This article was downloaded by: [155.246.103.35] On: 05 April 2017, At: 17:49
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA

Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Gain Without Pain: The Extended Effects of a Behavioral Health Intervention

Daniel Mochon, Janet Schwartz, Josiase Maroba, Deepak Patel, Dan Ariely

To cite this article:
Daniel Mochon, Janet Schwartz, Josiase Maroba, Deepak Patel, Dan Ariely (2017) Gain Without Pain: The Extended Effects of a Behavioral Health Intervention. Management Science 63(1):58-72. http://dx.doi.org/10.1287/mnsc.2015.2322

Full terms and conditions of use: http://pubsonline.informs.org/page/terms-and-conditions

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2017, INFORMS

Please scroll down for article—it is on subsequent pages

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

RIGHTSLINK

http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Vol. 63, No. 1, January 2017, pp. 58–72
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Gain Without Pain: The Extended Effects of a Behavioral Health Intervention

Daniel Mochon,[a] Janet Schwartz,[a] Josiase Maroba,[b] Deepak Patel,[b] Dan Ariely[c]

[a] A.B. Freeman School of Business, Tulane University, New Orleans, Louisiana 70118; [b] Discovery Vitality, Sandton 2146, South Africa; [c] Duke University, Durham, North Carolina 27705

Contact: dmochon@tulane.edu (DM); jschwa@tulane.edu (JS); josiasem@discovery.co.za (JM); deepakp@discovery.co.za (DP); dan@danariely.com (DA)

Received: September 12, 2014
Accepted: July 23, 2015
Published Online in Articles in Advance: February 16, 2016

https://doi.org/10.1287/mnsc.2015.2322

Copyright: © 2017 INFORMS

Abstract. We examine the extended effects of an incentive-based behavioral health intervention designed to improve nutrition behavior. Although the intervention successfully improved the target behavior, less is known about any spillovers, positive or negative, that impacted the program's net benefit. This novel examination presents an opportunity to advance our knowledge of this important question, particularly because many theories predict that balancing behaviors in other domains (e.g., reduced exercise) can occur. Our results show a positive and long-lasting persistence effect for the treatment group, even after the incentive was removed. Moreover, we observe no negative spillover effects into related domains such as exercise, and no negative impact on customer loyalty. These results support the use of incentive-based interventions and highlight the importance, for both theory and practice, of examining their extended effects.

History: Accepted by Uri Gneezy, behavioral economics.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2015.2322.

Keywords: behavioral interventions • self-control • health • precommitment • incentives

1. Introduction

Governments and private firms have shown increased interest in motivating healthy behavior as a means of improving health and controlling the escalation of healthcare costs. Indeed, almost all large U.S. employers (more than 200 employees) offer some type of wellness program (Claxton et al. 2013, Mattke et al. 2013), over two-thirds of which include financial incentives (Mattke et al. 2013). Behavioral studies have shown some promise to this approach in that incentives can effectively improve the behaviors they target (Gneezy et al. 2011). For example, financial incentives have been shown to increase exercise frequency (Acland and Levy 2015, Charness and Gneezy 2009, Royer et al. 2015), improve medication adherence (Volpp et al. 2008a), lead to weight loss (John et al. 2011, Volpp et al. 2008b), and aid smoking cessation (Donatelle et al. 2004, Volpp et al. 2006, 2009). More recently, studies have shown that negative incentives, in the form of commitment contracts that carry the threat of a financial loss for failure to improve a healthy behavior, can also be effective (Giné et al. 2010, Royer et al. 2015, Schwartz et al. 2014).

As more private firms and governments offer incentive programs for healthy living, it is essential to understand their broader impact on both the targeted and closely related behaviors. This question is crucial because the net benefit of incentivizing any one healthy behavior over some period of time must also consider the persistence of the intervention after it ends, how it affects other related behaviors, and how it might influence the relationship between the parties offering and receiving the intervention. If health interventions have some unintended negative effects, such as shifting bad behavior from one domain to another, it reduces their usefulness as mechanisms for improving overall health. For many good practical reasons previous research has rarely examined these questions. More recently, however, researchers have highlighted the importance of doing so as a means of truly understanding the policy implications of behavioral economics interventions (Frey and Rogers 2014, Gneezy et al. 2011).

To address the important question of the broader impact of incentive-based interventions on persistence and spillover into secondary behaviors, we present analyses on combined archival and experimental data gathered over a two-year period (calendar years 2012 and 2013) of Discovery Health's comprehensive Vitality Rewards Program—a longstanding points-based financial rewards system that incentivizes a variety of healthy behaviors. During this period (July–December 2012), Vitality introduced a short-term health intervention designed to improve self-control in the nutrition domain by imposing a financial penalty in the form of a precommitment contract. At a first glance, the intervention was successful and significantly improved the

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

target behavior; those who were offered the opportunity to penalize themselves improved more than those who were not (Schwartz et al. 2014). Little is known, however, about any long-term consequences or spillover effects impacting the intervention's overall success. In the following subsections, we outline the three possible extended effects and discuss the theoretical reasons for concern about a spillover effect for each. We then present the data and results testing these extended effects within the context of this field experiment. Our results show no negative extended effects of the intervention, and they even demonstrate some positive spillovers. Taken together, these findings mitigate some concerns over unintended negative consequences, particularly from such penalty-based schemes. We conclude by discussing the broader implications for incentive-based and behavioral interventions.

1.1. Persistence

One question regarding incentive-based health interventions is what will happen with the target behavior when the intervention ends. This is an essential question from a policy and practical perspective, because an intervention that gets people to behave more healthily for some period of time, but becomes ineffective or leads to worse outcomes once the intervention ends, might have dubious effects on overall health. There is a theoretical basis for this concern, because extrinsic rewards have been shown to lead to an overjustification effect (Deci et al. 1999, Lepper et al. 1973), which can "crowd out" intrinsic motivation (Benabou and Tirole 2003, Frey and Jegen 2001, Frey and Oberholzer-Gee 1997). If crowding out takes place, incentives may have to remain in place indefinitely or even increase, which may not be economically or logistically feasible. Conversely, if the effects persist beyond the intervention period, the value of the intervention may be greater than originally computed.

In spite of the these theoretical predictions, little evidence in the field supports negative postintervention effects. The relatively few studies that have tested persistence find neither positive nor negative postintervention effects (John et al. 2011; Volpp et al. 2008a, b, 2006). Some evidence for persistence exists, especially in the domain of exercise. For example, Charness and Gneezy (2009) found that incentivizing exercise increased the desired behavior both during an intervention and for several weeks after it was stopped, a result they attributed to habit formation (see also Acland and Levy 2015, Royer et al. 2015). In the domain of smoking cessation, both positive (Volpp et al. 2009) and negative (Giné et al. 2010) incentive schemes have demonstrated rates of prolonged abstinence, although this is atypical for smoking cessation studies (Cahill and Perera 2011). Taken together, these few studies show mixed results across relatively few domains. Many questions thus remain about the persistence of successful effects and the conditions under which these may occur (Frey and Rogers 2014, Rogers and Frey 2016).

1.2. Health Spillovers

Another potential concern with using incentives to improve health is the emergence of compensatory effects in other related domains, where encouraging better behavior in one domain leads to less desirable behavior in another. The topic warrants particular attention because various theories make this prediction. First, to the extent that two behaviors (such as healthy eating and exercise) are seen as substitute health behaviors, an increase in one may lead to the decrease in the other. Second, research has shown that self-control is a limited resource (Baumeister et al. 1998, Muraven and Baumeister 2000) and incentivizing people to improve health in a domain that requires ongoing self-control, such as eating healthier, may leave them less self-control for behaviors in other domains, such as exercising. Third, research on licensing (Khan and Dhar 2006, Monin and Miller 2001) argues that behaving well in one domain gives people a sense of freedom to be lax in another. That is, someone who recently purchased a salad may rationalize a break from the gym as a reward. Finally, Dhar and Simonson (1999) argued that when people are faced with competing goals, such as being healthy and indulging, engaging in a behavior that satisfies one of the goals may increase the likelihood that a subsequent behavior will satisfy the other one (see also Fishbach and Dhar 2005). In sum, considerable evidence from laboratory studies suggests that getting people to behave healthier in one domain could lead them to indulgence in others. Of course, it is also possible that the opposite would hold true—behaviors such as healthy eating and exercise could be seen as complements where improving one leads to an improvement in the other. Such an effect would also follow from self-perception theory (Bem 1972), which posits that improving one's behavior in one domain leads people to think of themselves as healthier overall, and subsequently more likely to make healthy choices in other domains. Given the conflicting theories and results, it is particularly important to carry out field experiments to determine which of the effects is stronger.

Some behavioral field interventions have started examining compensation effects within the same domain and show mixed evidence. Wisdom et al. (2010), for instance, found that people compensated for lower calorie sandwiches by ordering highly caloric beverages and side dishes, essentially wiping out the intervention effects. Alternatively, Schwartz et al. (2012) found that when fast food customers were

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

prompted to cut calories by downsizing high calorie side dishes they did not compensate by ordering more indulgent entrees. The results of studies examining immediate substitution effects within a domain raise more general questions about compensation across domains. An intervention that gets people to purchase less junk food, but as a consequence also leads them to exercise less, would likely do little to improve people's overall health and could come at a considerable financial cost to the organization footing the bill. Conversely, an intervention that has positive spillover effects may greatly amplify the effectiveness of the intervention.

1.3. Loyalty Spillovers

The final concern with regard to incentive-based interventions we address is their effect on the overall relationship between participants and the intervention's host firm or government. As private firms increasingly offer health programs to both their employees and their customers, they must balance the health benefits with the costs in terms of customer sentiment, engagement, and loyalty. Although an intervention may successfully improve a specific health behavior, it could come at the expense of customers becoming less engaged or even switching to a different brand. This would make most companies unwilling to implement such programs. A similar concern exists for public programs, where a successful intervention that somehow angers the population could threaten future participation and may not be worth implementing.

There are good reasons for firms to be concerned about alienating customers. Customer loyalty is often conceptualized as a function of a customer's attitude toward the firm (Dick and Basu 1994), and health programs that guide people into behaviors they would not necessarily do on their own, or impose a negative incentive or punishment (Giné et al. 2010, Royer et al. 2015, Schwartz et al. 2014), may hurt customer attitudes and weaken their brand loyalty. Indeed, research has shown that a negative experience with a firm can hurt customer relations (Luo 2007, Smith et al. 1999), an effect that could be even worse among a firm's most loyal and valuable customers (Grégoire and Fisher 2008). Penalty-based schemes may be particularly prone to such alienation. Although research shows that people are especially motivated to avoid losses relative to achieving equivalent gains (Kahneman and Tversky 1979), such loss aversion similarly suggests that people's negative reactions to even small penalties could be inflated. Therefore, as organizations examine the economic viability of different types of interventions, particularly those that leverage loss aversion, they must consider any negative backlash in terms of customer sentiment and loyalty. Firms will likely pay special attention to their most loyal and valuable customers, since the cost of alienating a loyal customer can be much higher than the cost of alienating one who is not very engaged.

1.4. The Current Research

Questions regarding the extended effects of incentive-based behavioral programs are an important component in determining the success of this approach to improving outcomes. In the current research, we examine these questions in the context of one such intervention (Schwartz et al. 2014), where participants were randomly assigned either to take part in a six-month voluntary commitment device designed to help them increase the nutritional quality of their supermarket groceries, or to a control group. Among the committed group, failure to meet the goal resulted in the forfeiture of a bonus that is routinely paid as part of the Vitality program. In the current paper, we present new data about members' long-term grocery shopping behavior, exercise behavior, and engagement with the Vitality program. This allows us to test for persistence effects, health spillovers, and loyalty spillovers, respectively.

The precommitment intervention offers a good testing ground for extended effects for multiple reasons. First, although committed participants knew when the commitment ended, and were even sent an email reminding them so, they were unaware that their shopping behavior was monitored even after the commitment was over. Participants were also unaware that their behaviors in other domains were examined as a function of the initial random assignment into experimental groups or their decision to precommit. However, because these data are captured and recorded automatically as part of the rewards system, we were able to examine grocery shopping and other behaviors, free from concern for demand effects.

Second, the commitment itself required participants to make active healthier choices for an extended time period. This differentiates it from other behavioral interventions where participants can choose a one-time course of action, but do not have to think about their behavior again (e.g., Thaler and Benartzi 2004) or might not even be aware that an intervention is taking place (e.g., Wansink 2004) and has altogether circumvented the need for active choices (e.g., Wansink and Chandon 2014). As such, this is a setting where previous research suggests compensation effects would be most likely to occur, as the sustained effort to be healthy under the threat of a penalty could deplete self-control (Baumeister et al. 1998) or provide justification (Khan and Dhar 2006) for making less healthy choices in other domains.

Finally, this particular intervention involved only the risk of losing money, with no possibility of gain. In fact, committed households collectively forfeited around $30,000 in cash-back rewards, which created a

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

real possibility of negative consequences with regard to customer loyalty. A failure to find negative effects on loyalty would consequently have implications for similar interventions involving precommitment, as well as any intervention using incentives.

In summary, the intervention's structure offers an ideal setting to test for the extended effects that are so crucial to understanding how to design interventions that promote desired behaviors. Indeed, the features described above make it more likely that negative spillover effects will emerge in this context than in many other incentive-based interventions. Thus, failure to find negative effects would have implications more broadly for the incentive-based approach toward health, and possibly other important domains such as savings. Below, we outline the original intervention, and then describe the novel data sets obtained to analyze the previously unreported extended effects.

2. Method

The current paper examines the extended effects of a successful health intervention (Schwartz et al. 2014) done in collaboration with Discovery Vitality. We begin by describing the Discovery Vitality program, the study sample, and the original intervention. Finally, we introduce the new data obtained to examine the intervention's extended effects.

2.1. Discovery Vitality Program

This study was run in collaboration with Discovery Vitality, a wellness program that offers rewards for a variety of healthy behaviors (for a full description of the program, see https://www.discovery.co.za/portal/individual/vitality). The breadth of activities covered by the Vitality program provides a comprehensive and centralized accounting of members' health engagement, which represents an optimal setting for examining the broader impact of a targeted intervention. Vitality members earn points for behaviors such as exercise, purchasing healthy groceries, vaccinations, and other preventive and routine healthcare. An accumulation of points leads to higher status levels, which, in turn, lead to greater benefits and rewards. (See Table A1 in the online appendix, available as supplemental material at https://doi.org/10.1287/mnsc.2015.2322, for examples of points activities, the different status levels, and some key benefits.) Discovery Vitality is a branch of Discovery Health, a large private health insurer in South Africa, covering close to 2.5 million people. All rewards are processed at the household level, making that the unit of analysis.

2.2. Sample

The original sample consisted of 6,570 Discovery Vitality member households participating in the HealthyFood program, a 25% discount on healthy food purchases at the Pick n Pay supermarket (Sturm et al. 2013). Analyses were restricted to the 4,073 households who opened the invitation email. Since the invitation email was identical across conditions (see the online appendix for exact text), this selection criteria does not introduce any confounds but rather eliminates participants who were likely unaware of the program or whose emails were incorrect and were never reached.

2.3. Experimental Intervention

The original field experiment randomly assigned the households to one of two conditions: One-third in the control condition, and two-thirds in the treatment condition. On June 19, 2012, all 6,570 households received an email invitation to complete a survey about their grocery shopping habits. Participants in both the control and treatment groups were asked whether they would be interested in signing up for a "Precommitment Programme." The commitment was to increase the percentage of healthy food items purchased at Pick n Pay by five percentage points over their historical baseline for each of the following six months (with ability to opt-out after the first month). Failure to achieve this goal in any month meant forfeiting their HealthyFood discount for that month. For the control group, this was a hypothetical commitment, but for the treatment group, the commitment was binding. Consenting households were given feedback each month regarding their percentage of healthy food and whether the goal was achieved. The message that followed the final month of precommitment reminded members that the HealthyFood cash back would return to normal (see Schwartz et al. 2014 for a complete description).

The intervention's structure offered no incentive to cheat. At best, committed households could get the standard discount that would have been awarded without signing up for the precommitment. Therefore, rational agents would simply choose not to sign up. In addition, because they could opt out after the first month, households who committed, but discovered that it was too difficult for them, could opt out by sending an email to Vitality, which was easier than trying to game the system. Very few (less than 15%) of the committed households ever opted out. These households continued to be treated as committed in the analyses so as to avoid any selection bias.

The original intervention successfully improved the committed households' grocery shopping behavior. Thirty-six percent of households who were offered the binding commitment agreed to the challenge and subsequently showed an average 3.5 percentage-point increase in healthy grocery item purchases. These changes occurred as a result of increased healthy food purchases and decreased unhealthy food purchases. This effect was persistent over each of the six intervention months and was measurable at the population

RIGHTSLINK

Table 1. (a) Descriptive Statistics for the Main Analysis Variables; (b) Distribution of Status Level (at the Beginning of 2012)

(a)

Variable	Mean	Median	SD	Min	Max	*N*
Healthy items count	26.79	23.00	18.97	0	152.50	4,073
Healthy spend (South African Rand)	510.28	422.26	378.46	0	3,402.79	4,073
Percentage of healthy items	31.93	31.38	12.01	0	74.06	4,073
Number of exercise events	6.97	4.83	7.69	0	65.33	4,073
Log number of exercise events	1.46	1.60	1.09	0	4.19	4,073
Points	5,772.81	4,645.50	4,650.92	0	35,629.00	4,073
Log points	7.29	7.69	1.51	0	10.18	4,073

(b)

	Blue	Bronze	Silver	Gold	Diamond
N	892	1,111	326	918	805
%	22.0	27.4	8.0	22.7	19.9

Note. The statistics in panel (a) represent the monthly average per policy during the preintervention period.

level (see Schwartz et al. 2014 for a full description of results). Although it was promising that the intervention was successful at improving the target behavior while the penalty was in place, subsequent questions emerged about any extended effects that impact the intervention's overall viability. For the current project, we petitioned Discovery Vitality for the data specifically to test for these effects.

2.4. Data

We obtained the data from two main sources. One was collected by the Pick n Pay supermarket chain, which provides members' itemized grocery data, and the other was collected by Discovery Vitality, which records all Vitality members' activity points. This data is recorded automatically as part of the Vitality program and was made available for all households.

Pick n Pay reports a monthly total itemization of grocery items purchased, which are subsequently designated by Vitality as healthy, neutral, and unhealthy. Members are aware of the healthy designation (store labels or grocery receipt indicators) because these items receive the 25% HealthyFood discount, but are unaware of the neutral and unhealthy categories. For each household, we have the number of items purchased and the total amount spent in each of those three categories. For analysis purposes, we obtained the data for each month from January 2012 (six months before the precommitment intervention began) until June 2013 (six months after the precommitment intervention ended). These data allow us to examine the critical question concerning whether the intervention effects persisted after the commitment device was removed.

In addition to the grocery item data, Discovery Vitality logs each points-earning activity recorded by each member for each policy. These data include the date the activity was completed, an activity category and subcategory, the status level of the household at the time the activity was recorded, and the points accrued because of the activity. The data set includes all points-accrual activity for 2012 and 2013 and is used to test our hypotheses about health and loyalty spillovers. Table 1 presents descriptive statistics of the main variables of interest.

3. Results

For all of the reported analyses (unless otherwise noted) we estimate a linear difference-in-differences regression model, where each observation in the panel is a specific policy in a specific month of the study. We report two types of analyses. To examine the effect of the commitment on the group of participants who signed up for the precommitment, we estimate the average treatment effect on the treated (ATT). For these analyses we include dummies for intervention period (before, during, and after the intervention), commitment status (committed versus not committed), and the critical committed by intervention period interaction term, with experimental condition as an instrument for commitment status. In addition, we also report the intention-to-treat effect (ITT), which examines the overall effect on the population of being randomly assigned to the treatment condition, independent of whether they chose to commit or not. For these analyses we also include dummies for intervention period (as above), a dummy for experimental condition (control versus treatment), and the critical interaction between these. Finally, for robustness purposes, we estimate all models with either fixed effects or robust clustered errors to account for the multiple observations per policy.

In addition to examining the main effect of the commitment, we run analyses where we split the sample

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

RIGHTSLINK

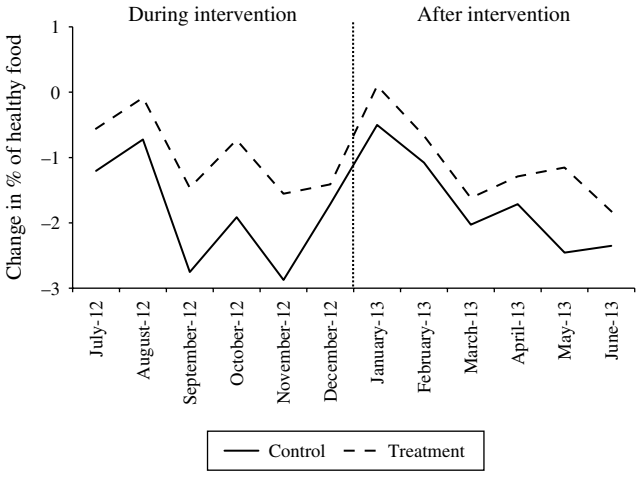
Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

into high status members (Gold and Diamond[1]) and low status members (Blue, Bronze, and Silver), as measured at the beginning of 2012, before the intervention began. Status level is a managerially relevant variable, because firms are likely to be very interested in understanding the extended effects on their most loyal and valuable customers, in addition to the population as a whole. Moreover, status level is theoretically interesting because it allows us to test whether people who are more engaged show a different response to the intervention than those who are not.

3.1. Persistence

We first examine the persistence of the effect on the target grocery shopping behavior after the commitment contract ended. We examine the households' supermarket shopping data for January through June 2013, the six months after the intervention ended and the threat of incentive loss was removed. Figure 1 shows the average percentage of healthy food items purchased by month for the control and treatment groups relative to their baseline behavior (the negative numbers reflect seasonality where baseline period purchases were healthier for all groups).[2] The left half of the graph shows the intention to treat effect of the commitment while it was in place (previously reported by Schwartz et al. 2014). The right half of the graph shows what happened after the intervention ended: households in the treatment condition continued to purchase healthier food. Although this effect is a bit smaller than during the commitment period, it is robust and fairly stable over each of the six months in the post-commitment period and shows no sign of declining in magnitude. Table 2 presents the corresponding formal analyses for the postintervention period. We find that there was a positive effect for both the committers (2.84 percentage-points increase in healthy grocery purchases) and the treatment population as a whole (0.64 percentage-points) after the commitment period ended.

Figure 1. Change from Baseline in Percentage of Healthy Food Purchases by Condition and Month



These results contribute to the nascent literature demonstrating persistence effects for incentive-based health interventions. Most notably, our results are the first to demonstrate persistence outside of exercise (Acland and Levy 2015, Charness and Gneezy 2009, Royer et al. 2015) and smoking cessation (Giné et al. 2010, Volpp et al. 2009) and suggest that incentivizing healthy food purchasing is a fruitful domain for developing healthier habits with short-term incentives. In addition, our results contribute to the growing literature demonstrating that negative incentives, particularly in the form of commitment contracts, can be a catalyst for lasting change (Giné et al. 2010, Royer et al. 2015).

As previously discussed, it is not only important to understand whether there are extended effects of an intervention, but also whether they are stronger for some members than others. To better understand how the effect persisted for different subpopulations, we split the data by status. Table 3 shows the ATT analyses for high status (Gold and Diamond) versus low status (Blue, Bronze, and Silver) members. (Table A2 in the online appendix presents the corresponding ITT analyses.) These results demonstrate strong persistence among high-status members, who showed about a five percentage point increase in their healthy shopping habits both during and after the intervention ended. However, we find no persistence effect for low-status members, who also showed a smaller effect during the intervention. (Table A7 in the online appendix presents the full factorial model.) These results suggest that commitment devices are more effective and more persistent with highly engaged customers (as measured by their status level). Note that engagement (i.e., Vitality Status) is not necessarily a predictor of health status, which we are unable to measure, so the results should not be interpreted as commitment devices are only attractive to or successful for healthy people.

To better understand the observed pattern, we split the commitment period into the first three months (July–September) and the last three months (October–December) to determine whether there were differences between these groups during the intervention itself that might predict the persistence effect. The first column of Table 4 shows the effect during each half of the intervention period for the entire sample. As shown, the effect is a bit stronger in the second half of the intervention period, suggesting some amount of learning. Moreover, we find stark differences during the intervention depending on which group we condition on. High-status members show an increase in their healthy purchases while the incentives are in place,

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

Table 2. The Effect of the Intervention on Percentage of Healthy Food Items Purchased During and After the Commitment Ended

Variable	ATT (Fixed effects)	ATT (Robust SE)	ITT (Fixed effects)	ITT (Robust SE)
Committed × During intervention	3.76***	3.37***		
	(0.92)	(1.21)		
Committed × After intervention	2.84***	2.45*		
	(0.95)	(1.49)		
Treatment × During intervention			0.85***	0.76***
			(0.21)	(0.27)
Treatment × After intervention			0.64***	0.55*
			(0.21)	(0.33)
Committed		−1.17		
		(1.77)		
Treatment condition				−0.27
				(0.40)
During intervention	−1.86***	−1.59***	−1.86***	−1.59***
	(0.17)	(0.22)	(0.17)	(0.22)
After intervention	−1.74***	−1.38***	−1.74***	−1.38***
	(0.17)	(0.27)	(0.17)	(0.27)
Constant	32.18***	32.19***	32.17***	32.19***
	(0.07)	(0.34)	(0.07)	(0.34)
Observations	67,608	67,608	67,608	67,608
Clusters	4,073	4,073	4,073	4,073

Notes. The dependent variable is the percentage of healthy food items purchased per month. Experimental condition is used as the instrument for commitment status in the ATT analyses.
*$p < 0.1$; ***$p < 0.01$.

which then persists after the commitment ends. Low-status members show the opposite pattern whereby they start off showing an effect of the commitment, but over time their response declines (even while the incentives are in place), which leads to no effect after the incentives are removed. These results highlight some interesting insights about the persistence of incentive-based interventions by suggesting that people who are more engaged with a particular intervention will show different responses to the incentives during the inter-

Table 3. Average Treatment on the Treated Effect of the Intervention on Percentage of Healthy Items Purchased During and After the Commitment Ended Split by Member Status Level

Variable	Non-Gold members (Fixed effects)	Gold members (Fixed effects)	Non-Gold members (Robust SE)	Gold members (Robust SE)
Committed × During intervention	2.35**	5.81***	1.75	5.70***
	(1.16)	(1.50)	(1.50)	(2.01)
Committed × After intervention	0.95	5.48***	0.43	5.24**
	(1.21)	(1.53)	(1.87)	(2.41)
Committed			−0.21	−2.22
			(2.22)	(2.85)
During intervention	−1.69***	−2.08***	−1.39***	−1.92***
	(0.23)	(0.26)	(0.29)	(0.34)
After intervention	−1.56***	−1.96***	−1.16***	−1.80***
	(0.23)	(0.26)	(0.36)	(0.40)
Constant	30.78***	33.97***	30.65***	34.20***
	(0.09)	(0.10)	(0.45)	(0.49)
Observations	38,094	29,514	38,094	29,514
Clusters	2,350	1,723	2,350	1,723

Notes. The dependent variable is the percentage of monthly healthy food items purchased. Experimental condition is used as the instrument for commitment status in the analyses. "Gold members" includes all Gold and Diamond members as of January 1, 2012.
**$p < 0.05$; ***$p < 0.01$.

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

Table 4. Average Treatment on the Treated Effect of the Intervention on Percentage of Healthy Items Purchased with the Commitment Period Split (July–September, October–December)

Variable	Full sample (Fixed effects)	Non-Gold members (Fixed effects)	Gold members (Fixed effects)	Full sample (Robust SE)	Non-Gold members (Robust SE)	Gold members (Robust SE)
Committed × *During intervention* (July–September 2012)	3.50*** (1.12)	2.88** (1.41)	4.42** (1.84)	2.90** (1.34)	2.06 (1.66)	4.12* (2.22)
Committed × *During intervention* (October–December 2012)	4.03*** (1.15)	1.77 (1.46)	7.22*** (1.86)	3.85** (1.53)	1.42 (1.93)	7.30*** (2.51)
Committed × *After intervention*	2.84*** (0.95)	0.95 (1.21)	5.48*** (1.53)	2.45* (1.49)	0.43 (1.87)	5.24** (2.41)
Committed				−1.17 (1.77)	−0.21 (2.22)	−2.22 (2.85)
During intervention (July–September 2012)	−1.55*** (0.21)	−1.51*** (0.27)	−1.61*** (0.32)	−1.30*** (0.25)	−1.21*** (0.32)	−1.45*** (0.39)
During intervention (October–December 2012)	−2.18*** (0.21)	−1.89*** (0.28)	−2.55*** (0.32)	−1.90*** (0.28)	−1.58*** (0.37)	−2.39*** (0.42)
After intervention	−1.74*** (0.17)	−1.56*** (0.23)	−1.96*** (0.26)	−1.38*** (0.27)	−1.16*** (0.36)	−1.80*** (0.40)
Constant	32.18*** (0.07)	30.78*** (0.09)	33.97*** (0.10)	32.19*** (0.34)	30.65*** (0.45)	34.20*** (0.49)
Observations	67,608	38,094	29,514	67,608	38,094	29,514
Clusters	4,073	2,350	1,723	4,073	2,350	1,723

Notes. The dependent variable is the percentage of monthly healthy food items purchased. Experimental condition is used as the instrument for commitment status in the analyses. "Gold members" includes all Gold and Diamond members as of January 1, 2012.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

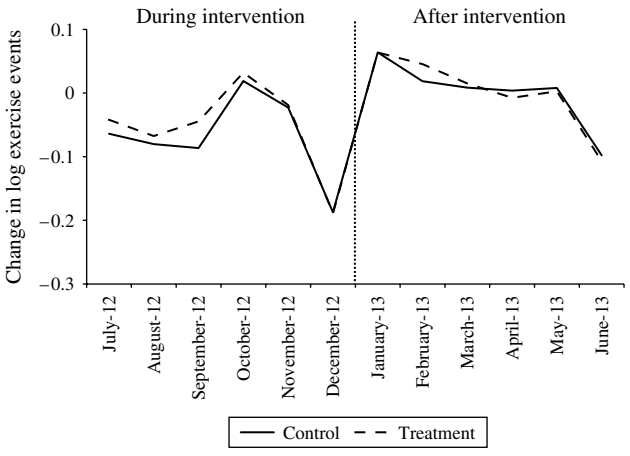
vention itself. This subsequently predicts some degree of learning and habit formation even after the intervention ends and thus indicates that persistence may be amplified by changing the incentives' structure for some subpopulations.

3.2. Health Spillovers

We test for health spillovers by examining whether committed households, who bought healthier food at the supermarket, altered their behavior in other observable health domains. Specifically, we examine the participants' exercise activities. Exercise is an attractive secondary behavior to investigate for many reasons. First, gym memberships are a primary draw to the Vitality program, where members receive an 80% discount on partner gym memberships, and earn 150 Vitality points for every visit, which provides an extra incentive to swipe their loyalty card each exercise session. In addition to these partnerships, Vitality offers other tools for tracking and reporting physical activity outside of the gym. For example, members also receive 150 points for exercise using a tracker such as a Polar heart rate monitor or a Nike+ fitness device to record activity. We therefore have a fairly good estimate of the amount of exercise engagement among the sample before, during, and after the intervention. Moreover, because exercise is a trackable behavior that occurs on the same time scale as grocery shopping (i.e., up to a few times per week), it creates a good opportunity to detect spillover effects. That is, diet and exercise are so inexorably linked when people think of regulating caloric intake and output, that slacking off at the gym after buying a salad may be particularly tempting. Finally, exercise is also an eligible activity for all Vitality members regardless of age and gender, unlike some other preventive activities, which helps preserve the sample size.

Figure 2 shows the average log number of exercise events by month for the control and treatment groups relative to their baseline behavior during and after the intervention.[3] We examine log number of events because the data are positively skewed.[4] Table 5

Figure 2. Change from Baseline in the Log Number of Exercise Events by Condition and Month



RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

Table 5. The Effect of the Intervention on Exercise During and After the Commitment Ended

Variable	ATT (Fixed effects)	ATT (Robust SE)	ITT (Fixed effects)	ITT (Robust SE)
Committed × *During intervention*	0.07	0.07		
	(0.05)	(0.07)		
Committed × *After intervention*	0.01	0.01		
	(0.05)	(0.10)		
Treatment × *During intervention*			0.02	0.02
			(0.01)	(0.02)
Treatment × *After intervention*			0.001	0.001
			(0.01)	(0.02)
Committed		−0.02		
		(0.16)		
Treatment condition				−0.004
				(0.04)
During intervention	−0.07***	−0.07***	−0.07***	−0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
After intervention	0.001	0.001	0.001	0.001
	(0.01)	(0.02)	(0.01)	(0.02)
Constant	1.46***	1.46***	1.46***	1.46***
	(0.004)	(0.03)	(0.004)	(0.03)
Observations	73,314	73,314	73,314	73,314
Clusters	4,073	4,073	4,073	4,073

Notes. The dependent variable is the log number of monthly exercise events. Experimental condition is used as the instrument for commitment status in the ATT analyses.

***$p < 0.01$.

presents the corresponding formal analyses testing whether the commitment to purchase healthier food leads members to worse behavior in another domain during the six-month commitment period and for the six months after the commitment ended. As the results show, there was a positive (but nonsignificant) effect of the commitment on exercise during the commitment period, and no effect at all after the commitment ended. These results suggest that there were no spillover effects such that getting people to eat healthier food leads them to exercise less.[5]

Failure to reject the null hypothesis that the intervention had no significant impact on exercise behavior does not mean that there was zero effect in the population. Nonetheless, our analysis allows us to bound the likely magnitude of this compensation effect if it exists. Indeed, the 95% confidence interval for the change in log number of exercise events during the intervention for the committed group is [−0.03 0.16], which amounts to about 0.11 fewer gym visits per month—under the worst-case scenario (for an average person in the sample). Thus, although we cannot conclude with certainty that there was no negative spillover effect from precommitment to healthier groceries to exercise, if there was one, it was likely too small to be of theoretical or practical importance.

Another potential concern with accepting a null hypothesis that relies on a means-based analysis is the possibility that some committed participants exercised more as a result of the intervention but others less. So although on average there was no effect, there was actually a mixed response in the population. To examine this possibility, we computed a difference score for each policy between the log of their preintervention and during intervention exercise activity, and then plotted the distributions of the control versus treatment conditions (Figure A1.1, online appendix) and of the committed participants versus all other participants (Figure A1.2, online appendix). Tests for the equality of these pairs of distributions show no significant difference (K-S $D = 0.02$, $p = 0.91$ for Control versus Treatment) and (K-S $D = 0.03$, $p = 0.76$ for Committed versus Not Committed). These results further suggest that the improvement in healthy shopping behavior did not come at the cost of exercising less. To further examine the heterogeneity of the effect, we again split our data by status. Table 6 shows that both high and low status members showed a similar pattern, whereby there was a positive but nonsignificant effect of the commitment on exercise behavior. (See Tables A4 and A7 in the online appendix for the corresponding ITT analyses and the full factorial model.) This further reinforces the lack of negative spillover effects.

3.3. Loyalty Spillovers

Thus far we have seen no negative extended health effects of the intervention, and some positive spillovers. In spite of this, firms may still be reluctant to offer such incentive-based programs to their consumers if they have negative effects in terms of customer engagement and loyalty. This would be of particular concern over the most loyal and valuable customers—such as those who are Gold and Diamond status with Vitality.

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

Table 6. Average Treatment on the Treated Effect of the Intervention on Exercise During and After the Commitment Ended Split by Status Level

Variable	Non-Gold members (Fixed effects)	Gold members (Fixed effects)	Non-Gold members (Robust SE)	Gold members (Robust SE)
Committed × During intervention	0.05	0.09	0.05	0.09
	(0.06)	(0.09)	(0.09)	(0.13)
Committed × After intervention	0.002	0.01	0.002	0.01
	(0.06)	(0.09)	(0.13)	(0.16)
Committed			−0.18	0.32
			(0.20)	(0.23)
During intervention	−0.07***	−0.07***	−0.07***	−0.07***
	(0.01)	(0.01)	(0.02)	(0.02)
After intervention	−0.01	0.01	−0.01	0.01
	(0.01)	(0.01)	(0.03)	(0.03)
Constant	1.16***	1.88***	1.18***	1.84***
	(0.005)	(0.01)	(0.04)	(0.04)
Observations	42,300	31,014	42,300	31,014
Clusters	2,350	1,723	2,350	1,723

Notes. The dependent variable is the log number of monthly exercise events. Experimental condition is used as the instrument for commitment status in the analyses. "Gold members" includes all Gold and Diamond members based as of January 1, 2012.
***$p < 0.01$.

We tested for loyalty spillovers by examining the points accumulated in the Vitality program during the intervention and for six months afterward. Vitality points offer a good test of customer loyalty, because they capture both churn (households with zero points) and decreased involvement (fewer points following the intervention). Similar to the exercise analyses, we examine log number of points because the data are positively skewed.[6] Figure 3 shows the average log number of points by month for the control and treatment groups relative to their baseline behavior[7] during and after the intervention. Table 7 presents the corresponding formal analyses testing whether the commitment to purchase healthier food leads members to engage less (or more) with the Vitality program during the six-month commitment period, and during the six months that followed the commitment. We find significantly higher engagement with the Vitality program for committers during the intervention, and a positive but nonsignificant effect after the commitment ended.[8] These positive effects are likely to be partially driven by the accumulation of extra points from additional healthy food items. Importantly, they show that this increase in healthy food shopping did not lead them to disengage more broadly with the program. As a robustness check, we also tested the effects of the commitment on customer loyalty by examining the members' change in status from the beginning of 2012 (before the intervention) until the end of 2013 (a year after the intervention ended). Similar to the points analyses, we find a positive, but nonsignificant effect of commitment on loyalty for the year following the intervention (see Table A5 in the online appendix).

Figure 3. Change from Baseline in the Log Number of Points by Condition and Month

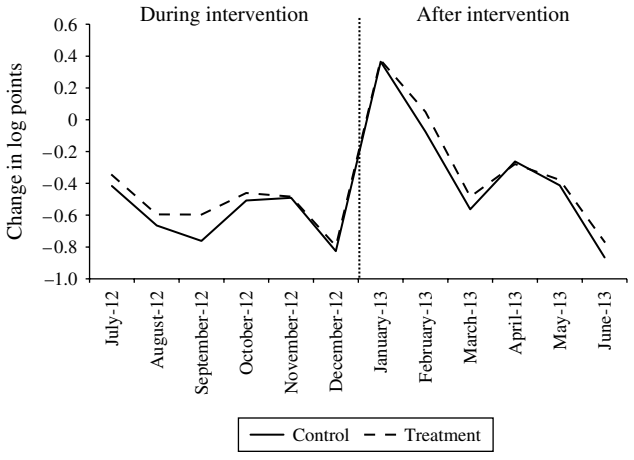


Table 8 shows the loyalty analyses split by baseline status level. We find that the most valuable (Gold and Diamond) members are the ones who show the largest positive effect, and non-Gold members show no negative consequences of this intervention, further reinforcing the claim that the commitment contract did not hurt Vitality's relationship with their customers. (See Tables A6 and A7 in the online appendix for the corresponding ITT analyses and the full factorial model.)

It is especially interesting that we observe no negative effects on customer loyalty, given that unlike most incentive-based health interventions, this one had no possibility of making more money and imposed real

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

Table 7. The Effect of the Intervention on Customer Loyalty During and After the Commitment Ended

Variable	ATT (Fixed effects)	ATT (Robust SE)	ITT (Fixed effects)	ITT (Robust SE)
Committed × During intervention	0.28*	0.28*		
	(0.15)	(0.17)		
Committed × After intervention	0.24	0.24		
	(0.15)	(0.22)		
Treatment × During intervention			0.07*	0.07*
			(0.04)	(0.04)
Treatment × After intervention			0.05	0.05
			(0.04)	(0.05)
Committed		−0.08		
		(0.22)		
Treatment condition				−0.02
				(0.05)
During intervention	−0.61***	−0.61***	−0.61***	−0.61***
	(0.03)	(0.03)	(0.03)	(0.03)
After intervention	−0.30***	−0.30***	−0.30***	−0.30***
	(0.03)	(0.04)	(0.03)	(0.04)
Constant	7.29***	7.30***	7.29***	7.30***
	(0.01)	(0.04)	(0.01)	(0.04)
Observations	73,314	73,314	73,314	73,314
Clusters	4,073	4,073	4,073	4,073

Notes. The dependent variable is the log number of monthly points. Experimental condition is used as the instrument for commitment status in the ATT analyses.

*$p < 0.1$; ***$p < 0.01$.

losses on those failing to meet their goal. In fact, committed households forfeited around $30,000 in cash-back reversals over the course of the intervention, which could have translated into negative feelings toward Vitality. Such findings highlight that penalties can be an attractive alternative to increasing incentive amounts because offering more money for good behavior may provide further discounts only to those who are already performing well or might become prohibitively expensive.

Table 8. Average Treatment on the Treated Effect of the Intervention on Customer Loyalty During and After the Commitment Ended Split by Member Status Level

Variable	Non-Gold members (Fixed effects)	Gold members (Fixed effects)	Non-Gold members (Robust SE)	Gold members (Robust SE)
Committed × During intervention	0.21	0.40*	0.21	0.40
	(0.21)	(0.22)	(0.22)	(0.26)
Committed × After intervention	−0.09	0.77***	−0.09	0.77**
	(0.21)	(0.22)	(0.31)	(0.30)
Committed			0.03	−0.09
			(0.29)	(0.22)
During intervention	−0.59***	−0.64***	−0.59***	−0.64***
	(0.04)	(0.04)	(0.04)	(0.05)
After intervention	−0.35***	−0.24***	−0.35***	−0.24***
	(0.04)	(0.04)	(0.06)	(0.05)
Constant	6.75***	8.03***	6.74***	8.04***
	(0.02)	(0.02)	(0.06)	(0.04)
Observations	42,300	31,014	42,300	31,014
Clusters	2,350	1,723	2,350	1,723

Notes. The dependent variable is the log number of monthly points. Experimental condition is used as the instrument for commitment status in the analyses. "Gold members" includes all Gold and Diamond members based as of January 1, 2012.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

Table 9. One-Hundred-Point Constant Sum for Assigned Credit/Blame for Precommitment Success/Failure

	Success	Failure
All survey completers		
My own personal choices	31.9	33.4
My household's choices	22.6	24.2
The availability of healthy food at Pick n Pay	27.5	26.5
The vitality program	18.0	15.8
N	1,362	1,362
Committed households only		
My own personal choices	33.2	32.7
My household's choices	22.4	23.6
The availability of healthy food at Pick n Pay	27.5	27.4
The vitality program	16.8	16.3
N	235	235

To better understand the intervention's effects on Vitality engagement, and confirm that there were no negative feelings among households who were invited to participate, we conducted a follow-up survey (N = 1,362, 33% response rate) with the original sample. With regard to the precommitment program, participants were asked to distribute 100 points among the four entities to whom they would assign credit/blame in a given month where they met/failed to meet their goal. As shown in Table 9, the survey respondents assigned most of the credit/blame to themselves, not to Vitality.[9] This suggests that the intervention did not hurt customer loyalty because participants took ownership of the responsibility to exercise self-control, meaning that any success or failure to keep the cash-back bonus was not attributed to an unfair or arbitrary penalty imposed by the firm.

Knowing that people take personal responsibility for their performance is an important finding, because a behavioral intervention that leads to better health outcomes at the expense of reduced customer loyalty or higher customer churn is ultimately unviable. The current research suggests that such backlash is less problematic than firms intuitively suspect, and, moreover, reinforces that the marketplace can tolerate stricter self-control devices without compromising customer loyalty. In addition to taking responsibility for meeting their precommitment goals, the majority of survey responders said they would welcome commitment devices in a variety of health domains. This was equally true of households who committed and those who did not.

4. Discussion and Conclusions

People often engage in unhealthy behaviors—such as smoking, overeating, and lack of exercise—that threaten their quality of life and longevity. On top of their direct health effects, these unhealthy habits also have considerable societal costs. For example, almost 35% of U.S. adults are obese (Ogden et al. 2014), which adds approximately $150 billion in annual medical costs (Finkelstein et al. 2009) and leads to over $40 billion annually in lost workplace productivity (Finkelstein et al. 2010). Governments and private firms, who largely foot the bill for medical care, have responded with increased interest in using financial incentives to improve health (Mattke et al. 2013), a practice that is likely to become more prevalent under new healthcare laws (Claxton et al. 2013).

Although there is evidence demonstrating that incentive-based health interventions can change a target behavior while the incentives are in effect (Gneezy et al. 2011), much less is known about their impact on nontargeted secondary behaviors. If incentives merely shift bad behavior from one domain to another or lead to some other negative consequence, then an overall improvement in health or reduction in healthcare costs is unlikely and firms will be unwilling to implement them.

In this paper, we examined the extended effects of an incentive-based health intervention to determine whether balancing behavior undermined the program's success and viability. Our results showed long-lasting positive persistence effects, and no evidence of negative spillover effects in terms of health or customer loyalty. These findings both illuminate the overall impact of a targeted health intervention and provide important evidence that the extended consequences of programs that motivate health with financial incentives may not be as negative as some theories predict.

Our first question focused on what happened once the penalty and threat of a loss ended. Although some studies have shown persistence effects in exercise (Acland and Levy 2015, Charness and Gneezy 2009, Royer et al. 2015) and smoking cessation (Giné et al. 2010, Volpp et al. 2009), these effects are not the norm (John et al. 2011; Volpp et al. 2008a, b, 2006) and the circumstances under which such health interventions persist are not well known. Our results show that the improvement in healthy shopping persisted for the six months following the intervention, which adds to the nascent literature showing that incentive-based interventions can have long-lasting positive effects (Gneezy et al. 2011). We also expand this finding into the new, and increasingly important, health domain of nutrition. Nutritional choice has become an essential health behavior to target because increased caloric intake is the main cause of the rise in obesity (Cutler et al. 2003). Of course, even the nonobese benefit from more nutritious food intake, which further highlights that targeting food purchases can be broadly appealing. The fact that the effect persisted suggests to interventionist firms that targeting nutrition behaviors is a ripe area for health interventions because it can lead to long-lasting change at a potentially reasonable cost.

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

Unlike some previous demonstrations of persistence, we found that the magnitude of the effect was very similar during and after the intervention period. Our estimate of the average effect on the treated was 2.84 percentage points during the six postintervention months, which is 76% of the change observed while the six-month incentive was in place. This differs markedly from exercise-based interventions, which have also demonstrated persistent effects, but of much smaller magnitude once the incentives are removed. For example, Acland and Levy (2015) found an increase of 1.21 gym visits per week while the incentives were in place, but this effect was reduced to 0.26 gym visits per week during the two immediate months following the intervention, an effect 21% the size of that while the incentives were in place. Our results raise the intriguing possibility that changes in food purchasing behavior may persist more easily than changes in exercise behavior. Although habits can obviously be formed in both domains, the habit to exercise still requires some degree of ongoing self-control because each exercise event requires separate time and effort from other behaviors. Grocery shopping habits, on the other hand, may be more routine and automatic once they have gone through an initial change. For example, once a shopper gets in the habit of purchasing skim milk rather than whole milk, she or he may continue to do this without much thought, and even after there is no financial incentive to do so. Such "mindlessness" is often seen as a challenge to healthy nutrition (Chandon and Wansink 2012, Wansink 2007), but here we have seen that it can be leveraged into making healthier choices part of the script.

This discussion raises the broader question of why the effect persisted in the first place. Prior research has suggested that persistence may result from various psychological mechanisms such as habits and changes in beliefs (Frey and Rogers 2014, Rogers and Frey 2016). Although we did not design the intervention specifically to test these theories, our results do provide some interesting insights about persistence. Specifically, we found that the effect persisted among the most active participants (high status). Moreover, we found that behavior while the incentives were in place was indicative of whether the effects would persist or not. It appears that the effect did not persist for some subpopulations because they did not respond to the incentives in the first place, not because the incentive worked while it was place but failed to create long-lasting change. This interpretation has a number of interesting implications for designing interventions with persistence in mind. For example, low-status participants may have been attracted to the commitment device but found that the incentive was not large enough to actually motivate a behavior change. In light of this, a firm could institute a larger penalty that is set to make sure everyone is sufficiently motivated during the intervention and consequently leads to habit change and persistence. If even setting larger penalty amounts fails to improve behavior, firms may be better aware of who does and does not respond to financial incentives in any form—which can lead to the development of different approaches for different subgroups. Future research should continue to examine not only whether incentives work on average, but for whom they work best or perhaps do not work at all.

Providing a more in-depth analysis of the extended and lasting effects of an incentive-based intervention begins to address the question of optimal incentive design. This intervention lasted a fixed amount of time (six months), but if the goal is lasting change, it could conceivably last longer. Indeed, Charness and Gneezy (2009) showed that an intervention needs to last long enough to create a habit in order for it to persist. The results here build on that finding by showing that the incentive also needs to be strong enough to motivate the initial change in behavior, and that once this has been achieved, the habit can persist for quite some time. The question of optimal intervention length is especially interesting for negative incentive schemes, which likely cost the firm nothing, and therefore could remain in place indefinitely if firms and participants wanted them to.[10] Our results suggest that whether firms should offer lengthy, indefinite, or extendable commitment contracts depends on both the impact of the contract's length on initial take-up and whether it is long and strong enough to create persistent effects. Moreover, even though our experiment was implemented in the health domain, these insights for encouraging lasting change could easily be applied to other domains as well.

In addition to examining persistence, we tested for two types of spillover effects that concern firms and policymakers, since their existence could greatly undermine this approach toward health interventions. First, in spite of the various theories suggesting that negative spillover effects may intrude into related domains, our results showed no evidence of compensation in the domain of exercise. Moreover, even though the effects were not significant, we observed a consistent positive effect of the intervention. This raises the intriguing possibility that incentives may not only positively impact a target behavior, but act as a catalyst for broader health improvement behaviors. Future research should continue to examine the important question of health spillovers in addition to questions about which related behaviors are most likely to be affected by them, and whether the framing of the intervention could be manipulated to encourage positive spillovers.

Finally, we examined the implications for the firm in terms of customer loyalty, and again found no negative effects (and some evidence of positive effects) of

RIGHTSLINK

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

running a voluntary commitment device that penalized customers for failing to meet their goals. This is also an important and previously unexamined finding; most firms would be unwilling to provide programs that improved health behaviors at the cost of customer loyalty. Interestingly, the effects of the intervention were strongest and most persistent among high-status members—essentially those most valuable to the firm and who themselves had the most to lose. These findings really highlight how people develop strategies for better self-control, including leveraging the threat of a loss, and for many members, an actual loss. The fact that the effects of penalizing people for failing to meet their own goals did not affect customer loyalty is an important step in broadening our understanding of how sophisticated individuals (O'Donoghue and Rabin 1999, 2001) self-regulate. As the results of the target intervention showed, there is substantial appetite for a commitment device. The results reported here show that appealing to the self-aware with such strategies not only avoids negative reactance to penalties (Brehm 1966, Clee and Wicklund 1980), but actually helps people develop and maintain healthier habits over time.

4.1. Conclusion

Financial incentive programs are an increasingly attractive way to improve health. Although research from the behavioral sciences shows that many such interventions effectively improve a specific health behavior while in place, less is known about the extended effects for both the participants and the host organizations. These effects are essential to understand when bringing behavioral science insights to scale. Our results suggest that these secondary consequences for the self and the firm may not be as dire as some theories predict. As such, these results should offer some reassurance for firms and governments interested in implementing incentive-based behavioral interventions. Moreover, we hope that these results will encourage future research to explore extended effects to better understand the circumstances that lead to positive or negative spillover effects.

Acknowledgments

The authors thank department editor Uri Gneezy, the anonymous associate editor, and the reviewers for their helpful feedback during the review process. The authors also thank Shaun Matisonn, Cindy Mtsweni, Craig Nossel, Jithen Paramanund, and Lauren Wyper for help executing the research.

Endnotes

[1] Diamond and Gold status members do not differ in the number of total points earned in a given year. Diamond status simply represents members who have achieved Gold status three or more consecutive years and are thus most loyal to the program. Vitality program status is not necessarily an indicator of health.

[2] This is a graphical representation of the intention to treat analysis in Table 2. We plot the monthly average percentage of healthy food purchased minus the preintervention average.

[3] This is a graphical representation of the intention to treat analysis in Table 5. We plot the monthly average log exercise events minus their preintervention average.

[4] Log of number of events plus one, so that zero is transformed to zero, and not dropped from the data.

[5] We also examined spillover effects by testing the effect of mean change in percent of healthy shopping from baseline to intervention period on the mean change in log exercise from baseline to intervention period, using experimental condition as an instrument for change in percent of healthy shopping. Table A3 in the online appendix shows the results of these analyses, which mirror those in Table 5.

[6] Log of number of points plus one, so that zero is transformed to zero, and not dropped from the data.

[7] This is a graphical representation of the intention to treat analysis in Table 7. We plot the monthly average log points minus their preintervention average.

[8] Figure A2 in the online appendix shows the distributions and nonparametric tests comparing these.

[9] Heckman corrections for survey response bias based on condition, commitment status, Vitality status, baseline healthy %, commitment success rate, choice to opt out, and money lost during the commitment did not substantively alter these results.

[10] An indefinite precommitment contract may reduce take-up. It may also lead unsuccessful committers to persist on an ineffective path.

References

Acland D, Levy MR (2015) Naiveté, projection bias, and habit formation in gym attendance. *Management Sci.* 61(1):146–160.

Baumeister RF, Bratslavsky E, Muraven M, Tice DM (1998) Ego depletion: Is the active self a limited resource? *J. Personality Soc. Psych.* 74(5):1252–1265.

Bem DJ (1972) *Self-Perception Theory* (Academic Press, New York).

Benabou R, Tirole J (2003) Intrinsic and extrinsic motivation. *Rev. Econom. Stud.* 70(3):489–520.

Brehm JW (1966) A Theory of Psychological Reactance (Academic Press, New York).

Cahill K, Perera R (2011) Competitions and incentives for smoking cessation. *Cochrane Database Syst. Rev.* 4:CD004307.

Chandon P, Wansink B (2012) Does food marketing need to make us fat? A review and solutions. *Nutrition Rev.* 70(10):571–593.

Charness G, Gneezy U (2009) Incentives to exercise. *Econometrica* 77(3):909–931.

Claxton G, Rae M, Panchal N, Damico A, Whitmore H, Bostick N, Kenward K (2013) Health benefits In 2013: Moderate premium increases in employer-sponsored plans. *Health Affairs* 32(9): 1667–1676.

Clee MA, Wicklund RA (1980) Consumer behavior and psychological reactance. *J. Consumer Res.* 389–405.

Cutler DM, Glaeser EL, Shapiro JM (2003) Why have Americans become more obese? *J. Econom. Perspectives* 17(3):93–118.

Deci EL, Koestner R, Ryan RM (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psych. Bull.* 125(6):627–668.

Dhar R, Simonson I (1999) Making complementary choices in consumption episodes: Highlighting versus balancing. *J. Marketing Res.* 36(1):29–44.

Dick AS, Basu K (1994) Customer loyalty: Toward an integrated conceptual framework. *J. Acad. Marketing Sci.* 22(2):99–113.

Donatelle RJ, Hudson D, Dobie S, Goodall A, Hunsberger M, Oswald K (2004) Incentives in smoking cessation: Status of the field and implications for research and practice with pregnant smokers. *Nicotine and Tobacco Res.* 6(Suppl 2):S163–S179.

RIGHTSLINK

Finkelstein EA, daCosta DiBonaventura M, Burgess SM, Hale BC (2010) The costs of obesity in the workplace. *J. Occupational Environ. Medicine* 52(10):971–976.

Finkelstein EA, Trogdon JG, Cohen JW, Dietz W (2009) Annual medical spending attributable to obesity: Payer-and service-specific estimates. *Health Affairs* 28(5):w822–w831.

Fishbach A, Dhar R (2005) Goals as excuses or guides: The liberating effect of perceived goal progress on choice. *J. Consumer Res.* 32(3):370–377.

Frey BS, Jegen R (2001) Motivation crowding theory. *J. Econom. Surveys* 15(5):589–611.

Frey BS, Oberholzer-Gee F (1997) The cost of price incentives: An empirical analysis of motivation crowding-out. *Amer. Econom. Rev.* 87(4):746–755.

Frey E, Rogers T (2014) Persistence. How treatment effects persist after interventions stop. *Policy Insights from the Behavioral and Brain Sci.* 1(1):172–179.

Giné X, Karlan D, Zinman J (2010) Put your money where your butt is: A commitment contract for smoking cessation. *Amer. Econom. J.: Appl. Econom.* 2(4):213–235.

Gneezy U, Meier S, Rey-Biel P (2011) When and why incentives (don't) work to modify behavior. *J. Econom. Perspectives* 25(4): 191–209.

Grégoire Y, Fisher RJ (2008) Customer betrayal and retaliation: When your best customers become your worst enemies. *J. Acad. Marketing Sci.* 36(2):247–261.

John LK, Loewenstein G, Troxel AB, Norton L, Fassbender JE, Volpp KG (2011) Financial incentives for extended weight loss: A randomized, controlled trial. *J. General Internal Medicine* 26(6): 621–626.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica: J. Econometric Soc.* 47(2): 263–292.

Khan U, Dhar R (2006) Licensing effect in consumer choice. *J. Marketing Res.* 43(2):259–266.

Lepper MR, Greene D, Nisbett RE (1973) Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *J. Personality Soc. Psych.* 28(1): 129–137.

Luo X (2007) Consumer negative voice and firm-idiosyncratic stock returns. *J. Marketing* 71(3):75–88.

Mattke S, Liu H, Caloyeras JP, Huang CY, Van Busum KR, Khodyakov D, Shier V (2013) Workplace wellness programs study. Research report, RAND Corporation, Santa Monica, CA.

Monin B, Miller DT (2001) Moral credentials and the expression of prejudice. *J. Personality Soc. Psych.* 81(1):33–43.

Muraven M, Baumeister RF (2000) Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psych. Bull.* 126(2):247–259.

O'Donoghue T, Rabin M (1999) Doing it now or later. *Amer. Econom. Rev.* 89(1):103–124.

O'Donoghue T, Rabin M (2001) Choice and procrastination. *Quart. J. Econom.* 116(1):121–160.

Ogden CL, Carroll MD, Kit BK, Flegal KM (2014) Prevalence of childhood and adult obesity in the United States, 2011–2012. *J. Amer. Medical Assoc.* 311(8):806–814.

Rogers T, Frey E (2016) Changing behavior beyond the here and now. Keren G, Wu G, eds. *Blackwell Handbook of Judgment and Decision Making* (Wiley-Blackwell, Hoboken, NJ), 725–748.

Royer H, Stehr MF, Sydnor JR (2015) Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a Fortune-500 company. *Amer. Econom. J.: Appl. Econom.* 7(3):51–84.

Schwartz J, Riis J, Elbel B, Ariely D (2012) Inviting consumers to downsize fast-food portions significantly reduces calorie consumption. *Health Affairs* 31(2):399–407.

Schwartz J, Mochon D, Wyper L, Maroba J, Patel D, Ariely D (2014) Healthier by precommitment. *Psych. Sci.* 25(2):538–546.

Smith AK, Bolton RN, Wagner J (1999) A model of customer satisfaction with service encounters involving failure and recovery. *J. Marketing Res.* 36(August):356–372.

Sturm R, An R, Segal D, Patel D (2013) A cash-back rebate program for healthy food purchases in South Africa: Results from scanner data. *Amer. J. Preventive Medicine* 44(6):567–572.

Thaler RH, Benartzi S (2004) Save More Tomorrow™: Using behavioral economics to increase employee saving. *J. Political Econom.* 112(S1):S164–S187.

Volpp KG, John LK, Troxel AB, Norton L, Fassbender J, Loewenstein G (2008a) Financial incentive—based approaches for weight loss. *J. Amer. Medical Assoc.* 300(22):2631–2637.

Volpp KG, Loewenstein G, Troxel AB, Doshi J, Price M, Laskin M, Kimmel SE (2008b) A test of financial incentives to improve warfarin adherence. *BMC Health Services Res.* 8(1):272.

Volpp KG, Levy AG, Asch DA, Berlin JA, Murphy JJ, Gomez A, Sox H, Zhu J, Lerman C (2006) A randomized controlled trial of financial incentives for smoking cessation. *Cancer Epidemiology Biomarkers and Prevention* 15(1):12–18.

Volpp KG, Troxel AB, Pauly MV, Glick HA, Puig A, Asch DA, Galvin R, Zhu J, Wan F, DeGuzman J (2009) A randomized, controlled trial of financial incentives for smoking cessation. *New England J. Medicine* 360(7):699–709.

Wansink B (2004) Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annual Rev. Nutrition* 24:455–479.

Wansink B (2007) *Mindless Eating: Why We Eat More Than We Think* (Bantam, New York).

Wansink B, Chandon P (2014) Slim by design: Redirecting the accidental drivers of mindless overeating. *J. Consumer Psych.* 24(3):413–431.

Wisdom J, Downs JS, Loewenstein G (2010) Promoting healthy choices: Information versus convenience. *Amer. Econom. J.: Appl. Econom.* 2(2):164–178.

Downloaded from informs.org by [155.246.103.35] on 05 April 2017, at 17:49 . For personal use only, all rights reserved.

RIGHTSLINK