M., and Hale, B. C. (2010). "The costs of obesity in the workplace", *Journal of Occupational and Environmental Medicine* **52**(10):971–976.

Finkelstein, E. A., Linnan, L., Tate, D., and Birken, B. (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.

Finucane, M. M., Stevens, G. A., Cowan, M. J. , Danaei, G., Lin, J. K. , Paciorek, C. J. , Singh, G. M., Gutierrez, H. R., Lu, Y., Bahalim, A. N., Farzadfar, F., Riley, L. M., and Ezzati, M. (2001). "National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants", *Lancet* **377**(9765):557–567.

Gates, D. M., Succop, P., Brehm, B. J., Gillespie, G. L., and Sommers, B. D. (2008). "Obesity and presenteeism: the impact of body mass index on workplace productivity", *Journal of Occupational and Environmental Medicine* **50**(1):39–45.

Halkjær, I., Holst, C., and Sørensen, T. I. A. (2003). "Intelligence Test Score and Educational Level in Relation to BMI Changes and Obesity", *Obesity Research* **11**(10):1238–1245.

Hansen, B. B. and Bowers, J. (2008). "Covariate balance in simple, stratified and clustered comparative studies", *Statistical Science* **23**(2):219–236.

Hebl, M. R. and Turchin, J. M. (2005). "The Stigma of Obesity: What About Men?", *Basic and Applied Social Psychology* **27**(3):267–275.

Kurth, B.-M. and Schaffrath Rosario, A. (2007). "Die Verbreitung von Übergewicht und Adipositas bei Kindern und Jugendlichen in Deutschland. Ergebnisse des bundesweiten Kinder- und Jugendgesundheitssurveys (KiGGS)", *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* **50**(5-6):736–743.

Lehmer, F. and Ludsteck, J. (2011). "The Immigrant Wage Gap in Germany: Are East Europeans Worse Off?", *International Migration Review* **45**(4):872–906.

Max-Rubner-Institut (2008). Nationale Verzehrs-Studie II. Ergebnisbericht, Teil 1. Karlsruhe: Max-Rubner-Institut.

Morris, S. (2007). "The impact of obesity on employment", *Labour Economics* **14**(3):413–433.

Paloyo, A., A. R. Reichert, H. Reinermann, and Tauchmann, H. (2011). "The Causal Link Between Financial Incentives and Weight Loss – An Evidence-based Survey of the Literature", Ruhr Economic Paper No. 290.

Paul-Ebhohimhen, V. and Avenell, A. (2008). "Systematic review of the use of financial incentives in treatments for obesity and overweight", *Obesity Reviews* **9**(4):355–367.

Puhl, R. M. and Heuer, C. A. (2009). "The Stigma of Obesity: A Review and Update", *Obesity* **17**(5):941–964.

Rashad, I. and Grossman, M. (2004). "The economics of obesity", *Public Interest* **156**:104–112.

Relton, C., Strong, M., and Li, J. (2011). "The 'Pounds for Pounds' weight loss financial incentive scheme: an evaluation of a pilot in NHS Eastern and Coastal Kent", *Journal of Public Health* **33**(4):536–542.

Rosin, O. (2008). "The economic causes of obesity: A survey", *Journal of Economic Surveys* **22**(4):617–647.

Saccone, A.J. and Israel, A.C. (1978). "Effects of experimenter versus significant other-controlled reinforcement and choice of target behavior on weight loss", *Behavior Therapy* **9**(2):271–278.

Vidal, J. (2002). "Updated review on the benefits of weight loss", *International Journal of Obesity Related Metabolic Disorders* **26**(Suppl 4):25–28.

Volpp, K., John, L., Troxel, A., Norton, L., Fassbender, J., and Loewenstein, G. (2008). Financial incentive-based approaches for weight loss, *Journal of the American Medical Association* **300**(22):2631–2637.

WHO (2012). Factsheet. World Health Organization.
URL: http://www.who.int/mediacentre/factsheets/en/ (Accessed October 05, 2012).