Information Systems Research

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

Gamified Challenges in Online Weight-Loss Communities

Behnaz Bojd, Xiaolong Song, Yong Tan, Xiangbin Yan

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

INFORMATION SYSTEMS RESEARCH
Articles in Advance, pp. 1–19
ISSN 1047-7047 (print), ISSN 1526-5536 (online)

Gamified Challenges in Online Weight-Loss Communities

Behnaz Bojd,[a] Xiaolong Song,[b] Yong Tan,[c] Xiangbin Yan[d,]*

[a] Paul Merage School of Business, University of California, Irvine, California 92697; [b] School of Management Science and Engineering, Dongbei University of Finance and Economics, Dalian 116025, China; [c] Michael G. Foster School of Business, University of Washington, Seattle, Washington 98195; [d] School of Economics and Management, University of Science and Technology Beijing, Beijing 100083, China
*Corresponding author

Contact: behnazb@uci.edu, https://orcid.org/0000-0002-6115-9774 (BB); xlsong@dufe.edu.cn, https://orcid.org/0000-0001-7248-2764 (XS); ytan@uw.edu, https://orcid.org/0000-0001-8087-3423 (YT); xbyan@ustb.edu.cn, https://orcid.org/0000-0002-8115-6191 (XY)

Received: April 5, 2018
Revised: August 6, 2019; August 27, 2020; September 18, 2021
Accepted: October 3, 2021
Published Online in Articles in Advance: February 1, 2022

https://doi.org/10.1287/isre.2021.1081

Abstract. Gamified challenges, one of the most popular features of online weight-loss communities, enable users to set weight-loss goals and compete with other challenge participants via leaderboards. In this paper, using the data from a leading online weight-loss community, we study the effect of gamified challenges on the weight-loss outcome. We employ a dynamic model, using a system generalized method of moments estimator combined with an inverse probability weighting approach, to address endogeneity issues. Our findings indicate that participation in gamified challenges has a positive and significant effect on weight loss. We found that, on average, the participants achieved a weight loss of 0.742 kilograms (kg) by participating in at least one challenge a month. We demonstrate that not all gamified challenges are equally effective; effective challenges do not include a numeric weight goal (e.g., lose 5 kg), focus on exercise-only behavioral goals, and have a large active group size. Further, the results show that the absence (presence) of a numeric weight goal benefits users in exercise (diet) challenges. Moreover, a small active group size can help (hurt) users in exercise (diet) challenges. We discuss, as a potential underlying mechanism, the role of leaderboards to induce social comparison and motivate (discourage) users in exercise (diet) challenges. Our findings have implications for designing gamified systems with competition elements in online weight-loss communities.

History: Ram Gopal, Senior Editor; Sanjukta Smith, Associate Editor.
Funding: This work was supported in part by the National Natural Science Foundation of China [Grants 71729001, 72025101, 71601034, and 71771040].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/isre.2021.1081.

Keywords: weight-loss challenge • gamification • social comparison • weight-loss goals • system generalized method of moments • inverse probability weighting

1. Introduction

Overweight and obesity have been associated with many adverse physical and psychological health conditions (National Institutes of Health 2002, Wilson et al. 2002). These conditions have become highly prevalent, costing healthcare systems billions of dollars per year and resulting in a significant economic crisis (Hill and Peters 1998, Flegal et al. 2010, Cawley and Meyerhoefer 2012). With the rise of the Health 2.0 movement, online weight-loss communities have emerged as a new cost-effective tool to support individuals to lose weight. These communities connect users with similar weight-loss goals and provide them with access to many online tools, including self-monitoring tools, forums for social support, online educational materials, virtual coaches, reminder emails or text messages, and goal-setting tools. More recently, online weight-loss communities have incorporated gamification tools. *Gamification* is defined as the use of game design elements in nongame contexts (Deterding et al. 2011). Design elements, such as leaderboards, points, and badges, are used to increase users' engagement in the community and enhance users' motivation to pursue their weight-loss goals.

In this paper, we focus on measuring the effect of gamified challenges, a popular but understudied feature of online weight-loss communities. Gamified challenges let users set weight-loss goals and compete with other challenge participants, using gamification elements such as leaderboards that rank users based on their weight-loss outcome. Gamified challenges became popular after the hit television reality show *The Biggest Loser*, in which contestants compete to lose weight and win a cash prize. Despite the popularity of gamified challenges in online weight-loss communities, no research has examined or quantified their effectiveness on weight-loss outcomes. A few papers have shown the positive effect of walking competitions on users' physical activity, as measured by number of steps (Chen and Pu 2014, Shameli et al. 2017). There is, however, an important difference between walking competitions and weight-loss challenges. Based on goal-setting theory, walking is a simple task in which learning new strategies is not necessary to control the outcome (Locke 1968). In a simple task, effort and concentration have a direct effect on the outcome. Thus, ranking individuals based on numbers of steps can motivate them to exert more effort and achieve better outcomes. Weight loss, in contrast, is a complex task that requires skilled individuals to discover appropriate strategies before they can successfully manage the outcome (Locke and Latham 2002). Thus, rankings based on weight-loss outcomes may at times discourage individuals of slow progress. Users may lose their focus on adopting healthy lifestyle changes and, thus, try different strategies in an unsystematic way and fail to learn which strategy is effective. Therefore, it is not clear whether participation in gamified challenges can be effective for weight-loss outcomes. In this paper, we seek to determine the effect of gamified challenges on users' weight-loss outcomes.

We empirically examine the effect of gamified challenges using the data from a leading online weight-loss community in the United States. With an emphasis on user-generated content, this online weight-loss community lets users track their weight, nutrition, and physical activity and engages them with a supportive community via general forums and specific groups. In mid-2008, gamified challenges were introduced as a new feature on this weight-loss community. We randomly chose users and tracked them for eight months, from April to November 2008 (four months before and four months after the introduction of challenges). In our data, we have access to the average weight that each user reports during a month, and we can calculate their monthly weight-loss performance. Further, we can observe whether a user participated in a challenge. We also can follow users' engagement with other features in the platform, tenure on the platform, and personal weight-loss goals.

In this study, we focus on two main components of gamified challenges: goal setting and gamification. First, gamified challenges enable users to set two types of goals: performance goals or a numeric weight-loss target, for example, "lose 5 kilograms (kg)," and behavioral goals or a set of dietary or physical activity strategies, such as "restrict daily caloric intake by 1,250 calories" or "walk 10,000 steps a day." In our data, all challenges have a behavioral goal; however, they may not have a numeric weight goal. In addition, the behavioral goals in a challenge can focus on exercise, diet, or a combination of exercise and diet. Second, the leaderboards, as a gamification tool, can motivate users to pursue their goals in the weight-loss challenges. In our setting, when users report their weight, the leaderboard shows their weight-loss outcome to the other challenge participants and ranks them based on their performance. Therefore, leaderboards can induce social comparison and competition. Research has shown that competitive motivation diminishes with the number of participants (Garcia and Tor 2009). Thus, we observe the number of participants in each challenge and the number of active members who reported their weight at least once during the challenge. The effect of weight-loss challenges is based on each of these two components and their characteristics and interaction. In this study, we are interested in estimating the heterogeneous effect of challenges with different characteristics.

For our empirical analysis, we have four estimation challenges in regard to the causal effect of users' participation in challenges on their weight-loss outcome. First, an individual's weight at each point in time is highly dependent on its previous value. Second, weight loss/gain is determined by a variety of unobserved elements, such as individuals' motivation, genetic predisposition, and gender. Third, the unobserved fixed or time-varying factors could be correlated with users' decision to participate in a challenge each month. To overcome these three concerns, we employ a dynamic model, using a system generalized method of moments (GMM) estimator (Blundell and Bond 1998). The system GMM estimation utilizes instruments *within* the model, and the validity of these instruments is verified using Hansen and Arellano-Bond tests. Moreover, the system GMM model is useful for accurately estimating the inertial effects of the lagged dependent variable in the dynamic model and to forecast the users' weight change for a duration after they finish a challenge. The final estimation challenge concerns the fact that some individuals do not participate in any challenge during our panel time frame. These nonadopters do not appear in our system GMM analysis. As a result, our estimation may be biased due to incidental sample truncation if nonadopters' decision to participate is nonrandom. To address this bias, we combine the system GMM estimation with the inverse probability weighting (IPW) approach. The validity of the IPW approach relies on the ignorability assumption. Although it is not possible to directly test this assumption, we use an indirect verification test. We provide many robustness checks, and we compare our method with a difference-in-differences model with time-varying treatment, coupled with propensity score matching.

Our main findings indicate that participation in gamified challenges has a positive and significant effect on weight loss. Users can achieve a weight loss of 0.742 kg by participating in at least one challenge a month. We demonstrate that not all gamified challenges are equally effective; effective challenges do not include a numeric weight goal, focus on exercise-only goals, and have a large active group size. Further, the results show that the absence (presence) of a numeric weight goal benefits users in exercise (diet) challenges. Moreover, a small active group size can help (hurt) users in exercise (diet) challenges. These results can be explained by the role of leaderboards in inducing social comparison. Consistent with the goal-setting and social comparison literature, we find that the absence of a numeric goal and small group size can induce higher levels of social comparison. Thus, we can infer that higher levels of social comparison can motivate (discourage) users in exercise (diet) challenges.

The main contributions of this study are as follows. To begin, we are the first to quantify the effect of participation in gamified challenges on weight-loss outcomes using a system GMM estimator combined with an IPW approach. We find that, on average, the participants achieved a weight loss of 0.742 kg by participating in at least one challenge per month. Further, we show that not all challenges are equally effective; specifically, the effect of weight-loss challenges depends on the type of promoted goal and the competitive structure in the challenge. Second, we contribute to the literature that measures the effect of dietary and physical activity goals on weight loss. In contrast to earlier findings, our results indicate that challenges with physical activity goals have a positive effect on weight loss, whereas challenges with diet goals or combined goals are not effective, on average. Third, we contribute to the literature on the effect of gamification on health outcomes by showing that high levels of social comparison induced by leaderboards may not benefit all behavioral goals.

Our findings have implications for designing successful gamified systems. Our results suggest that gamification elements that induce competition should be used with caution in goal-setting environments, particularly when gamifying dietary goals. Online weight-loss communities can recommend a useful combination of numeric weight goals, behavioral goals, and optimal number of participants in each challenge to induce an encouraging level of social comparison.

The remainder of this paper is organized as follows. In Section 2, we discuss the related literature. We introduce the setting and data in Sections 3 and 4. In Section 5, we discuss our empirical model. We present our results and a discussion of the underlying mechanisms in Section 6. In Section 7, we provide robustness checks. Finally, in Section 8, we present a summary of our main findings, limitations, and suggestions for future research.

2. Literature Review

First, our paper relates to the long-standing literature on the effect of eHealth interventions on weight loss. eHealth interventions employ websites, mobile apps, emails, text messages, and digital games, using devices such as PCs, personal digital assistants, tablets, mobile/smartphones, and smart wears (Neve et al. 2010, Greene et al. 2013, Napolitano et al. 2013, Steinberg et al. 2013). In a meta-analysis, Hutchesson et al. (2015) evaluated studies from 1995 to 2014 on the effectiveness of eHealth weight-loss interventions that demonstrated significant weight loss. Goal setting, self-monitoring, education via virtual coach or online material, expert feedback (Annesi 1998, McKay et al. 2001), social support (Hwang et al. 2010, Poncela-Casasnovas et al. 2015, Yan 2018), and reminder emails or text messages (Stevens et al. 2008, Patrick et al. 2009) are among the tools used in online weight-loss interventions.

In this study, we focus on measuring the effect of weight-loss challenges, a popular but understudied online weight-loss feature. Goal setting and gamification are the main components of a weight-loss challenge. In such a challenge, users pursue weight-loss goals with other challenge participants over a fixed period. Further, challenge leaderboards rank users based on their performance and induce competition and motivate challenge participants. In this study, we are the first to observe a positive effect of online challenges on weight loss. We find that, on average, the participants achieved a weight loss of 0.742 kg by taking part in at least one challenge per month. Further, we show that not all challenges are equally effective; rather, their effect depends on the type of promoted goal and the competitive structure in the challenge.

Second, our paper is related to a stream of literature that focuses on the effect of goal setting on weight loss. Research has shown that the overweight and obesity epidemic is caused largely by environmental factors that promote excessive food intake and a sedentary lifestyle (Hill and Peters 1998).[1] Thus, it is not surprising that most weight-loss interventions aim to set dietary and physical activity behavioral goals (Alexy 1985, Burke et al. 2002, Jeffery et al. 2003). Table A1 in the online appendix provides a detailed list of studies that used random assignment to groups with and without a goal-setting component or random assignment to groups with different goal characteristics for changing dietary and physical activity behaviors in adults. Although goal setting has been a part of many weight-loss interventions, an examination of the characteristics of an effective goal rarely has been the focus of the weight-loss intervention research (Strecher et al. 1995, Shilts et al. 2004, Pearson 2012). Further, among diet-only, exercise-only, and combined interventions, it is unclear which behavioral goals are the most effective. One study suggested that combined programs were more effective for weight loss than were diet-only programs (Avenell et al.

2004), whereas another study found no significant differences (Curioni and Lourenco 2005). Further, Johns et al. (2014) showed that weight loss is similar in the short term for diet-only and combined intervention programs; however, in the longer term, weight loss is greatest in combined interventions. Moreover, they found that physical activity goals alone are less effective than are combined goals in both the short and long term.

In our study, each challenge promotes a set of weight-loss goals. The weight-loss goal can be a numeric weight-loss target (e.g., lose 5 kg), a behavioral goal related to physical activity or diet (e.g., walk 10,000 steps a day, restrict daily caloric intake by 1,250 calories), or a combination of both. In contrast to earlier findings, our results indicate that challenges with physical activity goals have a positive effect on weight loss, whereas challenges with diet goals or combined goals are not effective, on average. Further, we show that the presence (absence) of a numeric weight goal benefits users in exercise (diet) challenges. We surmise that the differences in our results as compared with earlier findings are due to the use of leaderboards and induced social comparison processes in our study.

Third, our paper is related to a stream of literature on the impact of gamification on health and well-being. Gamification has been defined as the use of game design elements in nongame contexts (Deterding et al. 2011). Gamification objects, such as points, badges, levels, and leaderboards, are used to improve experiential outcomes, for example, user engagement, as well as the context-related outcomes of the gamified system, for example, increasing physical activity in an exercise app (Liu et al. 2016). Gamification research covers health-related outcomes, including physical activity (Chen and Pu 2014, Riva et al. 2014, Zuckerman and Gal-Oz 2014, Allam et al. 2015, Hamari and Koivisto 2015, Maher et al. 2015, Shameli et al. 2017), nutrition (Jones et al. 2014, Berger and Jung 2021), substance use (alcohol; Boendermaker et al. 2015), medication use (Riva et al. 2014), healthcare utilization (Allam et al. 2015), and anxiety and depression (Dennis and O'Toole 2014). This body of research has established either a positive, neutral, or mixed effect of gamification on the health outcomes and lacks high-quality studies (Johnson et al. 2016).

In this study, we examine the effect of gamification on enhancing the weight-loss outcome. Since most weight-loss interventions aim to change dietary and physical activity behaviors, we provide a detailed list of studies that used random assignment to groups with and without gamification for changing dietary and physical activity behaviors in Table A2 in the online appendix. As summarized in Table A2, some studies have established a positive effect of gamification on behavior change and user engagement. For example, research has shown that incorporation of gamification elements, such as points, badges, medals, and leaderboards, in a rheumatoid arthritis website not only increased the users' engagement with the website but also increased their physical activity and decreased their healthcare utilization (Allam et al. 2015). Notably, several examples of unsuccessful gamification attempts highlight that gamification is not just a matter of using points, badges, and leaderboards and that different dynamics can emerge from the interaction between these gamification elements and user actions.

In our setting, gamified challenges use leaderboards to rank participants based on their weight-loss outcomes. Leaderboards can create social comparison and harness individuals' competitive instincts (Festinger 1954, Bui et al. 2015, Liu et al. 2016). The literature suggests that social comparison and competition can yield both positive and negative effects on different outcomes. In educational settings, research has shown that competition can draw students' attention, improve their motivation, and produce better results for practical applications of concepts (Cheng et al. 2009, Domínguez et al. 2013, Garcia et al. 2006, Hanus and Fox 2015), whereas other studies have found that competition can produce worse results on written assignments and participation (Domínguez et al. 2013). Competition may not always foster all of the desired outcomes, because, when competitive evaluations are emphasized, they can reduce individuals' intrinsic motivation, create anxiety, and impede performance (Kohn 1992, Reeve and Deci 1996, Tauer and Harackiewicz 2004). Moreover, studies have found that the effect of competition on different outcomes depends on the skill level of one's peers (Liu et al. 2013, Santhanam et al. 2016). Santhanam et al. (2016) showed that competing with lower-skilled individuals can improve self-efficacy and learning outcomes and that facing equal or higher-skilled competitors can increase engagement.

In weight-loss settings, a few papers have shown that a weight-loss competition between employees at work sites is effective when monetary incentives are coupled with a physical leaderboard that shows employees' weekly progress (Brownell et al. 1984, Stunkard et al. 1989). Further, relative to targets of the same weight, weight-focused social comparisons to both thinner and heavier individuals lead to increased thoughts of dieting and exercising, and comparisons to thinner targets increases the likelihood of engaging in actual dieting and exercising behaviors (Rancourt et al. 2015). Uetake and Yang (2020) showed that the average and worst performers have a negative (i.e., discouraging) effect and the top performers have a positive (i.e., encouraging) effect on an individual's weight-loss performance. They also found that group size is a moderator of the negative competition effect on weight loss;

that is, the average performer effect is less harmful as the group size increases. Unlike these studies, no financial incentives are available in our study. Further, unlike work sites and in-person meetings, most users do not know each other in the online platform.

In this paper, we provide support for the positive (negative) effect of social comparison on weight loss when individuals pursue exercise (diet) behavioral goals in online weight-loss challenges. Here, we rely on different challenge characteristics to infer the level of social comparison and competition intensity in different weight-loss challenges. In goal-setting theory, competition is viewed as a unique form of goal setting, whereby the goal is related to the performance of other people rather than a standard target and is dynamic rather than static, as it changes due to the performance of other people (Locke and Latham 1985). Thus, we believe that in the absence of a numeric weight-loss goal, challenge leaderboards can create a higher level of social comparison. Further, prior research shows that competitive motivation diminishes with the number of participants (Garcia and Tor 2009). This is because the social comparison process is stronger in smaller groups and gets diluted in larger groups. Based on earlier findings, we use the existence of a numeric weight-loss goal and the active group size to determine the intensity level of social comparison in a challenge. Our results show that the presence (absence) of a numeric weight goal benefits users in exercise (diet) challenges. Moreover, a small active group size can help (hurt) users in exercise (diet) challenges. As noted, the absence of a numeric goal and a small group size can induce higher levels of social comparison. Thus, we can infer that higher levels of social comparison can motivate (discourage) users in exercise (diet) challenges. Our work contributes to the gamification literature by showing that social comparison may not benefit all behavioral goals.

3. Setting

We focus on one of the largest noncommercial online weight-loss communities, launched in 2006. The platform has over 45 million users, who want to either lose or maintain their weight. The platform lets users track their weight, nutrition, and physical activity and engage with a supportive community. To join and use this online weight-loss community, users need to create a profile and provide their current weight and goal weight. They also can update their weight changes. If users set their profile as public, their weight update will be posted on the platform's home page. If users choose a private profile, their weight update will be shared only on their friends' home page. Those who see the weight-loss update can like or comment on it.

This online community provides multiple weight-loss features. There is a "journal" feature on users' profile pages in which they can log in their calorie intake, report exercise activities, write about their daily weight-loss journey, and post pictures. The journal updates are accessible to everyone if the user's profile is public and accessible to only the user's friends if the user's profile is private. The "general forums" are another feature, where users can ask/provide answers to general questions, such as, "How often do you weigh yourself?" Further, users with common interests or concerns can create/join "group" forums, share their experiences with weight management, and exchange diet or exercise tips. An example of a group forum is a group for "low-carb lovers." In mid-2008, weight-loss "challenges" were introduced as a new feature on the platform. This feature enables users to set different weight-loss goals. Any user can create a challenge or join the challenges created by the other users on the platform. The creator of the challenge sets a fixed start date and duration, a name, and a set of instructions for the challenge. Other users can join the challenge only before its starting time. Users can create or join any number of challenges. The name and the instructions of a challenge may include a numeric weight-loss goal, such as "lose 5 kg." In addition, the name and the instructions of a challenge may include exercise or diet behavioral goals, such as, "Exercise 30 minutes a day" or, "Quit soda." Examples of challenge names and their instructions are presented in Table 1.

In our setting, weight-loss challenges are gamified via leaderboards. Leaderboards show all participants' weight-loss performance and rank it based on its progress percentage, which is calculated based on the weight-loss amount divided by the user's weight at the beginning of the challenge. As shown in Figure 1, the performance of those who have lost (gained) weight is shown in green (red). If challenge participants do not report their weight, then their names are listed at the end of the leaderboard with blank weight-loss information. Note that the challenge leaderboard does not show any direct information on how far the participants are from the challenge's numeric target (if any target is defined, e.g., lose 5 kg). The platform utilizes leaderboards as a gamification object to motivate challenge participants. Other gamification objects, such as status points, badges, or monetary prizes, are not used on this platform.

Finally, each challenge is integrated with a challenge forum, in which participants can interact with each other, share their experiences, ask questions, provide answers, and encourage each other. These challenge forums are similar to general forums, but nonchallenge participants do not have access to writing or replying. Challenge participants can initiate any number of

Table 1. Examples of Challenge Names and Instructions

Challenge name	Challenge instructions
Lose 5 kg	• Try to go without sugar. • Consume zero-calorie sweetener only. • Exercise 30 minutes a day.
Avoid the major fast-food chains	• If you absolutely need to eat at a fast-food restaurant, order a salad instead of a sandwich. • Substitute the major unhealthy chains with a relatively healthy chain.
Walk-walk-walk	• Do not use the elevator; use the stairs. • Park farther away and walk.

posts in the challenge forum. Each post may receive multiple replies from the challenge participants. This online weight-loss platform has a free website in more than 30 countries as well as a free mobile app version. The platform's mobile app enables users to log in their weight, count their intake calories, and upload pictures of their food. The general and group forums and challenges, however, are not accessible via the mobile app. Our data come from the platform's website in the United States.

4. Data

4.1 User Data

Online weight-loss challenges were introduced on the platform in August 2008. We randomly chose 4,208 users who had joined the platform at least four months before the introduction of challenges, and we tracked their online activity for eight months, from April 2008 to the end of November 2008. Users can choose to report their weight any number of times on the platform. We have access, however, to only the average weight that each user reports during a month, from which we can calculate their monthly weight-loss outcome. We calculate the monthly weight-loss outcomes for only 1,045 out of 4,208 users who have reported their weight in at least two consecutive months during our study period. Therefore, we focus on these 1,045 users in our study.

We define $weight_{it}$ as the average weight that user i reports during month t (in kilograms), and we define our dependent variable as $weightLoss_{it}$, which equals $weight_{it-1} - weight_{it}$. As summarized in Table 2, we have 4,719 observations for $weight_{it}$ and 3,159 observations for $weightLoss_{it}$, which shows that we have an unbalanced panel; that is, the weight-loss outcome is missing for some users at some periods. Further, the statistics show that the median user is around 80.75 kg and has lost 0.7 kg per month. Figure 2 shows that the distribution of $weightLoss_{it}$ is close to normal, and users report weight loss as well as weight gain.

We can observe the users' decision to join a challenge each month. We define our main independent variable $challenge_{it}$ as a binary variable, which indicates whether user i participates in at least one challenge at month t. As shown in Table 2, $challenge_{it}$ is highly right-skewed, with a median of zero. Therefore, we use its logarithm transformation in our analysis. Further, we define $numChallenge_{it}$ to count the number of challenges that user i participates in at month t. As summarized in Table 2, users have participated in a minimum of zero and a maximum of seven challenges simultaneously per month. In Section 4.2, we provide a detailed description of the different types of challenges created in this online weight-loss community. Note that we do not observe the users' rankings on the challenge leaderboards.

Figure 1. (Color online) Leaderboard on a Challenge Page

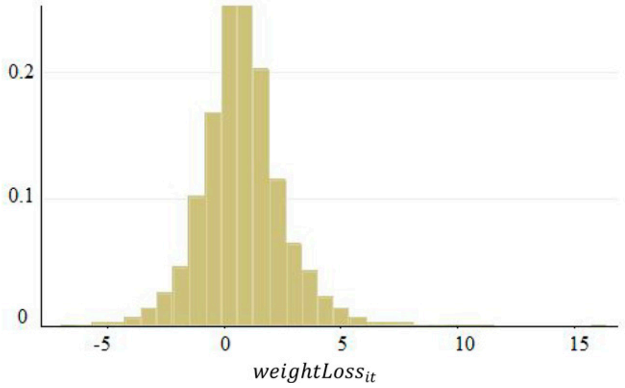
Table 2. Summary Statistics for 1,045 Users

Variable	Mean	Standard deviation	Median	(Min, Max)	Sample size
$weight_{it}$	84.452	21.296	80.75	(41.1, 199.6)	4,719
$weightLoss_{it}$	0.801	1.798	0.700	(−7, 16.425)	3,159
$challenge_{it}$	0.062	0.241	0	(0, 1)	8,360
$numChallenge_{it}$	0.085	0.382	0	(0, 7)	8,360
$numReport_{it}$	1.637	2.781	1	(0, 31)	8,360
$journal_{it}$	0.578	2.292	0	(0, 32)	8,360
$generalForum_{it}$	0.357	2.704	0	(0, 75)	8,360
$groupForum_{it}$	0.043	0.631	0	(0, 24)	8,360
$challengeForum_{it}$	0.084	1.309	0	(0, 93)	8,360
$tenure_{i1}$	5.041	4.254	4	(0, 19)	1,045
$goal_i$	68.587	12.595	66.7	(43, 124.7)	1,045

Moreover, we can observe each user's engagement on the platform, such as the number of times that user i reports weight during month t ($numReport_{it}$) and the number of posts that user i writes in the journal, group forums, general forums, and challenge forums ($journal_{it}$, $groupForum_{it}$, $generalForum_{it}$, and $challenge Forum_{it}$). As shown in Table 2, these variables are highly right-skewed, with a median of zero. Therefore, we use their logarithm transformation in our analysis. Such engagements on the platform can facilitate self-monitoring, communication, and social support, positively influencing weight-loss performance (Khaylis et al. 2010). Thus, we control for these variables in our analysis.

We also observe and control for the number of months since user i first joined the online weight-loss community at time t ($tenure_{it}$). As shown in Table 2, the median user had been on the platform for four months at the beginning of our data collection period. We also observe and control for users' personal weight goals on their profile ($goal_i$). As shown in the table, a median user has a weight goal of 65.8 kg. As previous literature suggests, individuals' goal weight has an impact on their weight-loss journey (Elfhag and Rössner 2005, Nelissen et al. 2011, De Vet et al. 2013).

Figure 2. (Color online) Average Monthly Weight-Loss Distribution



4.2. Challenge Data

Users participated in 96 different challenges, and, in our data, we can observe the challenge creator and challenge duration. On average, 25 challenges are created every month. We denote challenge duration by $duration_c$. As shown in Table 3, a median challenge is 42 days. The challenge duration is indirectly captured by $challenge_{it}$, which shows whether user i participates in any challenge, either a new challenge started at month t or an ongoing one started in a previous month but not finished at month t.[2] Moreover, we can observe the creation date and start date of a challenge. The duration between the challenge creation date and the start date is denoted by $join\_duration_c$. As shown in the table, a median challenge has five days between its creation and start date. Note that users can join a challenge only before the challenge start date.

Next, we observe the number of members (or participants) in a challenge (denoted by $total\_member_c$). As summarized in Table 3, the number of challenge participants varies between one and 169 members, with a median challenge as having 24 members. We can also observe the number of active members who reported their weight at least once during the challenge (denoted by $active\_member_c$). As shown in Table 3, the median challenge has 12 active members. Moreover, we can observe the number of posts initiated in a challenge forum (denoted by $post_c$). As shown in Table 3, the median challenge has three posts initiated on its forum. Next, we observe whether the challenge objective includes a numeric weight target (e.g., lose 5 kg). There are 15 (16%) challenges with a weight target and 81 (84%) challenges without a target. Finally, we observe that 21 challenges (22%) focus on exercise-only goals, 32 challenges (33%) focus on diet-only goals, 35 challenges (36%) include both exercise and diet goals, and 8 challenges (9%) have instructions that are not related to exercise or diet.[3]

As noted earlier, we define a user-month-level variable ($challenge_{it}$) to study the effect of users' participation in at least one challenge during a month on their weight-loss performance. Similarly, we define user-

Table 3. Summary Statistics for 96 Challenges

Variable	Mean	Standard deviation	(Min, Max)	25th	Median	75th
$duration_c$	43.052	23.357	(7, 84)	28	42	56.0
$join\_duration_c$	6.052	3.385	(1, 15)	4	5	8.0
$total\_member_c$	30.510	29.885	(1, 169)	9.5	24	36.5
$active\_member_c$	16.458	15.291	(0, 77)	5.5	12.5	22.5
$post_c$	5.614	6.404	(0, 40)	2	3	7.0
$target_c$		Target: 16%	No target: 84%			
$instruction_c$		Exercise: 22%	Diet: 33%		Both: 36%	Other: 9%

month-level variables for different types of challenges to study the heterogeneous effect of challenges. To evaluate the heterogeneous effect of challenges based on the type of the behavioral goals, we define four binary variables—$Exercise_{it}$, $Diet_{it}$, $Exercise\_Diet_{it}$, and $Other_{it}$—that indicate whether user i participates in at least one challenge with the corresponding goal type at month t. Note that users can participate in all four types simultaneously. Thus, these four binary variables are not exclusive.

Next, to study the heterogeneous effect of challenges with and without a target, we define Tar_{it} as a binary variable that indicates whether user i participates in at least one challenge with a numeric weight target at month t. Similarly, we define $NTar_{it}$ as a binary variable that indicates whether user i participates in at least one challenge without a numeric weight target at month t. Note that users can participate in both of these challenges simultaneously. Thus, Tar_{it} and $NTar_{it}$ are not exclusive binary variables. Finally, we define $HAMem_{it}$ as a binary variable that indicates whether user i participates in at least one challenge with more than 12 active members at month t. Similarly, we define $LAMem_{it}$ as a binary variable that indicates whether user i participates in at least one challenge with less than or equal to 12 active members at month t. Note that $HAMem_{it}$ and $LAMem_{it}$ are not exclusive binary variables.

5. Empirical Methodology

5.1. System GMM (Blundell-Bond) Estimator

To analyze the effect of participation in challenges on the user's weight loss, we employ a dynamic model in which the dependent variable, user's weight at any given time ($weight_{it}$), is modeled as a linear function of

$$weight_{it} = c + \alpha weight_{it-1} + \beta challenge_{it} + X_{it}\Phi + \gamma z_i + m_t + \eta_i + e_{it} \qquad t = 2, \ldots, 8, \tag{1}$$

where $challenge_{it}$ is the main variable of interest, a binary variable that indicates whether user i participates in at least one challenge at month t. The time-varying variables are denoted by the vector X_{it}, which captures user i's engagement on the platform during month t, including the log of number of times that the user reports his or her weight ($numReport_{it}$); the log of the number of posts that user i writes in a journal, group forums, general forums, and challenge forums ($journal_{it}$, $groupForum_{it}$, $generalForum_{it}$, and $challengeForum_{it}$); and the number of months since user i first joined the online weight-loss community at time t ($tenure_{it}$) and its squared term. User i's fixed weight goal is denoted by z_i. The month dummies are denoted by m_t. The user-specific fixed effect is denoted by η_i. Finally, we indicate a mean-zero error term by e_{it}.

In the estimation of Equation (1), several econometric problems may arise: (1) the main variable of interest, $challenge_{it}$, and other explanatory variables are potentially correlated with an individual's unobserved time-invariant characteristics (η_i), such as age, gender, and height; (2) the lagged dependent variable ($weight_{it-1}$), although not correlated with the current error term (e_{it}), is a predetermined variable and correlated with previous shocks;[4] and (3) the main variable of interest, $challenge_{it}$, and the explanatory variables that capture the users' activity on the platform, X_{it}, are potentially correlated with time-varying shocks (e_{it}). For example, a random shock at the user's motivation level may result in the user's self-selecting him or herself into participating in a challenge.

Problem (1) results in a biased ordinary least squares (OLS) estimate.[5] An initial remedy to problem (1) is to use the fixed-effects model. Due to problem (2), however, and because we have a short panel (T = 8) with many users (n = 253), the fixed-effects model will be biased (Nickell 1981).[6] To resolve this issue, we use a system GMM estimation that employs valid instrumental variables (IVs) within the model (Blundell and Bond 1998). The system GMM starts by using transformed regressors as valid instruments in the level Equation (1)—for example, $\Delta weight_{it-1}$ (or deeper lags) as an IV for $weight_{it-1}$, and $\Delta challenge_{it-1}$ (or deeper lags) as an IV for $challenge_{it}$.[7] In addition to the level Equation (1), the system GMM uses a difference equation by transforming all regressors, usually by "first-differences," to eliminate fixed effects (η_i) as follows:[8]

$$\Delta weight_{it} = \alpha \Delta weight_{it-1} + \beta \Delta challenge_{it} + \Delta X_{it}\Phi + \Delta m_t + \Delta e_{it} \qquad t = 2, \ldots, 8. \tag{2}$$

In the differenced Equation (2), we can use the lags of regressor as a valid instrument; that is, we can use lag 1 and deeper for predetermined variables (e.g., $weight_{it-2}$ as an IV for $\Delta weight_{it-1}$) and lag 2 or deeper for endogenous variables (e.g., $challenge_{it-2}$ as an IV for $\Delta challenge_{it}$). Finally, using Equations (1) and (2), we can apply GMM to calculate a consistent and efficient estimator. Thus, using valid IVs can mitigate both endogeneity concerns in problems (2) and (3).

Utilizing instruments within the model for the GMM estimator is useful when exogenous IVs are not available. Compared with the traditional instrumental variable approach, however, we should proceed with caution when making a causal interpretation. The validity of the system GMM estimation relies on the validity of its IVs. For example, in the level equation, $\Delta challenge_{it-1}$ is a valid IV for $challenge_{it}$ if we assume that it is correlated with $challenge_{it}$ but not correlated with η_i and e_{it}. Similarly, in the difference equation, $challenge_{it-2}$ is a valid IV for $\Delta challenge_{it}$ if we assume that it is correlated with $\Delta challenge_{it}$ but not correlated with Δe_{it}. We use the Hansen test to examine the validity of the group of IVs. Moreover, we assume that the e_{it}'s are independent and identically distributed across i and across t (no serial correlation), that is, $e_{it} \sim iid(0, \sigma_i^2)\ \forall i, t$. The validity of assuming no serial correlation is tested and verified by using the Arellano-Bond test (Arellano and Bond 1991).[9] In addition, the e_{it}'s can be heteroskedastic across individuals (σ_i^2). We can ensure the robustness of the results to this heteroscedasticity using a two-step system GMM. The full set of assumptions of system GMM is summarized in Section A2 of the online appendix.

In our analysis, it is easier to interpret the dependent variable $weightLoss_{it}$ instead of $weight_{it}$, which is defined as

$$weightLoss_{it} = weight_{it-1} - weight_{it}. \tag{3}$$

Therefore, we can easily transform Equation (1) to

$$weightLoss_{it} = -c + (1-\alpha)weight_{it-1} - \beta challenge_{it} - X_{it}\Phi - \gamma z_i - m_t - \eta_i - e_{it}. \tag{4}$$

This transformation does not affect the use of the system GMM approach explained earlier. To run the system GMM model, we use *xtabond2*, a Stata command (Roodman 2009a). With this command, we use a *two-step* GMM and use the *robust* option to apply the Windmeijer (2005) finite-sample correction to fix the *two-step* GMM standard errors' downward bias.

5.2. Inverse Probability Weighting (IPW)

In our data, 235 (out of 1,045) users participated in at least one challenge during the four months after the introduction of challenges. We categorize them as challenge adopters. Thus, 792 users (approximately 76%) never participated in any challenge during the data collection period (i.e., nonadopters). Note that $challenge_{it}$ equals zero for all nonadopters at all periods. Thus, when we estimate the effect of participation in challenges, the nonadopters' data drop from the system GMM analysis. Using only the adopters' data can result in a biased estimation, as an important kind of nonrandom selection, incidental sample truncation, arises when certain individuals do not appear in a random sample due to individual choices or behaviors (Wooldridge 2002).

An approach to consistent estimation in the presence of incidental sample truncation is based on IPW.[10] In the IPW approach, by considering the ignorability assumption, we assume that, conditioned on the observed variables, no unobserved variable exists that can affect both challenge adoption and weight-loss outcome; that is, adoption is exogenous. Although this assumption is not directly testable, in Section 7.4, we discuss a way to assess it indirectly. In the IPW approach, we use observed variables that are not affected by the introduction of challenges (pretreatment variables), including the user's goal, tenure on the platform, and the average engagement level with other features of the platform during the four months before the introduction of challenges, including the average frequency of platform use to report weight ($avgNumReoprt_i$), write in a journal ($avgJournal_i$), or post or respond to a comment on general forums ($avgGeneralForum_i$) and group forums ($avgGroupForum_i$). As presented in Table 4, we compare adopters' and nonadopters' (pretreatment) observed variables. There are significant differences between adopters' and nonadopters' online activity and tenure on the platform.

To employ IPW, we define $adopt_i$ as a dummy variable that indicates whether user i is a challenge adopter. Then, we use a *probit* model to calculate the probability of adoption based on users' observed variables before the introduction of challenges, using the following model:

$$\begin{aligned} adopt_i = {} & c_0 + c_1 avgNumReoprt_i + c_2 avgJournal_i \\ & + c_3 avgGeneralForum_i + c_4 avgGroupForum_i \\ & + c_5 tenure_{i1} + c_6 goal_i + e_i, \end{aligned} \tag{5}$$

The results of this *probit* model are summarized in Section A3 of the online appendix. The results show that users with high engagement in the platform (those who frequently report their weight and post in journals and group forums) are more likely to adopt challenges. Next, we use the inverse of the probability of adoption to weight the observations. We give less (more) weight to adopters who are more likely (less likely) to adopt the challenges by using the inverse probability of adoption as observation weights. The IPW approach ensures that the sample of adopters used in the model and nonadopters dropped from the model are similar (Table 4). Next, we incorporate

Table 4. Comparison of Adopters and Nonadopters Before the Introduction of Challenges

	Mean			Mean (weighted sample)		
Variable	Adopter (n = 253)	Nonadopter (n = 792)	$p > \|t\|$	Adopter (n = 253)	Nonadopter (n = 792)	$p > \|t\|$
$avgNumReoprt_i$	0.883	0.755	0.001	0.786	0.784	0.954
$avgJournal_i$	0.431	0.176	0.000	0.246	0.243	0.948
$avgGeneralForum_i$	0.287	0.099	0.000	0.145	0.143	0.937
$avgGroupForum_i$	0.016	0.001	0.000	0.005	0.003	0.539
$tenure_{i1}$	5.549	4.879	0.029	5.182	5.101	0.782
$goal_i$	68.399	68.648	0.785	68.760	69.683	0.934

these weights in the system GMM model, utilizing the Stata xtabond2 command.[11]

6. Results

6.1. Effect of Participation in Challenges

The results from our estimation exercise are presented in Table 5. We start with an OLS model and then modify it step by step to address all concerns discussed in Section 5.1, ending with the estimator of interest. The observations are weighted according to the IPW approach in all of these models. By applying OLS in Model 1, and a fixed-effects estimation in Model 2, we find a positive and significant effect of $challenge_{it}$, which suggests that participation in a challenge has a positive and significant effect on weight loss. As explained in Section 5.1, however, both OLS and fixed-effects models are biased due to the Nickell bias and the endogeneity concerns. In the table, we also report the autoregressive parameter (α) as the coefficient for the lagged dependent variable ($weight_{it-1}$). The unbiased estimation of α should fall within the range of OLS and fixed-effects models, that is, within the range (0.681, 0.983) (Roodman 2009a). The popular solution for the Nickell bias is to apply difference GMM (Arellano and Bond 1991). This method starts by transforming the data by first differencing and then instrumenting differenced regressors with their lags.[12] Applying the difference GMM method in Model 3, we find that the α's estimation (0.684) is inside the range. However, it is very close to the estimated α in the fixed-effects model (0.681), and it reflects the poor performance of the difference GMM model. Although the difference GMM method is shown to be suitable for "small T, large N" panels, it performs poorly in a small panel (T) with high inertial effects (α) (Blundell and Bond 1998).[13] In this situation, the system GMM estimator performs better than the difference GMM estimator (Blundell and Bond 1998).

Applying system GMM in Model 4, we find that the α's estimation (0.855) falls within the expected OLS fixed-effects range. This shows the higher performance of the system GMM model as compared with the difference GMM model. In addition, the Hansen test p-value (0.350) confirms that all of the instruments as a group are exogenous and valid in Model 4. Further, the autocorrelation AR(2) p-value (0.122) supports the assumption of no serial correlation. In Model 4, the estimated effect of $challenge_{it}$ is positive and significant.

In Model 5, we add the user's control variables, including the user's personal weight goal, tenure, and different activities on the platform. One of the advantages of the system GMM is that we can include time-invariant regressors in the model, such as $goal_i$, which would disappear in difference GMM. It is important to note that adding control variables requires more instruments and that too many instruments in the system GMM models can result in overfitting the endogenous variables (Roodman 2009b). One solution to limit the number of instruments in small samples is to collapse instrument sets by creating one instrument for each variable and lag distance, rather than one for each period, variable, and lag distance.[14] In Model 5 in Table 5, we reduce the number of instruments from 94 to 25 by collapsing them. Here, we find that the effect of $challenge_{it}$ remains positive and significant. There is, however, severe multicollinearity between $weight_{it-1}$ and $goal_i$. Excluding $goal_i$ in Model 6 does not change our results, and it is not expected to do so because our model controls for the fixed effects very well (see Section 5.1).

One advantage of the dynamic model is that it allows us to forecast the duration for which each user can maintain the lost weight after participating in a challenge. Consider a simulated example for which a user, who is around 75.742 kg, participates in at least one challenge at month 5. As shown in Figure 3, the user's weight will go down to 75 kg at month 6, but he or she will gain the lost weight back over the next 10 months if he or she does not participate in another challenge.[15]

6.2. The Effect of Different Challenge Characteristics

In our setting, we have 96 different challenges. Participation in all challenges may not have the same effect. As explained in Section 4.2, we divide challenges into

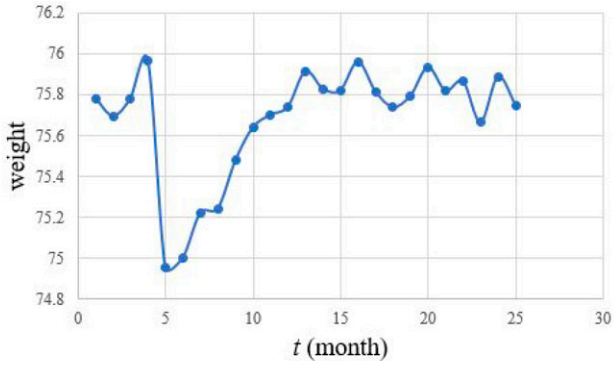
Table 5. Effect of Participation in Challenges on Weight Loss

Variable	OLS (Model 1)	FE (Model 2)	Difference GMM (Model 3)	System GMM (Model 4)	System GMM (Model 5)	System GMM (Model 6)
$challenge_{it}$	0.369***	0.314**	0.642*	0.742**	0.912***	0.934**
	(0.102)	(0.131)	(0.351)	(0.312)	(0.335)	(0.369)
$weight_{it-1}$[a]	0.983***	0.681***	0.684***	0.819***	0.830**	0.869***
	(0.002)	(0.021)	(0.107)	(0.035)	(0.028)	(0.023)
$goal_i$					−0.181***	—
					(0.033)	
$numReport_{it}$					1.581***	1.625***
					(0.521)	(0.510)
$journal_{it}$					−0.047	−0.210
					(0.360)	(0.339)
$generalForum_{it}$					−0.320	−0.107
					(0.423)	(0.423)
$groupForum_{it}$					0.952	0.828
					(0.751)	(0.788)
$challengeForum_{it}$					−1.483*	−1.042
					(0.791)	(0.789)
$tenure_{it}$					0.046	0.016
					(0.031)	(0.029)
$tenure^2_{it}$					−0.001	0.000
					(0.004)	(0.003)
constant	−0.993***	−26.347***	—	−14.434***	−2.681**	−12.065***
	(0.184)	(1.808)		(3.005)	(0.817)	(2.004)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,159	3,159	2,114	3,159	3,159	3,159
Individuals	1,045	1,045	689	1,045	1,045	1,045
IVs	—	—	14	24	25	24
AR(2) pvalue	—	—	0.180	0.122	0.461	0.474
Hansen pvalue	—	—	0.217	0.350	0.100	0.727

Note. Robust standard errors are reported in parentheses.
[a]α is reported.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

different groups based on the types of behavioral goals (exercise, diet, or a combination of both), the existence or absence of a performance goal (a numeric weight target), and the number of active members. To study the effect of different challenge characteristics, as shown in Table 6, we replace $challenge_{it}$ by different challenge categories and use the system GMM approach similar to Model 6.

Figure 3. (Color online) Weight Forecast for a User Who Participates in at Least One Challenge at Month 5



In Table 6, we start by studying the effect of different behavioral goals on weight-loss. In Model 7, we find a positive and significant effect for $Exercise_{it}$, showing that challenges with exercise-only goals are effective for weight loss. The effect of challenges with diet-only goals ($Diet_{it}$) is negative but insignificant. The effect of the combination of diet and exercise ($Exercise\_Diet_{it}$) is positive but significant only at the 0.1 level. Next, we study the effect of the existence or absence of a numeric weight target in a weight-loss challenge. As shown in Model 8, the effect of $NTar_{it}$ is positive and significant, indicating that participation in challenges without a numeric target is effective for weight loss. We then study the interaction effect between the existence of a numeric target and the type of behavioral goals. As seen in Model 9, the effect of exercise-only challenges remains positive, however, only when there are no numeric targets; that is, $NTar\_Exercise_{it}$ is positive and significant. We see the opposite pattern, however, for diet-only challenges. The effect of diet-only challenges is positive and significant only when a numeric target exists; that is, Tar_Diet_{it} is positive and significant. The effect of combination of diet and exercise is positive but significant

Table 6. Effect of Different Challenge Characteristics on Weight Loss

Variable	Model 7	Model 8	Model 9	Model 10	Model 11
$Exercise_{it}$	0.813**				
	(0.364)				
$Diet_{it}$	−0.300				
	(0.580)				
$Exercise\_Diet_{it}$	0.666*				
	(0.375)				
Tar_{it}		0.788			
		(0.901)			
$NTar_{it}$		0.676**			
		(0.336)			
$Tar\_Exercise_{it}$			0.382		
			(0.909)		
Tar_Diet_{it}			1.801**		
			(0.892)		
$Tar\_Exercise\_Diet_{it}$			0.676		
			(2.114)		
$NTar\_Exercise_{it}$			0.877**		
			(0.401)		
$NTar\_Diet_{it}$			−0.823		
			(0.695)		
$NTar\_Exercise\_Diet_{it}$			0.767*		
			(0.420)		
$HAMem_{it}$				0.818**	
				(0.357)	
$LAMem_{it}$				0.755	
				(0.644)	
$HAMem\_Excersie_{it}$					0.457
					(0.342)
$HAMem\_Diet_{it}$					0.290
					(0.711)
$HAMem\_Excersie\_Diet_{it}$					0.158
					(0.413)
$LAMem\_Excersie_{it}$					1.839**
					(0.861)
$LAMem\_Diet_{it}$					−2.407**
					(1.042)
$LAMem\_Exercise\_Diet_{it}$					2.135**
					(0.910)
Controls	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Observations	3,159	3,159	3,159	3,159	3,159
Individuals	1,045	1,045	1,045	1,045	1,045
IVs	30	26	35	26	36
AR(2) *pvalue*	0.524	0.613	0.847	0.586	0.355
Hansen pvalue	0.590	0.553	0.444	0.753	0.211

Notes. Controls include $weight_{it-1}$, $numReport_{it}$, $journal_{it}$, $generalForum_{it}$, $groupForum_{it}$, $other_{it}$ and interactions, $challengeForum_{it}$, $tenure_{it}$, $tenure^2_{it}$, and the *constant* term. Robust standard errors are reported in parentheses.

$^*p < 0.10$; $^{**}p < 0.05$.

only at the 0.1 level when there are no numeric targets. In sum, the presence of a numeric target can help participants in diet-only challenges, but the absence of a numeric target helps participants in exercise-only challenges.

Next, we study the effect of the number of active members, who reported their weight at least once during the challenge. In Model 10, we find that participation in challenges with a high number of active members ($HAMem_{it}$) has a positive and significant effect on weight loss. Next, we examine the interaction effect between the number of active members and the type of behavioral goals. As seen in Model 11, the effect of exercise challenges with a low number of active members is positive and significant ($LAMem\_Execise_{it}$). We see the opposite pattern, however, for the diet challenges. The effect of diet challenges with a low number of active members is negative and significant ($LAMem\_Diet_{it}$). Further, the effect of the combination of diet and exercise is positive but only significant with a low number of active members. In sum, the small active group size can hurt users in the diet-only challenges but is helpful for participants in the exercise-only and exercise-and-diet combination challenges.[16]

6.3. Discussion of Mechanism

In our analysis, in Table 5, we found a positive and significant effect for participation in challenges on the weight-loss outcome. As explained in Section 5, bias related to self-selection into a challenge can be strong. Those who are motivated to lose weight might be more likely to participate in a challenge. Thus, we combined an IPW approach with the system GMM method to reduce self-selection bias. This can help us to interpret the causal effect of participation in challenges versus nonparticipation.

In Table 6, we examined the effect of different challenge characteristics on weight-loss outcome to shed light on the underlying mechanisms driving the weight-loss challenges' effectiveness. In Model 7, we focused on different weight-loss behavioral goals. We found that participating in challenges with exercise-only goals has a positive and significant effect on weight loss. The effects of diet-only and combination challenges, however, are insignificant. There is evidence in the literature that goal setting for increasing exercise is more effective than is goal setting for dietary changes in the short term. In addition, the combination of exercise and diet might be too demanding for the average user (Booth et al. 2008).

We argue that bias related to self-selection into a challenge with exercise-only goals should be weak. We have no reason to believe that users with a higher motivation to lose weight may choose exercise challenges over diet challenges. The nature of motivation, however, can differ among users who choose physical activity versus dietary behavioral goals (Teixeira et al. 2012). Physical activities can be driven by intrinsic motivation, because most sports and physical activities can be a great source of enjoyment and provide a source of optimal challenge (Csikszentmihalyi 1990). According to self-determination theory, engaging in activities that are intrinsically motivated is central to fulfilling human beings' fundamental needs for competence and autonomy (Ryan and Deci 2000). Research has shown that physical activity intrinsic motivation is a strong predictor of successful weight management (Silva at al. 2008, Silva et al. 2010). Most dietary goals, however, are driven by extrinsic motivations (e.g., getting a result, appearance, compliance with others' expectations, and not feeling guilty). For example, a low-calorie or a reduced-carbohydrate diet are less likely to be explored for their enjoyable experience and, instead, valued only for their results. Indeed, it has been shown that rigid control of eating behavior negatively predicts weight-loss success (Elfhag and Rössner 2005, Teixeira et al. 2010).

Leaderboards can create social comparison and competitive instincts (Festinger 1954, Bui et al. 2015, Liu et al. 2016). Studies have shown that competition can create anxiety and impede performance (Kohn 1992, Reeve and Deci 1996). Users in exercise challenges may benefit from higher levels of intrinsic motivation to overcome anxiety and adhere to pursuing their exercise goals. Users in diet challenges with lower levels of intrinsic motivation, however, may be less likely to manage competition anxiety. Carter and Jansen (2012) show that anxiety can lead to maladaptive eating habits (e.g., eating prompted by stress or negative emotions). Maladaptive eating habits are the main cause of "yo-yo" dieting and a barrier to losing weight (Cooper and Fairburn 2001). Indeed, surveys have shown that leaderboards are a less-desired gamification element in nutrition apps versus physical activity apps, indicating that the competitive spirit is more pronounced in a sports context (Berger and Jung 2021). Thus, next, we shed light on the potential underlying mechanism, namely, social comparison induced via leaderboards, that drives the effect of exercise versus diet challenges. We use two different challenge characteristics to infer the level of social comparison intensity in different weight-loss challenges.

First, we focus on the existence of a performance goal, namely, a numeric weight-loss target. For Model 9, the results show that the presence (absence) of a numeric target can help participants in diet-only (exercise-only) challenges. One potential mechanism can be tied to how gamified challenges display the participants' performance via leaderboards and induce social comparison. In the absence of a target, leaderboards can induce a higher level of social comparison, because users may focus more on other participants' performance to evaluate their own progress. In the presence of a target, however, users can evaluate their progress toward the target and focus less on other participants' performance. Thus, the induced social comparison can be weak in challenges with a target. Based on goal-setting literature, we believe that a higher (lower) level of social comparison is induced by the absence (presence) of a target (Locke and Latham 1985). Therefore, we can infer that a lower (higher) level of social comparison can be helpful for diet-only (exercise-only) challenges. Note that the majority of available challenges do not include a specific numeric target in our setting. Thus, bias related to self-selection into challenges without a target would be weak.

Second, we focus on active group size, that is, the number of users who reported their weight at least once during the challenge. For Model 11, we find that participating in challenges with a small active group size can hurt (help) users in diet-only (exercise-only) challenges. The effect of active group size on weight loss can be linked to how leaderboards work in a challenge. In our study, leaderboards reveal the users' weight-loss outcome only if the users report their weight. Thus, leaderboards can create social comparison processes by exposing challenge participants to the weight-loss performance of the active members. Therefore, regardless of the total number of members in a challenge, only active members can induce social comparison processes via leaderboards. Research has shown that the process of social comparison becomes diluted by a large number of participants (Garcia and Tor 2009). In other words, leaderboards can induce a higher level of social comparison in challenges with a smaller active group size. Thus, we can infer that a higher level of social comparison can hurt (help) users in diet-only (exercise-only) challenges. These results are consistent with what we found earlier for Model 9.

We argue that bias related to self-selection into challenges with a specific number of active members should be weak. As explained in Section 4.2, for a median challenge, there are five days between the challenge creation and start date, and users cannot join the challenge after the start date. Therefore, it is less likely that users with higher motivation would self-select into a challenge with a specific size, as the number of members in a challenge can change until the start date of the challenge. Similarly, individuals cannot choose to join a challenge with a specific active group size, because before users choose to join a challenge, they do not know the number of active members who will report their weight at least once during the challenge. Thus, we argue that bias related to self-selection into a challenge with a specific active group size should not be a concern.

Finally, the results for Model 11 show that a small active group size benefits users in challenges with a combination of exercise and diet goals. We can infer that a higher level of social comparison, induced in challenges with a small active group size, encourages the participants in combination challenges. This result is consistent with those in Model 9, for which we find that the absence of a numeric target and potentially higher levels of social comparison can benefit users in combination challenges (significant at the 0.1 level). As discussed earlier, we have no reason to believe that users with higher motivation may choose exercise-only over diet-only challenges. Users with higher motivation or self-efficacy beliefs, however, may be more likely to choose a more difficult challenge that combines dietary and exercise goals over exercise-only or diet-only challenges. Research has shown that, when goals are self-set, people with high self-efficacy set higher goals than do those with lower self-efficacy (Seijts and Latham 2001) and that users with stronger self-efficacy beliefs can benefit from competition (Santhanam et al. 2016).

7. Robustness Checks

7.1. Difference in Differences

We compare our system GMM results with a difference-in-differences model with time-varying treatment coupled with propensity score matching. The difference-in-differences approach and results are explained in detail in Section A4 of the online appendix. The difference-in-differences results show that the effect of challenges on weight loss is positive and marginally significant ($p < 0.10$). In addition, the magnitude of the effect is smaller and closer to the magnitude of a biased fixed-effect model. We argue that a dynamic model is more suitable for our study due to the high inertial effects of the lagged dependent variable.

7.2. Ignorability Assumption

We use the IPW approach to address the incidental sample truncation bias. The main assumption of the IPW approach is the ignorability assumption. By considering the ignorability assumption, we assume that, conditioned on the observed variables, no unobserved variable exists that can affect both challenge adoption ($adopt_i$) and weight-loss outcome; that is, adoption is exogenous. For example, in our setting, one major unobserved factor that may affect both challenge adoption and the weight-loss outcome is the user's unobserved motivation level.

The ignorability assumption is not directly testable. There are, however, ways to assess it indirectly. One test relies on estimating the effect of the treatment on a variable known to be unaffected by it, typically because its value is determined before the treatment. The lagged (pretreatment) outcome variables are best for this test, because they are closely related to the outcome of interest (Imbens and Wooldridge 2007). If the treatment effect on the lagged outcome variable is estimated to be close to zero, it is more plausible that the unconfoundedness assumption holds. Implementing this test, we estimate the effect of $adopt_i$ on $weightLoss_{it}$ for periods before the introduction of challenges on the platform ($t \leq 4$), using the weighted sample. The intuition behind this test is that, if ignorability does not hold and if being an adopter is positively correlated with high motivation to lose weight, then we should see a significantly higher weight loss for adopters before the introduction of challenges.

As shown in Table 7, for Model 12, $adopt_i$ does not significantly affect the weight-loss outcome before the introduction of challenges. Thus, we can conclude that the ignorability assumption is plausible. Note that, in Model 12, due to the small number of periods, the Hansen test p-value and the autocorrelation AR(2) p-value are not calculated. As a robustness check, in Model 13, we run the same analysis using OLS and found qualitatively similar results.

7.3. Challenge Creators

The challenge creators might be more motivated than others to obtain a higher performance in their challenge. Fourteen users created all of the 96 challenges in our data. We can remove the challenge creators from our analysis to check our results' robustness. As shown in Table 7, Model 14, the results remain qualitatively similar to those of Table 5, Model 6.

7.4. Challenge Fixed Effects

Including the challenge dummies in our analysis can control the correlations across individuals who choose to participate in a specific challenge. Controlling for 96 different challenge dummies in a system GMM model, however, generates too many instruments and less-reliable results. It is, however, easy to add these many dummies in a fixed-effects model. As shown in

Table 7. Robustness Checks for the Estimation of the Effect of Participation in Challenges on Weight Loss

Variable	Ignorability Model 12	Ignorability Model 13	Noncreators Model 14	Challenge FE Model 15
$challenge_{it}$	—	—	0.977**	0.642**
			(0.393)	(0.246)
$adopt_i$	−0.614	−0.119	—	—
	(0.982)	(0.108)		
Model	System GMM	OLS	System GMM	Fixed effects
Controls	Yes	Yes	Yes	Yes
Challenge FE	No	No	No	Yes
Month FE	Yes	Yes	Yes	Yes
Observations	1,633	1,633	3,091	3,159
Individuals	841	841	1,029	1,045
IVs	15	—	24	—
AR(2) pvalue	—	—	0.554	—
Hansen pvalue	—	—	0.653	—

Notes. Controls include $weight_{it-1}$, $numReport_{it}$, $journal_{it}$, $generalForum_{it}$, $groupForum_{it}$, $challengeForum_{it}$, $tenure_{it}$, $tenure^2_{it}$, and the *constant* term. Robust standard errors are reported in parentheses.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table 7, Model 15, the effect of $challenge_{it}$ remains positive and significant after controlling for the challenge dummies in a fixed-effects model. Thus, we can infer that the correlation across participants in a challenge is not a concern in our analysis.

7.5. Self-Reported Weight Outcomes

One potential source of endogeneity arises from the fact that users self-report their weight, and each user can misreport a lower weight. If this decision is not random and correlated with users' challenge participation, it can bias the estimated effect of challenge participation. This decision can create a nonrandom measurement error. For example, if those who experience a weight gain during participation in a challenge choose to misreport a weight loss, then the estimated effect of challenge participation will be biased. To address this concern about the user's decision to misreport a lower weight, we consider two cases. First, if an individual has a motivation to misreport his or her weight, and his or her behavior is consistent throughout the time, then this error can be captured by the unobserved fixed effect η_i in Equation (1). As our estimation method controls for η_i, the estimation of β will be unbiased. Second, if an individual's motivation to misreport varies over time, then we can model this behavior by breaking the error (e_{it}) to Δ_{it} and e'_{it}:

$$weight_{it} = c + \alpha weight_{it-1} + \beta challenge_{it} + X_{it}\Phi + \gamma z_i + m_t + \eta_i + e'_{it} - \Delta_{it} \tag{6}$$

where Δ_{it} captures the amount of weight that the individual reports lower than his or her true weight. In this case, if an individual misreports regardless of participating in a challenge, that is, if Δ_{it} is independent of $challenge_{it}$, then this error does not create a bias in the estimation of β. If, however, an individual systematically misreports higher weight loss when he or she is participating in a challenge, that is, if Δ_{it} is correlated with $challenge_{it}$, then the effect of participating will be biased and overestimated. We believe, however, that this correlation is unlikely for two reasons. First, individuals who participate in a challenge may misreport their weight to show that they have achieved the challenge target or stand in a better ranking position on the challenge leaderboard. With these two incentives, untruthful challenge participants may never report a "weight gain." Although we cannot directly measure whether the challenge participants are truthfully reporting their weight "amount," we can examine how often they report a weight gain. Among 253 challenge adopters, 76 users (approximately 30%) have reported a weight gain at least once during the challenge participation. The fact that a high number of challenge participants report a weight gain at least once increases the reliability of their self-reporting behavior. Moreover, we use a difference-in-differences approach to compare the behavior of reporting a weight gain between adopters and nonadopters before and after the introduction of challenges. As shown in Section A5 of the online appendix, there is no significant difference between the reporting behavior between adopters and nonadopters before and after the introduction of challenges. Thus, we can infer that users are less likely to misreport during challenge participation. One potential reason might be the absence of a prize for the winners of these challenges.

8. Conclusions

In this paper, we empirically examine the effectiveness of gamified weight-loss challenges, using the data from a leading online weight-loss community in the United States. Our findings indicate that participation in gamified challenges has a positive and significant effect on the weight-loss outcome. Users can achieve a weight loss of 0.742 kg by participating in at least one challenge a month, a healthy weight-loss rate, according to the CDC guidelines.[17] We discuss and clarify the methodological strategies required to analyze the dynamic nature of the weight-loss outcome and overcome endogeneity problems in nonexperimental settings, using a system GMM model combined with an IPW approach. We also compare our system GMM results with a difference GMM dynamic approach and a difference-in-differences model with time-varying treatment, coupled with propensity score matching. We argue that a dynamic system GMM model is more suitable for our study due to the high inertial effects of the lagged dependent variable.

Focusing on the characteristics of an effective challenge, we show that not all gamified weight-loss challenges are the same. Effective challenges do not include a numeric weight-loss target; they focus on exercise-only goals, and they have a large active group size. We also show interaction effects between the challenge goal characteristics and the challenge size. The results show that the absence (presence) of a numeric weight goal benefits users in exercise (diet) challenges. Moreover, a small active group size can help (hurt) users in exercise (diet) challenges. We discuss, as a potential underlying mechanism, the role of leaderboards to induce social comparison and motivate (discourage) users in exercise (diet) challenges.

The main contributions of this study are as follows. First, we quantify the effect of participation in gamified challenges on weight-loss outcome, using a system GMM method combined with an IPW approach. Despite the popularity of gamified challenges in online weight-loss communities, no research examines or quantifies their effectiveness on the weight-loss outcome. Second, we contribute to the literature that focuses on the effect of dietary and physical activity goals on weight loss. In contrast to earlier findings, our results indicate that challenges with physical activity goals have a positive effect on weight loss, whereas challenges with diet goals or combined goals are not effective, on average. We surmise that the differences in our results as compared with earlier findings are due to the use of leaderboards and induced social comparison processes in our study. Third, we contribute to the literature on the effect of gamification on health outcomes by showing that high levels of social comparison induced by leaderboards may not benefit all behavioral goals.

Our findings have implications for designing successful gamified systems. Our results suggest that gamification elements, such as leaderboards, that induce competition should be used with caution in goal-setting environments, particularly when gamifying dietary goals. To induce an encouraging level of social comparison, online weight-loss communities can recommend a useful combination of numeric weight goals, behavioral goals, and an optimal number of participants in each challenge based on the existence of a numeric weight goal and the type of behavioral goals.

Our study raises valuable questions for future research. First, based on our estimates from the dynamic system GMM model, we show the positive effect of participation in gamified challenges on short-term weight-loss performance. Based on our system GMM model estimates, we can roughly forecast the duration in which each user can maintain the lost weight after participating in a challenge. Due to our short panel time frame and because many users do not report their weight regularly in consecutive months, we have a data limitation related to introducing deeper lags in our dynamic model and more accurately evaluating the long-term effect of participation in challenges on weight-loss performance. Based on prior research, we believe that adherence to gamified challenges can breed long-term weight-loss achievement (Uetake and Yang 2018). It also has been shown that most people who succeed in losing weight over a short period regained a substantial amount of their lost weight after participation in the weight-loss competitions (Fothergill et al. 2016). Thus, it is important to quantify the long-term effect of participation in online gamified challenges. Further, in our study, less than 10% of the challenge adopters participated in challenges in all four months. The remainder of the challenge adopters skipped one or more months and did not continue to participate in new challenges. Investigating the effect of continuous engagement in gamified challenges on long-term weight-loss performance is, thus, an area for future research.

Second, future research can evaluate the heterogeneous effect of gamified challenges across different groups of users. Our system GMM model is robust to all of the fixed unobserved effects, such as gender and age. Due to the unavailability of such individual-level measures, however, we could not shed light on the heterogeneous effect of gamified challenges across different users. Third, we could not directly measure social comparison. Future research could investigate the effect of users' ranking on their future performance. Finally, in our data, we did not observe the network connection between users who were participating in a challenge. The presence of online friends in a weight-loss challenge might strengthen the effect of challenge rankings. Further, participating in online challenges can allow users to get to know successful users and follow them after the challenge. Future research can provide recommendations on the optimal user configuration of challenges, using such individual and online social network data.

Acknowledgments

The authors thank the senior editor, the associate editor, and the anonymous reviewers for constructive comments throughout the review process.

Endnotes

[1] Although studies have shown the importance of genetic factors in determining individual susceptibility to obesity, genetic factors cannot explain the obesity epidemic, as our genes have not changed substantially during recent decades (Hill and Peters 1998).

[2] If a challenge is finished within the first week of month t, then we do not consider it in $challenge_{it}$, but we count it in $challenge_{it-1}$. Similarly, if a challenge is started within the last week of month t, then we do not consider it in $challenge_{it}$, but we count it in $challenge_{it+1}$. There is no challenge in our data with a starting date in the last week of month t and a finishing date in the first week of the next month.

[3] For example, in a challenge called "Lean by Halloween," the instruction is to pick a costume and use it as motivation. Or, in another challenge, the instruction is to break down the excuses and recognize the truth.

[4] We also consider $tenure_{it}$ to be a predetermined variable, that is, independent of e_{it} but correlated with e_{i1}. Note that $tenure_{it}$ can be written as $tenure_{it} = tenure_{i1} + (t-1)$; thus, $tenure_{it}$ is a function of $tenure_{i1}$ and a deterministic time increment t. If users join the platform when they have higher levels of motivation, then $tenure_{i1}$ can be correlated with e_{i1}. Hence, $tenure_{it}$ can be correlated with e_{i1}.

[5] In an OLS model, the estimate of α is biased upward (downward) if we assume a positive (negative) correlation between $weight_{it-1}$ and η_i. Other OLS coefficient estimates also are biased (Trognon 1978).

[6] The within estimator is inconsistent, because the mean-differencing makes $y_{it-1} - \overline{y}_i$ correlated with the error $e_{it} - \overline{e}_i$, because y_{it-1} is correlated with e_{it-1} and with $\overline{e}_i$. Due to the negative correlation of y_{it-1} with the error term $e_{it} - \overline{e}_i$, the estimation of α in the fixed-effects model will be biased downward.

[7] We consider time dummies and challenge fixed effects as strictly exogenous variables in our setting, and we use them as their own IVs in the level equations.

[8] "First-differences" subtracts the previous observation from the contemporaneous one. Note that, in equation (2), "forward

orthogonal deviations" are used for month dummies. Forward orthogonal deviation subtracts the average of all future "available" observations of a variable (Arellano and Bover 1995). In our analysis, to minimize data loss, we employ orthogonal deviations for all variables.

[9] This test has a null hypothesis of no autocorrelation and is applied to differenced residuals. AR(1) tests the autocorrelation between Δe_{it} and Δe_{it-1}. Usually, AR(1) rejects the null, because Δe_{it} and Δe_{it-1} have e_{it-1} in common. AR(2) tests the autocorrelation between Δe_{it} and Δe_{it-2}. The AR(2) p-value is more important than that of AR(1), because a rejected null hypothesis in AR(2) reveals a serial correlation between errors, which shows that the validity of IVs is violated.

[10] Unlike Heckman's (1976) approach, the IPW approach does not require identifying exogenous variables to satisfy the exclusion restrictions.

[11] These weights are incorporated in the system GMM model by using the equation below:

$$\hat{\beta} = (X'WZAZ'WX)^{-1}X'ZAZ'WY$$

where, for N observations, Y is the outcome matrix, X is the matrix of regressors, Z is the matrix of instruments, and W is the matrix that holds the weights. See Roodman (2009a) for a more detailed explanation of how to use the Stata xtabond2 command and its option to add weights to incorporate weights in a GMM estimation.

[12] Lag 2 or deeper is for endogenous variables (e.g., $challenge_{it-2}$ for $\Delta challenge_{it}$) and lag 1 and deeper is for predetermined variables (e.g., $weight_{it-2}$ for $\Delta weight_{it-1}$). We use only lag 1 to limit the number of instruments.

[13] In dynamic panel models in which the autoregressive parameter (α) is moderately large and the number of time series observations is moderately small, the IVs in difference GMM (i.e., past levels of the regressors) convey little information about the transformed regressors (i.e., future changes). Therefore, the weak instruments make the difference GMM estimator perform poorly (Blundell and Bond 1998).

[14] For example, for the predetermined variable y_{it-1}, we can collapse the set of instruments from

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ y_{i1} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & y_{i2} & y_{i1} & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & y_{i3} & y_{i2} & y_{i1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ to } \begin{pmatrix} 0 & 0 & 0 & \cdots \\ y_{i1} & 0 & 0 & \cdots \\ y_{i2} & y_{i1} & 0 & \cdots \\ y_{i3} & y_{i2} & y_{i1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where the first row of the matrix corresponds to $t = 1$.

[15] Based on Equation (4), although the effect of $challenge_{it}$ on $weightLoss_{it}$ is $-\beta$ (i.e., 0.742 kg), its effect on $weightLoss_{it+1}$ is $-(1-\alpha)\beta$ (i.e., 0.134 kg), on $weightLoss_{it+2}$ is $-\alpha(1-\alpha)\beta$ (i.e., 0.110 kg), on $weightLoss_{it+3}$ is $\alpha^2(1-\alpha)\beta$ (i.e., 0.090 kg), and so on.

[16] We examined other challenge characteristics, such as the total number of members, the number of posts on challenge forums, and the number of instructions. We did not, however, find any significant interaction effects between the type of behavioral goals and those characteristics.

[17] See https://www.cdc.gov/healthyweight/losing_weight/index.html.

References

Alexy B (1985) Goal setting and health risk reduction. *Nursing Res.* 34(5):283–238.

Allam A, Kostova Z, Nakamoto K, Schulz PJ (2015) The effect of social support features and gamification on a web-based intervention for rheumatoid arthritis patients: Randomized controlled trial. *J. Medical Internet Res.* 17(1):e14.

Annesi JJ (1998) Effects of computer feedback on adherence to exercise. *Perceptual Motor Skills* 87(2):723–730.

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev. Econom. Stud.* 58(2):277–297.

Arellano M, Bover O (1995) Another look at the instrumental variable estimation of error-components models. *J. Econometrics* 68(1):29–51.

Avenell A, Brown TJ, McGee MA, Campbell MK, Grant AM, Broom J, Jung RT, Smith WCS (2004) What interventions should we add to weight reducing diets in adults with obesity? A systematic review of randomized controlled trials of adding drug therapy, exercise, behaviour therapy or combinations of these interventions. *J. Human Nutrition Dietetics* 17(4):293–316.

Berger M, Jung C (2021) Gamification in nutrition apps—users' gamification element preferences: A best-worst-scaling approach. *Proc. 54th Hawaii Internat. Conf. System Sci.*, 1335–1344.

Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. *J. Econometrics* 87(1):115–143.

Boendermaker WJ, Boffo M, Wiers RW (2015) Exploring elements of fun to motivate youth to do cognitive bias modification. *Games Health J.* 4(6):434–443.

Booth AO, Nowson CA, Matters H (2008) Evaluation of an interactive, internet-based weight loss program: A pilot study. *Health Ed. Res.* 23(3):371–381.

Brownell KD, Cohen RY, Stunkard AJ, Felix MR, Cooley NB (1984) Weight loss competitions at the work site: Impact on weight, morale and cost-effectiveness. *Amer. J. Public Health* 74(11): 1283–1285.

Bui A, Veit D, Webster J (2015) Gamification—a novel phenomenon or a new wrapping for existing concepts? *Proc. 36th Internat. Conf. Inform. Systems, Fort Worth, TX*, December 13–16.

Burke V, Mori TA, Giangiulio N, Gillam HF, Beilin LJ, Houghton S, Cutt HE, Mansour J, Wilson A (2002) An innovative program for changing health behaviours. *Asia Pacific J. Clinical Nutrition* 11:S586–S597.

Carter FA, Jansen A (2012) Improving psychological treatment for obesity. Which eating behaviours should we target? *Appetite* 58(3):1063–1069.

Cawley J, Meyerhoefer C (2012) The medical care costs of obesity: An instrumental variables approach. *J. Health Econom.* 31(1):219–230.

Chen Y, Pu P (2014) HealthyTogether: Exploring social incentives for mobile fitness applications. *Proc. Second Internat. Sympos. Chinese CHI* (ACM, New York), 25–34.

Cheng HNH, Wu WMC, Liao CCY, Chan TW (2009) Equal opportunity tactic: Redesigning and applying competition games in classrooms. *Comput. Ed.* 53(3):866–876.

Cooper Z, Fairburn CG (2001) A new cognitive behavioural approach to the treatment of obesity. *Behav. Res. Therapy* 39(5):499–511.

Csikszentmihalyi M (1990) *Flow: The Psychology of Optimal Experience* (Harper and Row, New York).

Curioni CC, Lourenco PM (2005) Long-term weight loss after diet and exercise: A systematic review. *Internat. J. Obesity* 29(10):1168–1174.

Dennis TA, O'Toole LJ (2014) Mental health on the go: Effects of a gamified attention-bias modification mobile application in trait-anxious adults. *Clinical Psych. Sci.* 2(5):576–590.

Deterding S, Dixon D, Khaled R, Nacke L (2011) From game design elements to gamefulness: Defining gamification. *Proc. 15th Internat. Academic MindTrek Conf.: Envisioning Future Media Environments* (ACM, New York), 9–15.

De Vet E, Nelissen RM, Zeelenberg M, De Ridder DT (2013) Ain't no mountain high enough? Setting high weight loss goals predict effort and short-term weight loss. *J. Health Psych.* 18(5):638–647.

Domínguez A, Saenz-de-Navarrete J, De-Marcos L, Fernández-Sanz L, Pagés C, Martínez-Herráiz JJ (2013) Gamifying learning experiences: Practical implications and outcomes. *Comput. Ed.* 63: 380–392.

Elfhag K, Rössner S (2005) Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obesity Rev.* 6(1):67–85.

Festinger L (1954) A theory of social comparison processes. *Human Relations* 7(2):117–140.

Flegal KM, Carroll MD, Ogden CL, Curtin LR (2010) Prevalence and trends in obesity among US adults, 1999–2008. *J. Amer. Medical Assoc.* 303(3):235–241.

Fothergill E, Guo J, Howard L, Kerns JC, Knuth ND, Brychta R, Chen KY, et al. (2016) Persistent metabolic adaptation 6 years after "The Biggest Loser" competition. *Obesity* 24(8):1612–1619.

Garcia SM, Tor A (2009) The N-effect: More competitors, less competition. *Psych. Sci.* 20(7):871–877.

Garcia SM, Tor A, Gonzalez R (2006) Ranks and rivals: A theory of competition. *Personality Soc. Psych. Bull.* 32(7):970–982.

Greene J, Sacks R, Piniewski B, Kil D, Hahn JS (2013) The impact of an online social network with wireless monitoring devices on physical activity and weight loss. *J. Primary Care Comm. Health* 4(3):189–194.

Hamari J, Koivisto J (2015) "Working out for likes": An empirical study on social influence in exercise gamification. *Comput. Human Behav.* 50:333–347.

Hanus MD, Fox J (2015) Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Comput. Ed.* 80(January):152–161.

Heckman JJ (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Ann. Econom. Soc. Measurement* 5(4):475–492.

Hill JO, Peters JC (1998) Environmental contributions to the obesity epidemic. *Sci.* 280(5368):1371–1374.

Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, Callister R, Plotnikoff R, Collins CE (2015) eHealth interventions for the prevention and treatment of overweight and obesity in adults: A systematic review with meta-analysis. *Obesity Rev.* 16(5):376–392.

Hwang KO, Ottenbacher AJ, Green AP, Cannon-Diehl MR, Richardson O, Bernstam EV, Thomas EJ (2010) Social support in an internet weight loss community. *Internat. J. Medical Informatics* 79(1):5–13.

Imbens G, Wooldridge J (2007) What's new in econometrics. Lecture notes 1, July 30, National Bureau of Economic Research, Cambridge, MA.

Jeffery RW, Wing RR, Sherwood NE, Tate DF (2003 Physical activity and weight loss: Does prescribing higher physical activity goals improve outcome? *Amer. J. Clinical Nutrition* 78(4):684–689.

Johns DJ, Hartmann-Boyce J, Jebb SA, Aveyard P (2014) Diet or exercise interventions vs combined behavioral weight management programs: A systematic review and meta-analysis of direct comparisons. *J. Acad. Nutrition Dietetics* 114(10):1557–1568.

Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L (2016) Gamification for health and wellbeing: A systematic review of the literature. *Internet Interventions* 6:89–106.

Jones BA, Madden GJ, Wengreen HJ (2014) The FIT game: Preliminary evaluation of a gamification approach to increasing fruit and vegetable consumption in school. *Preventive Medicine* 68: 76–79.

Khaylis A, Yiaslas T, Bergstrom J, Gore-Felton C (2010) A review of efficacious technology-based weight-loss interventions: Five key components. *Telemedicine e-Health* 16(9):931–938.

Kohn A (1992) *No Contest: A Case Against Competition* (Houghton Mifflin, New York).

Liu D, Li X, Santhanam R (2013) Digital games and beyond: What happens when players compete? *MIS Quart.* 111–124.

Liu D, Santhanam R, Webster J (2016) Toward meaningful engagement: A framework for design and research of gamified information systems. *MIS Quart.* 41(4):1011–1034.

Locke EA (1968) Toward a theory of task motivation and incentives. *Organ. Behav. Human Performance* 3(2):157–189.

Locke EA, Latham GP (1985) The application of goal setting to sports. *J. Sport Exercise Psych.* 7(3):205–222.

Locke EA, Latham GP (2002) Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *Amer. Psychologist* 57(9):705–717.

Maher C, Ferguson M, Vandelanotte C, Plotnikoff R, De Bourdeaudhuij I, Thomas S, Nelson-Field K, Olds T (2015) A web-based, social networking physical activity intervention for insufficiently active adults delivered via Facebook app: Randomized controlled trial. *J. Medical Internet Res.* 17(7):e4086.

McKay HG, King D, Eakin EG, Seeley JR, Glasgow RE (2001) The diabetes network internet-based physical activity intervention: A randomized pilot study. *Diabetes Care* 24(8):1328–1334.

Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD (2013) Using Facebook and text messaging to deliver a weight loss program to college students. *Obesity* 21(1):25–31.

National Institutes of Health (2002) The practical guide: Identification, evaluation, and treatment of overweight and obesity in adults. NIH publication 00-4084, National Heart, Lung, and Blood Institute, Bethesda, MD.

Nelissen R, de Vet E, Zeelenberg M (2011) Anticipated emotions and effort allocation in weight goal striving. *British J. Health Psych.* 16(1):201–212.

Neve M, Morgan PJ, Jones PR, Collins CE (2010) Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: A systematic review with meta-analysis. *Obesity Rev.* 11(4):306–321.

Nickell S (1981) Biases in dynamic models with fixed effects. *Econometrica* 49(6):1417–1426.

Patrick K, Raab F, Adams M, Dillon L, Zabinski M, Rock CL, Griswold WG, Norman GJ (2009) A text message-based intervention for weight loss: Randomized controlled trial. *J. Medical Internet Res.* 11(1):e1.

Pearson ES (2012) Goal setting as a health behavior change strategy in overweight and obese adults: A systematic literature review examining intervention components. *Patient Ed. Counseling* 87(1): 32–42.

Poncela-Casasnovas J, Spring B, McClary D, Moller AC, Mukogo R, Pellegrini CA, Coons MJ, Davidson M, Mukherjee S, Nunes Amaral LA (2015) Social embeddedness in an online weight management program is linked to greater weight loss. *J. Roy. Soc. Interface* 12(104):20140686.

Rancourt D, Leahey TM, LaRose JG, Crowther JH (2015) Effects of weight-focused social comparisons on diet and activity outcomes in overweight and obese young women. *Obesity* 23(1):85–89.

Reeve J, Deci EL (1996) Elements of the competitive situation that affect intrinsic motivation. *Personality Soc. Psych. Bull.* 22(1): 24–33.

Riva S, Camerini AL, Allam A, Schulz PJ (2014) Interactive sections of an internet-based intervention increase empowerment of chronic back pain patients: Randomized controlled trial. *J. Medical Internet Res.* 16(8):e3474.

Roodman D (2009a) How to do Xtabond2: An introduction to difference and system GMM in Stata. *Stata J.* 9(1):86–136.

Roodman D (2009b) A note on the theme of too many instruments. *Oxford Bull. Econom. Statist.* 71(1):135–158.

Ryan RM, Deci EL (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and wellbeing. *Amer. Psychologist* 55(1):68–78.

Santhanam R, Liu D, Shen WCM (2016) Gamification of technology-mediated training: Not all competitions are the same. *Inform. Systems Res.* 27(2):453–465.

Seijts GH, Latham GP (2001) The effect of distal learning, outcome, and proximal goals on a moderately complex task. *J. Organ. Behav.* 22(3):291–307.

Shameli A, Althoff T, Saberi A, Leskovec J (2017) How gamification affects physical activity: Large-scale analysis of walking challenges in a mobile application. *Proc. 26th Internat. World Wide Web Conf., Perth, Australia, April 3–7*, 455–463.

Shilts MK, Horowitz M, Townsend MS (2004) Goal setting as a strategy for dietary and physical activity behavior change: A review of the literature. *Amer. J. Health Promotion* 19(2):81–93.

Silva MN, Vieira PN, Coutinho SR, Minderico CS, Matos MG, Sardinha LB, Teixeira PJ (2010) Using self-determination theory to promote physical activity and weight control: A randomized controlled trial in women. *J. Behav. Med.* 33(2):110–122.

Silva MN, Markland D, Minderico CS, Vieira PN, Castro MM, Coutinho SR, Santos TC, Matos MG, Sardinha LB, Teixeira PJ (2008) A randomized controlled trial to evaluate self-determination theory for exercise adherence and weight control: Rationale and intervention description. *BMC Public Health* 8(1):1–13.

Steinberg DM, Tate DF, Bennett GG, Ennett S, Samuel-Hodge C, Ward DS (2013) The efficacy of a daily self-weighing weight loss intervention using smart scales and e-mail. *Obesity* 21(9): 1789–1797.

Stevens V, Funk K, Brantley P, Erlinger T, Myers V, Champagne C, Bauck A, Samuel-Hodge C, Hollis J (2008) Design and implementation of an interactive website to support long-term maintenance of weight loss. *J. Medical Internet Res.* 10(1):e1.

Strecher VJ, Seijts GH, Kok GJ, Latham GP, Glasgow R, DeVellis B, Meertens RM, Bulger DW (1995) Goal setting as a strategy for health behavior change. *Health Ed. Quart.* 22(2):190–200.

Stunkard AJ, Cohen RY, Felix MR (1989) Weight loss competitions at the worksite: How they work and how well. *Preventive Medicine* 18(4):460–474.

Tauer JM, Harackiewicz JM (2004) The effects of cooperation and competition on intrinsic motivation and performance. *J. Personality Soc. Psych.* 86(6):849–861.

Teixeira PJ, Silva MN, Mata J, Palmeira AL, Markland D (2012) Motivation, self-determination, and long-term weight control. *Internat. J. Behavioral Nutrition Physical Activity* 9(1):1–3.

Teixeira PJ, Silva MN, Coutinho SR, Palmeira AL, Mata J, Vieira PN, Carraça EV, Santos TC, Sardinha LB (2010) Mediators of weight loss and weight loss maintenance in middle-aged women. *Obesity* 18(4):725–735.

Trognon A (1978) Miscellaneous asymptotic properties of ordinary least squares and maximum likelihood estimators in dynamic error components models. *Ann.* l'INSEE 30–31:631–657.

Uetake K, Yang N (2018) Success breeds success: Weight loss dynamics in the presence of short-term and long-term goals. Working paper 170002, Canadian Centre for Health Economics, Toronto.

Uetake K, Yang N (2020) Inspiration from the "Biggest Loser": Social interactions in a weight loss program. *Marketing. Sci.* 39(3):487–499.

Wilson PW, D'Agostino RB, Sullivan L, Parise H, Kannel WB (2002) Overweight and obesity as determinants of cardiovascular risk: The Framingham experience. *Arch. Internal Medicine* 162(16): 1867–1872.

Windmeijer F (2005) A finite sample correction for the variance of linear efficient two-step GMM estimators. *J. Econometrics* 126(1): 25–51.

Wooldridge JM (2002) Inverse probability weighted M-estimators for sample selection, attrition, and stratification. *Portuguese Econom. J.* 1(2):117–139.

Yan LL (2018) Good intentions, bad outcomes: The effects of mismatches between social support and health outcomes in an online weight loss community. *Production Oper. Management* 27(1): 9–27.

Zuckerman O, Gal-Oz A (2014) Deconstructing gamification: Evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. *Personal Ubiquitous Comput.* 18(7):1705–1719.