What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study*

Damon Jones
University of Chicago
and NBER

David Molitor
University of Illinois
and NBER

Julian Reif
University of Illinois
and NBER

June 2018

Abstract

Workplace wellness programs cover over 50 million workers and are intended to reduce medical spending, increase productivity, and improve well-being. Yet, limited evidence exists to support these claims. We designed and implemented a comprehensive workplace wellness program for a large employer with over 12,000 employees, and randomly assigned program eligibility and financial incentives at the individual level. Over 56 percent of eligible (treatment group) employees participated in the program. We find strong patterns of selection: during the year prior to the intervention, program participants had lower medical expenditures and healthier behaviors than non-participants. However, we do not find significant causal effects of treatment on total medical expenditures, health behaviors, employee productivity, or self-reported health status in the first year. Our 95% confidence intervals rule out 83 percent of previous estimates on medical spending and absenteeism. Our selection results suggest these programs may act as a screening mechanism: even in the absence of any direct savings, differential recruitment or retention of lower-cost participants could result in net savings for employers.

JEL Classification: I1, M5, J3

Keywords: Health care, workplace wellness, selection, productivity, field experiments

---

*This research was supported by the National Institute on Aging of the National Institutes of Health under award number R01AG050701; the National Science Foundation under Grant No. 1730546; the Abdul Latif Jameel Poverty Action Lab (J-PAL) North America U.S. Health Care Delivery Initiative; Evidence for Action (E4A), a program of the Robert Wood Johnson Foundation; and the W.E. Upjohn Institute for Employment Research. This study was pre-registered with the American Economics Association RCT Registry (AEARCTR-0001368). We are grateful to Andy de Barros for thoroughly replicating our analysis and to J-PAL for coordinating this replication effort. We thank our co-investigator Laura Payne for her vital contributions to the study, Lauren Geary for outstanding project management, Michele Guerra for excellent programmatic support, and Illinois Human Resources for invaluable institutional support. We are also thankful for comments from Kate Baicker, Jay Bhattacharya, Tatyana Deryugina, Joseph Doyle, Amy Finkelstein, Eliza Forsythe, Drew Hanks, Bob Kaestner, David Meltzer, Michael Richards, Justin Sydnor, Richard Thaler, and seminar participants at AHEC, ASHEcon, Harvard, Junior Health Economics Summit, MHEC, NBER Summer Institute, Ohio State University, University of Chicago AFE Conference, University of Zurich, UPenn Behavioral Economics and Health Symposium, SEA, and SIEPR. The findings and conclusions expressed are solely those of the authors and do not represent the views of the National Institutes of Health, any of our funders, or the University of Illinois. Email: damonjones@uchicago.edu, dmolitor@illinois.edu, and jreif@illinois.edu.

1 Introduction

Sustained growth in medical spending has prompted policymakers, insurers, and employers to search for ways to reduce medical spending. One widely touted solution is to increase the use of "wellness programs," interventions designed to encourage preventive care and discourage unhealthy behaviors such as inactivity or smoking. The 2010 Affordable Care Act (ACA) encourages firms to adopt wellness programs by permitting them to offer participation incentives up to 30 percent of the total cost of health insurance coverage, and the House Ways and Means Committee recently advanced a bill to make gym memberships tax deductible (Phillips Erb, 2018). Workplace wellness industry revenue has more than tripled in size to $8 billion since the passage of the ACA, wellness programs now cover over 50 million U.S. workers, and recent studies have investigated expanding wellness programs into Medicare and Medicaid (Mattke, Schnyer and Van Busum, 2012; Fout et al., 2013; Kaiser, 2016b; Askelson et al., 2017). A meta-analysis by Baicker, Cutler and Song (2010) finds large medical and absenteeism cost savings, but some studies find only limited benefits (e.g., Gowrisankaran et al., 2013; Baxter et al., 2014). As these authors have noted, identification is limited in prior studies because employee participation, along with the firm's decision to adopt a wellness program, is voluntary.

Moreover, the prior literature has overlooked important questions regarding selection into wellness programs. The increasing use of large financial incentives now permitted by the ACA may redistribute resources across employees in a manner that runs counter to the intentions of policymakers.[1] For example, wellness incentives may shift costs onto unhealthy or lower-income employees if these groups are less likely to participate in wellness programs. Furthermore, wellness programs may act as a screening device by encouraging employees who benefit most from these programs to join or remain at the firm—perhaps by earning rewards for behaviors they already enjoy.

[1] Kaiser (2017) estimates that 13 percent of large firms (at least 200 employees) offer incentives that exceed $500 dollars per year, and 4 percent of large firms offer incentives that exceed $1,000 per year.

To improve our understanding of what workplace wellness programs do, we designed and implemented the Illinois Workplace Wellness Study, a large-scale, randomized controlled trial (RCT) conducted at the University of Illinois at Urbana-Champaign (UIUC).[2] In conjunction with the director of Campus Wellbeing Services, we developed a comprehensive workplace wellness program that included an on-site biometric health screening, an online health risk assessment, and a wide variety of wellness activities (e.g., smoking cessation, stress management, and recreational classes). We invited 12,459 benefits-eligible university employees to participate in our study.[3] Study participants ($N = 4,834$) assigned to the treatment group ($N = 3,300$) were invited to take paid time off to participate in our workplace wellness program. Those who successfully completed the entire program earned rewards ranging from from \$50 to \$350, with the amounts randomly assigned and communicated at the start of the program. The remaining subjects ($N = 1,534$) were assigned to a control group, which was not permitted to participate. Our analysis combines individual-level data from online surveys, university employment records, health insurance claims, campus gym visit records, and administrative records from a popular community running event. We can therefore examine outcomes commonly studied by the prior literature (namely, medical spending and employee absenteeism) as well as a large number of novel outcomes.

In this paper, we provide the first set of findings from the Illinois Workplace Wellness Study. We address three key research questions. First, how do financial incentives affect the level of participation in wellness programs? Theory generally predicts that incentives should increase participation, but the magnitude of this increase, which matters for understanding whether these programs shift costs onto non-participants, is an empirical question.

[2]Supplemental materials, datasets, and additional publications from this project will be made available on the study website at http://www.nber.org/workplacewellness.

[3]UIUC administration provided access to university data and guidance to ensure our study conformed with university regulations, but did not otherwise influence the design of our intervention. Each component of the intervention, including the financial incentives paid to employees, was paid for entirely by our external funders. Participation required electronically signing an informed consent form and completing a 15-minute online survey. Because the consent form made subjects aware of the research project, our RCT can be classified as a "framed field experiment," in the parlance of Harrison and List (2004). The study was approved by the UIUC and University of Chicago Institutional Review Boards.

If employee participation is price elastic, then increasing the size of incentives reduces compensation gaps between participants and non-participants; if it is price inelastic, then larger incentives exacerbate those gaps. Second, what types of employees select into wellness programs? The expected direction of the effect is ambiguous. For example, while healthy employees may have low costs of participating in these programs, employees in poor health may have the most to gain from participating. Third, what are the causal effects of workplace wellness programs on medical spending, employee productivity, health behaviors, and well-being after one year? Again, the expected signs of these effects are uncertain. For example, medical spending could decrease if wellness programs improve health, but it could increase if wellness programs and primary care are complements.

In turn, we have three main sets of results. First, 56 percent of employees in our treatment group completed the initial major component of our study, which included an on-campus health screening. Completion depended on the size of the monetary incentive assigned to an employee: increasing the screening completion reward from $0 to $100 boosted the completion rate by 12 percentage points, from 47 to 59, but further increasing the reward to $200 only increased completion by 4 percentage points, to 63 percent. When combined with our accounting records, these participation rates imply that the marginal cost of using financial incentives to induce additional screening participation reaches $1,750 at the highest screening incentive level ($200). This rapidly diminishing effect implies that—at least in our setting—increasing a large financial incentive to even greater levels will transfer large sums of money to workplace wellness program participants, but will have little effect on their composition. We also find that incentives tied to completing downstream wellness activities are more cost-effective than up-front incentives tied to completing the initial health screening.

Second, we find evidence of significant advantageous selection into our program: at baseline, average annual medical spending among participants was $1,393 less than among non-participants. A more detailed investigation reveals that this selection effect is concentrated in the middle of the spending distribution: employees in the upper *and* lower tails of the medical spending distribution were least likely to participate. Because spending is right-skewed, the net result is that average, baseline spending among participants is lower than that of non-participants. Our estimate is economically significant: considering only medical spending, if our program increased the share of participating (i.e. low-spending) workers employed at the university by 4.5 percentage points or more, then our result implies that this change in composition alone would offset the entire costs of our intervention.[4] We also find that participants were more likely to have visited campus recreational facilities prior to our study, and were more likely to have participated in prior community running events. Thus, a primary benefit of these programs to employers may be their potential to attract and retain healthy workers with low medical spending.

Third, we do not find significant effects of our intervention on 37 out of the 39 outcomes we examine in the first year following random assignment.[5] These 37 outcomes include all our measures of medical spending, productivity, health behaviors, and self-reported health. We investigate the effect on medical expenditures in detail, but fail to find significant effects on different quantiles of the spending distribution or on any major subcategory of medical expenditures (pharmaceutical drugs, office, or hospital). We also do not find any effect of our intervention on the number of visits to campus gym facilities or on the probability of participating in a popular annual community running event, two health behaviors that are relatively simple for a motivated employee to change over the course of one year.

These null estimates are meaningfully precise, particularly for two key outcomes of interest in the literature: medical spending and absenteeism. Our 95 percent confidence intervals rule out 83 percent of the effects reported in 115 prior studies, and the 99 percent confidence intervals for the return on investment (ROI) of our intervention rule out the widely cited medical spending and absenteeism ROI's reported in the meta-analysis of Baicker, Cutler and Song (2010). In addition, we show that our OLS (non-RCT) estimate for medical spending

[4]We estimate positive, albeit small and insignificant, effects of the intervention on retention after one year. Our study, which focuses on an employee cohort, was not designed to examine recruitment effects.

[5]Participants were assigned to treatment and control groups in August 2016. Health screenings occurred in August and September, and wellness activities ran from October 2016 to April 2017.

is in line with estimates from prior observational studies, but is ruled out by the 95 percent confidence interval of our IV (RCT) estimate. This demonstrates the value of employing an RCT design in this literature.

We do find two robust, positive treatment effects from the intervention, both based on follow-up survey responses.[6] First, employees in the treatment group were more likely than employees in the control group to report that they had ever received a health screening. This indicates that the health screening component of our program did not merely crowd out health screenings that otherwise would have occurred in the absence of our intervention. Second, treatment group employees were much more likely to report that management places a high priority on worker health and safety.

Our study contributes to the economics literature on selection in labor and insurance markets. It is well known that signaling (Spence, 1973) and screening (Rothschild and Stiglitz, 1976; Wilson, 1977) can be effective responses to asymmetric information about worker productivity (e.g. Mas-Colell et al., 1995, Ch. 13; Lazear and Oyer, 2012). Because health insurance represents an increasingly large component of firm costs, prior studies have also focused on asymmetric information about worker health status (Cutler and Zeckhauser, 2000; Bhattacharya and Vogt, 2014). Our results suggest that workplace wellness programs may be an effective way to encourage workers with low medical spending to join or remain at firms, which is a novel example of a "self-selection" device (Salop and Salop, 1976). We complement prior studies that show compensation packages may be used to attract specific types of workers (Lazear, 2000; Liu et al., 2017) and provide an additional economic justification for the prevalent and growing use of non-wage employment benefits (Oyer, 2008). Moreover, because enrollment into wellness programs is often linked to discounts on insurance premiums, our work is related to a broader literature on adverse selection in insurance markets (see Chiappori and Salanié, 2013, and Geruso and Layton, 2017, for reviews).

Our results also speak directly to the effects of workplace wellness on worker equity.

[6] We address the multiple inference concern that arises when testing many hypotheses by controlling for the family-wise error rate. We discuss our approach in greater detail in Section 3.4.

When incentives are linked to pooled expenses such as health insurance premiums, wellness programs can have distributional consequences. A concern is that wellness programs may effectively increase insurance premiums for low-income workers in poor health (Volpp et al., 2011; Horwitz, Kelly and DiNardo, 2013; McIntyre et al., 2017). The results of our selection analysis provide support for these concerns: non-participating employees are more likely to be in the bottom quartile of the salary distribution, are less likely to engage in healthy behaviors, and have higher medical spending, on average.

We also contribute to the large health literature evaluating the causal effects of workplace wellness programs. Our randomized controlled design allows us to establish reliable causal effects by comparing outcomes across the treatment and control groups. By contrast, most existing studies rely on observational comparisons between participants and non-participants (see Pelletier, 2011, and Chapman, 2012, for reviews). Reviews of the literature have called for additional research on this topic and have also noted the potential for publication bias to skew the set of existing results (Baicker, Cutler and Song, 2010; Abraham and White, 2017). To that end, our intervention, empirical specifications, and outcome variables were pre-specified and publicly archived.[7] In addition, the analyses in this paper were independently replicated by a J-PAL affiliated researcher. A number of RCTs have focused on components of workplace wellness, such as wellness activities (Volpp et al., 2008; Charness and Gneezy, 2009; Royer, Stehr and Sydnor, 2015; Handel and Kolstad, 2017), health risk assessments (Haisley et al., 2012), or particular biometric outcomes such as obesity (Meenan et al., 2010). To our knowledge, no RCTs of comprehensive workplace wellness programs exist.

The rest of the paper proceeds as follows. Section 2 provides a background on workplace wellness, a description of our experimental design, and a summary of our datasets. Section 3 outlines our empirical methods, while Section 4 presents the results of our analysis and discussion. Finally, section 5 offers concluding observations.

[7] Our pre-analysis plan is available at http://www.socialscienceregistry.org/trials/1368. We indicate in the paper the few instances in which we deviate from our pre-analysis plan. A small number of pre-specified analyses have been omitted from the main text for the sake of brevity and because their results are not informative. For completeness, we will report those omitted results in a separate appendix.

2 Experimental Design

2.1 Background

Workplace wellness programs are employer-provided efforts to "enhance awareness, change behavior, and create environments that support good health practices" (Aldana, 2001, p. 297). For the purposes of this study, "wellness programs" encompass three major types of interventions: (1) biometric screenings, which provide clinical measures of health; (2) health risk assessments (HRA), which identify potential health issues; and (3) wellness activities, which promote a healthy lifestyle by encouraging behaviors such as smoking cessation, stress management, or fitness. Best practice guides advise employers to let employees take paid time off to participate in wellness programs, and to combine wellness program components to maximize their effectiveness (Ryde et al., 2013). In particular, it is recommended that information from a biometric screening and HRA inform the selection of wellness activities (Soler et al., 2010). Among firms with 200 or more employees, the share offering a biometric screening, HRA, or wellness activities in 2016 was 53 percent, 59 percent, and 83 percent, respectively (Kaiser, 2016a). These benefits are often coupled with financial incentives for participation, such as cash compensation or discounted health insurance premiums. A 2015 survey estimates an average cost of $693 per employee for these programs (Jaspen, 2015) and a recent industry analysis estimates annual revenues of $8 billion (Kaiser, 2016b).

A number of factors may explain the increasing popularity of workplace wellness programs. First, some employers believe that these programs reduce medical spending and increase productivity. For example, Safeway famously attributed its low medical spending to its wellness program (Burd, 2009) (although this evidence was subsequently disputed (Reynolds, 2010)), and recent work suggests wellness programs may increase productivity (Gubler, Larkin and Pierce, 2017). Second, if employees have a high private value of wellness-related benefits, then labor market competition may drive employers to offer wellness programs in order to attract and retain workers. Third, the Affordable Care Act (ACA) has relaxed constraints on the maximum size of financial incentives offered by employers. Prior to the ACA, health-contingent incentives could not exceed 20 percent of the cost of employee health coverage. The ACA increased that general limit to 30 percent, and raised it to 50 percent for tobacco cessation programs (Cawley, 2014). The average premium for a family insurance plan in 2017 was $18,764 (Kaiser, 2017), which means that many employers are permitted to offer wellness rewards or penalties in excess of $5,000.

Like other large employers, many universities also have workplace wellness programs. Of the nearly 600 universities and liberal arts colleges ranked by *U.S. News & World Report*, over two-thirds offer an employee wellness program.[8] Prior to our intervention, UIUC's campus wellness services were run by the University of Illinois Wellness Center, which has one staff member. The Wellness Center coordinates smoking cessation resources for employees and provides a limited number of wellness activities, many of which are not free. Importantly for our study, the campus did not offer any health screenings or HRAs and did not provide monetary incentives to employees in exchange for participating in wellness activities. Therefore, our intervention effectively represents the introduction of all major components of a wellness program at this worksite.

2.2 The Illinois Workplace Wellness Study and iThrive

The Illinois Workplace Wellness Study is a large-scale randomized controlled trial designed to investigate the effects of workplace wellness programs on employee medical spending, productivity, and well-being. As part of the study, we designed a comprehensive wellness program named "iThrive" at the University of Illinois at Urbana-Champaign. We summarize the program here and provide full details in Appendix D.

Figure 1 illustrates the experimental design of our study. In July 2016 we invited 12,459 benefits-eligible university employees to enroll in our study by completing a 15-minute online survey designed to measure baseline health and wellness. The invitations were sent by

[8] Source: authors' tabulation of data collected from university and colleges via website search and phone inquiry.

postcard and email. Employees were offered a $30 Amazon.com gift card to complete the survey, as well as a chance "to participate in a second part of the research study." Over the course of three weeks, 4,834 employees completed this baseline survey. Study participants, whom we define as anybody completing the 15-minute baseline survey, were then randomly assigned to either a control group (N=1,534), or one of six treatment groups (N=3,300). Members of the control group were notified that they may be contacted for follow-up surveys in the future, and further contact with this group was thereafter minimized. Members of the treatment group were offered the opportunity to participate in iThrive.

The first step of iThrive included a biometric health screening and an online HRA. For a period of 5 weeks in August and September 2016, participants had an opportunity to schedule a screening at one of many locations on campus. They had to make an appointment in advance and fast for 12 hours prior to the screening, where a clinician measured their height, weight, waist circumference, and blood pressure. The clinician also performed a fingerstick test to measure blood cholesterol, triglycerides, and glucose levels. Finally, participants met with a health coach, who explained their health measurements to them. The entire screening process lasted about 20 minutes. A few days later, participants received an email invitation to complete an online HRA designed to assess their lifestyle habits. Upon completion of the HRA, participants were given a score card incorporating the results of their biometric screening and providing them with recommended areas of improvement. The HRA was available as early as one week after the beginning of biometric screening and remained open until two weeks after the last biometric screening. Only participants who completed both the screening and HRA were eligible to participate in wellness activities.

The second step of iThrive consisted of wellness activities. Eligible participants were offered the opportunity to participate in one of several activities in the fall and then again in the spring. Eligibility to participate in spring wellness activities was not contingent on enrollment or completion of fall activities. In the fall, activities included in-person classes on chronic disease management, weight management, tai chi, physical fitness, financial wellness, and healthy workplace habits; a tobacco cessation hotline; and an online, self-paced wellness challenge. A similar set of activities was offered in the spring. Classes ranged from 6 to 12 weeks in length, and "completion" of a class was generally defined as attending at least three-fourths of the sessions. Participants were given two weeks to enroll in wellness activities and were encouraged to incorporate their HRA feedback when choosing a class.

Study participants were offered monetary rewards for completing each step of the iThrive program, and these rewards varied depending on the treatment group to which an individual was assigned. Individuals in treatment groups labeled A, B, and C were offered a screening incentive of \$0, \$100, or \$200, respectively, for completing the biometric screening and the HRA. Treatment groups were further split based on an activity incentive of either \$25 or \$75 for each wellness activity completed (up to one per semester). Thus, there were six treatment groups in total: A25, A75, B25, B75, C25, and C75 (see Figure 1). The total reward for completing all iThrive components—the screening, the HRA, and a wellness activity during both the fall and spring—ranged from \$50 to \$350, depending on the treatment group. These amounts are in line with typical wellness programs (Mattke, Schnyer and Van Busum, 2012). The probability of assignment to each group was equal across participants, and randomization was stratified by employee class (faculty, staff, or civil service), sex, age, quartile of annual salary, and race (see Appendix D.1.2 for additional randomization details). We privately informed participants about their screening and wellness activity rewards at the start of the intervention (August 2016), and did not disclose information about rewards offered to others.

To help guide participants through iThrive, we developed a secure online website that granted access to information about the program. At the onset of iThrive in August, the website instructed participants to schedule a biometric screening and then to take the online HRA. Beginning in October, and then again in January, the website provided a menu of wellness activities and online registration forms for those activities. The website also provided information on a participant's current progress and rewards earned to date, answers to frequently asked questions, and contact information for participant support.

2.3 Data

Our analysis employs a combination of self-reported survey data and a number of administrative data sources, all merged together at the individual level. We briefly describe each data source below. Appendix Table A.7 provides a definition for each variable used in our analysis. Additional details are provided in Appendix D.2.

2.3.1 University Administrative Data

We obtained university administrative data on 12,486 employees who as of June 2016 were (1) working at the Urbana-Champaign campus of the University of Illinois and (2) eligible for part-time or full-time employee benefits from the Illinois Department of Central Management Services. We excluded 27 people who did not have a university email address or who were substantially involved with our study, yielding a final sample size of 12,459 employees.

The initial denominator file includes the employee's name, university identification number, contact information (email and home mailing address), date of birth, sex, race, salary, and employee class (faculty, academic staff, or civil service). We used the email and home mailing address to invite employees to participate in our study, and we used the sex, race, date of birth, salary, and employee class variables to generate the strata for random sampling.

A second file includes employment history information as of July 31, 2017. This provides two employee productivity outcomes that are measured over the first 12 months of our study: job termination and salary raises. All employees in our sample were eligible for a mid-year, merit-based salary increase that occurred in February 2017.

A third file provides data on sick leave. The number of sick days taken is available at the monthly level for Civil Service employees. For academic faculty and staff, the number of sick days taken is available biannually, on August 15 and May 15. We first calculate the total number of sick days taken during our pre-period (August 2015 - July 2016) and post-period

(August 2016 - July 2017) for each employee. We then normalize by the number of days employed to make this measure comparable across employees. All specifications that include sick days taken as an outcome variable are weighted by the number of days employed.

A fourth file contains data on exact attendance dates for the university's gym and recreational facilities. Entering one of these facilities requires swiping an ID card, which creates a database record linked to the individual's university ID. We calculate the total number of visits per year for the pre-period (August 2015 - July 2016) and the post-period (August 2016 - July 2017).

2.3.2 Online Survey Data

As described in Section 2.2, all study participants took a 15-minute online survey in July 2016 as a condition of enrollment in the study. The survey covered topics including health status, health care utilization, job satisfaction, and productivity.

Our survey software recorded that, out of the 12,459 employees invited to take the survey, 7,468 employees clicked on the link to the survey, 4,918 employees began the survey, and 4,834 employees completed the survey. Although participants were allowed to skip questions, response rates for the survey were very high: 4,822 out of 4,834 participants (99.7 percent) answered every one of the questions used in our analysis. To measure the reliability of the survey responses, we included a question about age at the end of the survey and compared participants' self-reported ages with the ages available in the university's administrative data. Of the 4,830 participants who reported an age, only 24 (<0.5 percent) reported a value that differed from the university's administrative records by more than one year.

All study participants were also invited via postcard and email to take a one-year, follow-up survey online in July 2017.[9] In addition to the questions asked on the baseline survey, the follow-up survey included additional questions on productivity, presenteeism, and job satisfaction. A total of 3,568 participants (74 percent) successfully completed the 2017

[9]Invitations to the follow-up survey were sent regardless of current employment status with the university.

follow-up survey. The completion rates for the control and treatment groups were 75.4 and 73.1 percent, respectively. This difference in completion rates is marginally significant ($p = 0.079$). The full texts of our 2016 baseline and 2017 follow-up online surveys are available on the study website and as part of our supplementary materials.[10]

2.3.3 Health Insurance Claims Data

We obtained health insurance claims data for the time period January 1, 2015, through July 31, 2017, for the 67 percent of employees who subscribe to the university's most popular insurance plan. We use the total payment due to the provider to calculate average total monthly spending. We also use the place of service code on the claim to break total spending into four major subcategories: pharmaceutical, office, hospital, and other.[11] Our spending measures include all payments from the insurer to providers, as well as any deductibles or copays paid by individuals. We merged these data at the individual level with our other datasets for those employees who consented to participate in our study. In addition, we have access to anonymized panel data on health claims for non-participating employees who subscribe to this same plan.

Employees choose their health plan annually during the month of May, and plan changes become effective July 1. Participants were informed of their treatment assignment on August 9, 2016. We therefore define baseline medical spending to include all allowed amounts with dates of service corresponding to the 13-month time period July 1, 2015, through July 31, 2016. We define spending in the post period to correspond to the 12-month time period August 1, 2016, through July 31, 2017.

In our health claims sample, 11 percent of employees are not continuously enrolled

[10] Interactive examples of the surveys administered for the study are available at http://www.nber.org/workplacewellness.

[11] Pharmaceutical and office-based spending each have their own place of service codes. Hospital spending is summed across the following four codes: "Off Campus - Outpatient Hospital," "Inpatient Hospital," "On Campus - Outpatient Hospital," and "Emergency Room - Hospital." All remaining codes are assigned to "other" spending, which serves as the omitted category in our analysis. We did not pre-specify subcategories of spending in our pre-analysis plan.

throughout the 13-month pre-period, and 9 percent are not continuously enrolled throughout the 12-month post-period. This is primarily due to job turnover. Because measures of average monthly spending are less noisy for employees with more months of claims data, we weight our regressions by the number of covered months whenever the outcome variable is average spending.

2.3.4 Illinois Marathon/10K/5K Data

The Illinois Marathon is a running event held annually in Champaign, Illinois. The individual races offered include a marathon, a half marathon, a 5K, and a 10K. When registering for a race, a participant must provide her name, age, sex, and hometown. That information, along with the results of the race, are published online after the races have concluded. We downloaded those data for the 2014-2017 races and matched it to individuals in our dataset using name, age, sex, and hometown.

2.4 Baseline Summary Statistics and Balance Tests

Tables 1a and 1b provide summary statistics at baseline for the employees in our sample. Columns (2)-(8) report means for those who were assigned to our control group and to each of our six treatment groups. Column (1) additionally reports summary means for employees not enrolled in our study, where available. The variables are grouped into four panels, based on the source and type of data. Panel A presents means of the university administrative data variables used in our stratified randomization, Panel B presents means of variables from our 2016 baseline survey, Panel C presents means of medical spending variables from our health insurance claims data for the July 2015 - July 2016 time period, and Panel D presents baseline means of administrative data variables used to measure health behaviors and employee productivity.

Our experimental framework relies on the random assignment of study participants to the treatment and control groups. To evaluate the validity of this assumption, we first compare the means of the variables displayed in Tables 1a and 1b. For each row, we regress the study variable on seven indicators, one for the control and each of six treatment groups, and test for the joint equality of the seven coefficients. Column (9) reports the p-value from that test. We also estimate a seemingly unrelated regression model to test whether the variables listed within each panel predict enrollment into either the control or any of the six treatment groups. The bottom of Tables 1a and 1b reports the p-value from jointly testing whether all regression coefficients across all seven groups are equal to 0, within each panel.

By construction, we find no evidence of differences in means among the variables used for stratification (Panel A): all p-values in column (9) are greater than 0.97. Among all other variables listed in Panels B, C, and D, we find statistically significant differences at a 10 percent or lower level in 2 out of 34 cases, which is approximately what one would expect from random chance. This is confirmed by our joint balance tests, which fail to reject the null hypothesis that the variables in Panel B ($p = 0.165$), Panel C ($p = 0.220$), or Panel D ($p = 0.437$) are not predictive of group assignment.

A unique feature of our study is our ability to characterize the employees who declined to participate in our experiment. We investigate the extent of this selection into our study by comparing means for study participants, reported in columns (2)-(9) of Tables 1a and 1b, to the means for non-participating employees who did not complete our baseline survey, reported in column (1). Study participants are younger, are more likely to be female, are more likely to be white, have lower incomes on average, are more likely to be administrative staff, and are less likely to be faculty. They also have lower baseline medical spending, are more likely to have participated in one of the Illinois Marathon/10K/5K running events, and have a higher rate of monthly gym visits. These selection effects mirror the ones we report below in Section 4.2, suggesting that the factors governing the decision to participate in a wellness program are similar to the ones driving the decision to participate in our study.

3 Empirical Methods

3.1 Participation

We begin by estimating the effect of our wellness program incentives on participation outcomes among employees randomly assigned to a treatment group. We exclude members of the control group, for whom participation is mechanically zero. First, we jointly estimate the average effects of being assigned a positive screening incentive (groups B and C) or being assigned the $75 wellness activity incentive using the following ordinary least squares (OLS) regression:

$$P_i = \alpha + \beta_{BC} T_{i,BC} + \beta_{75} T_{i,75} + \Gamma X_i + \varepsilon_i. \tag{1}$$

Here, $T_{i,BC}$ is an indicator for membership in treatment groups B or C, and $T_{i,75}$ is an indicator for receiving the $75 wellness activity incentive. The omitted category includes members of treatment group A with a $25 wellness activity incentive.

Second, we augment equation (1) to estimate participation effects for groups B and C separately, as follows:

$$P_i = \alpha + \beta_B T_{i,B} + \beta_C T_{i,C} + \beta_{75} T_{i,75} + \Gamma X_i + \varepsilon_i. \tag{2}$$

Here, the independent variables $T_{i,B}$ and $T_{i,C}$ are indicators for membership in treatment groups B and C, respectively.

In equations (1) and (2), the outcome P_i is an indicator for one of the following three participation outcomes: completing a screening and HRA, completing a fall wellness activity, or completing a spring wellness activity. The coefficients of interest—β_{BC}, β_B, β_C, and β_{75}—represent the causal effect of increased incentives on participation. We estimate results with and without the inclusion of strata fixed effects, X_i. The identifying assumption requires that treatment be uncorrelated with unobservable determinants of participation, ε_i, which is delivered by virtue of random assignment. This assumption is supported by the balance tests across the treatment groups, reported in Section 2.4.

3.2 Selection

Next, we characterize the types of employees who are most likely to participate in or complete the various stages of our wellness program. We pool data across the six treatment groups and estimate the following OLS regression:

$$X_i = \alpha + \theta P_i + \varepsilon_i. \tag{3}$$

The left-hand side variable, X_i, is a pre-determined covariate. The regressor, P_i, is an indicator for one of the following three participation outcomes: completing a screening and HRA, completing a fall wellness activity, or completing a spring wellness activity. The coefficient θ represents the correlation between participation and the baseline characteristic, X_i; it should not be interpreted causally.

3.3 Causal Effects

In our final analysis, we estimate the one-year effect of our wellness intervention on a number of outcomes, including medical spending from health claims data, employment and productivity variables measured in administrative and survey data, health behaviors measured in administrative data, and self-reported health status and behaviors. We compare outcomes in the treatment group to those in the control group using the following specification:

$$Y_i = \alpha + \gamma T_i + \Gamma X_i + \varepsilon_i. \tag{4}$$

Here, T_i is an indicator variable for membership in one of our six treatment groups, and Y_i is an outcome of interest. We estimate equation (4) with and without the inclusion of controls, X_i. In one control specification, X_i includes baseline strata fixed effects. One could also include a much broader set of controls, but doing so comes at the cost of reduced degrees of freedom. Thus, our second control specification implements the Lasso double-selection method of Belloni, Chernozhukov and Hansen (2014), as outlined by Urminsky, Hansen and Chernozhukov (2016), which selects controls that predict either the dependent variable or the focal independent variable.[12] The set of potential controls includes baseline values of the outcome variable, strata variables, the baseline survey variables reported in Table 1a, and all pairwise interactions. We then estimate a regression that includes only the controls selected by double-Lasso. In our tables, we follow convention and refer to this third control strategy as "post-Lasso." As before, our main identifying assumption requires treatment to be uncorrelated with unobserved determinants of the outcome. The key parameter of interest, γ, is the intent-to-treat effect of our intervention on the outcome Y_i.

3.4 Inference

We report conventional robust standard errors in all tables. We do not cluster standard errors because randomization was performed at the individual level (Abadie et al., 2017). Because we estimate equations (3) and (4) for many different outcome variables, the probability that we incorrectly reject at least one null hypothesis is greater than the significance level used for each individual hypothesis test. When appropriate, we address this multiple inference concern by controlling for the family-wise error rate, i.e. the probability of incorrectly rejecting one or more null hypotheses belonging to a family of hypotheses.

To control for the family-wise error rate, we first define seven mutually exclusive families of hypotheses that encompass all of our outcome variables. Each family contains all variables belonging to one of our four outcome domains (strata variables, medical spending, employment/productivity, or health) and one of our two types of data (administrative or

[12]No control variable will be predictive of a randomly assigned variable, in expectation. Thus, when implementing the double-selection method with randomly assigned treatment status as the focal independent variable, we only select controls that are predictive of the dependent variable. When implementing Lasso, we use the penalty parameter that minimizes 10-fold cross-validated mean squared error.

survey).[13] When testing multiple hypotheses using equations (3) and (4), we then calculate family-wise adjusted p-values based on 10,000 bootstraps of the free step-down procedure of Westfall and Young (1993).[14]

4 Results

4.1 Participation

We begin by summarizing the effect of incentives on participation. Figure 2 reports that 56.0 percent of participants in the treatment group completed both the health screening and online HRA, which together comprise the first major step of our workplace wellness program. These participants earned their assigned rewards ($0, $100, or $200), and were subsequently allowed to sign up for wellness activities; the remaining 44 percent were excluded. In the fall, 39.5 percent of the treatment group registered for an activity, and 27.4 percent completed enough of the activity to earn their assigned activity reward. Registration and completion rates were slightly lower for the spring wellness activity. By way of comparison, a survey of employers with workplace wellness programs found that less than 50 percent of their eligible employees complete health screenings, and that most firms have wellness activity participation rates of less than 20 percent (Mattke et al., 2013).

Figure 3 reports participation rates for different levels of incentives, first for the screening and HRA stage and then for the fall activities.[15] The first set of three dark bars in Figure 3a show how screening participation varies as a function of the screening incentive. Increasing the screening incentive from $0 to $100 boosts participation from 46.9 percent to 58.5 percent.

[13] One could assign all variables to a single family of hypotheses. This is unappealing, however, because it assigns equal importance to all outcomes when in fact some outcomes (e.g., total medical spending) are of much greater interest than others. Instead, our approach groups together variables that measure related outcomes and that originate from similar data sources.

[14] We have made our generalized Stata code module publicly available for other interested researchers to use. It can be installed by typing "`ssc install wyoung, replace`" at the Stata prompt. We provide additional documentation of this multiple testing adjustment in Appendix C.

[15] We report the results for spring activities, which are very similar to those for the fall, in Appendix A.

This difference is statistically significant at the 5 percent level. Increasing the screening incentive to $200 increases turnout further, to 62.5 percent. The second set of two dark bars in Figure 3a shows screening participation as a function of the wellness activity incentives. Increasing the activity incentive from $25 to $75 increases turnout from 53.6 percent to 58.4 percent, indicating that at least some participants were forward looking: they understood that they needed to first complete the screening and HRA in order to later be eligible to sign up for a wellness activity.

Table 2 provides formal statistical testing of the patterns described above for health screening participation. Panel A reports estimates of equations (1) and (2), using the completion of the screening and HRA as the outcome variable. Columns (1) and (2) pool together groups B and C, while columns (3) and (4) estimate the effects separately for groups B and C. The omitted group in each specification is group A25: members who were assigned a $0 screening incentive and a $25 wellness incentive.

As reported in Panel A, the baseline participation rate for the screening and HRA in the omitted group is 44.5 percent (see column (1) or (3)). Column (3) of Panel A shows that the screening/HRA completion rates of treatment groups B and C are larger than those of group A by 11.6 ($p < 0.001$) and 15.6 ($p < 0.001$) percentage points, respectively. In addition, the difference between group B and C is marginally significant ($p = 0.05$). We also estimate that a $75 wellness incentive increases screening and HRA completion by 4.9 percentage points relative to a $25 wellness incentive ($p < 0.01$). Comparing columns (1) and (3) to columns (2) and (4), respectively, shows that controlling for baseline stratification variables has very little effect on the point estimates.

We find consistently positive, but marginally diminishing, effects of monetary rewards on screening and HRA participation. The optimal reward amount depends on the marginal cost and marginal benefit associated with additional participation. Using our participation results, it is straightforward to provide some basic estimates of marginal cost using data on the field costs of our study.

The lightly shaded bars in Figure 3a report the realized average variable costs for treatment groups with different monetary rewards. The average variable costs are equal to the average monetary incentives paid to the group plus the costs of providing the health screening, the HRA, and the wellness activities.[16] We calculate the marginal cost of the additional participation induced by each reward by dividing the increase in average cost for each group by the corresponding increase in participation.[17] The results of those calculations are plotted in Figure 3c. The marginal cost is increasing in the share of employees participating and is largest (at $1,750) for group C, whose members received $200 if they completed a screening and HRA. All else equal, this estimate implies that the optimal screening incentive is less than $200 if the marginal benefit associated with additional participation in group C is less than $1,750. Interestingly, the marginal cost of using activity incentives to increase screening participation lines up closely with that of the screening incentives.

We repeat this exercise for fall activity participation in Figures 3b and 3d. Here, a different pattern emerges. Screening incentives have only a small effect on fall activity completion, and, as a result, generate a relatively steep marginal cost curve. On the other hand, wellness activity incentives have a sizeable effect on activity completion, and exhibit a much flatter marginal cost.[18]

Panels B and C of Table 2 report that the screening incentives for groups B and C increase the completion probability for the fall or spring wellness activity by about 4-5 percentage points ($0.004 \leq p \leq 0.03$). Finally, the $75 wellness incentive, as compared to a $25 incentive, generates a 12 percentage point increase in the likelihood of completing a fall or spring wellness activity ($p < 0.001$). This last effect is sizeable when compared to a

[16]Our variable cost measure does not account for paid time off or the fixed costs of managing the iThrive intervention. The health screening and HRA cost $78.22 per participant. This includes the costs of purchasing a fingerstick blood test, hiring nurses to administer the test, and licensing the HRA. The wellness activities cost an average of $26.07 per enrollee per semester. Employees who declined to participate in the health screening are assigned a variable cost of $0.

[17]For the $25 activity incentive and $0 screening incentive groups, the marginal cost is calculated relative to a baseline of 0 percent participation and $0 average variable cost. Thus, the marginal cost for these two groups is simply the group's average variable cost divided by its participation rate.

[18]We find qualitatively similar patterns for spring activity participation, which we present in Appendix Figure A.1.

baseline completion rate of 18.2 percent in the fall and 13.7 percent in the spring for group A (see column (1) or (3)).

Overall, we find that financial incentives have a significant, but diminishing, effect on health screening participation. This suggests that when screening incentives are large, further increases in reward sizes will result in larger transfers to existing participants but little change in total participation. By contrast, we find that screening incentives have little effect on subsequent wellness activity participation, while wellness incentives have a relatively large effect. For this reason, the back-loaded wellness activity incentives are arguably more cost-effective than the upfront screening incentives: they are about as effective as screening incentives in increasing screening participation—as evidenced by similar marginal cost curves (Figure 3c)—and at the same time are more efficient at increasing wellness activity completion—i.e., they have a flatter marginal cost curve (Figure 3d).

4.2 Selection

4.2.1 Average Selection

Next, we characterize the types of workers most likely to participate in our wellness program. We focus on medical spending and health behaviors, which are primary targets of wellness programs, and on salary, which is useful for understanding the redistribution effects of these programs. Selection results for the full set of pre-specified observables are presented in Appendix Tables A.1a through A.1d.

Table 3 reports our main selection results, as estimated by equation (3). We test for selection at three different, sequential points in the study: completing the health screening and HRA; completing a fall wellness activity; and completing a spring wellness activity. Column (1) reports the mean of the selection variable of interest for employees assigned to one of our study's treatment groups. Columns (3)-(5) report the difference in means between those employees who successfully completed the participation outcome of interest and those who did not. We also report family-wise p-values in brackets that account for the number of selection variables in each "family."[19]

Column (3) of the first row of Table 3 reports that employees who completed the screening and HRA spent, on average, \$116.1 per month less on health care in the 13 months prior to our study than employees who did not participate. This pattern of advantageous selection is strongly significant using conventional inference ($p = 0.026$), and remains marginally significant even after adjusting for the five outcomes in this family (family-wise $p = 0.080$). The magnitude is also economically significant, representing about 25 percent of the \$479 in average monthly spending (column (1)). Columns (4) and (5) present further evidence of advantageous selection into the fall and spring wellness activities, although in these cases the magnitude of selection falls by half and becomes statistically insignificant.

In contrast, the second row of Table 3 reports that employees participating in our wellness program were *more* likely to have non-zero medical spending at baseline than non-participants, by about 5 percentage points (family-wise $p \leq 0.021$), for all three participation outcomes. When combined with our results from the first row on average spending, this suggests that our wellness program is more attractive to employees with moderate spending than to employees in either tail of the spending distribution.

We investigate these results further in Figure 4, which displays the empirical distributions of prior spending for those employees who participated in screening and for those who did not. We perform two tests of the equality of the spending distributions across these two samples: Pearson's chi-squared test and the non-parametric Kolmogorov-Smirnov test.[20] Both tests strongly reject the null hypothesis that these two samples were drawn from the same distribution (Chi-squared $p < 0.001$; Kolmogorov-Smirnov $p = 0.007$). More specifically, Figure 4 reveals a "tail-trimming" effect: participating (screened) employees are less likely to be high spenders ($> \$2,338$ per month), but they are also less likely to be low spenders (\$0

[19]The seven families of outcome variables are defined in Section 3.4. The family-wise p-values reported in Table 3 account for all the variables in the family, including ones that are not reported in the main text. An expanded version of Table 3 that reports estimates for all pre-specified outcomes is provided in Appendix Tables A.1a through A.1d.

[20]These tests were not specified in our pre-analysis plan.

per month). Because medical spending is highly skewed to the right, the overall effect on the mean among participants is negative, which explains the advantageous selection effect reported in the first row of Table 3.

Panel B of Table 3 reports selection estimates for income. The first row reports that the average annual salary of participants is lower than that of non-participants, significantly so for the fall and spring wellness activities (family-wise $p \leq 0.012$). This initially suggests that participants are disproportionately lower-income. Yet, the second row of Panel B reports that the share of screening participants in the first (bottom) quartile of income is actually 6.9 percentage points *lower* than the share among non-participants (family-wise $p < 0.001$). Columns (4) and (5) also report negative, albeit smaller, selection effects for the fall and spring wellness activities. We again delve deeper by comparing the entire empirical distributions of income for participants and non-participants in Figure 5. We can reject that these two samples came from the same distribution ($p \leq 0.02$). As in Figure 4, we again find a tail-trimming effect: participating employees are less likely to come from either tail of the income distribution.

Lastly, we test for differences in baseline health behaviors as measured by our administrative data variables. The first row of Panel C in Table 3 reports that the share of screening participants who had previously participated in one of the IL Marathon/5K/10K running events is 8.9 percentage points larger than the share among non-participants (family-wise $p < 0.001$), a sizeable difference that represents over 75 percent of the mean participation rate of 11.8 percent (column (1)). This selection effect is even larger for the fall and spring wellness activities. The second row of Panel C reports that participants also visited the campus gym facilities more frequently, although these selection effects are only statistically significant for screening and HRA completion (family-wise $p = 0.013$).

Prior studies have raised concerns that the benefits of wellness programs accrue primarily to higher-income employees with lower health risks (Horwitz, Kelly and DiNardo, 2013). Our results are broadly consistent with these concerns: participating employees are less likely to have very high medical spending, less likely to be in the bottom quartile of income, and more likely to engage in healthy activities such as running or visiting the gym. At the same time, participating employees are also less likely to have very low medical spending or have very high incomes, which suggests a more nuanced story.

4.2.2 Marginal Selection

Our study design allows us to characterize not only how participants differ from non-participants on average, but also how the *marginal* participant varies as we increase incentives. As reported previously in Table 3, screening participants had *lower* baseline medical spending than non-participants, on average. Figure 6a (orange bars) shows how this pattern of selection varies by screening incentive size. For example, participants in the treatment groups with \$100 and \$200 screening incentives spent, on average, \$79 *more* per month ($p = 0.06$) than participants in the treatment group with a \$0 screening incentive. At low levels of screening incentives, wellness programs attract below-average spenders, but as incentive levels increase, the marginal participants have spending levels that are higher than the average participant. Thus, over the range of incentives we offer, increasing the size of the screening incentive reduces the average amount of selection.

By contrast, Figure 6a (blue bars) illustrates a different pattern for wellness activity incentives: as we increase activity incentives, the marginal participant has significantly *lower* spending ($p = 0.03$). While we have less power for other outcomes, we find similar selection patterns when using pre-intervention health behaviors as a proxy for health status.[21] As we increase screening incentives, the marginal participant is potentially less likely to have participated in a prior marathon or have used the campus gym. Conversely, increasing wellness activity incentives potentially draws in marginal participants with a higher propensity for gym use. Thus, the selection patterns are potentially heterogeneous across type of incentive. As was the case when we examined the marginal cost of increasing participation, the type

[21] Marginal selection patterns with respect to income and non-zero health spending are provided in Appendix A.

of incentive matters when examining selection into wellness programs.

4.2.3 Health Care Cost-Savings via Selection

The selection patterns we have uncovered may provide, by themselves, a potential motive for firms to offer wellness programs. We have shown that wellness participants have lower medical spending on average than non-participants. If wellness programs differentially increase the recruitment or retention of these types of employees, then the accompanying reduction in health care costs will save firms money.[22]

A simple back-of-the-envelope calculation demonstrates this possibility. In our setting, 39 percent ($= 4,834/12,459$) of eligible employees enrolled into our study, and 56 percent of the treatment group completed a screening and health assessment (Figure 2). Participating employees spent on average \$132.7 per month less than non-participants in the post-period (Table 5, column 4), which translates into an annual spending difference of \$1,592. When combined with average program costs of \$271 per participant, this implies that the employer would need to increase the share of employees who are similar to wellness participants by 4.5 ($= 0.39 \times 0.56 \times 271/(1592 - 271)$) percentage points in order for the resulting reduction in medical spending to offset the entire cost of the wellness program. To be clear, this calculation does not imply that adoption of workplace wellness programs is socially beneficial. But, it does provide a profit-maximizing rationale for firms to adopt wellness programs, *even in the absence of any direct effects on health, productivity, or medical spending.*

4.3 Causal Effects

4.3.1 Intent-to-Treat

Finally, we estimate the causal, intent-to-treat (ITT) effect of our intervention on three domains of outcomes: medical spending, employment and productivity, and health behaviors.

[22]Wellness participants differ from non-participants along other dimensions as well (e.g., health behaviors). Because it is difficult in many cases to sign, let alone quantify, a firm's preferences over these other dimensions, we focus our cost-savings discussion on the medical spending consequences.

Table 4 reports estimates of equation (4) for all administratively measured outcomes, as well as a select set of outcomes from the one-year follow-up survey. An expanded version of this table reporting 39 administrative and survey outcomes is provided in Appendix Tables A.2a through A.2f.

We report ITT estimates using three specifications. The first includes no control variables. Our second specification includes fixed effects for the 69 strata used for stratified random assignment at baseline. Because the probability of treatment assignment was constant across strata, these controls are included not to reduce bias, but to improve the precision of the treatment effect estimates (Bruhn and McKenzie, 2009). Our third specification includes a set of baseline outcomes and covariates chosen via Lasso, as described in Section 3.3.

Medical spending We do not detect statistically significant effects of treatment on average medical spending over the first 12 months (August 2016 - July 2017) of the wellness intervention in any of our specifications. Column (2) of the first row of Table 4 shows that the difference in average spending between treatment and control was only $4.1 per month. The point estimate increases slightly when using either of our control strategies (columns (3) or (4)) but remains small and statistically indistinguishable from zero. The post-Lasso specification generates a significant improvement in precision, with a standard error about 25 percent smaller than that of either the no-control or strata fixed effects specifications. In the rest of Panel A, we continue to find small and insignificant results for different subcategories of spending, as well as the probability of any spending over this 12-month period.

Panels (a) and (b) of Figure 7—which reproduce the basic results for total and non-zero spending presented in Panel A, column (2) of Table 4—reveal no significant differences in average spending or probability of any spending between treatment and control. However, these results do not rule out mean-preserving treatment effects that alter other moments of the distribution. We investigate this possibility in Panel (c) of Figure 7, which displays the empirical distributions of spending for the treatment and control groups, but fail to observe any clear differences between these two groups. This is confirmed formally by Pearson's chi-squared test and the Kolmogorov-Smirnov test, which both fail to reject the null hypothesis that the control and treatment samples were drawn from the same spending distribution ($p = 0.867$ and $p = 0.458$, respectively).

Finally, we investigate the potential for spending treatment effects to vary by treatment arm. Those results, which are available in Appendix Tables A.4a and A.4b, show no evidence of meaningful differences in spending effects across treatment arms.

Employment and productivity Next, we estimate the effect of treatment on a variety of employment and productivity outcomes. As reported in Panel B of Table 4, we do not detect statistically significant effects on any of the three outcomes that are administratively measured: annual salary, the probability of job termination after 12 months of the wellness intervention, and sick leave taken. Turning to variables measured during the one-year follow-up survey, we find no statistically significant effects on most self-reported employment and productivity measures, including being happier at work than last year or feeling very productive at work. The only exception is that individuals in the treatment group are 5.7 percentage points (7.2 percent) more likely (family-wise $p < .001$) to believe that management places a priority on health and safety (column (2), Table 4). Appendix Tables A.2c and A.2d report ITT estimates for all pre-specified administrative and survey productivity measures.

Health behaviors Finally, we investigate health behaviors, which may respond more quickly to a wellness intervention than medical spending and productivity outcomes. Our main results are reported in Panel C of Table 4. We find very small and statistically insignificant effects of treatment assignment on participation in any running event of the April 2017 Illinois Marathon (i.e. 5K, 10K, and half/full marathons). Similarly, we do not find meaningful effects on the average number of days per month that an employee visits a campus recreation facility. However, we do find that individuals in the treatment group are nearly 4 percentage points more likely ($p < .01$) to report having a previous health screening. This provides evidence that our program provided biometric health information to a significant number of employees who report not previously being screened, and did not completely crowd out screenings that would have otherwise occurred within the first year of our study.

Discussion Across all 39 outcome variables we examine, we only find two statistically significant effects of our intervention: an increase in the number of employees who ever received a health screening, and an increase in the number who believe that management places a priority on health and safety.[23] The next section addresses the precision of our estimates by quantifying what effects we can rule out. But first, we mention two caveats.

First, our post-period only includes one year of data. While we do not find significant effects for most of the outcomes we examine, it is possible that longer-run effects may emerge in the second or third year following the intervention. Second, our analysis assumes that the control group was unaffected by the intervention. The research team's contact with the control group was confined to the communication procedures employed for the 2016 and 2017 online surveys. Although we never shared details of the intervention with the control group, some of them may have learned about it from their colleagues. To evaluate how often this occurred, we asked study participants on the 2017 follow-up survey whether they ever talked about the iThrive workplace wellness program with any of their coworkers. Only 3 percent of the control group responded affirmatively, compared to 44 percent of the treatment group.

4.3.2 Comparison to Prior Studies

We now compare our estimates to the prior literature, which has focused on medical spending and absenteeism. This exercise employs a spending estimate derived from a data sample that winsorizes (top-codes) medical spending at the one percent level (see Column 3 of Table 6). We do this to reduce the influence of a small number of extreme outliers on the precision of

[23] We show in the appendix that these two effects are driven by the health screening component of our intervention rather than the wellness activity component.

our estimate, as has been done in prior studies (e.g. Clemens and Gottlieb, 2014).[24]

Figure 8 illustrates how our estimates compare to the prior literature.[25] The top-left figure in Panel (a) plots the distribution of the intent-to-treat (ITT) point estimates for medical spending from 22 prior workplace wellness studies. The figure also plots our ITT point estimate for total medical spending from Table 4, and shows that our 95-percent confidence interval rules out 20 of these 22 estimates. For ease of comparison, all effects are expressed as percent changes. The bottom-left figure in Panel (a) plots the distribution of treatment-on-the-treated (TOT) estimates for health spending from 33 prior studies, along with the IV estimates from our study. In this case, our 95-percent confidence interval rules out 23 of the 33 studies. Overall, our confidence intervals rule out 43 of 55 (78 percent) prior ITT and TOT point estimates for health spending.[26] The two figures in Panel (b) repeat this exercise for absenteeism, and show that our estimates rule out 53 of 60 (88 percent) prior ITT and TOT point estimates for absenteeism. Across both sets of outcomes, we rule out 96 of 115 (83 percent) prior estimates.

We can also combine our spending and absenteeism estimates with our cost data to calculate a return on investment (ROI) for workplace wellness programs. The 99 percent confidence intervals for the ROI associated with our intervention rule out the widely cited savings estimates reported in the meta-analysis of Baicker, Cutler and Song (2010).[27]

[24] Winsorizing can introduce bias if there are heterogeneous treatment effects in the tails of the spending distribution. However, Figure 7c provides evidence of a consistently null treatment effect throughout the spending distribution. This evidence is further supported by Table 6, which shows that the point estimate of the medical spending treatment effect changes little after winsorization. For completeness, Appendix Figure A.3 illustrates the stability of the point estimate across a wide range of winsorization levels.

[25] Appendix B provides the sources and calculations underlying the point estimates reported in Figure 8.

[26] If we do not winsorize medical spending, we rule out 37 of 55 (67 percent) prior health studies.

[27] The first year of the iThrive program cost \$152 ($= \$271 \times 0.56$) per person assigned to treatment. This is a conservative estimate because it does not account for paid time off or the fixed costs of managing iThrive. Focusing on the first year of our intervention and assuming that the cost of a sick day equals \$240, we calculate that the lower bounds of the 99 percent confidence intervals for annual medical and absenteeism costs are -\$415 ($= (15.4 - 2.577 \times 19.4) \times 12$) and -\$74 ($= (0.195 - 2.577 \times 0.196) \times 240$), which imply ROI lower bounds of 2.73 and 0.49, respectively. By comparison, Baicker, Cutler and Song (2010) found that spending fell by \$3.27, and absenteeism costs fell by \$2.73, for every dollar spent on wellness programs.

4.3.3 IV versus OLS

Across a variety of outcomes, we find very little evidence that our intervention had any effect in its first year. As shown above, our results differ from many prior studies that find significant reductions in health expenditures and absenteeism. One possible reason for this discrepancy is the presence of advantageous selection bias in these other studies, which are generally not randomized controlled trials. A second possibility is that there is something unique about our setting. We investigate these competing explanations by performing a typical observational (OLS) analysis and comparing its results to those of our experimental estimates.[28] Specifically, we estimate

$$Y_i = \alpha + \gamma P_i + \Gamma X_i + \varepsilon_i, \tag{5}$$

where Y_i is the outcome variable as in (4), P_i is an indicator for participating in the screening and HRA, and X_i is a vector of variables that control for potentially non-random selection into participation.

We estimate two variants of equation (5). The first is an instrumental variables (IV) specification that includes observations for individuals in the treatment or control groups, and uses treatment assignment as an instrument for completing the screening and HRA. The second variant estimates equation (5) using OLS, restricted to individuals in the treatment group. For each of these two variants, we estimate three specifications similar to those used for the ITT analysis described above (no controls, strata fixed effects, and post-Lasso).[29] This generates six estimates for each outcome variable. Table 5 reports the results for our

[28] This observational analysis was not specified in our pre-analysis plan.

[29] To select controls for the post-Lasso IV specification, we follow the "triple" selection strategy proposed in Chernozhukov, Hansen and Spindler (2015). This strategy first estimates three Lasso regressions of (1) the (endogenous) focal independent variable on all potential controls and instruments; (2) the focal independent variable on all potential controls; and (3) the outcome on all potential controls. It then forms a 2SLS estimator using instruments selected in step (1) and all controls selected in any of the steps (1)-(3). When the instrument is randomly assigned, as it is in our setting, the set of controls selected in steps (1)-(2) above will be the same, in expectation. Thus, we form our 2SLS estimator using treatment assignment as the instrument and controls selected in Lasso steps (2) or (3) of this algorithm.

primary outcomes of interest. The results for all pre-specified administrative and survey outcomes are reported in Appendix Tables A.3e-A.3f.

As in our previous ITT analysis, the IV estimates reported in columns (1)-(3) are small and indistinguishable from zero for nearly every outcome. By contrast, the observational estimates reported in columns (4)-(6) are frequently large and statistically significant. Moreover, the IV estimate rules out the OLS estimate for several key outcomes. Based on our most precise and well-controlled specification (post-Lasso), the OLS monthly spending estimate of −$88.1 (row 1, column (6)) lies outside the 95 percent confidence interval of the IV estimate of $38.5 with a standard error of $58.8 (row 1, column (3)). For participation in the 2017 IL Marathon/10K/5K, the OLS estimate of 0.024 lies outside the 99 percent confidence interval of the corresponding IV estimate of -0.011 (standard error = 0.011). For campus gym visits, the OLS estimate of 2.160 lies just inside the 95 percent confidence interval of the corresponding IV estimate of 0.757 (standard error = 0.656). Under the assumption that the IV (RCT) estimates are unbiased, these difference imply that even after conditioning on a rich set of controls, participants selected into our workplace wellness program on the basis of lower-than-average contemporaneous spending and higher-than-average health activity. This is consistent with the evidence presented in Section 3.2 that pre-existing spending is lower, and pre-existing behaviors are healthier, among participants than among non-participants.

In addition, the observational estimates presented in columns (4)-(6) are in line with estimates from previous observational studies, which suggests that our setting is not particularly unique. In the spirit of LaLonde (1986), these estimates demonstrate that even well-controlled observational analyses can suffer from significant selection bias in our setting, suggesting that similar biases might be at play in other wellness program settings as well.

5 Conclusion

This paper presents a first set of findings from the Illinois Workplace Wellness Study. We find a large but diminishing effect of incentives on wellness participation. At large incentive levels, further increases have little effect on participation and thus primarily just increase compensation for inframarginal participants. We also find that employees who chose to participate in our wellness program were less likely to be in the bottom quartile of the income distribution, and already had lower medical spending and healthier behaviors than non-participants prior to our intervention. These selection results have two implications. First, they suggest that workplace wellness programs shift costs onto low-income employees with high health care spending and poor health habits. Second, the large magnitude of our spending estimate suggests the primary value of wellness programs to firms may be their potential to attract and retain workers with low health care costs. All else equal, reducing the share of non-participating employees by just 4.5 percentage points would lower total medical spending in our setting by an amount sufficient to pay for our entire wellness program.

After one year we find no significant effects of our wellness program on the many outcomes we examine, with two exceptions: employees are more likely to have received a health screening and to believe that the employer places a priority on worker health and safety. Our null results are economically meaningful: we can rule out 83 percent of the medical spending and absenteeism estimates from the prior literature, along with the average ROIs calculated by Baicker, Cutler and Song (2010) in a widely cited meta-analysis. Our OLS estimate is consistent with results from the prior literature, but ruled out by our IV estimate, suggesting that non-RCT studies in this literature suffer from selection bias.

Although we fail to find effects of our workplace wellness program on the majority of the outcomes in our analysis, we emphasize that we have only examined outcomes in the first year following randomization. It is possible that meaningful effects may emerge in later years, although if there is sufficient employee turnover then these benefits may not accrue to the employer who made the initial investment in workplace wellness. The net effect is therefore an empirical question. As a part of the Illinois Workplace Wellness Study, we will continue to collect data so that we can estimate long-run effects in future research.

References

Abadie, Alberto, Susan Athey, Guido Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?" *NBER Working Paper #24003*.

Abraham, Jean, and Katie M White. 2017. "Tracking The Changing Landscape Of Corporate Wellness Companies." *Health Affairs*, 36(2): 222–228.

Aldana, Steven G. 2001. "Financial impact of health promotion programs: a comprehensive review of the literature." *American Journal of Health Promotion*, 15(5): 296–320.

Askelson, Natoshia M, Brad Wright, Suzanne Bentler, Elizabeth T Momany, and Peter Damiano. 2017. "Iowa's Medicaid Expansion Promoted Healthy Behaviors But Was Challenging To Implement And Attracted Few Participants." *Health Affairs*, 36(5): 799–807.

Baicker, Katherine, David Cutler, and Zirui Song. 2010. "Workplace wellness programs can generate savings." *Health Affairs*, 29(2): 304–311.

Baxter, Siyan, Kristy Sanderson, Alison J Venn, C Leigh Blizzard, and Andrew J Palmer. 2014. "The relationship between return on investment and quality of study methodology in workplace health promotion programs." *American Journal of Health Promotion*, 28(6): 347–363.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *The Review of Economic Studies*, 81(2): 608–650.

Bhattacharya, Jayanta, and William B Vogt. 2014. "Employment and adverse selection in health insurance." Vol. 17, 79–104.

Bruhn, Miriam, and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, 1(4): 200–232.

Burd, Steven A. 2009. "How Safeway Is Cutting Health-Care Costs." *The Wall Street Journal*. http://www.wsj.com/articles/SB124476804026308603.

Cawley, John. 2014. "The Affordable Care Act permits greater financial rewards for weight loss: a good idea in principle, but many practical concerns remain." *Journal of Policy Analysis and Management*, 33(3): 810–820.

Chapman, Larry S. 2012. "Meta-evaluation of worksite health promotion economic return studies: 2012 update." *American Journal of Health Promotion*, 26(4): 1–12.

Charness, Gary, and Uri Gneezy. 2009. "Incentives to exercise." *Econometrica*, 77(3): 909–931.

Chernozhukov, Victor, Christian Hansen, and Martin Spindler. 2015. "Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments." *American Economic Review*, 105(5): 486–90.

Chiappori, Pierre-André, and Bernard Salanié. 2013. "Asymmetric information in insurance markets: Predictions and tests." In *Handbook of insurance*. 397–422. Springer.

Clemens, Jeffrey, and Joshua D. Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, 104(4): 1320–1349.

Cutler, David M, and Richard J Zeckhauser. 2000. "The anatomy of health insurance." *Handbook of health economics*, 1: 563–643.

Fout, Betty, Daniel Weinberg, Nicholas Bill, Katherine Kahn, and Cary Sennett. 2013. "Evaluation of the Senior Risk Reduction Demonstration (SRRD) Under Medicare: Final Evaluation Report." *IMPAQ International*.

Geruso, Michael, and Timothy J. Layton. 2017. "Selection in Health Insurance Markets and Its Policy Remedies." *Journal of Economic Perspectives*, 31(4): 23–50.

Gowrisankaran, Gautam, Karen Norberg, Steven Kymes, Michael E Chernew, Dustin Stwalley, Leah Kemper, and William Peck. 2013. "A hospital system's wellness program linked to health plan enrollment cut hospitalizations but not overall costs." *Health Affairs*, 32(3): 477–485.

Gubler, Timothy, Ian Larkin, and Lamar Pierce. 2017. "Doing well by making well: The impact of corporate wellness programs on employee productivity." *Management Science*.

Haisley, Emily, Kevin G Volpp, Thomas Pellathy, and George Loewenstein. 2012. "The impact of alternative incentive schemes on completion of health risk assessments." *American Journal of Health Promotion*, 26(3): 184–188.

Handel, Benjamin, and Jonathan Kolstad. 2017. "Wearable Technologies and Health Behaviors: New Data and New Methods to Understand Population Health." *American Economic Review: Papers and Proceedings*, 107(5): 481–85.

Harrison, Glenn W, and John A List. 2004. "Field experiments." *Journal of Economic literature*, 42(4): 1009–1055.

Horwitz, Jill R, Brenna D Kelly, and John E DiNardo. 2013. "Wellness incentives in the workplace: cost savings through cost shifting to unhealthy workers." *Health Affairs*, 32(3): 468–476.

Jaspen, Bruce. 2015. "Employers Boost Wellness Spending 17% From Yoga to Risk Assessments." *Forbes Online*. http://www.forbes.com/sites/brucejapsen/2015/03/26/employers-boost-wellness-spending-17-from-yoga-to-risk-assessments/#6a37ebf2350f.

LaLonde, Robert J. 1986. "Evaluating the econometric evaluations of training programs with experimental data." *The American economic review*, 604–620.

Lazear, Edward P. 2000. "Performance pay and productivity." *American Economic Review*, 90(5): 1346–1361.

Lazear, Edward P., and Paul Oyer. 2012. "Personnel Economics." *The Handbook of Organizational Economics*, 479–519. Princeton University Press.

Liu, Tim, Christos Makridis, Paige Ouimet, and Elena Simintzi. 2017. "Is Cash Still King: Why Firms Offer Non-Wage Compensation and the Implications for Shareholder Value." Available at SSRN: https://ssrn.com/abstract=3088067.

Mas-Colell, Andreu, Michael Dennis Whinston, Jerry R Green, et al. 1995. *Microeconomic theory.* Vol. 1, Oxford university press New York.

Mattke, Soeren, Christopher Schnyer, and Kristin R Van Busum. 2012. "A review of the US workplace wellness market." *The RAND Corporation*, Occasional Paper Series, https://www.dol.gov/sites/default/files/ebsa/researchers/analysis/health-and-welfare/workplacewellnessmarketreview2012.pdf.

Mattke, Soeren, Hangsheng Liu, John Caloyeras, Christina Y Huang, Kristin R Van Busum, Dmitry Khodyakov, and Victoria Shier. 2013. "Workplace wellness programs study: Final report." *RAND Health Quarterly*, 3(2).

McIntyre, Adrianna, Nicholas Bagley, Austin Frakt, and Aaron Carroll. 2017. "The dubious empirical and legal foundations of workplace wellness programs." *Health Matrix*, 27: 59.

Meenan, Richard T, Thomas M Vogt, Andrew E Williams, Victor J Stevens, Cheryl L Albright, and Claudio Nigg. 2010. "Economic evaluation of a worksite obesity prevention and intervention trial among hotel workers in Hawaii." *Journal of occupational and environmental medicine/American College of Occupational and Environmental Medicine*, 52(Suppl 1): S8.

Oyer, Paul. 2008. "Salary or benefits?" In *Work, Earnings and Other Aspects of the Employment Relation.* 429–467. Emerald Group Publishing Limited.

Pelletier, Kenneth R. 2011. "A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: update VIII 2008 to 2010." *Journal of occupational and environmental medicine*, 53(11): 1310–1331.

Phillips Erb, Kelly. 2018. "Congress Considers Bill To Make Gym Memberships, Fitness Classes Tax Deductible." *Forbes.* https://www.forbes.com/sites/kellyphillipserb/2018/07/14/congress-considers-bill-to-make-gym-memberships-fitness-classes-tax-deductible/#6572de1d4c25.

Reynolds, Chelsea. 2010. "Myth Surrounds Reform's 'Safeway Amendment." *Covering Health.* http://healthjournalism.org/blog/2010/01/myth-surrounds-reforms-safeway-amendment/.

Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information." *The quarterly journal of* *economics*, 629–649.

Royer, Heather, Mark Stehr, and Justin Sydnor. 2015. "Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company." *American Economic Journal: Applied Economics*, 7(3): 51–84.

Ryde, Gemma C, Nicholas D Gilson, Nicola W Burton, and Wendy J Brown. 2013. "Recruitment rates in workplace physical activity interventions: characteristics for success." *American Journal of Health Promotion*, 27(5): e101–e112.

Salop, Joanne, and Steven Salop. 1976. "Self-selection and turnover in the labor market." *The Quarterly Journal of Economics*, 619–627.

Soler, Robin E, Kimberly D Leeks, Sima Razi, David P Hopkins, Matt Griffith, Adam Aten, Sajal K Chattopadhyay, Susan C Smith, Nancy Habarta, Ron Z Goetzel, et al. 2010. "A systematic review of selected interventions for worksite health promotion: the assessment of health risks with feedback." *American Journal of Preventive Medicine*, 38(2): S237–S262.

Spence, Michael. 1973. "Job market signaling." *The Quarterly Journal of Economics*, 87(3): 355–374.

The Kaiser Family Foundation and Health Research and Educational Trust. 2016*a*. "Employer Health Benefits: 2016 Annual Survey." http://files.kff.org/attachment/Report-Employer-Health-Benefits-2016-Annual-Survey.

The Kaiser Family Foundation and Health Research and Educational Trust. 2016*b*. "Workplace Wellness Programs Characteristics and Requirements." http://files.kff.org/attachment/Issue-Brief-Workplace-Wellness-Programs-Characteristics-and-Requirements.

The Kaiser Family Foundation and Health Research and Educational Trust. 2017. "Employer Health Benefits: 2017 Annual Survey." http://files.kff.org/attachment/Report-Employer-Health-Benefits-Annual-Survey-2017.

Urminsky, Oleg, Christian Hansen, and Victor Chernozhukov. 2016. "Using Double-Lasso Regression for Principled Variable Selection."

Volpp, Kevin G, David A Asch, Robert Galvin, and George Loewenstein. 2011. "Redesigning employee health incentives—lessons from behavioral economics." *New England Journal of Medicine*, 365(5): 388–390.

Volpp, Kevin G, Leslie K John, Andrea B Troxel, Laurie Norton, Jennifer Fassbender, and George Loewenstein. 2008. "Financial incentive–based approaches for weight loss: a randomized trial." *Journal of the American Medical Association*, 300(22): 2631–2637.

Westfall, Peter H, and S Stanley Young. 1993. *Resampling-based multiple testing: Examples and methods for p-value adjustment.* Vol. 279, John Wiley & Sons.

Wilson, Charles. 1977. "A model of insurance markets with incomplete information." *Journal of Economic theory*, 16(2): 167–207.

Figure 1: Experimental Design of the Illinois Workplace Wellness Study

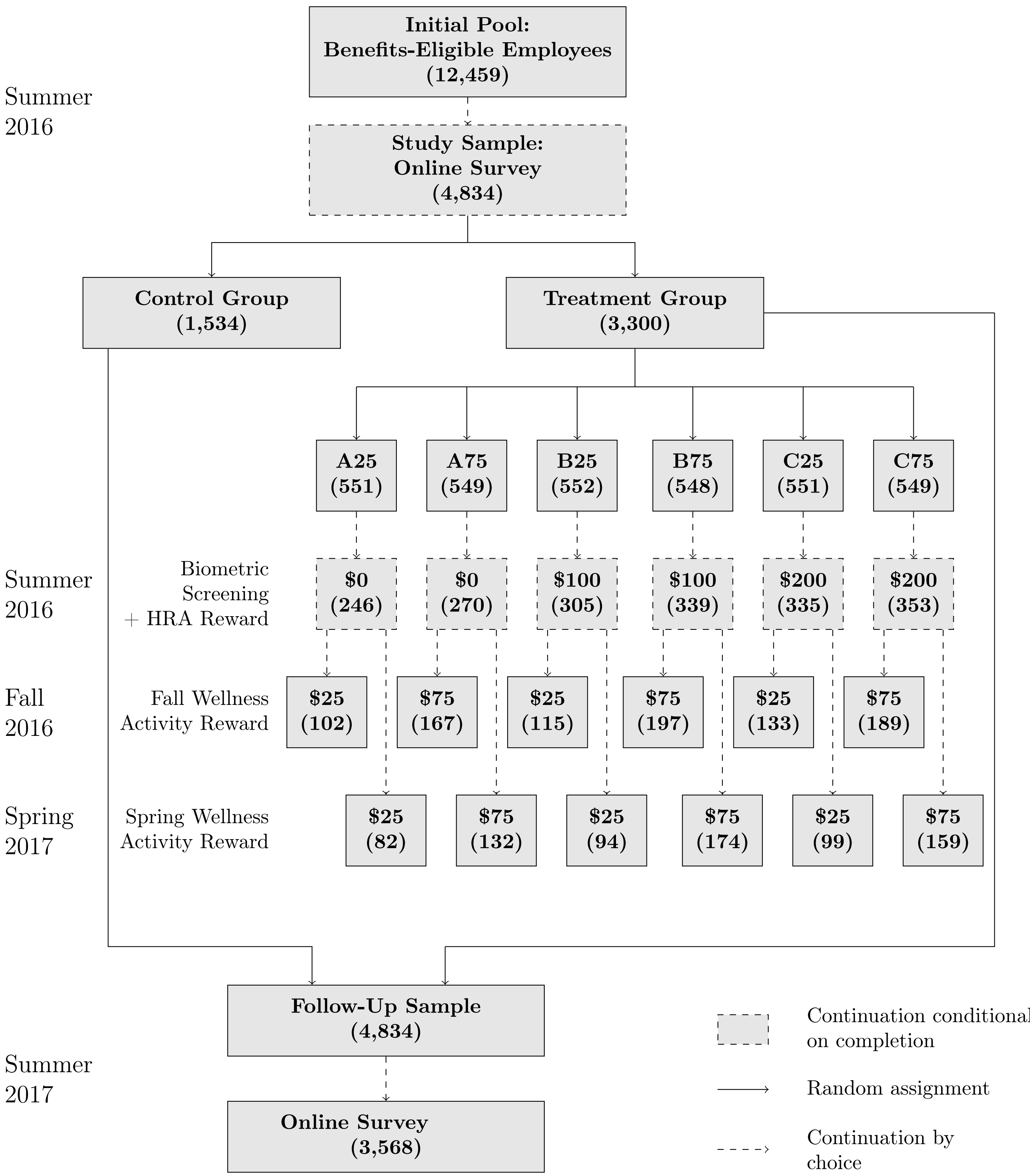
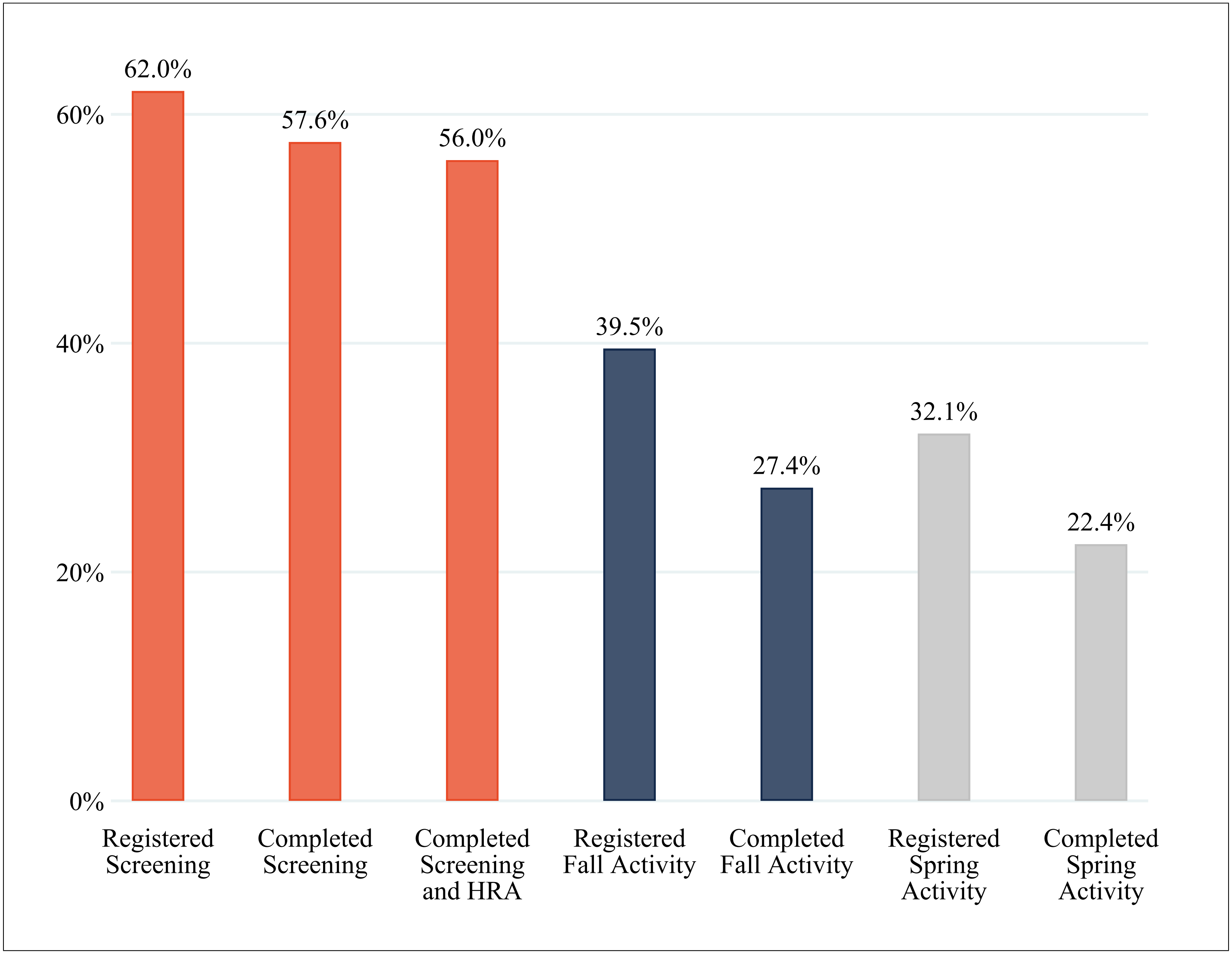
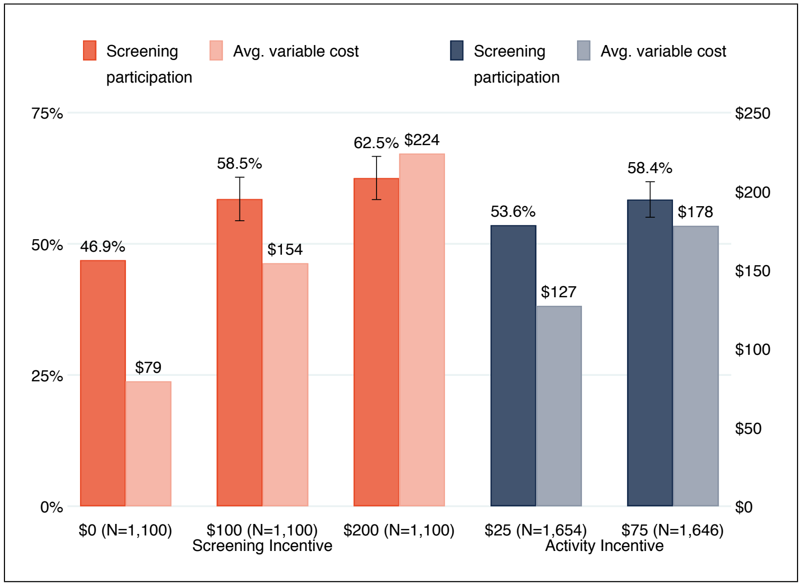


Figure 2: Employee participation rates in the first year of the workplace wellness program

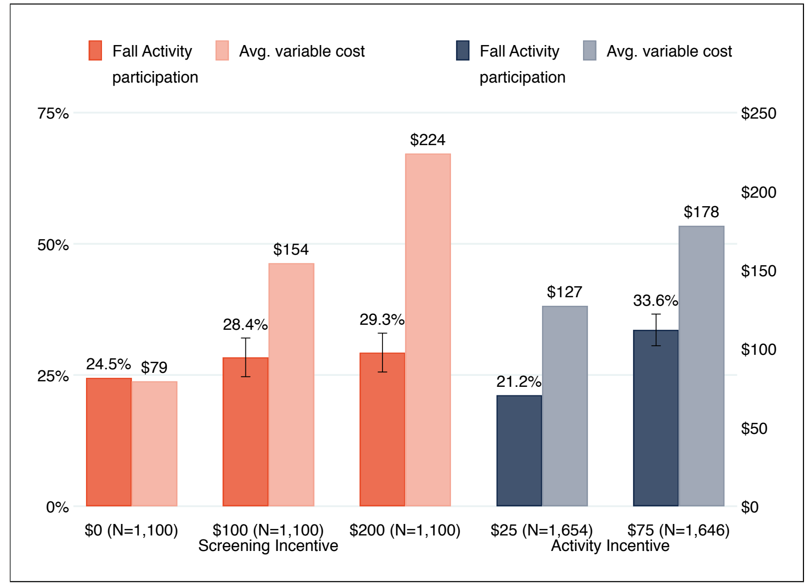


Notes: Participation rates are measured as a fraction of the treatment group ($N = 3,300$).

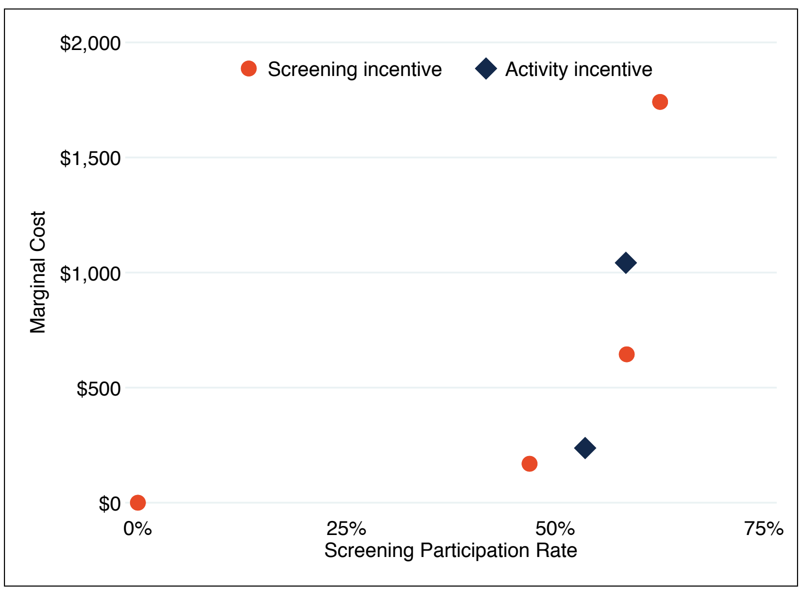
Figure 3: Marginal cost of inducing additional participation into health screening/HRA and fall activities



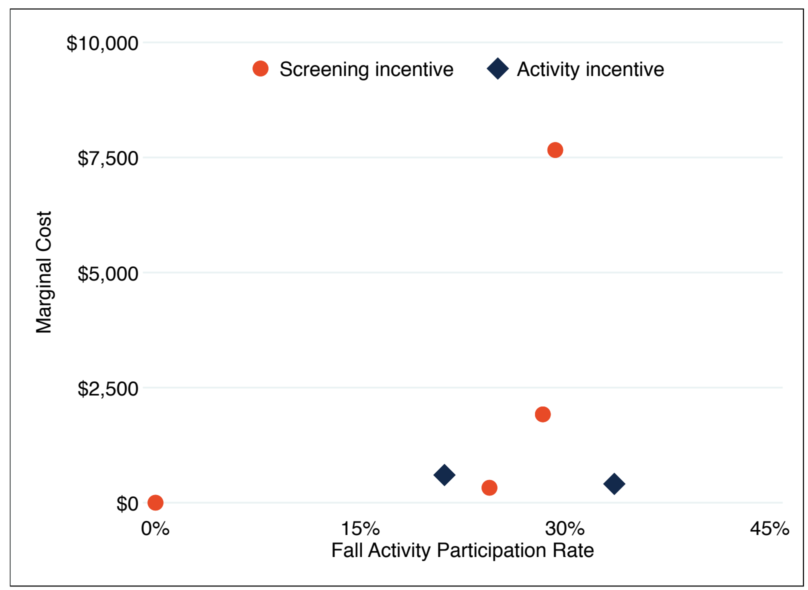
(a) Screening participation and costs



(b) Fall activities participation and costs



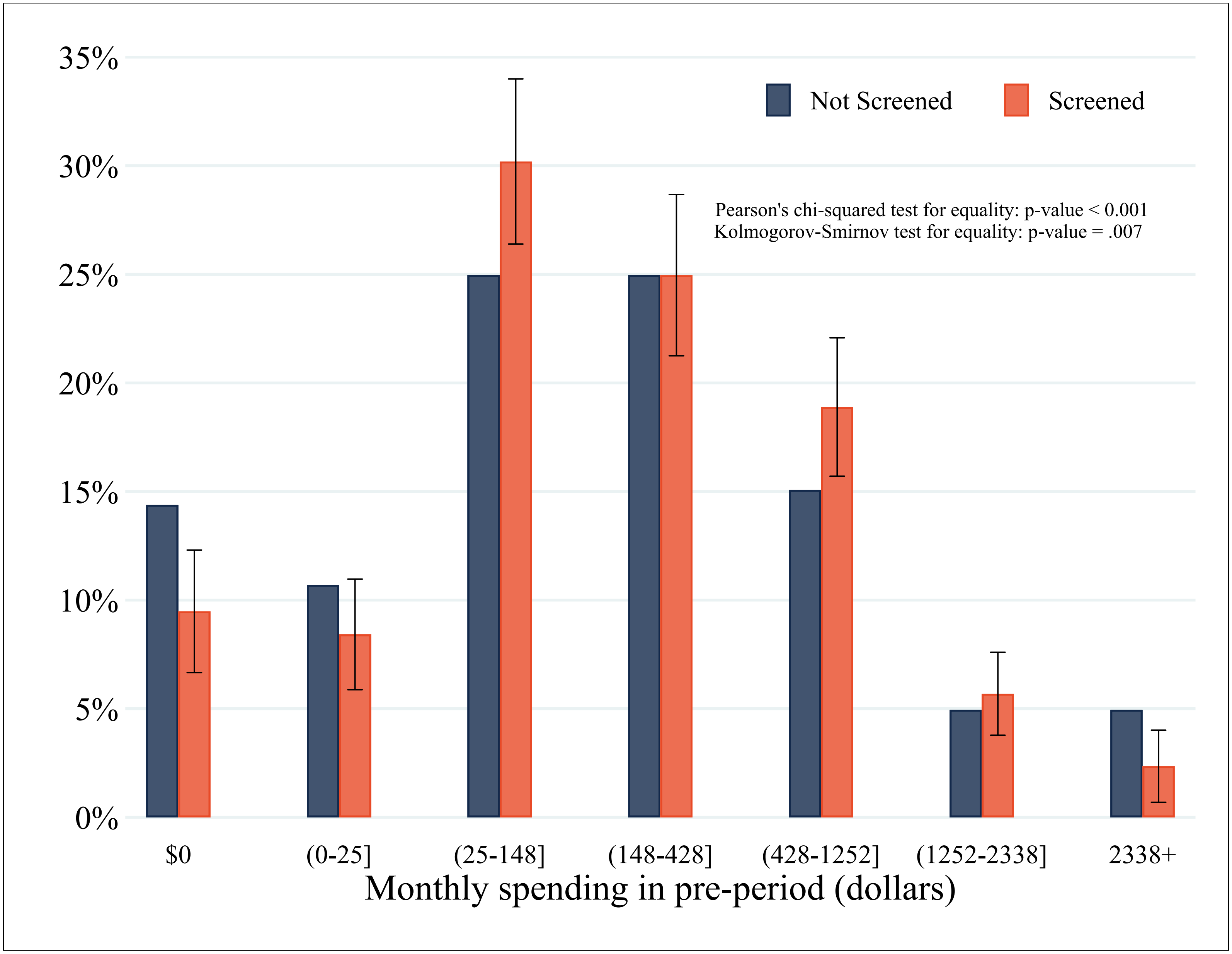
(c) Marginal cost of additional screening participation



(d) Marginal cost of additional fall activity participation

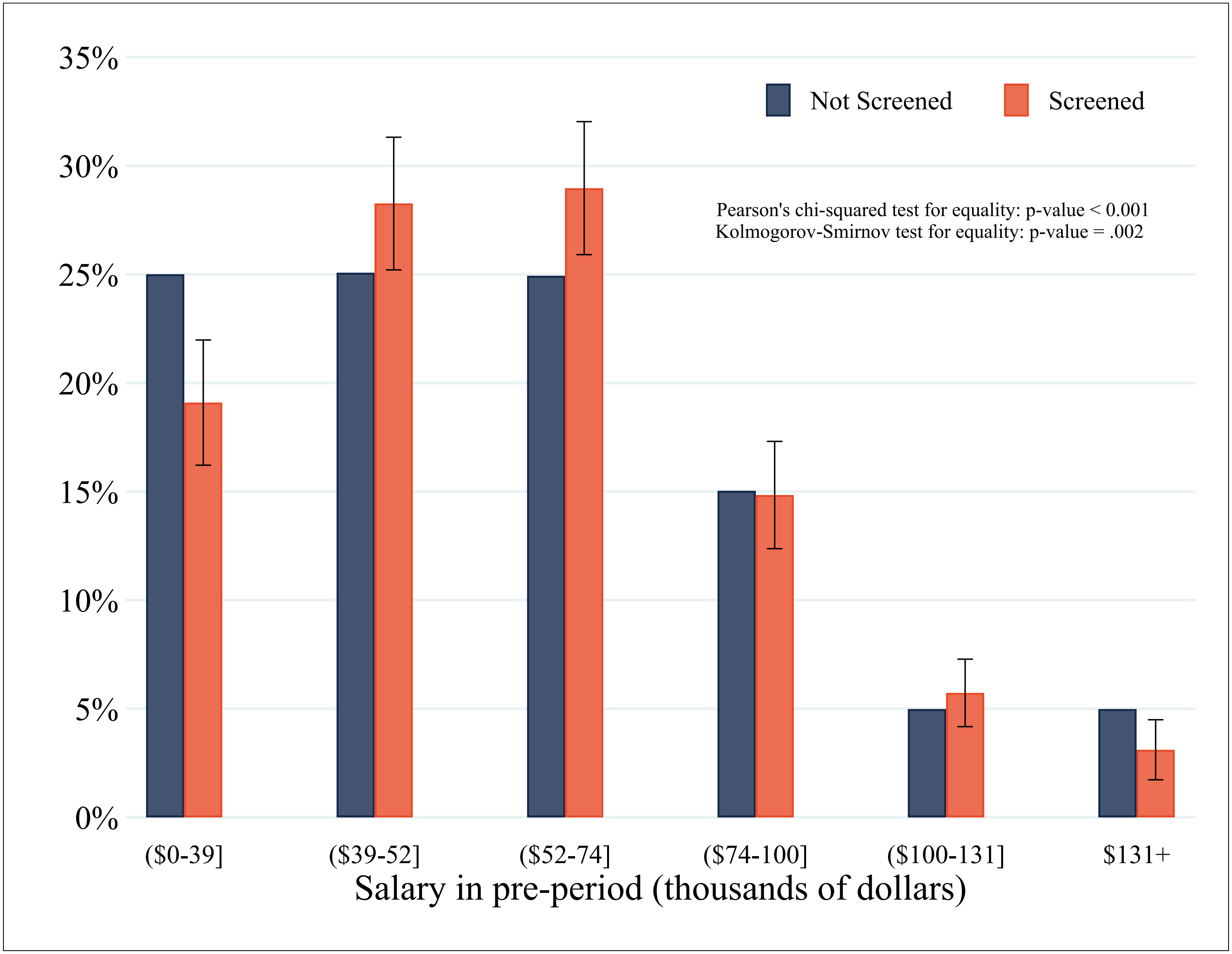
Notes: Panels (a) and (b) plot health screening participation rates (PR) and average variable costs (AVC) as a function of screening and activity incentives, separately for screening and fall activity particpation. Vertical bars display 95% confidence intervals on the difference in means relative to the lowest reward group. AVC includes costs of the health screening, HRA, and wellness activities. Panels (c) and (d) plot the implied marginal costs (MC), calculated as MC$=\frac{\Delta \text{AVC}}{\Delta \text{PR}}$. The MC of the control group (PR=0 percent) is set equal to 0.

Figure 4: Pre-intervention medical spending among treatment group, by participation status



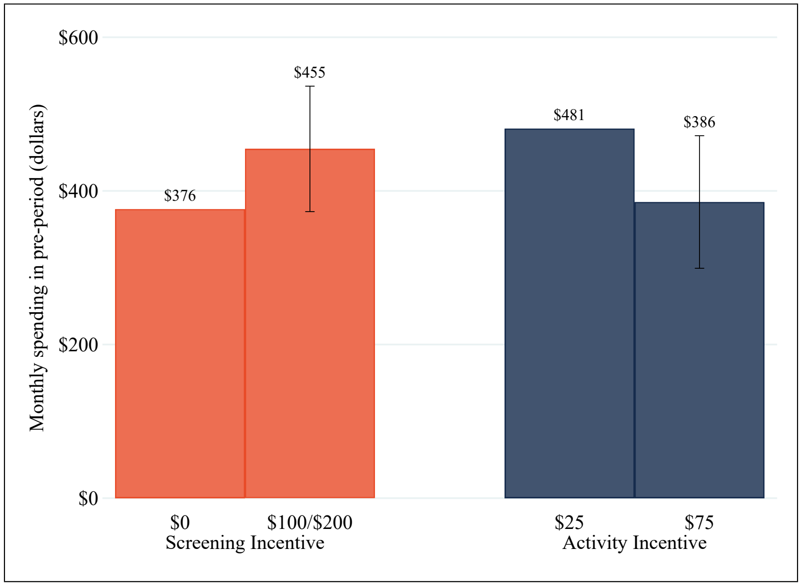
Notes: Data are from claims covering the period July 2015 - July 2016 ($N = 2,187$). The first two bins ($0 and $(0-25]$) include 25 percent of those not screened. The remaining five bins were defined to include 25, 25, 15, 5, and 5 percent of those not screened, respectively. The null hypothesis of the Pearson's chi-squared and the non-parametric Kolmogorov-Smirnov tests is that the two samples are drawn from the same distribution.

Figure 5: Pre-intervention salary among treatment group, by participation status

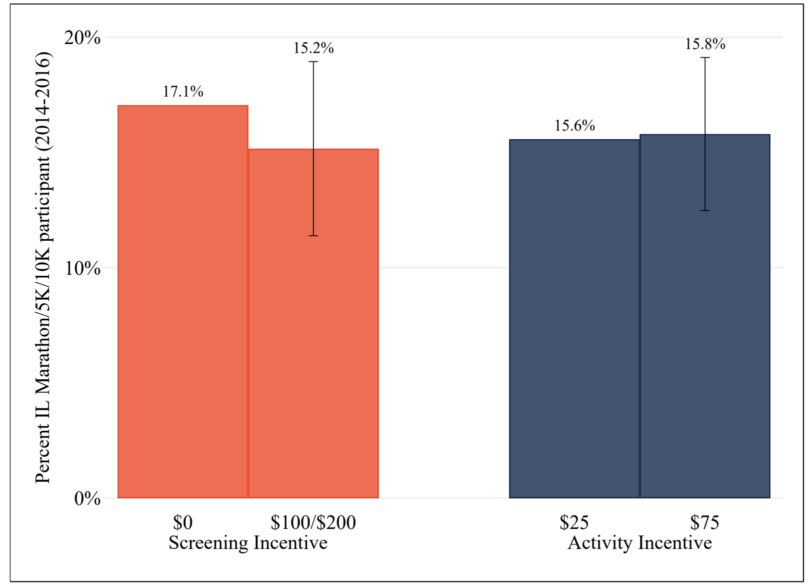


Notes: Salary was measured on June 1, 2016 ($N = 3,257$). The six bins were defined to include 25, 25, 25, 15, 5, and 5 percent of employees not screened, respectively. The null hypothesis of the Pearson's chi-squared and the non-parametric Kolmogorov-Smirnov tests is that the two samples are drawn from the same distribution.

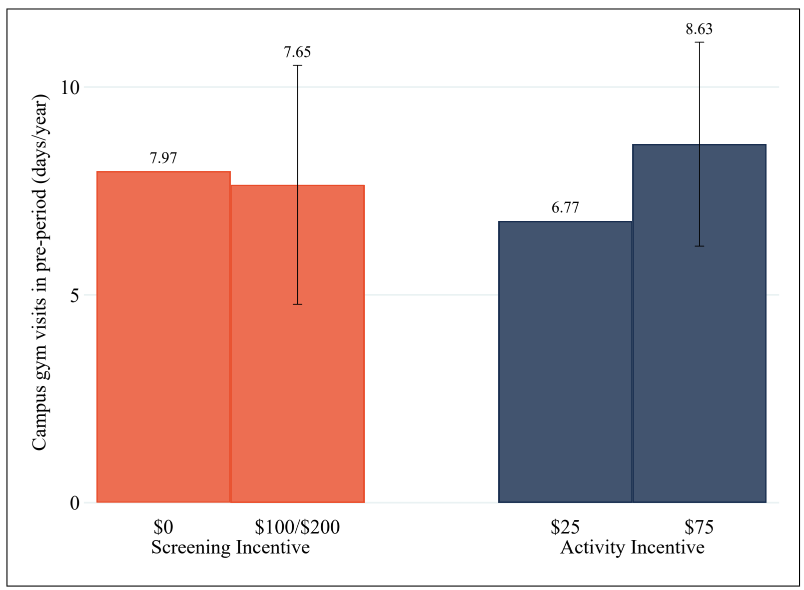
Figure 6: Marginal Selection on Medical Spending and Health Behaviors



(a) Total spending (dollars/month)



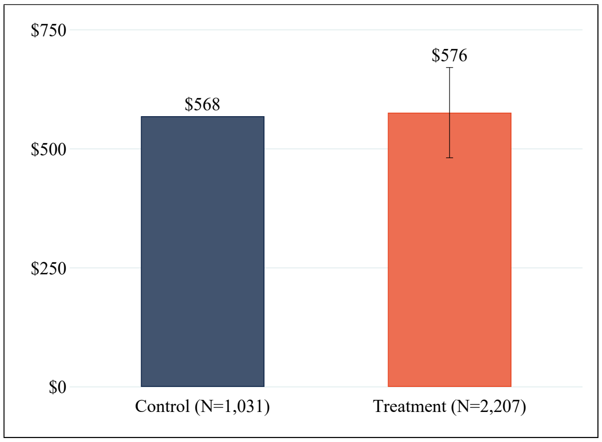
(b) IL Marathon/10K/5K participant (2014-2016)



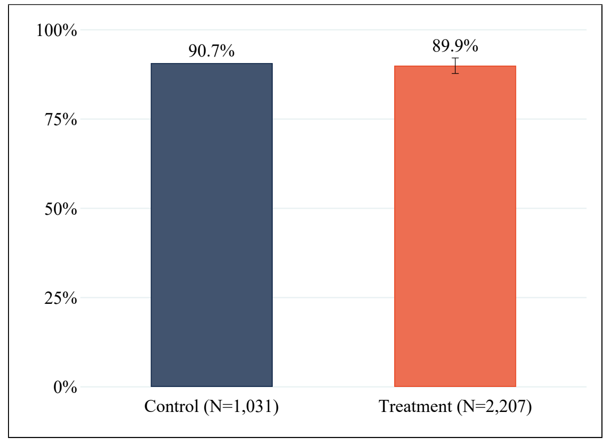
(c) Gym visits (days/year)

Notes: Each panel presents average characteristics of members of different treatment arms, conditional on having completed the screenings/HRA. The $100 and $200 treatment groups are combined. Vertical bars represent 95% confidence intervals on the difference in means between each pair of treatment groups.

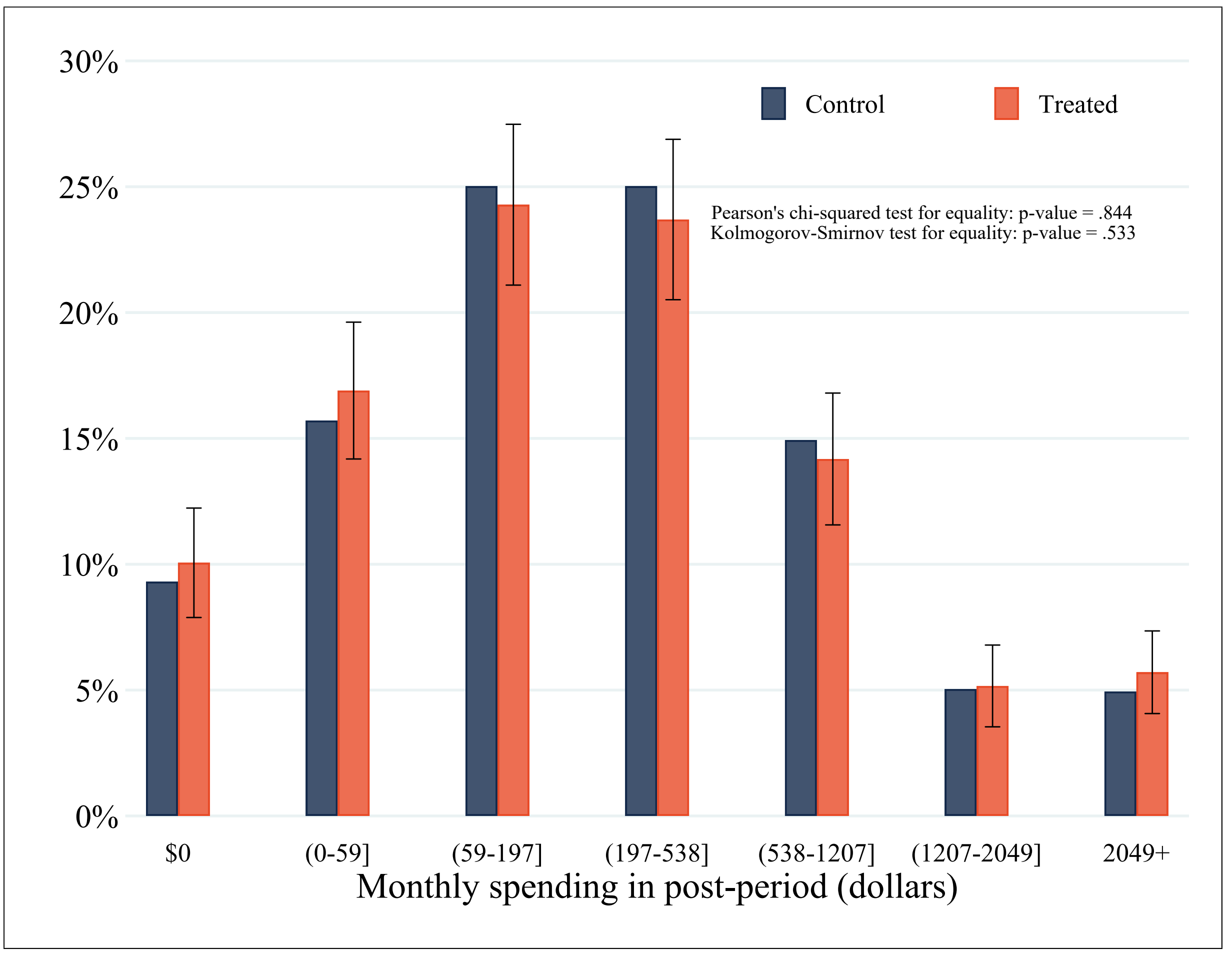
Figure 7: Post-intervention medical spending by treatment status



(a) Average monthly spending (dollars)



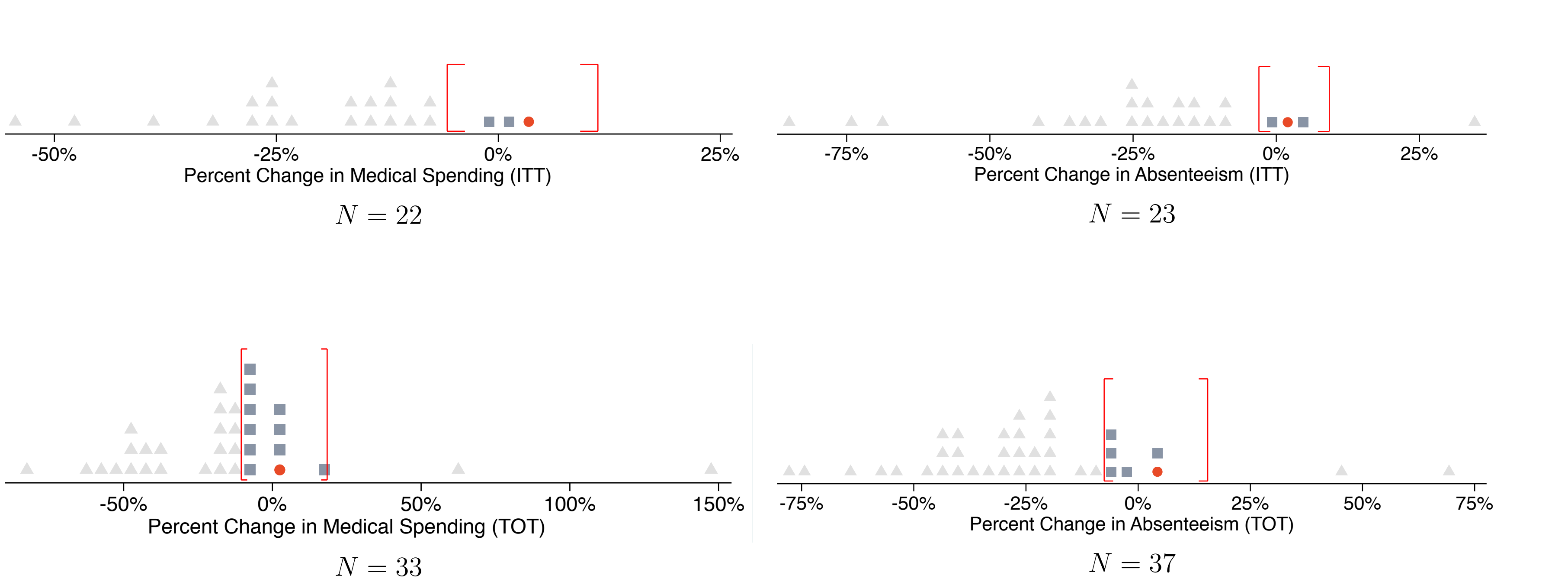
(b) Probability of any spending



(c) Histogram of average monthly spending, by quantile of control group spending ($N = 3,238$)

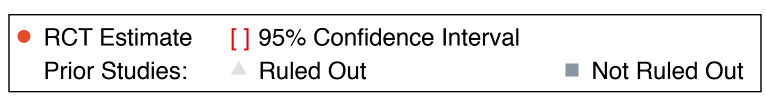
Notes: Results based on health care claims over the 12-month period August 2016 - July 2017. The null hypothesis of the Pearson's chi-squared and the non-parametric Kolmogorov-Smirnov tests is that the two samples are drawn from the same distribution.

Figure 8: Comparison of experimental estimates to prior studies



(a) Medical spending estimates

(b) Absenteeism estimates



Notes: Each figure shows the distribution of N point estimates from prior workplace wellness studies. Panel (a) plots intent-to-treat (ITT) and treatment-on-the-treated (TOT) estimates for medical spending. Panel (b) plots corresponding estimates for absenteeism. The point estimates from our own study ("RCT Estimate"), and their associated confidence intervals, are taken from Table 6, Column 3, for medical spending, and Table 4, Column 4 and Table 5, Column 3 for absenteeism. Our RCT estimates and confidence intervals are plotted in order to demonstrate the share of prior study point estimates we are able to rule out. Appendix Table B.1 provides the full details of this meta-analysis.

Table 1a: Means of Study Variables at Baseline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Enrolled in Study								
	Not in Study	Control	A25	A75	B25	B75	C25	C75	p-value	Sample size
A. Stratification Variables										
Male	0.536	0.426	0.423	0.434	0.429	0.427	0.421	0.432	1.000	12,459
Age 50+	0.430	0.323	0.332	0.322	0.326	0.325	0.328	0.326	1.000	12,459
Age 37-49	0.362	0.340	0.330	0.333	0.330	0.336	0.330	0.335	0.999	12,459
White	0.774	0.841	0.828	0.847	0.835	0.832	0.842	0.831	0.971	12,459
Salary Q1 (bottom quartile)	0.234	0.244	0.243	0.239	0.246	0.237	0.241	0.244	1.000	12,459
Salary Q2	0.189	0.255	0.254	0.259	0.255	0.261	0.258	0.266	0.999	12,459
Salary Q3	0.197	0.249	0.252	0.260	0.250	0.248	0.250	0.240	0.996	12,459
Faculty	0.298	0.196	0.198	0.202	0.199	0.203	0.198	0.204	1.000	12,459
Academic Staff	0.324	0.443	0.439	0.439	0.438	0.434	0.436	0.435	1.000	12,459
B. 2016 Survey Variables										
Ever screened		0.885	0.895	0.900	0.891	0.876	0.887	0.902	0.817	4,834
Physically active		0.359	0.350	0.397	0.399	0.392	0.370	0.381	0.387	4,834
Trying to be active		0.822	0.799	0.791	0.799	0.843	0.797	0.827	0.161	4,834
Current smoker (cigarettes)		0.072	0.051	0.060	0.062	0.075	0.071	0.075	0.513	4,833
Current smoker (other)		0.085	0.075	0.062	0.089	0.089	0.096	0.100	0.224	4,833
Former smoker		0.198	0.216	0.186	0.185	0.204	0.211	0.171	0.481	4,833
Drinker		0.657	0.641	0.658	0.636	0.625	0.656	0.656	0.836	4,830
Heavy drinker		0.050	0.051	0.035	0.054	0.044	0.056	0.055	0.553	4,829
Chronic condition		0.729	0.751	0.729	0.712	0.741	0.701	0.721	0.562	4,834
Excellent or v. good health		0.586	0.613	0.619	0.612	0.604	0.563	0.603	0.433	4,834
Not poor health		0.989	0.982	0.991	0.993	0.987	0.995	0.989	0.509	4,834
Physical problems		0.392	0.387	0.395	0.380	0.392	0.401	0.375	0.979	4,834
Lots of energy		0.310	0.339	0.324	0.346	0.327	0.323	0.321	0.790	4,834
Bad emotional health		0.308	0.247	0.326	0.292	0.288	0.279	0.299	0.078	4,834
Overweight		0.545	0.577	0.530	0.507	0.518	0.552	0.514	0.202	4,834
High BP/cholesterol/glucose		0.308	0.328	0.281	0.292	0.266	0.290	0.313	0.273	4,834
Sedentary		0.545	0.569	0.499	0.538	0.571	0.530	0.545	0.239	4,833
Pharmaceutical drug utilization		0.723	0.736	0.710	0.710	0.670	0.708	0.701	0.286	4,830
Physician/ER utilization		0.772	0.797	0.734	0.774	0.712	0.715	0.760	0.003	4,833
Hospital utilization		0.038	0.036	0.020	0.024	0.022	0.034	0.026	0.168	4,833
Any sick days in past year		0.618	0.628	0.622	0.580	0.607	0.583	0.581	0.325	4,828
Worked 50+ hours/week		0.187	0.162	0.168	0.192	0.175	0.176	0.164	0.711	4,831
Very satisfied with job		0.396	0.385	0.426	0.408	0.389	0.435	0.408	0.534	4,832
Very or somewhat satisfied with job		0.836	0.858	0.829	0.841	0.847	0.842	0.852	0.818	4,832
Management priority on health/safety		0.771	0.797	0.780	0.746	0.781	0.791	0.796	0.399	4,831
Sample size	7,625	1,534	551	549	552	548	551	549		
Joint balance test for panel A (p-value)									1.000	4,834
Joint balance test for panel B (p-value)									0.165	4,817

Notes: Columns (1)-(8) report unweighted means for different, nonoverlapping subsets of university employees. Column (9) reports the p-value from a joint test of equality of the seven coefficients reported in Columns (2)-(8). We also estimate a seemingly unrelated regression model to test whether the variables listed in a particular panel predict enrollment into any of the seven control or treatment groups. The joint balance test row reports the p-value from jointly testing whether all regression coefficients across all seven study groups are equal to 0.

Table 1b: Means of Study Variables at Baseline, Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Enrolled in Study								
	Not in Study	Control	A25	A75	B25	B75	C25	C75	*p*-value	Sample size
C. Health Claims Variables (2015-2016)										
Total spending (dollars/month)	579	505	452	393	486	458	502	494	0.570	8,095
Office spending	54	67	61	53	54	49	80	50	0.327	8,095
Hospital spending	345	283	242	231	281	239	263	299	0.711	8,095
Drug spending	105	103	97	75	113	124	95	103	0.843	8,095
Non-zero medical spending	0.888	0.899	0.911	0.886	0.901	0.862	0.869	0.886	0.311	8,095
D. Health Behavior and Productivity Variables										
Sick leave (days/year)	5.89	6.04	6.53	5.82	5.69	6.36	6.24	6.13	0.393	12,459
Annual salary (dollars)	73,927	61,528	62,774	60,579	60,906	62,719	61,042	62,407	0.875	12,221
IL Marathon/10K/5K (2014-2016)	0.072	0.107	0.120	0.120	0.118	0.111	0.102	0.137	0.597	12,459
Campus gym visits (days/year)	6.14	7.36	5.44	8.68	7.68	5.69	5.34	7.86	0.119	12,459
Sample size	7,625	1,534	551	549	552	548	551	549		
Joint balance test for panel C (*p*-value)									0.220	3,222
Joint balance test for panel D (*p*-value)									0.437	4,770

Notes: Columns (1)-(8) report unweighted means for different, nonoverlapping subsets of university employees. Column (9) reports the *p*-value from a joint test of equality of the seven coefficients reported in Columns (2)-(8). We also estimate a seemingly unrelated regression model to test whether the variables listed in a particular panel predict enrollment into any of the seven control or treatment groups. The joint balance test row reports the *p*-value from jointly testing whether all regression coefficients across all seven study groups are equal to 0.

Table 2: Wellness Program Participation by Treatment Group

	(1)	(2)	(3)	(4)
A. Screening and HRA Completion				
Group B* or C* (B25, B75, C25, C75)	0.136***	0.137***		
	(0.018)	(0.018)		
Group B* (B25, B75)			0.116***	0.117***
			(0.021)	(0.021)
Group C* (C25, C75)			0.156***	0.157***
			(0.021)	(0.021)
Group *75 (A75, B75, C75)	0.049***	0.050***	0.049***	0.049***
	(0.017)	(0.017)	(0.017)	(0.017)
Constant	0.445***	0.444***	0.445***	0.444***
	(0.017)	(0.017)	(0.017)	(0.017)
N	3,300	3,300	3,300	3,300
Strata FE	No	Yes	No	Yes
F Test	0.000	0.000	0.000	0.000
B. Fall 2016 Activity Completion				
Group B* or C* (B25, B75, C25, C75)	0.044***	0.044***		
	(0.016)	(0.016)		
Group B* (B25, B75)			0.039**	0.039**
			(0.019)	(0.018)
Group C* (C25, C75)			0.048**	0.049***
			(0.019)	(0.019)
Group *75 (A75, B75, C75)	0.124***	0.126***	0.124***	0.126***
	(0.015)	(0.015)	(0.015)	(0.015)
Constant	0.182***	0.182***	0.182***	0.182***
	(0.014)	(0.014)	(0.014)	(0.014)
N	3,300	3,300	3,300	3,300
Strata FE	No	Yes	No	Yes
F Test	0.000	0.000	0.000	0.000
C. Spring 2017 Activity Completion				
Group B* or C* (B25, B75, C25, C75)	0.045***	0.045***		
	(0.015)	(0.015)		
Group B* (B25, B75)			0.049***	0.049***
			(0.017)	(0.017)
Group C* (C25, C75)			0.040**	0.040**
			(0.017)	(0.017)
Group *75 (A75, B75, C75)	0.116***	0.117***	0.116***	0.117***
	(0.014)	(0.014)	(0.014)	(0.014)
Constant	0.137***	0.136***	0.137***	0.136***
	(0.013)	(0.013)	(0.013)	(0.013)
N	3,300	3,300	3,300	3,300
Strata FE	No	Yes	No	Yes
F Test	0.000	0.000	0.000	0.000

Notes: This table reports rates of completion for the three components of the wellness program tied to completion incentives. Each column in each panel reports estimates from a separate regression estimated over individuals in one of the six treatment groups (A25, A75, B25, B75, C25, and C75). The outcome in each regression is an indicator for completing the program component indicated by the panel, and the independent variables are indicators for inclusion in the specified treatment groups. The regressions reported in Columns (2) and (4) are the same as those reported in Columns (1) and (3), respectively, but with the addition of strata fixed effects. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference.

Table 3: Selection on Medical Spending, Income, and Health Behaviors

	(1)	(2)	(3)	(4)	(5)
Selection Variable	Mean	N	Completed Screening and HRA	Completed Fall Activity	Completed Spring Activity
A. Baseline Medical Spending					
Total spending (dollars/month) [admin]	479	2187	-116.1**	-60.9	-62.8
			(52.3)	(43.6)	(44.3)
			[0.080]	[0.401]	[0.271]
Non-zero medical spending [admin]	0.886	2187	0.049***	0.049***	0.045***
			(0.014)	(0.014)	(0.014)
			[0.009]	[0.006]	[0.021]
B. Baseline Income					
Annual salary (dollars) [admin]	61,736	3257	-782.7	-3363.9***	-3429.1***
			(1248.3)	(1191.6)	(1251.8)
			[0.519]	[0.009]	[0.012]
Salary Q1 (bottom quartile) [admin]	0.242	3300	-0.069***	-0.022	-0.036**
			(0.015)	(0.016)	(0.017)
			[0.000]	[0.398]	[0.121]
C. Baseline Health Behaviors					
IL Marathon/10K/5K (2014-2016) [admin]	0.118	3300	0.089***	0.111***	0.090***
			(0.011)	(0.014)	(0.016)
			[0.000]	[0.000]	[0.000]
Campus gym visits (days/year) [admin]	6.780	3300	2.178**	1.006	1.629
			(0.885)	(1.024)	(1.132)
			[0.013]	[0.328]	[0.153]

Notes: Column (1) reports the mean among subjects assigned to treatment. Columns (3)-(5) report the difference in means between those who completed the participation outcome and those who did not. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference, i.e., not adjusting for multiple outcomes. Family-wise p-values, reported in brackets, adjust for the number of outcome (selection) variables in each family and are estimated using 10,000 bootstraps.

Table 4: Treatment Effects (ITT)

Outcome Variable	(1) Mean	(2) No Controls	(3) Strata FEs	(4) Post-Lasso
A. Medical Spending				
Total spending (dollars/month) [admin]	573.6	7.6	17.8	30.9
		(48.4)	(48.5)	(36.7)
		[0.950]	[0.941]	[0.903]
	N=3,238	N=3,238	N=3,238	N=3,152
Drug spending [admin]	132.0	-8.4	-5.3	-6.1
		(26.5)	(25.7)	(12.0)
		[0.950]	[0.941]	[0.947]
	N=3,238	N=3,238	N=3,238	N=3,152
Office spending [admin]	69.5	-6.1	-5.7	-2.0
		(10.0)	(9.8)	(4.4)
		[0.950]	[0.941]	[0.947]
	N=3,238	N=3,238	N=3,238	N=3,152
Hospital spending [admin]	310.7	19.4	26.2	22.1
		(30.7)	(32.0)	(27.7)
		[0.950]	[0.899]	[0.903]
	N=3,238	N=3,238	N=3,238	N=3,152
Non-zero medical spending [admin]	0.902	-0.007	-0.007	0.002
		(0.011)	(0.011)	(0.010)
		[0.950]	[0.941]	[0.947]
	N=3,238	N=3,238	N=3,238	N=3,152
B. Employment and Productivity				
Annual salary (share of baseline salary) [admin]	0.059	-0.000	-0.002	-0.001
		(0.005)	(0.005)	(0.004)
		[0.969]	[0.687]	[0.771]
	N=4,146	N=4,146	N=4,146	N=4,130
Job terminated [admin]	0.112	-0.012	-0.013	-0.012
		(0.010)	(0.010)	(0.009)
		[0.538]	[0.395]	[0.467]
	N=4,834	N=4,834	N=4,834	N=4,753
Sick leave (days/year) [admin]	6.336	0.229	0.292	0.195
		(0.226)	(0.204)	(0.196)
		[0.538]	[0.395]	[0.546]
	N=4,782	N=4,782	N=4,782	N=4,711
Management priority on health/safety [survey]	0.790	0.057***	0.057***	0.050***
		(0.015)	(0.015)	(0.014)
		[0.001]	[0.001]	[0.003]
	N=3,566	N=3,566	N=3,566	N=3,514
C. Health Status and Behaviors				
IL Marathon/10K/5K 2017 [admin]	0.066	0.002	0.002	-0.005
		(0.008)	(0.008)	(0.006)
		[0.975]	[0.962]	[0.471]
	N=4,834	N=4,834	N=4,834	N=4,817
Campus gym visits (days/year) [admin]	5.839	-0.062	-0.068	0.401
		(0.733)	(0.721)	(0.360)
		[0.975]	[0.962]	[0.471]
	N=4,834	N=4,834	N=4,834	N=4,817
Ever screened [survey]	0.942	0.039***	0.042***	0.036***
		(0.009)	(0.009)	(0.008)
		[0.001]	[0.000]	[0.000]
	N=3,567	N=3,567	N=3,567	N=3,557

Notes: Each row and column reports estimates from a separate regression, where observations include individuals in the control or treatment groups. The outcome in each regression is specified by the table row. The focal independent variable is an indicator for inclusion in the treatment group, and the control strategy is specified by the column. Post-Lasso controls include covariates selected by Lasso to predict the dependent variable. The set of potential predictors include baseline values of all available variables in the same family of outcomes, strata variables, and the baseline (2016) survey variables reported in Table 1a, as well as all two-way interactions between these predictors. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference, i.e., not adjusting for multiple outcomes. Family-wise p-values, reported in brackets, adjust for the number of outcome variables in each family. See Appendix Tables A.2a-A.2f for results for all outcomes, categorized by family.

Table 5: Treatment Effects: Experimental vs. Observational Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Experimental (IV)			Observational (OLS)		
Outcome Variable	No Controls	Strata FEs	Post-Lasso	No Controls	Strata FEs	Post-Lasso
A. Medical Spending						
Total spending (dollars/month) [admin]	12.4 (78.8) N=3,238	29.1 (78.4) N=3,238	45.0 (59.1) N=3,152	-132.7* (68.0) N=2,207	-157.8** (65.5) N=2,207	-98.4 (61.1) N=2,140
Drug spending [admin]	-13.7 (43.2) N=3,238	-8.6 (41.6) N=3,238	-12.8 (20.4) N=3,152	-26.5 (27.3) N=2,207	-34.9 (26.9) N=2,207	-7.3 (12.0) N=2,140
Office spending [admin]	-9.9 (16.2) N=3,238	-9.3 (15.9) N=3,238	-3.2 (6.8) N=3,152	12.1 (7.5) N=2,207	9.4 (7.2) N=2,207	8.8* (5.1) N=2,140
Hospital spending [admin]	31.6 (50.0) N=3,238	42.8 (51.7) N=3,238	40.6 (45.0) N=3,152	-113.9** (55.1) N=2,207	-123.0** (52.1) N=2,207	-101.1* (54.2) N=2,140
Non-zero medical spending [admin]	-0.012 (0.018) N=3,238	-0.011 (0.018) N=3,238	0.004 (0.016) N=3,152	0.060*** (0.014) N=2,207	0.042*** (0.013) N=2,207	0.036*** (0.012) N=2,140
B. Employment and Productivity						
Annual salary (share of baseline salary) [admin]	-0.000 (0.008) N=4,146	-0.003 (0.008) N=4,146	-0.003 (0.008) N=4,130	0.004 (0.005) N=2,840	0.005 (0.005) N=2,840	0.006 (0.005) N=2,828
Job terminated [admin]	-0.022 (0.018) N=4,834	-0.023 (0.017) N=4,834	-0.023 (0.017) N=4,753	-0.082*** (0.011) N=3,300	-0.080*** (0.011) N=3,300	-0.068*** (0.011) N=3,244
Sick leave (days/year) [admin]	0.397 (0.391) N=4,782	0.506 (0.351) N=4,782	0.311 (0.336) N=4,711	0.266 (0.273) N=3,265	0.030 (0.254) N=3,265	-0.072 (0.249) N=3,216
Management priority on health/safety [survey]	0.087*** (0.023) N=3,566	0.087*** (0.023) N=3,566	0.077*** (0.021) N=3,514	-0.004 (0.017) N=2,410	-0.012 (0.017) N=2,410	-0.007 (0.016) N=2,376
C. Health Status and Behaviors						
IL Marathon/10K/5K 2017 [admin]	0.003 (0.014) N=4,834	0.003 (0.013) N=4,834	-0.011 (0.011) N=4,817	0.059*** (0.008) N=3,300	0.054*** (0.008) N=3,300	0.024*** (0.006) N=3,287
Campus gym visits (days/year) [admin]	-0.110 (1.309) N=4,834	-0.121 (1.276) N=4,834	0.757 (0.656) N=4,817	3.527*** (0.813) N=3,300	3.849*** (0.804) N=3,300	2.160*** (0.425) N=3,287
Ever screened [survey]	0.060*** (0.014) N=3,567	0.065*** (0.013) N=3,567	0.056*** (0.012) N=3,557	0.073*** (0.011) N=2,410	0.074*** (0.010) N=2,410	0.061*** (0.009) N=2,404

Notes: Each row and column reports estimates from a separate regression. The outcome in each regression is specified by the table row, and the (endogenous) focal independent variable is an indicator for completing the screening and HRA. For the IV specifications (columns (1)-(3)), the instrument is an indicator for inclusion in the treatment group, and observations include individuals in the control or treatment groups. For the OLS specifications (columns (4)-(6)), there is no instrument and observations are restricted to individuals in the treatment group. The control strategy is specified by the column. Post-Lasso controls include covariates selected by Lasso to predict either the dependent variable or the focal independent variable. The set of potential predictors include baseline values of all available variables in the same family of outcomes, strata variables, and the baseline (2016) survey variables reported in Table 1a, as well as all two-way interactions between these predictors. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference.

Table 6: Winsorized Medical Spending Treatment Effects

	(1)	(2)	(3)	(4)	(5)
A. ITT Estimates (Post-Lasso)					
Total spending (dollars/month) [admin]	30.9	13.7	15.4	14.6	10.4
	(36.7)	(23.1)	(19.4)	(13.5)	(9.7)
	[-41.0, 102.8]	[-31.7, 59.0]	[-22.7, 53.4]	[-11.8, 41.0]	[-8.7, 29.5]
N	3,152	3,152	3,152	3,152	3,152
Winsorization (percent)	0	0.5	1	2.5	5
B. IV Estimates (Post-Lasso)					
Total spending (dollars/month) [admin]	45.0	16.2	18.5	20.0	15.3
	(59.1)	(37.9)	(31.6)	(21.8)	(15.7)
	[-70.8, 160.8]	[-58.1, 90.5]	[-43.5, 80.5]	[-22.8, 62.8]	[-15.5, 46.0]
N	3,152	3,152	3,152	3,152	3,152
Winsorization (percent)	0	0.5	1	2.5	5

Notes: Each row and column reports estimates from a separate regression, where observations include individuals in the control or treatment groups. The outcome in each regression is winsorized (top-coded) average monthly medical spending over the first 12 months of the intervention, winsorized at the level indicated in each column. Regressions are weighted by the number of months of coverage. In Panel A (ITT), the focal independent variable is an indicator for inclusion in the treatment group, and all regressions include the same controls as the ITT post-Lasso specification reported in row 1 and column (4) of Table 4. In Panel B (IV), the (endogenous) focal independent variable is an indicator for completing the screening and HRA, the instrument is an indicator for inclusion in the treatment group, and all regressions include the same controls as the IV post-Lasso specification reported in row 1 and column (3) of Table 5. Column (1) replicates the (non-winsorized) ITT and IV post-Lasso results reported in Table 4 and Table 5. Robust standard errors are reported in parentheses, and 95% confidence intervals are reported in brackets. A */**/*** indicates significance at the 10/5/1% level using conventional inference.

Appendix A: For Online Publication Only

Figure A.1 reports how participation in the spring wellness activities varies as a function of rewards. It also reports the marginal cost of the additional participation induced by each reward. An increase in the size of screening incentives has modest, positive effects on participation rates for spring wellness activities; an increase in the size of participation incentives has a large, positive effect.

Panels (a)-(c) of Figure A.2 shows how selection on prior probability of nonzero medical spending, prior annual salary, and prior annual salary in the first quartile varies as a function of the monetary incentives assigned to study participants. Panel (a) shows that, at larger incentive levels, participants are slightly less likely to have non-zero medical spending in the prior year. Panels (b) and (c) show little effect of the size of incentives on selection with respect to annual salary

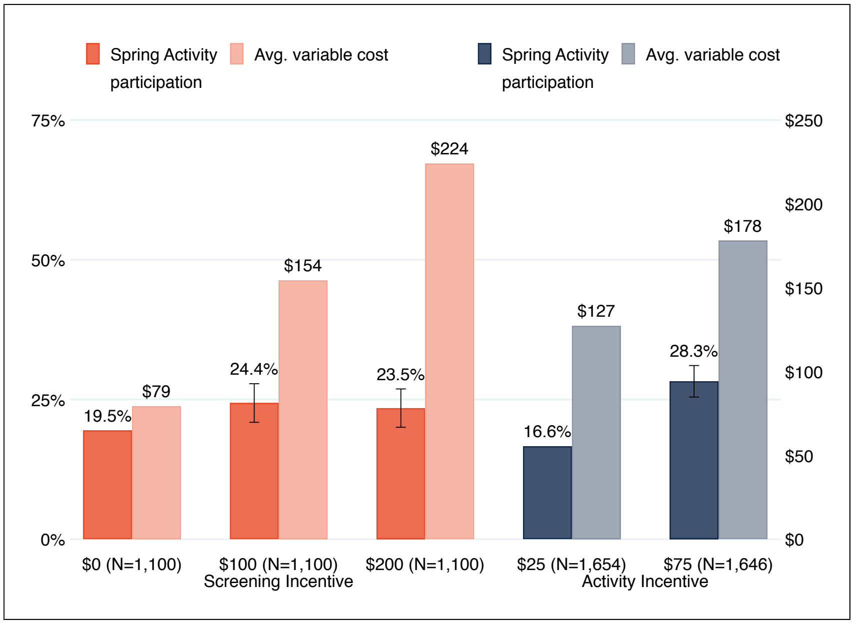
Tables A.1a - A.1d provide selection results for the full set of pre-specified variables shown in Tables 1a and 1b using equation (3). Tables A.2a - A.2f provide the causal, intent-to-treat (ITT) effect of our intervention on all pre-specified variables. In addition, Table A.2g provides results for different measures of medical utilization. Tables A.3a - A.3g provide the corresponding IV and OLS estimates of equation (5) for all pre-specified variables.

Tables A.4a and A.4b report intent-to-treat estimates for medical spending from a model that allows the treatment effect to vary by treatment group. We do not find statistically significant treatment effects for any treatment group in any of these specifications.

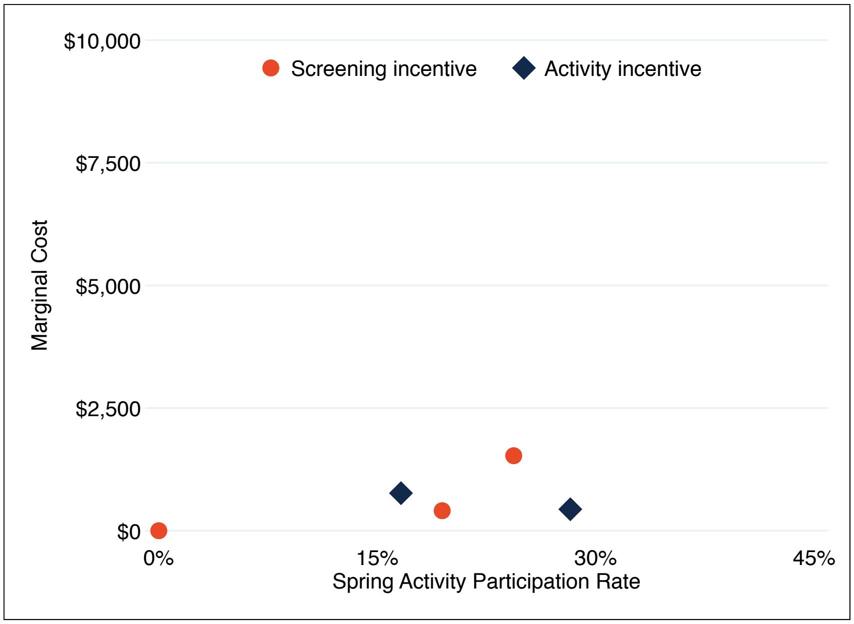
As discussed in the main text, we find two statistically significant effects of our intervention: an increase in the number of employees who ever received a health screening, and an increase in employees who believe that management places a priority on health and safety. Because our monetary incentives were varied independently across the health screening and wellness activity components of our study, these incentives can be used as instruments for participation in those components. Table A.5 reports estimates of those IV regressions. For both outcomes, the effects are driven by the health screening component of our intervention.

Finally, Table A.7 provides the definition, data source, and time period for every variable presented in the paper.

Figure A.1: Marginal cost of inducing additional participation into spring wellness activities



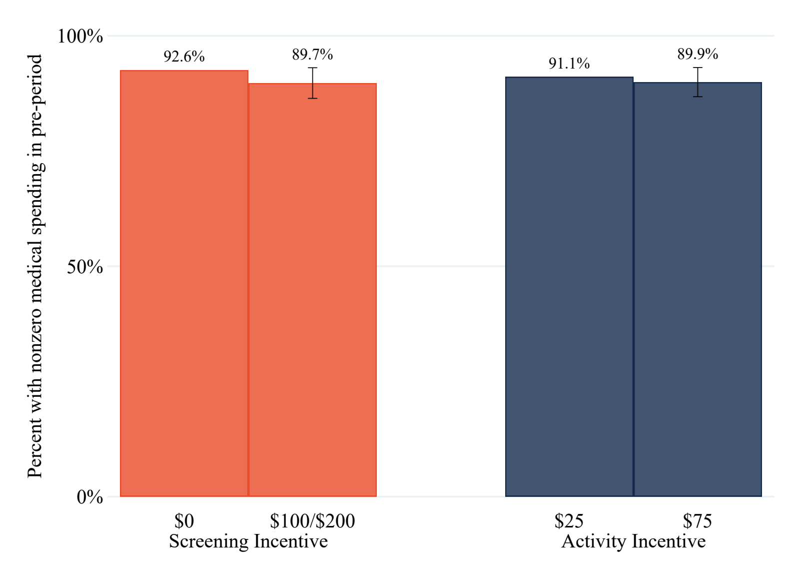
(a) Spring activities participation and costs



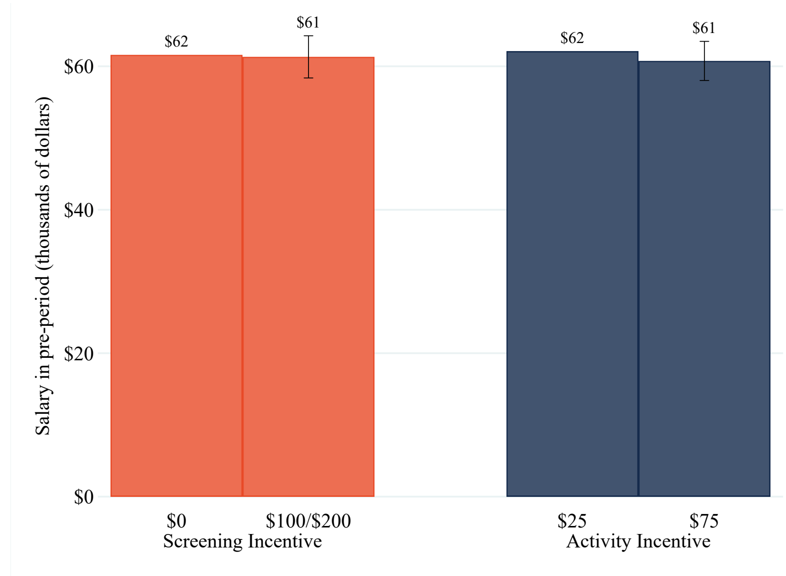
(b) Marginal cost of additional spring participation

Notes: Panel (a) plots health screening participation rates (PR) and average variable costs (AVC) as a function of screening and activity incentives. Vertical bars display 95% confidence intervals on the difference in means relative to the lowest reward group. AVC includes costs of the health screening, HRA, and wellness activities. Panel (b) plots the implied marginal costs (MC), calculated as MC$=\frac{\Delta \text{AVC}}{\Delta \text{PR}}$. The MC of the control group (PR=0 percent) is set equal to 0. We omit the MC for group C because its marginal PR is negative.

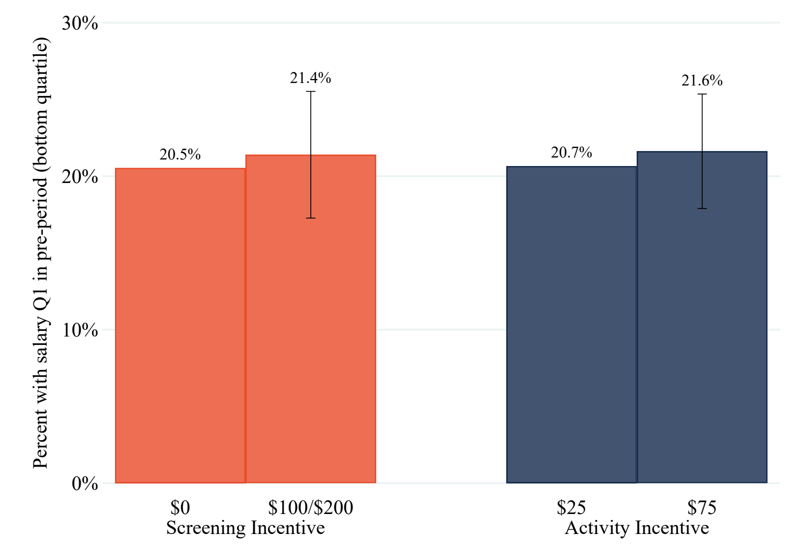
Figure A.2: Marginal Selection on Non-zero Spending and Income



(a) Non-zero Medical Spending



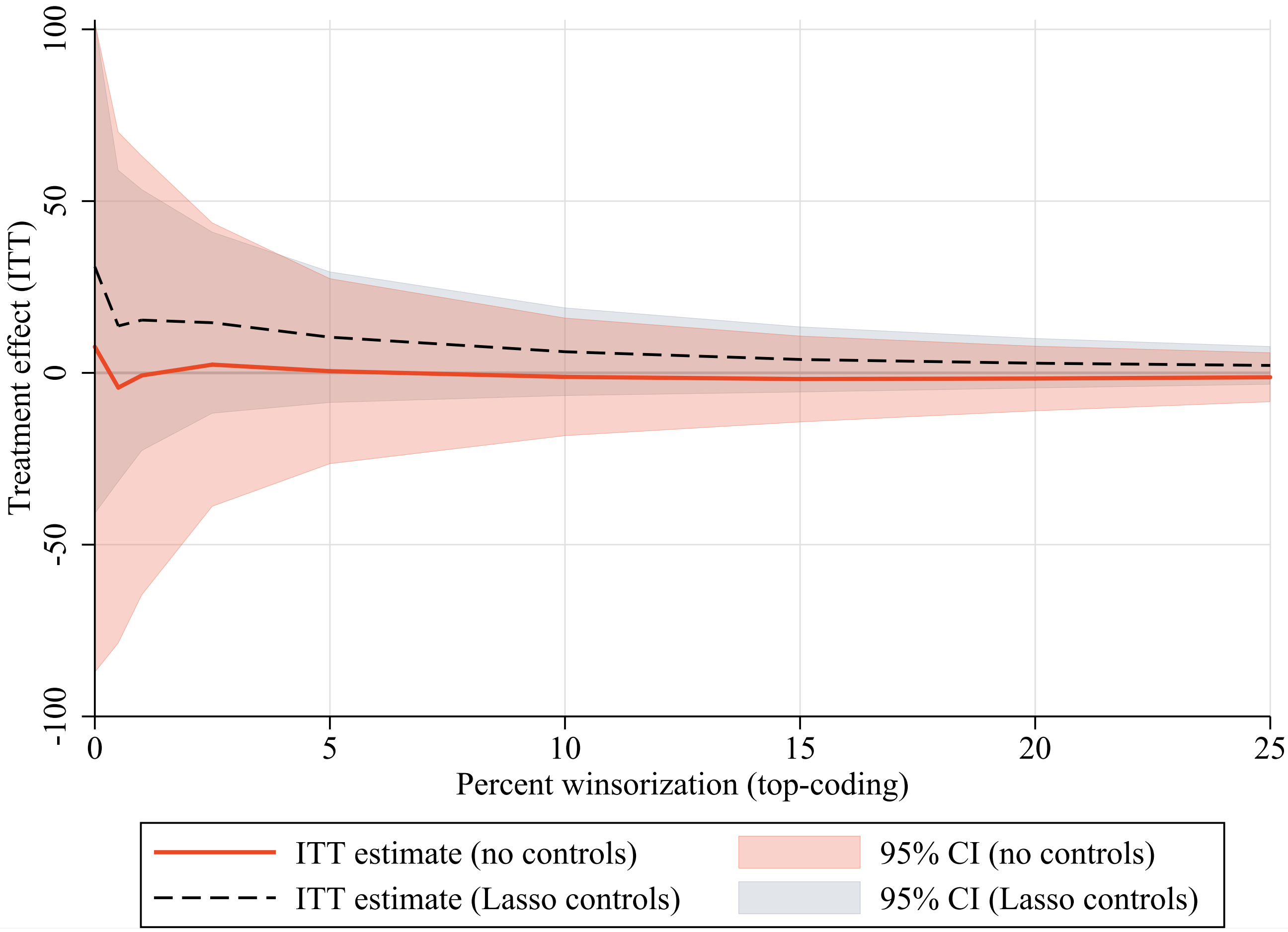
(b) Salary annual (thousands of dollars)



(c) Salary Q1 (bottom quartile)

Notes: Each panel presents average characteristics of members of different treatment arms, conditional on having completed the screenings/HRA. The $100 and $200 treatment groups are combined. Vertical bars represent 95% confidence intervals on the difference in means between each pair of treatment groups.

Figure A.3: Pre-intervention medical spending among treatment group, by participation status



Notes: The figure reports how intent-to-treat (ITT) medical spending effect estimates vary by the degree of winsorization (top-coding) of medical spending, calculated as the average monthly health care spending over the first 12 months of the wellness program (August 2016 - July 2017). Each ITT estimate is estimated from a separate regression of medical spending (winsorized at the level indicated by the horizontal axis) on an indicator for inclusion in the treatment group. Observations include individuals in the control or treatment groups, and regressions are weighted by the number of months of medical coverage. The solid orange line reports estimates from a specification that includes no controls. The dashed black line reports estimates from a specification that includes the same controls as the ITT post-Lasso specification reported in row 1 and column (4) of Table 4. Shaded regions indicate 95% confidence intervals based on robust standard errors. The values of the ITT point estimates and confidence intervals for selected levels of winsorization are reported in Panel A (no controls) and Panel B (post-Lasso controls) of Table A.6.

Table A.1a: Selection on Strata Variables

	(1)	(2)	(3)	(4)	(5)
Selection Variable	Mean	N	Completed Screening and HRA	Completed Fall Activity	Completed Spring Activity
Male [admin]	0.428	3300	-0.058***	-0.114***	-0.149***
			(0.017)	(0.019)	(0.020)
			[0.005]	[0.000]	[0.000]
Age 50+ [admin]	0.327	3300	-0.027	-0.015	-0.020
			(0.016)	(0.018)	(0.019)
			[0.270]	[0.399]	[0.473]
Age 37-49 [admin]	0.332	3300	0.008	0.026	0.017
			(0.017)	(0.019)	(0.020)
			[0.850]	[0.398]	[0.473]
White [admin]	0.836	3300	-0.001	0.046***	0.036**
			(0.013)	(0.014)	(0.015)
			[0.962]	[0.005]	[0.072]
Salary Q1 (bottom quartile) [admin]	0.242	3300	-0.069***	-0.022	-0.036**
			(0.015)	(0.016)	(0.017)
			[0.000]	[0.398]	[0.121]
Salary Q2 [admin]	0.259	3300	0.038**	0.028	0.058***
			(0.015)	(0.017)	(0.019)
			[0.052]	[0.346]	[0.012]
Salary Q3 [admin]	0.250	3300	0.044***	0.043**	0.040**
			(0.015)	(0.017)	(0.019)
			[0.019]	[0.067]	[0.121]
Faculty [admin]	0.201	3300	-0.051***	-0.098***	-0.097***
			(0.014)	(0.014)	(0.015)
			[0.002]	[0.000]	[0.000]
Academic Staff [admin]	0.437	3300	0.077***	0.077***	0.086***
			(0.017)	(0.019)	(0.021)
			[0.000]	[0.001]	[0.000]

Notes: Column (1) reports the mean among subjects assigned to treatment. Columns (3)-(5) report the difference in means between those who completed the participation outcome and those who did not. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference, i.e., not adjusting for multiple outcomes. Family-wise p-values, reported in brackets, adjust for the number of outcome (selection) variables in each family and are estimated using 10,000 bootstraps.

Table A.1b: Selection on Health Care Utilization Variables

	(1)	(2)	(3)	(4)	(5)
Selection Variable	Mean	N	Completed Screening and HRA	Completed Fall Activity	Completed Spring Activity
Total spending (dollars/month) [admin]	479	2187	-116.1**	-60.9	-62.8
			(52.3)	(43.6)	(44.3)
			[0.080]	[0.401]	[0.271]
Office spending [admin]	59	2187	2.3	-5.7	-12.5**
			(7.2)	(6.5)	(6.2)
			[0.750]	[0.637]	[0.144]
Hospital spending [admin]	268	2187	-104.1***	-47.4*	-62.9**
			(40.3)	(28.3)	(27.5)
			[0.045]	[0.295]	[0.102]
Drug spending [admin]	104	2187	-14.8	-4.3	14.5
			(20.6)	(25.5)	(28.9)
			[0.728]	[0.869]	[0.637]
Non-zero medical spending [admin]	0.886	2187	0.049***	0.049***	0.045***
			(0.014)	(0.014)	(0.014)
			[0.009]	[0.006]	[0.021]
Pharmaceutical drug utilization [survey]	0.706	3297	-0.001	0.029*	0.040**
			(0.016)	(0.018)	(0.019)
			[0.929]	[0.183]	[0.059]
Physician/ER utilization [survey]	0.748	3300	0.050***	0.070***	0.061***
			(0.015)	(0.016)	(0.017)
			[0.003]	[0.000]	[0.002]
Hospital utilization [survey]	0.027	3299	-0.012**	-0.005	-0.012**
			(0.006)	(0.006)	(0.006)
			[0.072]	[0.400]	[0.059]

Notes: Column (1) reports the mean among subjects assigned to treatment. Columns (3)-(5) report the difference in means between those who completed the participation outcome and those who did not. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference, i.e., not adjusting for multiple outcomes. Family-wise p-values, reported in brackets, adjust for the number of outcome (selection) variables in each family and are estimated using 10,000 bootstraps.

Table A.1c: Selection on Employment and Productivity Variables

	(1)	(2)	(3)	(4)	(5)
Selection Variable	Mean	*N*	Completed Screening and HRA	Completed Fall Activity	Completed Spring Activity
Sick leave (days/year) [admin]	6.274	3296	0.473*	0.705**	0.617**
			(0.267)	(0.290)	(0.312)
			[0.144]	[0.015]	[0.048]
Annual salary (dollars) [admin]	61,736	3257	-782.7	-3363.9***	-3429.1***
			(1248.3)	(1191.6)	(1251.8)
			[0.519]	[0.009]	[0.012]
Any sick days in past year [survey]	0.600	3296	0.043**	0.057***	0.051**
			(0.017)	(0.019)	(0.020)
			[0.049]	[0.008]	[0.046]
Worked 50+ hours/week [survey]	0.173	3297	-0.058***	-0.065***	-0.064***
			(0.013)	(0.014)	(0.014)
			[0.000]	[0.000]	[0.000]
Very satisfied with job [survey]	0.408	3299	0.002	0.002	0.002
			(0.017)	(0.019)	(0.021)
			[0.899]	[0.921]	[0.911]
Very or somewhat satisfied with job [survey]	0.845	3299	0.023*	0.043***	0.029**
			(0.013)	(0.013)	(0.014)
			[0.193]	[0.005]	[0.092]
Management priority on health/safety [survey]	0.782	3299	0.012	0.033**	0.035**
			(0.015)	(0.016)	(0.017)
			[0.618]	[0.062]	[0.092]

Notes: Column (1) reports the mean among subjects assigned to treatment. Columns (3)-(5) report the difference in means between those who completed the participation outcome and those who did not. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference, i.e., not adjusting for multiple outcomes. Family-wise p-values, reported in brackets, adjust for the number of outcome (selection) variables in each family and are estimated using 10,000 bootstraps.

Table A.1d: Selection on Health and Behavior Variables

	(1)	(2)	(3)	(4)	(5)
Selection Variable	Mean	N	Completed Screening and HRA	Completed Fall Activity	Completed Spring Activity
IL Marathon/10K/5K (2014-2016) [admin]	0.118	3300	0.089*** (0.011) [0.000]	0.111*** (0.014) [0.000]	0.090*** (0.016) [0.000]
Campus gym visits (days/year) [admin]	6.780	3300	2.178** (0.885) [0.013]	1.006 (1.024) [0.328]	1.629 (1.132) [0.153]
Ever screened [survey]	0.892	3300	0.033*** (0.011) [0.029]	0.042*** (0.011) [0.002]	0.035*** (0.012) [0.046]
Physically active [survey]	0.382	3300	-0.015 (0.017) [0.909]	0.013 (0.019) [0.964]	0.040* (0.020) [0.445]
Trying to be active [survey]	0.809	3300	0.045*** (0.014) [0.014]	0.033** (0.015) [0.293]	0.030* (0.016) [0.445]
Current smoker (cigarettes) [survey]	0.065	3299	-0.041*** (0.009) [0.000]	-0.047*** (0.008) [0.000]	-0.053*** (0.008) [0.000]
Current smoker (other) [survey]	0.085	3299	-0.034*** (0.010) [0.011]	-0.046*** (0.010) [0.000]	-0.066*** (0.009) [0.000]
Former smoker [survey]	0.196	3299	-0.009 (0.014) [0.909]	-0.004 (0.015) [0.964]	-0.019 (0.016) [0.770]
Drinker [survey]	0.645	3296	0.026 (0.017) [0.707]	0.021 (0.019) [0.889]	0.009 (0.020) [0.929]
Heavy drinker [survey]	0.049	3295	-0.010 (0.008) [0.798]	-0.005 (0.008) [0.964]	-0.006 (0.009) [0.929]
Chronic condition [survey]	0.726	3300	0.024 (0.016) [0.707]	0.038** (0.017) [0.293]	0.023 (0.018) [0.770]
Excellent or v. good health [survey]	0.602	3300	-0.022 (0.017) [0.798]	0.032* (0.019) [0.626]	0.060*** (0.020) [0.045]
Not poor health [survey]	0.989	3300	0.003 (0.004) [0.909]	0.005 (0.004) [0.703]	0.007* (0.003) [0.445]
Physical problems [survey]	0.388	3300	0.022 (0.017) [0.798]	-0.015 (0.019) [0.964]	-0.027 (0.020) [0.750]
Lots of energy [survey]	0.330	3300	-0.031* (0.017) [0.502]	0.006 (0.018) [0.964]	0.014 (0.020) [0.929]
Bad emotional health [survey]	0.288	3300	0.001 (0.016) [0.944]	-0.019 (0.018) [0.889]	-0.041** (0.018) [0.280]
Overweight [survey]	0.533	3300	0.057*** (0.017) [0.015]	0.015 (0.019) [0.964]	-0.008 (0.021) [0.929]
High BP/cholesterol/glucose [survey]	0.295	3300	-0.007 (0.016) [0.909]	-0.022 (0.018) [0.866]	-0.034* (0.019) [0.445]
Sedentary [survey]	0.542	3299	0.117*** (0.017) [0.000]	0.115*** (0.019) [0.000]	0.110*** (0.020) [0.000]

Notes: Column (1) reports the mean among subjects assigned to treatment. Columns (3)-(5) report the difference in means between those who completed the participation outcome and those who did not. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference, i.e., not adjusting for multiple outcomes. Family-wise p-values, reported in brackets, adjust for the number of outcome (selection) variables in each family and are estimated using 10,000 bootstraps.

Table A.2a: Treatment Effects (ITT)

Outcome Variable	(1) Mean	(2) No Controls	(3) Strata FEs	(4) Post-Lasso
A. Medical Spending [admin]				
Total spending (dollars/month) [admin]	573.6	7.6	17.8	30.9
		(48.4)	(48.5)	(36.7)
		[0.950]	[0.941]	[0.903]
	N=3,238	N=3,238	N=3,238	N=3,152
Drug spending [admin]	132.0	-8.4	-5.3	-6.1
		(26.5)	(25.7)	(12.0)
		[0.950]	[0.941]	[0.947]
	N=3,238	N=3,238	N=3,238	N=3,152
Office spending [admin]	69.5	-6.1	-5.7	-2.0
		(10.0)	(9.8)	(4.4)
		[0.950]	[0.941]	[0.947]
	N=3,238	N=3,238	N=3,238	N=3,152
Hospital spending [admin]	310.7	19.4	26.2	22.1
		(30.7)	(32.0)	(27.7)
		[0.950]	[0.899]	[0.903]
	N=3,238	N=3,238	N=3,238	N=3,152
Non-zero medical spending [admin]	0.902	-0.007	-0.007	0.002
		(0.011)	(0.011)	(0.010)
		[0.950]	[0.941]	[0.947]
	N=3,238	N=3,238	N=3,238	N=3,152

Notes: The outcomes in this table constitute a single family of outcomes for calculating family-wise p-values. Each row and column reports estimates from a separate regression, where observations include individuals in the control or treatment groups. The outcome in each regression is specified by the table row. The focal independent variable is an indicator for inclusion in the treatment group, and the control strategy is specified by the column. Post-Lasso controls include covariates selected by Lasso to predict the dependent variable. The set of potential predictors include baseline values of all available variables in the same family of outcomes, strata variables, and the baseline (2016) survey variables reported in Table 1a, as well as all two-way interactions between these predictors. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference, i.e., not adjusting for multiple outcomes. Family-wise p-values, reported in brackets, adjust for the number of outcome variables in the table.

Table A.2b: Treatment Effects (ITT)

Outcome Variable	(1) Mean	(2) No Controls	(3) Strata FEs	(4) Post-Lasso
A. Medical Spending [survey]				
Pharmaceutical drug utilization [survey]	0.725	-0.011	-0.009	-0.002
		(0.016)	(0.015)	(0.014)
		[0.851]	[0.864]	[0.894]
	N=3,567	N=3,567	N=3,567	N=2,433
Physician/ER utilization [survey]	0.745	0.003	0.002	0.018
		(0.016)	(0.015)	(0.017)
		[0.863]	[0.919]	[0.632]
	N=3,567	N=3,567	N=3,567	N=2,433
Hospital utilization [survey]	0.026	0.003	0.004	0.006
		(0.006)	(0.006)	(0.007)
		[0.851]	[0.864]	[0.632]
	N=3,567	N=3,567	N=3,567	N=2,433

Notes: The outcomes in this table constitute a single family of outcomes for calculating family-wise p-values. See notes to Appendix Table A.2a for additional details.

Table A.2c: Treatment Effects (ITT)

Outcome Variable	(1) Mean	(2) No Controls	(3) Strata FEs	(4) Post-Lasso
B. Employment and Productivity [admin]				
Annual salary (share of baseline salary) [admin]	0.059	-0.000	-0.002	-0.001
		(0.005)	(0.005)	(0.004)
		[0.969]	[0.687]	[0.771]
	N=4,146	N=4,146	N=4,146	N=4,130
Job terminated [admin]	0.112	-0.012	-0.013	-0.012
		(0.010)	(0.010)	(0.009)
		[0.538]	[0.395]	[0.467]
	N=4,834	N=4,834	N=4,834	N=4,753
Sick leave (days/year) [admin]	6.336	0.229	0.292	0.195
		(0.226)	(0.204)	(0.196)
		[0.538]	[0.395]	[0.546]
	N=4,782	N=4,782	N=4,782	N=4,711

Notes: The outcomes in this table constitute a single family of outcomes for calculating family-wise p-values. See notes to Appendix Table A.2a for additional details.

Table A.2d: Treatment Effects (ITT)

Outcome Variable	(1) Mean	(2) No Controls	(3) Strata FEs	(4) Post-Lasso
B. Employment and Productivity [survey]				
Any sick days in past year [survey]	0.576	0.005	0.007	0.012
		(0.018)	(0.017)	(0.016)
		[0.997]	[0.994]	[0.961]
	N=3,565	N=3,565	N=3,565	N=3,514
Worked 50+ hours/week [survey]	0.150	-0.004	-0.008	0.005
		(0.013)	(0.012)	(0.010)
		[0.997]	[0.991]	[0.961]
	N=3,566	N=3,566	N=3,566	N=3,515
Very satisfied with job [survey]	0.387	-0.025	-0.028	-0.029*
		(0.017)	(0.017)	(0.015)
		[0.749]	[0.631]	[0.376]
	N=3,564	N=3,564	N=3,564	N=3,512
Very or somewhat satisfied with job [survey]	0.835	-0.004	-0.006	-0.014
		(0.013)	(0.013)	(0.012)
		[0.997]	[0.994]	[0.876]
	N=3,564	N=3,564	N=3,564	N=3,512
Management priority on health/safety [survey]	0.790	0.057***	0.057***	0.050***
		(0.015)	(0.015)	(0.014)
		[0.001]	[0.001]	[0.003]
	N=3,566	N=3,566	N=3,566	N=3,514
Happier at work than last year [survey]	0.542	0.009	0.005	-0.003
		(0.018)	(0.018)	(0.018)
		[0.995]	[0.994]	[0.978]
	N=3,562	N=3,562	N=3,562	N=3,510
Presenteeism [survey]	23.900	-0.023	-0.050	-0.151
		(0.261)	(0.259)	(0.238)
		[0.997]	[0.994]	[0.961]
	N=3,567	N=3,567	N=3,567	N=3,515
Feel very productive at work [survey]	0.449	-0.018	-0.013	-0.021
		(0.018)	(0.018)	(0.017)
		[0.930]	[0.991]	[0.866]
	N=3,567	N=3,567	N=3,567	N=3,515
Received promotion [survey]	0.472	0.008	0.000	0.002
		(0.018)	(0.018)	(0.018)
		[0.995]	[0.994]	[0.978]
	N=3,562	N=3,562	N=3,562	N=3,511
Job search very likely [survey]	0.139	0.031**	0.026**	0.027**
		(0.012)	(0.012)	(0.011)
		[0.095]	[0.208]	[0.142]
	N=3,561	N=3,561	N=3,561	N=3,511
Job search somewhat/very likely [survey]	0.337	0.019	0.012	0.013
		(0.017)	(0.017)	(0.016)
		[0.908]	[0.991]	[0.961]
	N=3,561	N=3,561	N=3,561	N=3,511

Notes: The outcomes in this table constitute a single family of outcomes for calculating family-wise p-values. See notes to Appendix Table A.2a for additional details.

Table A.2e: Treatment Effects (ITT)

Outcome Variable	(1) Mean	(2) No Controls	(3) Strata FEs	(4) Post-Lasso
C. Health Status and Behaviors [admin]				
IL Marathon/10K/5K 2017 [admin]	0.066	0.002	0.002	-0.005
		(0.008)	(0.008)	(0.006)
		[0.975]	[0.962]	[0.471]
	N=4,834	N=4,834	N=4,834	N=4,817
Campus gym visits (days/year) [admin]	5.839	-0.062	-0.068	0.401
		(0.733)	(0.721)	(0.360)
		[0.975]	[0.962]	[0.471]
	N=4,834	N=4,834	N=4,834	N=4,817

Notes: The outcomes in this table constitute a single family of outcomes for calculating family-wise p-values. See notes to Appendix Table A.2a for additional details.

Table A.2f: Treatment Effects (ITT)

Outcome Variable	(1) Mean	(2) No Controls	(3) Strata FEs	(4) Post-Lasso
C. Health Status and Behaviors [survey]				
Ever screened [survey]	0.942	0.039***	0.042***	0.036***
		(0.009)	(0.009)	(0.008)
		[0.001]	[0.000]	[0.000]
	N=3,567	N=3,567	N=3,567	N=3,557
Physically active [survey]	0.381	0.015	0.016	-0.009
		(0.017)	(0.017)	(0.012)
		[0.991]	[0.981]	[0.977]
	N=3,567	N=3,567	N=3,567	N=3,557
Trying to be active [survey]	0.825	0.005	0.007	0.017
		(0.014)	(0.014)	(0.012)
		[1.000]	[0.996]	[0.723]
	N=3,567	N=3,567	N=3,567	N=3,557
Current smoker (cigarettes) [survey]	0.060	-0.023**	-0.022**	-0.009*
		(0.009)	(0.009)	(0.005)
		[0.139]	[0.159]	[0.589]
	N=3,566	N=3,566	N=3,566	N=3,556
Drinker [survey]	0.672	-0.012	-0.013	-0.003
		(0.017)	(0.016)	(0.013)
		[0.998]	[0.983]	[0.992]
	N=3,565	N=3,565	N=3,565	N=3,555
Heavy drinker [survey]	0.047	-0.003	-0.002	0.003
		(0.008)	(0.008)	(0.007)
		[1.000]	[0.999]	[0.992]
	N=3,563	N=3,563	N=3,563	N=3,553
Chronic condition [survey]	0.735	-0.004	0.003	0.001
		(0.016)	(0.015)	(0.012)
		[1.000]	[0.999]	[0.997]
	N=3,565	N=3,565	N=3,565	N=3,555
Excellent or v. good health [survey]	0.564	-0.004	-0.007	-0.024
		(0.018)	(0.017)	(0.015)
		[1.000]	[0.996]	[0.689]
	N=3,567	N=3,567	N=3,567	N=3,557
Not poor health [survey]	0.990	-0.004	-0.005	-0.005*
		(0.003)	(0.003)	(0.003)
		[0.952]	[0.863]	[0.675]
	N=3,567	N=3,567	N=3,567	N=3,557
Physical problems [survey]	0.403	-0.007	-0.003	0.001
		(0.018)	(0.017)	(0.015)
		[1.000]	[0.999]	[0.997]
	N=3,567	N=3,567	N=3,567	N=3,557
Lots of energy [survey]	0.309	0.040**	0.039**	0.027*
		(0.016)	(0.016)	(0.014)
		[0.176]	[0.166]	[0.530]
	N=3,566	N=3,566	N=3,566	N=3,556
Bad emotional health [survey]	0.311	0.017	0.015	0.021
		(0.016)	(0.016)	(0.015)
		[0.977]	[0.981]	[0.723]
	N=3,566	N=3,566	N=3,566	N=3,556
Overweight [survey]	0.562	0.009	0.018	0.027**
		(0.018)	(0.017)	(0.011)
		[0.999]	[0.980]	[0.162]
	N=3,567	N=3,567	N=3,567	N=3,557
High BP/cholesterol/glucose [survey]	0.324	0.005	0.015	0.020
		(0.017)	(0.016)	(0.013)
		[1.000]	[0.981]	[0.699]
	N=3,567	N=3,567	N=3,567	N=3,557
Sedentary [survey]	0.560	0.001	-0.002	-0.008
		(0.018)	(0.017)	(0.013)
		[1.000]	[0.999]	[0.977]
	N=3,565	N=3,565	N=3,565	N=3,555

Notes: The outcomes in this table constitute a single family of outcomes for calculating family-wise p-values. See notes to Appendix Table A.2a for additional details.

Table A.2g: Treatment Effects (ITT)

Outcome Variable	(1) Mean	(2) No Controls	(3) Strata FEs	(4) Post-Lasso
D. Medical Utilization (Quantity) [admin]				
Time to first claim <= 1 month [admin]	0.578	-0.029	-0.027	-0.011
		(0.019)	(0.018)	(0.015)
		[0.521]	[0.564]	[0.956]
	N=3,162	N=3,162	N=3,162	N=3,145
Time to first claim <= 2 months [admin]	0.689	-0.005	-0.002	0.011
		(0.018)	(0.017)	(0.014)
		[0.954]	[0.991]	[0.955]
	N=3,165	N=3,165	N=3,165	N=3,145
Time to first claim <= 3 months [admin]	0.758	0.007	0.009	0.022
		(0.016)	(0.016)	(0.014)
		[0.954]	[0.967]	[0.535]
	N=3,166	N=3,166	N=3,166	N=3,145
Time to first claim <= 6 months [admin]	0.842	-0.008	-0.006	-0.001
		(0.014)	(0.013)	(0.012)
		[0.954]	[0.977]	[0.984]
	N=3,175	N=3,175	N=3,175	N=3,147
Time to first claim <= 12 months [admin]	0.902	-0.007	-0.007	0.001
		(0.011)	(0.011)	(0.010)
		[0.954]	[0.967]	[0.984]
	N=3,238	N=3,238	N=3,238	N=3,152
Pharmaceutical events (days/month) [admin]	0.822	-0.022	-0.009	0.010
		(0.038)	(0.036)	(0.018)
		[0.954]	[0.991]	[0.956]
	N=3,238	N=3,238	N=3,238	N=3,152
Physician office visits (days/month) [admin]	0.308	0.032	0.032	0.010
		(0.025)	(0.026)	(0.018)
		[0.698]	[0.721]	[0.956]
	N=3,238	N=3,238	N=3,238	N=3,152
Hospital stays (days/month) [admin]	0.490	-0.011	-0.006	0.027
		(0.024)	(0.024)	(0.017)
		[0.954]	[0.991]	[0.543]
	N=3,238	N=3,238	N=3,238	N=3,152

Notes: The outcomes in this table constitute a single family of outcomes for calculating family-wise p-values. See notes to Appendix Table A.2a for additional details.

Table A.3a: Treatment Effects: Experimental vs. Observational Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Experimental (IV)			Observational (OLS)		
Outcome Variable	No Controls	Strata FEs	Post-Lasso	No Controls	Strata FEs	Post-Lasso
A. Medical Spending [admin]						
Total spending (dollars/month) [admin]	12.4 (78.8) N=3,238	29.1 (78.4) N=3,238	45.0 (59.1) N=3,152	-132.7* (68.0) N=2,207	-157.8** (65.5) N=2,207	-98.4 (61.1) N=2,140
Drug spending [admin]	-13.7 (43.2) N=3,238	-8.6 (41.6) N=3,238	-12.8 (20.4) N=3,152	-26.5 (27.3) N=2,207	-34.9 (26.9) N=2,207	-7.3 (12.0) N=2,140
Office spending [admin]	-9.9 (16.2) N=3,238	-9.3 (15.9) N=3,238	-3.2 (6.8) N=3,152	12.1 (7.5) N=2,207	9.4 (7.2) N=2,207	8.8* (5.1) N=2,140
Hospital spending [admin]	31.6 (50.0) N=3,238	42.8 (51.7) N=3,238	40.6 (45.0) N=3,152	-113.9** (55.1) N=2,207	-123.0** (52.1) N=2,207	-101.1* (54.2) N=2,140
Non-zero medical spending [admin]	-0.012 (0.018) N=3,238	-0.011 (0.018) N=3,238	0.004 (0.016) N=3,152	0.060*** (0.014) N=2,207	0.042*** (0.013) N=2,207	0.036*** (0.012) N=2,140

Notes: Each row and column reports estimates from a separate regression. The outcome in each regression is specified by the table row, and the (endogenous) focal independent variable is an indicator for completing the screening and HRA. For the IV specifications (columns (1)-(3)), the instrument is an indicator for inclusion in the treatment group, and observations include individuals in the control or treatment groups. For the OLS specifications (columns (4)-(6)), there is no instrument and observations are restricted to individuals in the treatment group. The control strategy is specified by the column. Post-Lasso controls include covariates selected by Lasso to predict either the dependent variable or the focal independent variable. The set of potential predictors include baseline values of all available variables in the same family of outcomes, strata variables, and the baseline (2016) survey variables reported in Table 1a, as well as all two-way interactions between these predictors. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference.

Table A.3b: Treatment Effects: Experimental vs. Observational Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Experimental (IV)			Observational (OLS)		
Outcome Variable	No Controls	Strata FEs	Post-Lasso	No Controls	Strata FEs	Post-Lasso
A. Medical Spending [survey]						
Pharmaceutical drug utilization [survey]	-0.017	-0.013	0.000	0.022	0.018	0.018
	(0.024)	(0.023)	(0.020)	(0.019)	(0.019)	(0.019)
	N=3,567	N=3,567	N=2,433	N=2,410	N=2,410	N=1,641
Physician/ER utilization [survey]	0.004	0.002	0.025	0.024	0.020	0.016
	(0.024)	(0.023)	(0.025)	(0.019)	(0.019)	(0.022)
	N=3,567	N=3,567	N=2,433	N=2,410	N=2,410	N=1,641
Hospital utilization [survey]	0.005	0.006	0.012	-0.009	-0.010	-0.015
	(0.008)	(0.008)	(0.010)	(0.007)	(0.008)	(0.010)
	N=3,567	N=3,567	N=2,433	N=2,410	N=2,410	N=1,641

Notes: See notes to Appendix Table A.3a.

Table A.3c: Treatment Effects: Experimental vs. Observational Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Experimental (IV)			Observational (OLS)		
Outcome Variable	No Controls	Strata FEs	Post-Lasso	No Controls	Strata FEs	Post-Lasso
B. Employment and Productivity [admin]						
Annual salary (share of baseline salary) [admin]	-0.000	-0.003	-0.003	0.004	0.005	0.006
	(0.008)	(0.008)	(0.008)	(0.005)	(0.005)	(0.005)
	N=4,146	N=4,146	N=4,130	N=2,840	N=2,840	N=2,828
Job terminated [admin]	-0.022	-0.023	-0.023	-0.082***	-0.080***	-0.068***
	(0.018)	(0.017)	(0.017)	(0.011)	(0.011)	(0.011)
	N=4,834	N=4,834	N=4,753	N=3,300	N=3,300	N=3,244
Sick leave (days/year) [admin]	0.397	0.506	0.311	0.266	0.030	-0.072
	(0.391)	(0.351)	(0.336)	(0.273)	(0.254)	(0.249)
	N=4,782	N=4,782	N=4,711	N=3,265	N=3,265	N=3,216

Notes: See notes to Appendix Table A.3a.

Table A.3d: Treatment Effects: Experimental vs. Observational Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Experimental (IV)			Observational (OLS)		
Outcome Variable	No Controls	Strata FEs	Post-Lasso	No Controls	Strata FEs	Post-Lasso
B. Employment and Productivity [survey]						
Any sick days in past year [survey]	0.007 (0.027) N=3,565	0.011 (0.026) N=3,565	0.021 (0.024) N=3,514	0.004 (0.021) N=2,409	-0.004 (0.021) N=2,409	-0.020 (0.019) N=2,376
Worked 50+ hours/week [survey]	-0.006 (0.020) N=3,566	-0.013 (0.018) N=3,566	0.008 (0.015) N=3,515	-0.037** (0.016) N=2,409	-0.034** (0.015) N=2,409	-0.009 (0.012) N=2,376
Very satisfied with job [survey]	-0.038 (0.027) N=3,564	-0.042 (0.026) N=3,564	-0.043* (0.023) N=3,512	-0.017 (0.021) N=2,407	-0.018 (0.021) N=2,407	-0.012 (0.018) N=2,373
Very or somewhat satisfied with job [survey]	-0.006 (0.020) N=3,564	-0.009 (0.020) N=3,564	-0.020 (0.018) N=3,512	0.003 (0.016) N=2,407	0.001 (0.016) N=2,407	0.005 (0.015) N=2,373
Management priority on health/safety [survey]	0.087*** (0.023) N=3,566	0.087*** (0.023) N=3,566	0.077*** (0.021) N=3,514	-0.004 (0.017) N=2,410	-0.012 (0.017) N=2,410	-0.007 (0.016) N=2,376
Happier at work than last year [survey]	0.014 (0.027) N=3,562	0.008 (0.027) N=3,562	-0.004 (0.027) N=3,510	0.022 (0.021) N=2,408	0.023 (0.022) N=2,408	0.013 (0.021) N=2,374
Presenteeism [survey]	-0.035 (0.397) N=3,567	-0.076 (0.391) N=3,567	-0.227 (0.361) N=3,515	-0.378 (0.312) N=2,410	-0.304 (0.314) N=2,410	-0.334 (0.289) N=2,376
Feel very productive at work [survey]	-0.027 (0.027) N=3,567	-0.020 (0.027) N=3,567	-0.031 (0.026) N=3,515	-0.040* (0.021) N=2,410	-0.043** (0.021) N=2,410	-0.036* (0.020) N=2,376
Received promotion [survey]	0.012 (0.027) N=3,562	0.000 (0.027) N=3,562	0.001 (0.027) N=3,511	0.032 (0.021) N=2,408	0.039* (0.021) N=2,408	0.024 (0.021) N=2,375
Job search very likely [survey]	0.047** (0.018) N=3,561	0.040** (0.018) N=3,561	0.039** (0.017) N=3,511	-0.011 (0.015) N=2,406	-0.013 (0.015) N=2,406	-0.001 (0.014) N=2,374
Job search somewhat/very likely [survey]	0.028 (0.026) N=3,561	0.019 (0.025) N=3,561	0.018 (0.024) N=3,511	-0.030 (0.021) N=2,406	-0.033* (0.020) N=2,406	-0.023 (0.019) N=2,374

Notes: See notes to Appendix Table A.3a.

Table A.3e: Treatment Effects: Experimental vs. Observational Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Experimental (IV)			Observational (OLS)		
Outcome Variable	No Controls	Strata FEs	Post-Lasso	No Controls	Strata FEs	Post-Lasso
C. Health Status and Behaviors [admin]						
IL Marathon/10K/5K 2017 [admin]	0.003	0.003	-0.011	0.059***	0.054***	0.024***
	(0.014)	(0.013)	(0.011)	(0.008)	(0.008)	(0.006)
	N=4,834	N=4,834	N=4,817	N=3,300	N=3,300	N=3,287
Campus gym visits (days/year) [admin]	-0.110	-0.121	0.757	3.527***	3.849***	2.160***
	(1.309)	(1.276)	(0.656)	(0.813)	(0.804)	(0.425)
	N=4,834	N=4,834	N=4,817	N=3,300	N=3,300	N=3,287

Notes: See notes to Appendix Table A.3a.

Table A.3f: Treatment Effects: Experimental vs. Observational Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Experimental (IV)			Observational (OLS)		
Outcome Variable	No Controls	Strata FEs	Post-Lasso	No Controls	Strata FEs	Post-Lasso
C. Health Status and Behaviors [survey]						
Ever screened [survey]	0.060***	0.065***	0.056***	0.073***	0.074***	0.061***
	(0.014)	(0.013)	(0.012)	(0.011)	(0.010)	(0.009)
	N=3,567	N=3,567	N=3,557	N=2,410	N=2,410	N=2,404
Physically active [survey]	0.023	0.025	-0.016	0.020	0.032	0.027*
	(0.026)	(0.026)	(0.019)	(0.021)	(0.020)	(0.015)
	N=3,567	N=3,567	N=3,557	N=2,410	N=2,410	N=2,404
Trying to be active [survey]	0.008	0.010	0.028	0.052***	0.049***	0.036**
	(0.021)	(0.020)	(0.018)	(0.017)	(0.017)	(0.015)
	N=3,567	N=3,567	N=3,557	N=2,410	N=2,410	N=2,404
Current smoker (cigarettes) [survey]	-0.035**	-0.034**	-0.014*	-0.033***	-0.032***	-0.005
	(0.014)	(0.013)	(0.008)	(0.010)	(0.010)	(0.006)
	N=3,566	N=3,566	N=3,556	N=2,410	N=2,410	N=2,404
Drinker [survey]	-0.018	-0.020	-0.007	0.010	0.015	-0.010
	(0.025)	(0.025)	(0.021)	(0.020)	(0.020)	(0.017)
	N=3,565	N=3,565	N=3,555	N=2,409	N=2,409	N=2,403
Heavy drinker [survey]	-0.004	-0.002	0.005	-0.003	-0.003	0.001
	(0.012)	(0.012)	(0.010)	(0.009)	(0.009)	(0.008)
	N=3,563	N=3,563	N=3,553	N=2,408	N=2,408	N=2,402
Chronic condition [survey]	-0.005	0.005	0.000	0.033*	0.037**	0.016
	(0.024)	(0.023)	(0.018)	(0.019)	(0.019)	(0.014)
	N=3,565	N=3,565	N=3,555	N=2,409	N=2,409	N=2,403
Excellent or v. good health [survey]	-0.007	-0.011	-0.034	-0.015	-0.018	0.005
	(0.027)	(0.026)	(0.023)	(0.021)	(0.021)	(0.018)
	N=3,567	N=3,567	N=3,557	N=2,410	N=2,410	N=2,404
Not poor health [survey]	-0.006	-0.007	-0.008	0.009*	0.008	0.009*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
	N=3,567	N=3,567	N=3,557	N=2,410	N=2,410	N=2,404
Physical problems [survey]	-0.010	-0.005	0.000	0.025	0.026	0.011
	(0.027)	(0.026)	(0.024)	(0.021)	(0.021)	(0.020)
	N=3,567	N=3,567	N=3,557	N=2,410	N=2,410	N=2,404
Lots of energy [survey]	0.060**	0.060**	0.036*	-0.030	-0.026	-0.013
	(0.025)	(0.024)	(0.022)	(0.020)	(0.020)	(0.018)
	N=3,566	N=3,566	N=3,556	N=2,410	N=2,410	N=2,404
Bad emotional health [survey]	0.026	0.022	0.035	-0.003	-0.005	0.003
	(0.025)	(0.025)	(0.023)	(0.020)	(0.020)	(0.019)
	N=3,566	N=3,566	N=3,556	N=2,410	N=2,410	N=2,404
Overweight [survey]	0.014	0.027	0.041**	0.031	0.029	-0.005
	(0.027)	(0.026)	(0.016)	(0.021)	(0.021)	(0.015)
	N=3,567	N=3,567	N=3,557	N=2,410	N=2,410	N=2,404
High BP/cholesterol/glucose [survey]	0.008	0.023	0.028	0.030	0.033*	0.032*
	(0.025)	(0.025)	(0.020)	(0.020)	(0.020)	(0.017)
	N=3,567	N=3,567	N=3,557	N=2,410	N=2,410	N=2,404
Sedentary [survey]	0.002	-0.003	-0.012	0.074***	0.056***	-0.003
	(0.027)	(0.026)	(0.020)	(0.021)	(0.021)	(0.016)
	N=3,565	N=3,565	N=3,555	N=2,408	N=2,408	N=2,402

Notes: See notes to Appendix Table A.3a.

Table A.3g: Treatment Effects: Experimental vs. Observational Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Experimental (IV)			Observational (OLS)		
Outcome Variable	No Controls	Strata FEs	Post-Lasso	No Controls	Strata FEs	Post-Lasso
D. Medical Utilization (Quantity) [admin]						
Time to first claim <= 1 month [admin]	-0.047 (0.031) N=3,162	-0.044 (0.029) N=3,162	-0.014 (0.024) N=3,145	0.034 (0.022) N=2,148	0.009 (0.021) N=2,148	0.027 (0.018) N=2,134
Time to first claim <= 2 months [admin]	-0.008 (0.029) N=3,165	-0.003 (0.027) N=3,165	0.018 (0.023) N=3,145	0.053** (0.021) N=2,151	0.029 (0.020) N=2,151	0.032* (0.017) N=2,134
Time to first claim <= 3 months [admin]	0.012 (0.027) N=3,166	0.016 (0.025) N=3,166	0.035 (0.022) N=3,145	0.060*** (0.019) N=2,152	0.034* (0.019) N=2,152	0.041** (0.017) N=2,134
Time to first claim <= 6 months [admin]	-0.013 (0.023) N=3,175	-0.010 (0.022) N=3,175	0.002 (0.020) N=3,147	0.069*** (0.017) N=2,159	0.048*** (0.016) N=2,159	0.051*** (0.015) N=2,136
Time to first claim <= 12 months [admin]	-0.012 (0.018) N=3,238	-0.011 (0.018) N=3,238	0.004 (0.016) N=3,152	0.060*** (0.014) N=2,207	0.042*** (0.013) N=2,207	0.036*** (0.012) N=2,140
Pharmaceutical events (days/month) [admin]	-0.036 (0.061) N=3,238	-0.014 (0.058) N=3,238	0.015 (0.030) N=3,152	-0.107** (0.045) N=2,207	-0.137*** (0.043) N=2,207	-0.043** (0.022) N=2,140
Physician office visits (days/month) [admin]	0.052 (0.041) N=3,238	0.052 (0.041) N=3,238	0.016 (0.029) N=3,152	0.057* (0.032) N=2,207	0.043 (0.034) N=2,207	0.042** (0.021) N=2,140
Hospital stays (days/month) [admin]	-0.018 (0.039) N=3,238	-0.009 (0.038) N=3,238	0.040 (0.028) N=3,152	-0.019 (0.030) N=2,207	-0.039 (0.030) N=2,207	-0.027 (0.023) N=2,140

Notes: See notes to Appendix Table A.3a.

Table A.4a: Treatment Effects (ITT) by Treatment Group: Total Health Care Spending

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Group (any)	7.59	17.80				
	(48.35)	(48.52)				
Group A* (A25, A75)			12.35	29.18		
			(61.74)	(62.12)		
Group B* (B25, B75)			76.58	84.23		
			(94.79)	(94.30)		
Group C* (C25, C75)			7.49	14.71		
			(59.36)	(59.22)		
Group *75 (A75, B75, C75)			-50.66	-51.64		
			(59.73)	(59.82)		
Group A25					15.51	24.39
					(65.91)	(66.05)
Group B25					150.61	163.06
					(128.19)	(127.81)
Group C25					-67.61	-57.35
					(58.55)	(57.16)
Group A75					-41.58	-17.37
					(69.61)	(69.37)
Group B75					-50.89	-49.48
					(70.27)	(70.53)
Group C75					38.61	41.80
					(80.94)	(80.74)
Constant	568.38***	561.40***	568.38***	561.51***	568.38***	561.50***
	(37.97)	(37.50)	(37.99)	(37.51)	(38.00)	(37.52)
N	3,238	3,238	3,238	3,238	3,238	3,238
Strata FE	No	Yes	No	Yes	No	Yes
F Test	0.88	0.71	0.91	0.91	0.56	0.59

Notes: Each column reports estimates from a separate regression estimated over individuals in the treatment and control groups in the claims sample. The outcome in each regression is average monthly health care spending over the first 12 months of the wellness program (August 2016 - July 2017), and regressions are weighted by the number of months of coverage. The independent variables are indicators for inclusion in the specified treatment groups. Regressions reported in columns (2), (4), and (6) are the same as those reported in columns (1), (3), and (5) respectively, but with the addition of strata fixed effects. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference.

Table A.4b: Treatment Effects (ITT) by Treatment Group: Any Health Care Spending

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment Group (any)	-0.007	-0.007				
	(0.011)	(0.011)				
Group A* (A25, A75)			0.005	0.005		
			(0.015)	(0.015)		
Group B* (B25, B75)			-0.012	-0.012		
			(0.016)	(0.015)		
Group C* (C25, C75)			0.000	-0.004		
			(0.016)	(0.015)		
Group *75 (A75, B75, C75)			-0.010	-0.006		
			(0.013)	(0.012)		
Group A25					0.017	0.012
					(0.017)	(0.016)
Group B25					-0.015	-0.014
					(0.018)	(0.018)
Group C25					-0.009	-0.010
					(0.018)	(0.017)
Group A75					-0.018	-0.009
					(0.019)	(0.018)
Group B75					-0.019	-0.017
					(0.019)	(0.018)
Group C75					-0.000	-0.003
					(0.018)	(0.018)
Constant	0.907***	0.906***	0.907***	0.906***	0.907***	0.906***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
N	3,238	3,238	3,238	3,238	3,238	3,238
Strata FE	No	Yes	No	Yes	No	Yes
F Test	0.500	0.523	0.684	0.749	0.581	0.799

Notes: Each column reports estimates from a separate regression estimated over individuals in the treatment and control groups in the claims sample. The outcome in each regression is an indicator for positive health care spending over the first 12 months of the wellness program (August 2016 - July 2017). The independent variables are indicators for inclusion in the specified treatment groups. Regressions reported in columns (2), (4), and (6) are the same as those reported in columns (1), (3), and (5) respectively, but with the addition of strata fixed effects. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level using conventional inference.

Table A.5: IV Treatment Effects: Screening and Wellness Participation

	(1)	(2)	(3)	(4)
	Ever screened	Ever screened	Management priority on health/safety	Management priority on health/safety
Completed Screening and HRA	0.097***	0.098***	0.124**	0.117**
	(0.026)	(0.026)	(0.050)	(0.050)
Completed Fall and Spring Wellness Activities	-0.098	-0.087	-0.122	-0.103
	(0.061)	(0.059)	(0.119)	(0.118)
N	3,567	3,567	3,566	3,566
Strata FE	No	Yes	No	Yes
First-stage F-statistic	12.580	12.814	12.580	12.814

Notes: Each column reports estimates from a separate regression. The outcome variable is specified by the column heading. We instrument for both regressors using six indicators for inclusion in the six treatment groups. Robust standard errors are reported in parentheses. A */**/*** indicates significance at the 10/5/1% level.

Table A.6: Winsorized Medical Spending Treatment Effects

	(1)	(2)	(3)	(4)	(5)
A. ITT Estimates (No Controls)					
Total spending (dollars/month) [admin]	7.6	-4.3	-0.7	2.4	0.5
	(48.4)	(38.0)	(32.6)	(21.1)	(13.8)
	[-87.2, 102.4]	[-78.7, 70.2]	[-64.7, 63.2]	[-38.9, 43.7]	[-26.5, 27.5]
N	3,238	3,238	3,238	3,238	3,238
Winsorization (percent)	0	0.5	1	2.5	5
B. ITT Estimates (Post-Lasso)					
Total spending (dollars/month) [admin]	30.9	13.7	15.4	14.6	10.4
	(36.7)	(23.1)	(19.4)	(13.5)	(9.7)
	[-41.0, 102.8]	[-31.7, 59.0]	[-22.7, 53.4]	[-11.8, 41.0]	[-8.7, 29.5]
N	3,152	3,152	3,152	3,152	3,152
Winsorization (percent)	0	0.5	1	2.5	5
C. IV Estimates (No Controls)					
Total spending (dollars/month) [admin]	12.4	-7.0	-1.2	3.9	0.8
	(78.8)	(61.8)	(53.1)	(34.3)	(22.4)
	[-142.0, 166.7]	[-128.2, 114.2]	[-105.3, 102.9]	[-63.3, 71.2]	[-43.1, 44.8]
N	3,238	3,238	3,238	3,238	3,238
Winsorization (percent)	0	0.5	1	2.5	5
D. IV Estimates (Post-Lasso)					
Total spending (dollars/month) [admin]	45.0	16.2	18.5	20.0	15.3
	(59.1)	(37.9)	(31.6)	(21.8)	(15.7)
	[-70.8, 160.8]	[-58.1, 90.5]	[-43.5, 80.5]	[-22.8, 62.8]	[-15.5, 46.0]
N	3,152	3,152	3,152	3,152	3,152
Winsorization (percent)	0	0.5	1	2.5	5

Notes: Each row and column reports estimates from a separate regression, where observations include individuals in the control or treatment groups. The outcome in each regression is winsorized (top-coded) average monthly health care spending over the first 12 months of the wellness program (August 2016 - July 2017), winsorized at the level indicated in each column. Regressions are weighted by the number of months of coverage. In Panels A and B (ITT), the focal independent variable is an indicator for inclusion in the treatment group. The specifications reported in Panel A do not include controls, while those reported in Panel B include the same controls as the ITT post-Lasso specification reported in row 1 and column (4) of Table 4. In Panels C and D (IV), the (endogenous) focal independent variable is an indicator for completing the screening and HRA and the instrument is an indicator for inclusion in the treatment group. The specifications reported in Panel C do not include controls, while those reported in Panel D include the same controls as the IV post-Lasso specification reported in row 1 and column (3) of Table 5. There is no winsorization of the outcome in column (1), and thus the ITT and IV estimates are identical to the total spending effects of the corresponding No Controls and Post-Lasso specifications reported in Table 4 and Table 5. Robust standard errors are reported in parentheses, and 95% confidence intervals are reported in brackets. A */**/*** indicates significance at the 10/5/1% level using conventional inference.

Table A.7: Variable Definitions

Variable Name	Data Source	Survey Question(s)	Formula	Time Period
Male	Human resources data (C)	N/A	Sex = Male	May 30, 2016
Age 50+	Human resources data (C)	N/A	$50 \leq$ Age	May 30, 2016
Age 37-49	Human resources data (C)	N/A	$37 \leq$ Age ≤ 49	May 30, 2016
White	Human resources data (C)	N/A	Race = White	May 30, 2016
Salary Q1 (bottom quartile)	Human resources data (C)	N/A	Salary $\leq$ 25th percentile	Pre-period: May 30, 2016 Post-period: August 15, 2017
Salary Q2	Human resources data (C)	N/A	25th pctile $\leq$ Salary $\leq$ 50th pctile	Pre-period: May 30, 2016 Post-period: August 15, 2017
Salary Q3	Human resources data (C)	N/A	50th pctile $\leq$ Salary $\leq$ 75th pctile	Pre-period: May 30, 2016 Post-period: August 15, 2017
Faculty	Human resources data (C)	N/A	Employment Class = Faculty	May 30, 2016
Academic Staff	Human resources data (C)	N/A	Employment Class = Academic Staff	May 30, 2016
Annual salary	Human resources data (C)	N/A	N/A	Pre-period: May 30, 2016 Post-period: August 15, 2017
Job terminated	Human resources data (C)	N/A	TerminationDate<= August 15, 2017	Pre-period: N/A Post-period: August 15, 2017
Sick leave (days/year)	Human resources data (C)	N/A	Sick days are measured monthly for CS employees, and biannually (August 15th and May 15th) for AP and Faculty employees. Number of sick days is normalized by fraction of year employed.	Pre-period:[30] 8/1/15 - 7/31/16 Post-period: 8/1/16 - 7/31/17

[30] Pre- and Post-period are offset by 15 days for AP and Faculty employees (see description in Formula).

Table A.7: Variable Definitions

Variable Name	Data Source	Survey Question(s)	Formula	Time Period
Ever screened	2016 Online survey (A) 2017 Online survey (G)	A1 (G1) Have you ever had your cholesterol checked? A2 (G2) Have you ever had a blood test for high blood sugar or diabetes, other than during pregnancy? A3 (G3) Have you ever had a blood test for high blood sugar or diabetes? A4 (G4) In the last 12 months, have you had a Pap test or Pap smear? A5 (G5) In the last 12 months, have you had a mammogram? A8 (G8) In the last 12 months, have you had a sigmoidoscopy or a colonoscopy? A9 (G9) In the last 12 months, have you had a blood test to check for prostate cancer?	Any of A1-A5, A8-A9 (G1-G5, G8-G9) = "Yes"	(A) July 2016 (G) July 2017
Physically Active	2016 Online survey (A) 2017 Online survey (G)	A11 (G11) Compared with most people your age, would you say you are more physically active, less physically active, or about the same?	A11 (G11) ="More active"	(A) July 2016 (G) July 2017
Trying to be active	2016 Online survey (A) 2017 Online survey (G)	A12 (G12) In the last 12 months, have you been told by a doctor or health professional to increase your physical activity or exercise? A13 (G13) Are you currently trying to increase your physical activity or exercise?	A12 (G12) = "Yes" or A13 (G13) = "Yes"	(A) July 2016 (G) July 2017
Current smoker (cigarettes)	2016 Online survey (A) 2017 Online survey (G)	A16 (G16) Have you smoked at least 100 cigarettes in your entire life? A17 (G17) Do you now smoke cigarettes every day, some days, or not at all?	A16 (G16) = "Yes" and A17 (G17) = "Every day" or "Some days"	(A) July 2016 (G) July 2017
Current smoker (other)	2016 Online survey (A) 2017 Online survey (G)	A22 (G22) Do you now smoke or use any other type of tobacco product, such as pipes, cigars, or chewing tobacco, every day, some days, or not at all? A23 (G23) Do you now use e-cigarettes (also known as vape-pens, hookah-pens, e-hookahs, or e-vaporizers) every day, some days, or not at all?	A22 (G22) & A23 (G23) != "Not at all"	(A) July 2016 (G) July 2017
Former smoker	2016 Online survey (A) 2017 Online survey (G)	A16 (G16) Have you smoked at least 100 cigarettes in your entire life? A17 (G17) Do you now smoke cigarettes every day, some days, or not at all?	A16 (G16) = "Yes" and A17 (G17) = "Not at all"	(A) July 2016 (G) July 2017
Drinker	2016 Online survey (A) 2017 Online survey (G)	A24 (G24) In the last 7 days, on how many days did you drink any type of alcoholic beverage?	A24 (G24) != 0	(A) July 2016 (G) July 2017
Heavy drinker	2016 Online survey (A) 2017 Online survey (G)	A25 (G25) In the last 7 days, on the days when you did drink alcohol, how many drinks did you usually have per day? One ?drink? is a 12 ounce can of beer, a 5 ounce glass of wine, or a 1.5 ounce shot of liquor.	A25 (G25) $\geq$ 4 if female A25 (G25) $\geq$ 5 if male	(A) July 2016 (G) July 2017
Chronic condition	2016 Online survey (A) 2017 Online survey (G)	A27 (G27) Have you ever been told by a doctor or other health professional that you have any of the following? Mark all that apply.	At least one box is checked	(A) July 2016 (G) July 2017

Table A.7: Variable Definitions

Variable Name	Data Source	Survey Question(s)	Formula	Time Period
Excellent or v. good health	2016 Online survey (A) 2017 Online survey (G)	A28 (G28) Overall, how would you rate your health during the past 4 weeks?	A28 (G28) = "Excellent" or "Very good"	(A) July 2016 (G) July 2017
Not poor health	2016 Online survey (A) 2017 Online survey (G)	A28 (G28) Overall, how would you rate your health during the past 4 weeks?	A28 (G28) != "Poor"	(A) July 2016 (G) July 2017
Physical problems	2016 Online survey (A) 2017 Online survey (G)	A29 (G29) During the past 4 weeks, how much did physical health problems limit your usual physical activities (such as walking or climbing stairs)? A30 (G30) During the past 4 weeks, how much difficulty did you have doing your daily work, both at home and away from home, because of your physical health? A31 (G31) How much bodily pain have you had during the past 4 weeks?	A29 (G29)="Somewhat?, "Quite a lot?, "Could not do physical activities? **or** A30 (G30) = "Some?, "Quite a lot?, "Could not do daily work? **or** A31 (G31) = "Mild?, "Moderate?, "Severe?, "Very severe?	(A) July 2016 (G) July 2017
Lots of energy	2016 Online survey (A) 2017 Online survey (G)	A32 (G32) During the past 4 weeks, how much energy did you have?	A32 (G32) = "An extraordinary amount", or "Quite a lot"	(A) July 2016 (G) July 2017
Bad emotional health	2016 Online survey (A) 2017 Online survey (G)	A33 (G33) During the past 4 weeks, how much have you been bothered by emotional problems (such as feeling anxious, depressed or irritable)?	A33 (G33) ov= "Moderately", "Quite a lot", "Extremely"	(A) July 2016 (G) July 2017
Overweight	2016 Online survey (A) 2017 Online survey (G)	A39 (39) How would you describe your body weight?	A39 (G39) = "Overweight" or "Very overweight"	
High BP / cholesterol / glucose	2016 Online survey (A) 2017 Online survey (G)	A40 (G40) How would you describe your blood pressure level? That is, if we measured it right now, do you think your blood pressure level would be: A41 (G41) How would you describe your cholesterol level? That is, if we measured it right now, do you think your cholesterol level would be: A42 (G42) How would you describe your blood glucose level? That is, if we measured it right now, do you think your blood glucose level would be:	A40 or A41 or A42 (G40 or G41 or G42) = "High" or "Very high"	(A) July 2016 (G) July 2017
Sedentary	2016 Online survey (A) 2017 Online survey (G)	A53 (G63) On an average day, how often does your job involve standing or walking around?	A53 (G63) = "None at all" or "Some, but less than 1 hour"	(A) July 2016 (G) July 2017
Pharmaceutical drug utilization	2016 Online survey (A) 2017 Online survey (G)	A34 (G34) How many different prescription medications are you currently taking? A35 (G35) How many different over-the-counter medications are you currently taking?	A34 (G34) > 0 or A35 (G35) > 0	(A) July 2016 (G) July 2017
Physician/ER utilization	2016 Online survey (A) 2017 Online survey (G)	A36 (G36) In the last 6 months, how many times did you go to a doctor?s office, clinic, emergency room, or other healthcare provider to get care for yourself? Do not include dental visits. Your best estimate is fine.	A36 (G36) != "None"	(A) July 2016 (G) July 2017

Table A.7: Variable Definitions

Variable Name	Data Source	Survey Question(s)	Formula	Time Period
Hospital utilization	2016 Online survey (A) 2017 Online survey (G)	A37 (G37) = In the last 6 months, how many different times were you a patient in a hospital at least overnight? Do not include hospital stays to deliver a baby. Your best estimate is fine.	A37 (G37) != "None"	(A) July 2016 (G) July 2017
Any sick days in past year	2016 Online survey (A) 2017 Online survey (G)	A45 (G46) In the last 12 months, about how many days of work have you missed because of disability or poor health? Your best estimate is fine.	A45 (G46) != 0	(A) July 2016 (G) July 2017
Worked 50+ hours/week	2016 Online survey (A) 2017 Online survey (G)	A44 (G45) About how many hours a week do you usually work at your current job or jobs?	A44 (G45) = "50 or more"	(A) July 2016 (G) July 2017
Very satisfied with job	2016 Online survey (A) 2017 Online survey (G)	A46 (G53) How satisfied are you with your current job?	A46 (G53) = "Very satisfied"	(A) July 2016 (G) July 2017
Very or somewhat satisfied with job	2016 Online survey (A) 2017 Online survey (G)	A46 (G53) How satisfied are you with your current job?	A46 (G53) = "Very satisfied" or "Somewhat satisfied"	(A) July 2016 (G) July 2017
Management priority on health/safety	2016 Online survey (A) 2017 Online survey (G)	A52 (G62) How much of a priority do you think your unit's management places on the health and safety of workers?	A52 (G62) = "Very high priority" or "Some priority"	(A) July 2016 (G) July 2017
Happier at work than last year	2017 Online survey (G)	G54 Do you feel happier at work this year than you did last year?	G54 = Yes	July 2017
Presenteeism	2017 Online survey (G)	G47 Despite having disability or poor health, I was able to finish hard tasks in my work. G48 At work, I was able to focus on achieving my goals despite disability or poor health. G49 Despite having disability or poor health, I felt energetic enough to complete all my work. G50 Because of disability or poor health, the stresses of my job were much harder to handle. G51 My disability or poor health distracted me from taking pleasure in my work. G52 I felt hopeless about finishing certain work tasks, due to my disability or poor health.	Stanford Presenteeim Scale (SPS-6), using G47-G52	July 2017
Feel very productive at work	2017 Online survey (G)	G56 How productive do you feel at work?	G56 = "Very productive"	July 2017
Received promotion	2017 Online survey (G)	G57 During the last 12 months, have you been given a promotion or more responsibility at work?	G57 = "Yes"	July 2017
Job search very likely	2017 Online survey (G)	G64 Taking everything into consideration, how likely are you to make a genuine effort to find a job with a new employer (outside the university) within the next year?	G64 = "Very likely"	July 2017
Job search somewhat / very likely	2017 Online survey (G)	G64 Taking everything into consideration, how likely are you to make a genuine effort to find a job with a new employer (outside the university) within the next year?	G64 = "Very likely" or "Somewhat likely"	July 2017
Total spending (dollars/month)	Health Insurance Claims Data (B)	N/A	Monthly Average	Pre-period: 7/1/15 - 7/31/16 Post-period: 8/1/16 - 7/31/17

Table A.7: Variable Definitions

Variable Name	Data Source	Survey Question(s)	Formula	Time Period
Drug spending	Health Insurance Claims Data (B)	N/A	Monthly Average	Pre-period: 7/1/15 - 7/31/16 Post-period: 8/1/16 - 7/31/17
Office spending	Health Insurance Claims Data (B)	N/A	Monthly Average	Pre-period: 7/1/15 - 7/31/16 Post-period: 8/1/16 - 7/31/17
Hospital spending	Health Insurance Claims Data (B)	N/A	Monthly Average	Pre-period: 7/1/15 - 7/31/16 Post-period: 8/1/16 - 7/31/17
Non-zero medical spending	Health Insurance Claims Data (B)	N/A	Monthly Average	Pre-period: 7/1/15 - 7/31/16 Post-period: 8/1/16 - 7/31/17
IL Marathon/10K/5K	Human Resources Data (C)	N/A	Pre-period: participated in *at least one* event during 2014 - 2016	Pre-period: April 2014 - April 2016 Post-period: April 2017
Campus gym visits (days/year)	Human Resources Data (C)	N/A	Number of visits to gym, measured by ID card swipe-in	Pre-period: 8/1/15 - 7/31/16 Post-period: 8/1/16 - 7/31/17

B For Online Publication: Comparison with prior literature — further details

We compiled all treatment effects estimates for health care costs and absenteeism from the studies included in the following review articles on wellness programs: Baicker, Cutler and Song (2010), Soler et al. (2010), Osilla et al. (2012), Lerner et al. (2013), and Baxter et al. (2014). There are two additional articles included below that are not featured in these review articles: Moore, LoGerfo and Inui (1980) and Bernacki, Tao and Yuspeh (2006). For each study, we identify the outcome of interest, i.e. health care costs (HCC) or absenteeism (ABS). We also indicate whether the study estimated a treatment-on-the-treated (TOT) or an intent-to-treat (ITT) effect.

If a study includes only a treatment and control group, we report the levels for each, T_1 and C_1, respectively. We use the level for the control group as the counterfactual level (CF Level). We then calculate the effect as $T_1 - C_1$, and the percent change as the effect divided by the counterfactual level.

Some studies also include pre and post levels for the treatment and control, T_0 and C_0, respectively. In those cases, we calculate the effect as $(T_1 - T_0) - (C_1 - C_0)$, and the counterfactual level as T_1 minus the effect. The percent change is still calculated as the effect divided by the counterfactual level.

Finally, some studies only include pre and post levels for the treatment group. In those cases, the effect is calculated as $T_1 - T_0$, the counterfactual level is T_0, and the percent change is again the effect divided by the counterfactual level.

For Entries with a "+" mark, we have taken the results as directly reported in an appendix table from Baicker, Cutler and Song (2010).

Table B.1: Detailed Description of Estimates from Figure 8

Title(Year)	Outcome	T_0	T_1	C_0	C_1	CF Level	Effect	% Change	Type
Aldana et al. (1993)$^{+}$	HCC	2,148	1,800	1,480	1,368	1,648	-189	-0.11	TOT
Aldana et al. (2005)	HCC		2,666.07		2,621	2,621	45.07	0.02	TOT
Aldana et al. (2005)	ABS		14.71		15.40	15.40	-0.69	-0.04	TOT
At'kov et al. (2011)	ABS		8.15		18.97	18.97	-10.82	-0.57	TOT
At'kov et al. (2011)	ABS		4.8		7.86	7.86	-3.06	-0.39	TOT
Baker et al. (2008)	HCC					4,090,978	-311,755	-0.08	TOT
Baun, Bemacki and Tsai (1986)$^{+}$	HCC		1,256		2,424	2,424	-1,168	-0.48	TOT
Baun, Bemacki and Tsai (1986)$^{+}$	ABS	8.7	9.0	10.0	12.4	11.1	-2.1	-0.19	TOT
Bernacki, Tao and Yuspeh (2005)	HCC		6,749		12,542	12,542	-5793	-0.46	TOT
Bernacki, Tao and Yuspeh (2005)	ABS		53.4		95.0	95.0	-41.6	-0.44	TOT
Bernacki, Tao and Yuspeh (2006)	HCC		12,554		20,400	20,400	-7846	-0.38	TOT
Bernacki, Tao and Yuspeh (2006)	ABS		53.0		99.0	99.0	-46.0	-0.46	TOT
Bertera (1990)$^{+}$	ABS	5.7	4.9	5.2	4.9	5.4	-0.5	-0.09	ITT
Bertera (1993)	ABS		3.0		2.9	2.9	0.1	0.03	ITT

Table B.1: Detailed Description of Estimates from Figure 8

Title(Year)	**Outcome**	**T_0**	**T_1**	**C_0**	**C_1**	**CF Level**	**Effect**	**% Change**	**Type**
Blair et al. (1986)+	ABS	5.6	5.5	6.0	6.2	5.8	-0.3	-0.05	TOT
Bly, Jones and Richardson (1986)+	HCC	247	655	253	1,234	1,228	-573	-0.47	ITT
Bridges et al. (2000) (1997–1998)	HCC		26.18		10.51	10.51	15.67	1.49	TOT
Bridges et al. (2000) (1997–1998)	ABS		0.60		0.41	0.41	0.19	0.45	TOT
Bridges et al. (2000) (1998–1999)	HCC		6.22		9.71	9.71	-3.49	-0.36	TOT
Bridges et al. (2000) (1998–1999)	ABS		0.18		0.24	0.24	-0.06	-0.26	TOT
Bunting and Cranor (2006)	HCC		1,585		3,050	3,050	-1,465	-0.48	TOT
Bunting and Cranor (2006)	ABS		16.80		66.50	66.50	-49.70	-0.75	TOT
Burton and Conti (2000)	ABS	29.3	23.2	22	23.3	30.60	-7.40	-0.24	ITT
Burton et al. (2005)	ABS		1.86		3.15	3.15	-1.29	-0.41	TOT
Campbell and Rumley (1997)	HCC		1,181		2,990	2,990	-1809	-0.61	TOT

Table B.1: Detailed Description of Estimates from Figure 8

Title(Year)	**Outcome**	**T_0**	**T_1**	**C_0**	**C_1**	**CF Level**	**Effect**	**% Change**	**Type**
Campbell and Rumley (1997)	ABS		50		109	109	-59	-0.54	TOT
Chenoweth and Garrett (2006)	HCC		1,351		1,580	1,580	-229	-0.14	TOT
Chenoweth et al. (2008)	HCC		11,165,777		13,344,709	13,344,709	-2,178,932	-0.16	TOT
Colombo et al. (2006)	ABS		294.2		366.82	366.82	-72.62	-0.20	TOT
Colombo et al. (2006)	ABS		161		231	231	-70	-0.30	TOT
Cousins and Liu (2003)	HCC		5,264		5,825	5,825	-561	-0.11	ITT
Davis et al. (2009)	HCC	24.6	-4.4	10.4	6.0	100	-24.60	-0.25	ITT
Davis et al. (2009)	ABS		7.6		10.1	10.1	-2.5	-0.25	ITT
Dille (1999)	HCC		946.27		6,177.52	6,177.52	-5231.24	-0.85	TOT
Dille (1999)	ABS		35		63	63	-28	-0.44	TOT
Fera, Bluml and Ellis (2009)	HCC		13,829		14,909	14,909	-1080	-0.07	TOT
Foote and Erfurt (1991) (1)	HCC	3,196	4,046	2,946	4,326	4,576	-530	-0.12	ITT
Foote and Erfurt (1991) (2)	HCC	2,579	3,407	2,946	4,326	3959	-552	0.24	ITT
Foote and Erfurt (1991) (3)	HCC	1,875	2,183	2,946	4,326	3,255	-1,072	-0.33	ITT

Table B.1: Detailed Description of Estimates from Figure 8

Title(Year)	**Outcome**	**T_0**	**T_1**	**C_0**	**C_1**	**CF Level**	**Effect**	**% Change**	**Type**
Fries et al. (1994)	HCC		354		497	497	-143	-0.08	ITT
Fries et al. (1994)	ABS		4.30		5.50	5.50	-1.20	-0.22	ITT
Fries and McShane (1998)[+]	HCC	1,616	1,185	500	419	1,535	-350	-0.23	TOT
Fries and McShane (1998)[+]	ABS	3.9	3.0	1.6	1.5	3.80	-0.80	-0.21	TOT
Gibbs et al. (1985)[+]	HCC	695	1,687	605	1,977	2,067	-380	-0.18	TOT
Goetzel et al. (1998)[+]	HCC		1,413		1,396	1,396	17	0.01	TOT
Green-McKenzie et al. (2002)	HCC		191,992		469,694	469,694	-277,702	-0.59	TOT
Groeneveld et al. (2011)	HCC		212		279	279	-67	-0.24	ITT
Groeneveld et al. (2011)	ACC		12.3		9.1	9.1	3.2	0.35	ITT
Groeneveld et al. (2011) (imputed)	ACC		14.4		15.7	15.7	-1.3	-0.08	ITT
Henke et al. (2011)	HCC		4,435		5,000	5,000	-565	-0.11	ITT
Herman et al. (2006)	ABS	0.052	0.051	0.065	0.077	0.06	-0.01	-0.20	TOT
Hochart and Lang (2011)	HCC	225.74	227.77	226.75	276.01	275.0	-47.23	-0.17	ITT
Hughes et al. (2007)	HCC		1,970		4,353	4,353	-2,383	-0.55	TOT
Hughes et al. (2007)	ABS		1.1		3.1	3.1	-2.0	-0.65	TOT

Table B.1: Detailed Description of Estimates from Figure 8

Title(Year)	Outcome	T_0	T_1	C_0	C_1	CF Level	Effect	% Change	Type
Jeffery et al. (1993)	ABS	18.0	13.5	19.1	18.2	18.0	-3.6	-0.2	ITT
Jones, Bly and Richardson (1990)[+]	ABS	5.9	5.6	5.3	6.0	6.6	-1.0	-0.15	ITT
Knight et al. (1994)[+]	ABS	9.1	10.2	9.1	10.8	10.8	-0.6	-0.06	TOT
Lechner and de Vries (1997)	ABS	12.4	11.0	14.3	14.2	12.3	-1.3	-0.11	TOT
Leigh et al. (1992)[+]	HCC	2,171	1,695	1,881	1,995	2,285	-590	-0.26	ITT
Leigh et al. (1992)	ABS	18.0	17.2	18.0	19.4	19.4	-2.2	-0.11	ITT
Linz et al. (2001)	ABS		2,137		3,702	3,702	-1,565	-0.42	ITT
Loeppke et al. (2008)	ABS		9.83		5.75	5.75	4.08	0.71	ITT
Lynch et al. (1990)[+]	ABS	4.4	3.7	5.6	5.5	4.3	-0.6	-0.14	TOT
Maes et al. (1998)	ABS	0.158	0.077	0.143	0.095	0.11	-0.03	-0.30	ITT
McCulloch et al. (2001)	ABS		56.4		73.5	73.5	-17.1	-0.23	TOT
McEachan et al. (2011)	HCC		17,900.0		17,979.4	17,979.4	-79.4	-0.004	ITT
Merrill et al. (2011)	HCC		3,441.3		5,969.3	5,969.3	-2,528.0	-0.42	TOT
Milani and Lavie (2009)	HCC	2,960	1,539	3,002	2,522	2,480	-941	-0.38	ITT
Mills et al. (2007)	ABS	0.38	0.35	0.58	0.76	0.56	-0.21	-0.38	TOT
Moore, LoGerfo and Inui (1980) (G1 vs G2)	HCC	7.8	6.2	7.0	5.9	6.70	-0.50	-0.07	ITT

Table B.1: Detailed Description of Estimates from Figure 8

Title(Year)	Outcome	T_0	T_1	C_0	C_1	CF Level	Effect	% Change	Type
Moore, LoGerfo and Inui (1980) (G1 vs G3)	HCC	7.9	6.0	7.0	5.9	6.70	-0.80	-0.12	ITT
Morales et al. (2004)	ABS		22.66		29.08	29.08	-6.42	-0.22	TOT
Musich, Adams and Edington (2000)$^{+}$	HCC	2,140	2,337	1,083	2,908	3,965	-1,628	-0.41	TOT
Naydeck et al. (2008)	HCC	1,531	2,907	1,427	3,429	3,533	-626	-0.18	TOT
Nilsson, Klasson and Nyberg (2001)	ABS	6.0	2.9	4.5	7.4	8.9	-6.0	-0.67	ITT
Nyman et al. (2012) (DM)	HCC	625.46	734.99	470.33	646.97	802.10	-67.11	-0.08	TOT
Nyman et al. (2012) (DM)	ABS	67.87	76.3	67.38	72.52	73.02	3.28	0.04	TOT
Nyman et al. (2012) (LM)	HCC	403.19	481.46	302.68	407.87		-26.93	-0.07	TOT
Nyman et al. (2012) (LM)	ABS	60.36	65.66	57.57	64.08	66.88	-1.22	-0.02	TOT
Osilla et al. (2010)	ABS		7.88		13.75	13.75	-5.87	-0.43	TOT
Ozminkowski et al. (1999)$^{+}$	HCC	2,736	3,411	2,896	4,136	3,976	-565	-0.14	TOT
Page et al. (2009)	HCC		169,780		105,220	105,220	64,560	0.61	TOT

Table B.1: Detailed Description of Estimates from Figure 8

Title(Year)	**Outcome**	**T_0**	**T_1**	**C_0**	**C_1**	**CF Level**	**Effect**	**% Change**	**Type**
Page et al. (2009)	ABS		600		800	800	-200	-0.25	TOT
Pegus et al. (2002)	ABS		0.33		0.49	0.49	-0.16	-0.32	
Pelletier, Boles and Lynch (2004)	ABS		0.01		0.015	0.015	-0.005	-0.33	TOT
Proper et al. (2004)	ABS		4,735		5,505	5,505	-770	-0.14	ITT
Proper et al. (2004)	ABS		21.0		27.25	27.25	-6.25	-0.23	ITT
Ringen et al. (2002)	HCC		236		325	325	-89	-0.27	ITT
Sacks et al. (2009)	HCC		2,413		2,327.86	2,327.86	85.14	0.04	TOT
Sacks et al. (2009) (High CV risk subgroup)	HCC		3,425		4,251.95	4,251.95	-826.95	-0.19	TOT
Samad et al. (2006)	ABS		14.22		67.44	67.44	-53.22	-0.79	TOT
Samad et al. (2006)	ABS		3.0		4.22	4.22	-1.22	-0.29	TOT
Schneider and Häck (2011)	HCC		134,700		289,141	289,141	-154,441	-0.53	ITT
Schultz et al. (2002)	ABS	6.6	17.2	6.6	23.3	23.3	-6.1	-0.26	TOT
Sciacca et al. (1993)	HCC	1,159	2,397	825	1,701	2,035	362	0.18	TOT
Serxner et al. (2001)	ABS	29.2	27.8	33.2	38.1	34.1	-6.3	-0.18	TOT
Serxner et al. (2003)+	HCC		4,176		4,454	4,454	-278	-0.06	TOT
Shephard et al. (1982)+	HCC	294	296	295	396	395	-99	-0.25	ITT

Table B.1: Detailed Description of Estimates from Figure 8

Title(Year)	Outcome	T_0	T_1	C_0	C_1	CF Level	Effect	% Change	Type
Shephard et al. (1982)+	ABS	0.3	0.1	0.1	0.5	0.7	-0.6	-0.86	ITT
Shi (1993)+ (G1 vs G2)	HCC	1,891	1,621	1,970	1,710	1,631	-10.0	-0.01	ITT
Shi (1993)+ (G1 vs G2)	ABS	4.96	4.69	5.05	4.78	4.69	0.0	0.0	ITT
Shi (1993)+ (G1 vs G3)	HCC	1,986	1,485	1,970	1,710	1,726	-241	-0.14	ITT
Shi (1993)+ (G1 vs G3)	ABS	5.15	4.08	5.05	4.78	4.88	-0.8	-0.16	ITT
Shi (1993)+ (G1 vs G4)	HCC	2,036	1,283	1,970	1,710	1,776	-493	-0.28	ITT
Shi (1993)+ (G1 vs G4)	ABS	5.22	3.24	5.05	4.78	4.95	-1.71	-0.35	ITT
Stave, Muchmore and Gardner (2003)	HCC		3,222		3,909	3,909	-687	-0.18	TOT
Stave, Muchmore and Gardner (2003)	ABS	3.1	2.3	3.1	3.3	3.3	-1.0	-0.3	TOT
Taimela et al. (2008)	HCC		925.1		1108.6	1108.6	-183.5	-0.17	ITT
Taimela et al. (2008)	HCC	17.4	19.3	17.1	29.9	30.2	-10.9	-0.36	ITT
Wang et al. (2007)	ABS		10.20		13.45	13.45	-3.25	-0.24	ITT
Wolf et al. (2009)	ABS	0.74	0.31	0.75	1.16	1.16	-0.85	-0.73	ITT
Wood, Olmstead and Craig (1989)	ABS	2.5	2.6	2.9	4.3	3.9	-1.3	-0.33	TOT
Golaszewski et al. (1992)	HCC	6,185	7,743	5,249	7,734	8,670	-927	-0.11	TOT

Prior Wellness Literature

Aldana, Steven G, Bert H Jacobson, Clifford J Harris, Patrick L Kelley, and William J Stone. 1993. "Influence of a mobile worksite health promotion program on health care costs." *American journal of preventive medicine*, 9(6): 378–383.

Aldana, Steven G, Ray M Merrill, Kristine Price, Aaron Hardy, and Ron Hager. 2005. "Financial impact of a comprehensive multisite workplace health promotion program." *Preventive medicine*, 40(2): 131–137. doi: 10.1016/j.ypmed.2004.05.008.

At'kov, O Yu, AV Azarov, DA Zhukov, N Nicoloyannis, and Laure Durand. 2011. "Influenza vaccination in healthy working adults in Russia." *Applied health economics and health policy*, 9(2): 89–99. doi: 10.2165/11538680-000000000-00000.

Baicker, Katherine, David Cutler, and Zirui Song. 2010. "Workplace wellness programs can generate savings." *Health Affairs*, 29(2): 304–311.

Baker, K.M., et al. 2008. "Using a return-on-investment estimation model to evaluate outcomes from on obesity management worksite health promotion program." *J Occup Environ Med*, 50: 981–990. doi: 10.1097/JOM.0b013e318184a489.

Baun, William B, Edward J Bemacki, and Shan P Tsai. 1986. "A preliminary investigation: effect of a corporate fitness program on absenteeism and health care cost." *Journal of Occupational and Environmental Medicine*, 28(1): 18–22. PMID: 3081697.

Baxter, Siyan, Kristy Sanderson, Alison J Venn, C Leigh Blizzard, and Andrew J Palmer. 2014. "The relationship between return on investment and quality of study methodology in workplace health promotion programs." *American Journal of Health Promotion*, 28(6): 347–363.

Bernacki, Edward J, Xuguang Grant Tao, and Larry Yuspeh. 2005. "A preliminary investigation of the effects of a provider network on costs and lost-time in workers' compensation." *Journal of occupational and environmental medicine*, 47(1): 3–10. PMID:15655406.

Bernacki, Edward J, Xuguang Grant Tao, and Larry Yuspeh. 2006. "An investigation of the effects of a healthcare provider network on costs and lost time in workers' compensation." *Journal of occupational and environmental medicine*, 48(9): 873–882.

Bertera, R. L. 1993. "Behavioral risk factor and illness day changes with workplace health promotion: two-year results." *American Journal of Health Promotion*, 7: 365–373. PMID: 10148712.

Bertera, Robert L. 1990. "The effects of workplace health promotion on absenteeism and employment costs in a large industrial population." *American journal of public health*, 80(9): 1101–1105. PMCID: PMC1404872.

Blair, Steven N, Mike Smith, Thomas R Collingwood, Roger Reynolds, Michael C Prentice, and Charles L Sterling. 1986. "Health promotion for educators: impact on absenteeism." *Preventive medicine*, 15(2): 166–175. PMID: 3714669.

Bly, Janet L, Robert C Jones, and Jean E Richardson. 1986. "Impact of worksite health promotion on health care costs and utilization: evaluation of Johnson & Johnson's Live for Life program." *JAMA*, 256(23): 3235–3240. doi:10.1001/jama.1986.03380230059026.

Bridges, Carolyn Buxton, William W Thompson, Martin I Meltzer, Gordon R Reeve, Walter J Talamonti, Nancy J Cox, Heather A Lilac, Henrietta Hall, Alexander Klimov, and Keiji Fukuda. 2000. "Effectiveness and cost-benefit of influenza vaccination of healthy working adults: a randomized controlled trial." *Jama*, 284(13): 1655–1663. doi:10.1001/jama.284.13.1655.

Bunting, Barry A, and Carole W Cranor. 2006. "The Asheville Project: long-term clinical, humanistic, and economic outcomes of a community-based medication therapy management program for asthma." *Journal of the American Pharmacists Association*, 46(2): 133–147. doi: 10.1331/154434506776180658.

Burton, Wayne N, and Daniel J Conti. 2000. "Disability management: corporate medical department management of employee health and productivity." *Journal of Occupational and Environmental Medicine*, 42(10): 1006–1012. PMID: 11039164.

Burton, Wayne N, Katherine T McCalister, Chin-Yu Chen, and Dee W Edington. 2005. "The association of health status, worksite fitness center participation, and two measures of productivity." *Journal of Occupational and Environmental Medicine*, 47(4): 343–351. PMID: 15824625.

Campbell, Douglas S, and Maria H Rumley. 1997. "Cost-effectiveness of the influenza vaccine in a healthy, working-age population." *Journal of Occupational and Environmental Medicine*, 39(5): 408–414. PMID: 9172085.

Chenoweth, David H, and Judy Garrett. 2006. "Cost-effectiveness analysis of a worksite clinic: is it worth the cost?" *Aaohn Journal*, 54(2): 84–91. doi: 10.1177/216507990605400206.

Chenoweth, David, Nanette Martin, Jared Pankowski, and Lawrence W Raymond. 2008. "Nurse practitioner services: Three-year impact on health care costs." *Journal of occupational and environmental medicine*, 50(11): 1293–1298. doi: 10.1097/JOM.0b013e318184563a.

Colombo, Giorgio L, Antonio Ferro, Marta Vinci, Maria Zordan, and Giulio Serra. 2006. "Cost-benefit analysis of influenza vaccination in a public healthcare unit." *Therapeutics and clinical risk management*, 2(2): 219. PMID: 18360596.

Cousins, Michael S, and Ying Liu. 2003. "Cost savings for a preferred provider organization population with multi-condition disease management: evaluating program impact using predictive modeling with a control group." *Disease Management*, 6(4): 207–217. doi: 10.1089/109350703322682522.

Davis, L, K Loyo, A Glowka, R Schwertfeger, L Danielson, C Brea, et al. 2009. "A comprehensive worksite wellness program in Austin, Texas: Steps to a Healthier Austin and Capital Metropolitan Transportation Authority." *Prev Chronic Dis*, 6(2). PMID: 19289003.

Dille, JoAnne Hein. 1999. "A worksite influenza immunization program: impact on lost work days, health care utilization, and health care spending." *AAOHN Journal*, 47(7): 301–309. PMID: 10661043.

Fera, Toni, Benjamin M Bluml, and William M Ellis. 2009. "Diabetes Ten City Challenge: final economic and clinical results." *Journal of the American Pharmacists Association*, 49(3): 383–391. doi: 10.1331/JAPhA.2009.09015.

Foote, Andrea, and John C Erfurt. 1991. "The benefit to cost ratio of work-site blood pressure control programs." *JAMA*, 265(10): 1283–1286. doi: 10.1001/jama.1991.03460100085029.

Fries, James F, and Dennis McShane. 1998. "Reducing need and demand for medical services in high-risk persons. A health education approach." *Western Journal of Medicine*, 169(4): 201. PMID:9795579.

Fries, James F, Harry Harrington, Robert Edwards, Louis A Kent, and Nancy Richardson. 1994. "Randomized controlled trial of cost reductions from a health education program: the California Public Employees' Retirement System (PERS) study." *American Journal of Health Promotion*, 8(3): 216–223.

Gibbs, James O, Dallas Mulvaney, Carol Henes, and Roger W Reed. 1985. "Work-site health promotion: five-year trend in employee health care costs." *Journal of Occupational and Environmental Medicine*, 27(11): 826–830. PMID: 4067688.

Goetzel, Ron Z, Bert H Jacobson, Steven G Aldana, Kris Vardell, and Leslie Yee. 1998. "Health care costs of worksite health promotion participants and non-participants." *Journal of Occupational and Environmental Medicine*, 40(4): 341–346. PMID: 9571525.

Golaszewski, Thomas, David Snow, Wendy Lynch, Louis Yen, and Debra Solomita. 1992. "A benefit-to-cost analysis of a work-site health promotion program." *Journal of Occupational and Environmental Medicine*, 34(12): 1164–1172. doi: 10.1108/17538351011054998.

Green-McKenzie, Judith, Sharon Rainer, Amy Behrman, and Edward Emmett. 2002. "The effect of a health care management initiative on reducing workers' compensation costs." *Journal of occupational and environmental medicine*, 44(12): 1100–1105. PMID: 12500451.

Groeneveld, Iris F, Marieke F van Wier, Karin I Proper, Judith E Bosmans, Willem van Mechelen, and Allard J van der Beek. 2011. "Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the construction industry at risk for cardiovascular disease." *Journal of occupational and environmental medicine*, 53(6): 610–617. doi: 10.1097/JOM.0b013e31821b9c24.

Henke, Rachel M, Ron Z Goetzel, Janice McHugh, and Fik Isaac. 2011. "Recent experience in health promotion at Johnson & Johnson: lower health spending, strong return on investment." *Health Affairs*, 30(3): 490–

499. doi: 10.1377/hlthaff.2010.0806.

Herman, Christopher W, Shirley Musich, Chifung Lu, Stewart Sill, Joyce M Young, and Dee W Edington. 2006. "Effectiveness of an incentive-based online physical activity intervention on employee health status." *Journal of Occupational and Environmental Medicine*, 48(9): 889–895. doi: 10.1097/01.jom.0000232526.27103.71.

Hochart, Cindy, and Michelle Lang. 2011. "Impact of a comprehensive worksite wellness program on health risk, utilization, and health care costs." *Population health management*, 14(3): 111–116. doi: 10.1089/pop.2010.0009.

Hughes, M Courtney, Teresa M Girolami, Allen D Cheadle, Jeffrey R Harris, and Donald L Patrick. 2007. "A lifestyle-based weight management program delivered to employees: examination of health and economic outcomes." *Journal of occupational and environmental medicine*, 49(11): 1212–1217. doi: 10.1097/JOM.0b013e318159489d.

Jeffery, Robert W, Jean L Forster, Bonny V Dunn, Simone A French, Paul G McGovern, and Harry A Lando. 1993. "Effects of work-site health promotion on illness-related absenteeism." *Journal of Occupational and Environmental Medicine*, 35(11): 1142–1146.

Jones, Robert C, Janet L Bly, and Jean E Richardson. 1990. "A study of a work site health promotion program and absenteeism." *Journal of Occupational and Environmental Medicine*, 32(2): 95–99.

Knight, Kevin K, Ron Z Goetzel, Jonathan E Fielding, Marvin Eisen, George W Jackson, Toby Y Kahr, Gayle M Kenny, Sally W Wade, and Shyun Duann. 1994. "An evaluation of Duke University's LIVE FOR LIFE health promotion program on changes in worker absenteeism." *Journal of Occupational and Environmental Medicine*, 36(5): 533–536.

Lechner, Lilian, and Hein de Vries. 1997. "Effects of an employee fitness program on reduced absenteeism." *Journal of occupational and environmental medicine*, 39(9): 827–831.

Leigh, J Paul, Nancy Richardson, Robert Beck, Clark Kerr, Harry Harrington, Charles L Parcell, and James F Fries. 1992. "Randomized controlled study of a retiree health promotion program: the Bank of America study." *Archives of Internal Medicine*, 152(6): 1201–1206.

Lerner, Debra, Angie Mae Rodday, Joshua T Cohen, and William H Rogers. 2013. "A systematic review of the evidence concerning the economic impact of employee-focused health promotion and wellness programs." *Journal of occupational and environmental medicine*, 55(2): 209–222.

Linz, Douglas H, Linda F Ford, Miriam J Nightingale, Pamela L Shannon, Jeffrey S Davin, Carol O Bradford, and Craig D Shepherd. 2001. "Care management of work injuries: results of a 1-year pilot outcome assurance program." *Journal of occupational and environmental medicine*, 43(11): 959–968.

Loeppke, Ron, Sean Nicholson, Michael Taitel, Matthew Sweeney, Vince Haufle, and Ronald C Kessler. 2008. "The impact of an integrated population health enhancement and disease management program on employee health risk, health conditions, and productivity." *Population health management*, 11(6): 287–296. doi: 10.1089/pop.2008.0006.

Lynch, Wendy D, Thomas J Golaszewski, Andrew F Clearie, David Snow, and Donald M Vickery. 1990. "Impact of a facility-based corporate fitness program on the number of absences from work due to illness." *Journal of Occupational and Environmental Medicine*, 32(1): 9–12. doi: 10.1097/00043764-199001000-00006.

Maes, Stan, Chris Verhoeven, France Kittel, and Hetty Scholten. 1998. "Effects of a Dutch work-site wellness-health program: the Brabantia Project." *American journal of public health*, 88(7): 1037–1041.

McCulloch, Joyce, Ronald J Ozminkowski, Brian Cuffel, Rodney L Dunn, William Goldman, Dolores Kelleher, and Andrea Comporato. 2001. "Analysis of a managed psychiatric disability program." *Journal of occupational and environmental medicine*, 43(2): 101–109.

McEachan, Rosemary RC, Rebecca J Lawton, Cath Jackson, Mark Conner, David M Meads, and Robert M West. 2011. "Testing a workplace physical activity intervention: a cluster randomized controlled trial." *International Journal of Behavioral Nutrition and Physical Activity*, 8(1): 29. doi: 10.1186/1479-5868-8-29.

Merrill, Ray M, Beverly Hyatt, Steven G Aldana, and Dan Kinnersley. 2011. "Lowering employee health care costs through the healthy lifestyle incentive program." *Journal of Public Health Management and Practice*, 17(3): 225–232. doi: 10.1097/PHH.0b013e3181f54128.

Milani, Richard V, and Carl J Lavie. 2009. "Impact of worksite wellness intervention on cardiac risk factors and one-year health care costs." *The American journal of cardiology*, 104(10): 1389–1392. doi: 10.1016/j.amjcard.2009.07.007.

Mills, Peter R, Ronald C Kessler, John Cooper, and Sean Sullivan. 2007. "Impact of a health promotion program on employee health risks and work productivity." *American Journal of Health Promotion*, 22(1): 45–53. doi: 10.4278/0890-1171-22.1.45.

Moore, Stephen H, James LoGerfo, and Thoma S Inui. 1980. "Effect of a self-care book on physician visits: a randomized trial." *JAMA*, 243(22): 2317–2320.

Morales, Alvaro, Maria M Martinez, Anne Tasset-Tisseau, Elena Rey, Florence Baron-Papillon, and Alain Follet. 2004. "Costs and benefits of influenza vaccination and work productivity in a Colombian company from the employer's perspective." *Value in Health*, 7(4): 433–441. doi: 10.1111/j.1524-4733.2004.74006.x.

Musich, Shirley A, Laura Adams, and Dee W Edington. 2000. "Effectiveness of health promotion programs in moderating medical costs in the USA." *Health Promotion International*, 15(1): 5–15. doi: 10.1093/heapro/15.1.5.

Naydeck, Barbara L, Janine A Pearson, Ronald J Ozminkowski, Brian T Day, and Ron Z Goetzel. 2008. "The impact of the highmark employee wellness programs on 4-year healthcare costs." *Journal of occupational and environmental medicine*, 50(2): 146–156. doi: 10.1097/JOM.0b013e3181617855.

Nilsson, Peter M, Eva-Birgitta Klasson, and Per Nyberg. 2001. "Life-style intervention at the worksite — reduction of cardiovascular risk factors in a randomized study." *Scandinavian journal of work, environment & health*, 57–62. PMID: 11266148.

Nyman, John A, Jean M Abraham, Molly Moore Jeffery, and Nathan A Barleen. 2012. "The effectiveness of a health promotion program after 3 years: evidence from the University of Minnesota." *Medical care*, 50(9): 772–778. doi: 10.1097/MLR.0b013e31825a8b1f.

Osilla, Karen Chan, Erin Dela Cruz, Jeremy NV Miles, Steven Zellmer, Katherine Watkins, Mary E Larimer, and G Alan Marlatt. 2010. "Exploring productivity outcomes from a brief intervention for at-risk drinking in an employee assistance program." *Addictive behaviors*, 35(3): 194–200. doi: 10.1016/j.addbeh.2009.10.001.

Osilla, Karen Chan, Kristin Van Busum, Christopher Schnyer, Jody Wozar Larkin, Christine Eibner, and Soeren Mattke. 2012. "Systematic review of the impact of worksite wellness programs." *The American journal of managed care*, 18(2): e68–81.

Ozminkowski, Ronald J, Rodney L Dunn, Ron Z Goetzel, Richard I Cantor, Jan Murnane, and Mary Harrison. 1999. "A return on investment evaluation of the Citibank, NA, Health Management Program." *American Journal of Health Promotion*, 14(1): 31–43. PMID: 10621522.

Page, Matthew J, L Clark Paramore, Dilesh Doshi, and Marcia FT Rupnow. 2009. "Evaluation of resource utilization and cost burden before and after an employer-based migraine education program." *Journal of occupational and environmental medicine*, 51(2): 213–220. doi: 10.1097/JOM.0b013e318192bcd1.

Pegus, Cheryl, Terry L Bazzarre, Jeffrey S Brown, and Joseph Menzin. 2002. "Effect of the Heart At Work program on awareness of risk factors, self-efficacy, and health behaviors." *Journal of Occupational and Environmental Medicine*, 44(3): 228–236. PMID: 11911024.

Pelletier, Barbara, Myde Boles, and Wendy Lynch. 2004. "Change in health risks and work productivity over time." *Journal of Occupational and Environmental Medicine*, 46(7): 746–754. PMID: 15247815.

Proper, Karin I, Martine C De Bruyne, Vincent H Hildebrandt, Allard J Van Der Beek, Willem Jan Meerding, and Willem Van Mechelen. 2004. "Costs, benefits and effectiveness of worksite physical activity counseling from the employer's perspective." *Scandinavian journal of work, environment & health*, 36–46. PMID: 15247815.

Ringen, Knut, Norman Anderson, Tim McAfee, Susan M Zbikowski, and Donald Fales. 2002. "Smoking cessation in a blue-collar population: Results from an evidence-based pilot program." *American journal of industrial medicine*, 42(5): 367–377. doi: 10.1002/ajim.10129.

Sacks, Naomi, Howard Cabral, Lewis E Kazis, Kelli M Jarrett, Delia Vetter, Russell Richmond, and

Thomas J Moore. 2009. "A web-based nutrition program reduces health care costs in employees with cardiac risk factors: before and after cost analysis." *Journal of Medical Internet Research*, 11(4). doi: 10.2196/jmir.1263.

Samad, Abu H, Mohd HBHJ Usul, Dalilah Zakaria, Raman Ismail, Anne Tasset-Tisseau, Florence Baron-Papillon, and Alain Follet. 2006. "Workplace vaccination against influenza in Malaysia: does the employer benefit?" *Journal of occupational health*, 48(1): 1–10.

Schneider, M, and HJ Häck. 2011. "Screening for colorectal cancer: a cost benefit analysis on a health prevention programme at the Boehringer Ingelheim Company." *Deutsche medizinische Wochenschrift (1946)*, 136(20): 1047–1052. doi: 10.1055/s-0031-1275840.

Schultz, Alyssa B, Chifung Lu, Tracey E Barnett, Louis Tze-ching Yen, Timothy McDonald, David Hirschland, and Dee W Edington. 2002. "Influence of participation in a worksite health-promotion program on disability days." *Journal of Occupational and Environmental Medicine*, 44(8): 776–780. PMID: 12185799.

Sciacca, John, Roger Seehafer, Roger Reed, and Dallas Mulvaney. 1993. "The impact of participation in health promotion on medical costs: a reconsideration of the Blue Cross and Blue Shield of Indiana study." *American Journal of Health Promotion*, 7(5): 374–395. PMID: 10148713.

Serxner, Seth A, Daniel B Gold, Jessica J Grossmeier, and David R Anderson. 2003. "The Relationship Between Health Promotion Program Participation and Medical Costs:: A Dose Response." *Journal of Occupational and Environmental Medicine*, 45(11): 1196–1200. doi: 10.1097/01.jom.0000095002.12772.6a.

Serxner, Seth, Daniel Gold, David Anderson, and David Williams. 2001. "The impact of a worksite health promotion program on short-term disability usage." *Journal of Occupational and Environmental Medicine*, 43(1): 25–29.

Shephard, Roy J, PAUL Corey, PETER Renzland, and MICHAEL Cox. 1982. "The influence of an employee fitness and lifestyle modification program upon medical care costs." *Canadian journal of public health= Revue canadienne de sante publique*, 73(4): 259–263.

Shi, Leiyu. 1993. "Health promotion, medical care use, and costs in a sample of worksite employees." *Evaluation Review*, 17(5): 475–487. doi: 10.1177/0193841X9301700501.

Soler, Robin E, Kimberly D Leeks, Sima Razi, David P Hopkins, Matt Griffith, Adam Aten, Sajal K Chattopadhyay, Susan C Smith, Nancy Habarta, Ron Z Goetzel, et al. 2010. "A systematic review of selected interventions for worksite health promotion: the assessment of health risks with feedback." *American journal of preventive medicine*, 38(2): S237–S262.

Stave, Gregg M, Lamont Muchmore, and Harold Gardner. 2003. "Quantifiable impact of the contract for health and wellness: health behaviors, health care costs, disability, and workers' compensation." *Journal of Occupational and Environmental Medicine*, 45(2): 109–117.

Taimela, Simo, Selina Justen, Pasi Aronen, Harri Sintonen, Esa Läärä, Antti Malmivaara, Jaakko Tiekso, and Timo Aro. 2008. "An occupational health intervention programme for workers at high risk for sickness absence. Cost effectiveness analysis based on a randomised controlled trial." *Occupational and environmental medicine*, 65(4): 242–248. doi: 10.1136/oem.2007.033167.

Wang, Philip S, Gregory E Simon, Jerry Avorn, Francisca Azocar, Evette J Ludman, Joyce McCulloch, Maria Z Petukhova, and Ronald C Kessler. 2007. "Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial." *Jama*, 298(12): 1401–1411. doi: 10.1001/jama.298.12.1401.

Wolf, Anne M, Mir S Siadaty, Jayne Q Crowther, Jerry L Nadler, Douglas L Wagner, Stephen L Cavalieri, Kurtis S Elward, and Viktor E Bovbjerg. 2009. "Impact of lifestyle intervention on lost productivity and disability: improving control with activity and nutrition (ICAN)." *Journal of occupational and environmental medicine/American College of Occupational and Environmental Medicine*, 51(2): 139. doi: 10.1097/JOM.0b013e3181965db5.

Wood, E Andrew, Gary W Olmstead, and James L Craig. 1989. "An evaluation of lifestyle risk factors and absenteeism after two years in a worksite health promotion program." *American Journal of Health Promotion*, 4(2): 128–133. doi: 10.4278/0890-1171-4.2.128.

C For Online Publication: Multiple Hypothesis Testing Methodology

Multiple hypotheses arise when there are multiple outcomes of interest, multiple subgroups of interest, multiple independent variables of interest, or some combination thereof. Consider testing $K > 1$ different null hypotheses. The family-wise error rate (FWER) is the probability of rejecting at least one true null hypothesis (i.e., a "false discovery") belonging to this "family" of K hypotheses. A procedure is said to provide *strong* control of the FWER if it does not depend on which of the K null hypotheses happen to be true.

We estimate the FWER using the free step-down resampling method of Westfall and Young (1993) (Algorithm 2.8, p. 66-67). The procedure consists of the following steps:[1]

1. Estimate $\{\widehat{\beta}_1, \widehat{\beta}_2, ..., \widehat{\beta}_K\}$. Estimate the conventional, unadjusted p-values $\{p_1, p_2, ..., p_K\}$ that correspond to separately testing each null hypothesis $\widehat{\beta}_k = 0$. Without loss of generality, assume the estimated p-values are indexed such that $p_1 \leq p_2 \leq ... \leq p_K$.

2. Draw with replacement from the dataset to create a bootstrap sample.

   (a) Estimate $\{\widehat{\beta}^*_{i1}, \widehat{\beta}^*_{i2}, ..., \widehat{\beta}^*_{iK}\}$. Estimate the conventional, unadjusted p-values $\{p^*_{i1}, p^*_{i2}, ..., p^*_{iK}\}$ that correspond to separately testing each null hypothesis $\widehat{\beta}^*_{ik} = \widehat{\beta}_k$. The k index here corresponds to the ranking computed in step 1. It will *not* generally be the case that $p^*_{i1} \leq p^*_{i2} \leq ... \leq p^*_{iK}$.

   (b) Enforce monotonicity with respect to the original ordering in step 1 by computing the successive minima:

[1] Our program was written in Stata and is easily applied to other settings. The module can be obtained by typing "`ssc install wyoung, replace`" at the Stata prompt, or downloaded directly from `ideas.repec.org/c/boc/bocode/s458440.html`.

$$q^*_{iK} = p^*_{iK}$$
$$q^*_{i,K-1} = \min(q^*_{iK}, p^*_{i,K-1})$$
$$q^*_{i,K-2} = \min(q^*_{i,K-1}, p^*_{i,K-2})$$
$$\vdots$$
$$q^*_{i1} = \min(q^*_{i2}, p^*_{i1})$$

3. Repeat step 2 N times. For each bootstrap sample i and hypothesis k, define the indicator $COUNT_{ik} = 1$ if $q^*_{ik} \leq p_k$ and 0 otherwise.[2]

4. For each hypothesis $k = 1, 2, ..., K$, calculate the fraction of successive minima that were lower than the original p-value:
$$r_k = \frac{1}{N}\sum_{i=1}^{N} COUNT_{ik}$$

5. Enforce monotonicity using successive maximization to calculate the adjusted p-value:
$$p^{adj}_1 = r_1$$
$$p^{adj}_2 = \max(r_1, r_2)$$
$$\vdots$$
$$p^{adj}_K = \max(r_{K-1}, r_K)$$

This resampling algorithm exhibits strong control of the FWER under subset pivotality, which is a multivariate generalization of pivotality.[3] This condition requires that the multivariate distribution of any subvector of p-values is unaffected by the truth or falsehood of hypotheses corresponding to p-values not included in the subvector. The condition is satisfied in many settings, including testing the significance of

[2]To compute "single-step" p-values instead of "step-down" p-values, define the indicator $COUNT_{ik} = 1$ if $\min\{p^*_{i1}, p^*_{i2}, ..., p^*_{iK}\} < p_k$ and 0 otherwise. Resampling-based single-step methods often control family-wise type 3 (sign) error rates. Whether their step-down counterparts also control type III error rates is unknown (Westfall and Young, 1993, p. 51).

[3]The sampling distribution of a pivotal statistic does not depend upon which distribution generated the data. The t-statistic is a common example.

coefficients in a general multivariate regression model with possibly non-normal or heteroskedastic errors (Westfall and Young, 1993, p. 122-123).

It is possible for this algorithm to produce adjusted p-values that are smaller than unadjusted p-values. For example, consider the extreme case where the number of bootstraps is equal to 1 (so that $N = 1$ in steps 3 and 4). Then all adjusted p-values are equal to either 0 or 1. The ones that are equal to 0 will of course be smaller than the unadjusted values. For this reason, we recommend employing a large number of bootstraps. (Westfall and Young (1993) recommend at least 10,000 bootstrap draws.) If adjusted p-values remain significantly smaller than the unadjusted p-values, even when the number of bootstraps is large, this may indicate model misspecification. For example, in simulations with clustered errors (described below), we found that adjusted p-values are frequently smaller than unadjusted values when we fail to employ a cluster bootstrap.

We ran four different sets of simulations to evaluate the effectiveness and statistical power of this resampling algorithm. Let μ be a 10-dimensional zero vector $(0, 0, ..., 0)'$. Let I be a 10×10 identity matrix. Let Σ be a 10×10 covariance matrix where all off-diagonal elements are equal to 0.9. The data generating process for each simulation scenario is described below:

1. Normal i.i.d. errors (10 outcomes)

   $e \sim \mathcal{N}(\mu, I)$

   $Y = e$

2. Normal i.i.d. errors (1 outcome, 10 subgroups)

   $e \sim \mathcal{N}(0, 1)$

   $Y = e$

3. Correlated errors (10 outcomes)

   $X \sim \mathcal{N}(\mu, I)$

   $e \sim \mathcal{N}(\mu, \Sigma)$

   $Y = 0.2X + e$

4. Lognormal, mean-zero i.i.d. errors (10 outcomes)[4]

   $e \sim \exp[\mathcal{N}(\mu, I)] - \sqrt{\exp[1]}$

   $Y = e$

We simulated 2,000 datasets for each of these four data generating processes. In each of these 2,000 simulations, we estimated a series of 10 regressions:

$$Y_i = \alpha + \beta_i X_i + \varepsilon_i, i = 1...10$$

The sample size for each regression was 100. The regressor $X_i \sim N(0, 1)$ in simulations 1, 2, and 3. In scenario 4, the regressor is just a constant equal to 1 (α is omitted). There are 10 null hypotheses that correspond to these 10 regressions: $\beta_i = 0, i = 1, ..., 10$. These 10 null hypotheses are all true in scenarios 1, 2, and 4, and all false in scenario 3 (correlated errors).

Table C.1 compares the effectiveness of the Westfall-Young resampling algorithm to other well-known multiple inference adjustment methods.[5] Each column in the table reports how often at least one null hypothesis was rejected using each adjustment method. When outcomes are independent and normally distributed, the probability that at least one of the 10 hypotheses is statistically significant is equal to $1 - (1 - .05)^{10} = 0.401$. This calculation accords well with the simulation: the first row of column (1) reports that at least one of the 10 hypotheses was rejected at $\alpha = 0.05$ in 39.8 percent of the 2,000 simulations when no adjustment was performed. By contrast, the Bonferroni-Holm, Sidak-Holm, and Westfall-Young adjustments reject at least one null hypothesis only about 4 percent of the time, thus achieving a family-wise error rate of less than 5 percent.

In column (2), the 10 hypotheses arise from examining multiple subgroups rather than multiple outcome variables. Failing to adjust the p-values again results in a high rejection rate of nearly 40 percent. The

[4] The mean of the standard lognormal distribution is $\sqrt{\exp[1]}$.

[5] The Bonferroni-Holm and Sidak-Holm (step-down) p-values are calculated as follows. Sort the K unadjusted p-values so that $p_1 \leq p_2 \leq ... \leq p_K$. The Bonferroni-Holm adjusted p-values are calculated as $\{p_1 K, \max[p_1, p_2(K-1)], ..., \max[p_{K-1}, p_K]\}$. The Sidak-Holm adjusted p-values are calculated as $\{1 - (1 - p_1)^K, \max[p_1, 1 - (1 - p_2)^{(K-1)}], ..., \max[p_{K-1}, p_K]\}$. If the calculation yields a value larger than 1, then the adjusted p-value is set equal to 1.

Bonferroni-Holm, Sidak-Holm, and Westfall-Young adjustment methods, however, all achieve rejection rates of around 5 percent.

The downside of the Bonferroni-Holm and Sidak-Holm adjustment methods is that they assume outcomes are independent, and therefore can be too conservative when outcomes are correlated. This is demonstrated in column (3), which reports rejection rates for a scenario where the 10 null hypotheses are all *false*. Here, the Bonferroni-Holm and Sidak-Holm methods reject at least one hypothesis only about 35 percent of the time. The Westfall-Young resampling algorithm, however, achieves a rejection rate in excess of 50 percent.

Although traditional adjustment methods such as Bonferroni-Holm and Sidak-Holm are generally thought to be conservative, Westfall and Young (1993) emphasize that these traditional methods can actually *over-reject* when the data-generating process is nonnormal. This is demonstrated in column (4): the resampling method of Westfall-Young achieves a family-wise error rate of under 6 percent, but the Bonferroni-Holm and Sidak-Holm methods reject at least one null hypothesis over 20 percent of the time.

Clustered standard errors

Westfall and Young (1993) do not discuss how to perform multiple inference in regression models where observations can be grouped into clusters, with model errors correlated within clusters. The presence of clustered errors does not violate subset pivotality, which is automatically satisfied in linear regression models. However, in this case it is important that the resampling in step 2 of the procedure be done over entire clusters, rather than individual observations. This is accomplished by specifying the **cluster()** option of the **wyoung** command.

We demonstrate the importance of resampling over clusters by performing another set of simulations. Again, let μ be a 10-dimensional zero vector $(0, 0, ..., 0)'$, and let I be a 10×10 identity matrix. The data generating process for this simulation scenario is:

5. Serially correlated errors (10 outcomes)

   $i = 1...100$ clusters

   $t = 1...10$ time periods

$\eta_i \sim \mathcal{N}(\mu, I)$

$e_{it} \sim \mathcal{N}(\mu, I)$

$Y_{it} = \eta_i + e_{it}$

We again simulated 2,000 datasets. In each simulation, we estimated the following 10 regressions:

$$Y_{it} = \alpha + \beta_i D_{it} + \varepsilon_{it}, i = 1...10$$

where the dummy variable $D_{it} = 1\{t > START_i\}$ and $START_i$ is a Poisson random variable with mean equal to 5. We estimated these regressions under two different assumptions about the standard errors (homoskedastic or clustered), and with and without a bootstrap cluster. Our results are reported in Table C.2.

Comparing column (2) to column (1) in the first row of Table C.2 shows that estimating the model using clustered standard errors results in a smaller family-wise error rate relative to a model that assumes errors are homoskedastic. Nevertheless, the rejection rate for the unadjusted value in column (2) still significantly exceeds five percent because this specification does not account for the number of hypotheses being tested.[6]

The second and third rows of Table C.2 show that the Bonferroni-Holm and Sidak-Holm corrections achieve a 5 percent rejection rate when the standard errors are clustered. This is unsurprising since the outcome variables in this simulation are independent.

The fourth row of Table C.2 demonstrates the importance of properly accounting for clustered standard errors when implementing the Westfall-Young correction. Column (2) shows that (erroneously) employing a simple bootstrap that resamples over individual observations causes the Westfall-Young correction to perform worse than even the unadjusted specification! However, column (3) shows that the Westfall-Young correction achieves a five percent rejection rate when the cluster bootstrap is employed.

[6]By construction, the values in columns (2) and (3) are identical in the first three rows, because these two columns vary only the bootstrapping methodology, which matters only for the Westfall-Young correction.

References

Westfall, Peter H, and S Stanley Young. 1993. *Resampling-based multiple testing: Examples and methods for p-value adjustment.* Vol. 279, John Wiley & Sons.

Table C.1: Family-wise rejection proportions at $\alpha = 0.05$

	(1)	(2)	(3)	(4)
Adjustment method	Normal errors	Multiple subgroups	Correlated errors	Lognormal errors
Unadjusted	0.398	0.387	0.685	0.577
Bonferroni-Holm	0.040	0.047	0.344	0.234
Sidak-Holm	0.040	0.051	0.347	0.237
Westfall-Young	0.041	0.045	0.513	0.058
Num. observations	100	100	100	100
Num. hypotheses	10	10	10	10
Hypotheses are true	Y	Y	N	Y

Notes: Table reports the fraction of 2,000 simulations where at least one null hypothesis in a family of 10 hypotheses was rejected. All hypotheses are true for the simulations reported in columns (1), (2), and (4), i.e., lower rejection rates are better. All hypotheses are false for the simulation reported in column (3), i.e., higher rejection rates are better. The Westfall-Young correction is performed using 1,000 bootstraps.

Table C.2: Family-wise rejection proportions at $\alpha = 0.05$, when the data generating process is serially correlated

	(1)	(2)	(3)
Unadjusted	0.652	0.401	0.401
Bonferroni-Holm	0.187	0.049	0.049
Sidak-Holm	0.188	0.049	0.049
Westfall-Young	0.191	0.498	0.046
Num. observations	1,000	1,000	1,000
Num. hypotheses	10	10	10
Model std. errors	Homoskedastic	Clustered	Clustered
Cluster bootstrap	N	N	Y

Notes: Table reports the fraction of 2,000 simulations where at least one null hypothesis in a family of 10 hypotheses was rejected. The difference between columns (1) and (2) is the assumption about the standard errors (homoskedastic or clustered). The difference between columns (2) and (3) is the method of bootstrapping (resampling over individual observations versus clusters), which matters only for the Westfall-Young correction. All null hypotheses are true, i.e., lower rejection rates are better. Each simulation generated 100 panels (clusters) with 10 time periods. The Westfall-Young correction is performed using 1,000 bootstraps.

D For Online Publication: Appendix for the Illinois Workplace Wellness Study

Table of Contents

D.1 Sample Selection and Study Overview

We designed and implemented a randomized controlled trial of an employee wellness program called iThrive at the University of Illinois at Urbana-Champaign. To participate in the study, university employees had to first digitally sign an informed consent form and complete an online baseline survey (described below). Employees who completed the baseline survey received a $30 Amazon.com gift card. Participants were subsequently randomly assigned either to a control group or to one of six different treatment groups. Treatment groups differed only in the amount of financial rewards that participants were offered: $0, $100, or $200 for completing a health screening and online health assessment, and $25 or $75 for each completed round of wellness activities. Treatment group participants were informed of their reward amounts at the time of their assignment.

Contact with members of the control group was minimized whenever possible. Participants in the control group were aware that they were participating in a study exploring "the link between wellness program incentives and program participation and health outcomes among employees", as stated in their informed consent form, but the details of the program and the size of the incentives for those in the treatment group was not revealed to them. Nevertheless, it is likely that many members of the control group were aware that others on campus were participating in wellness activities and receiving rewards for doing so.

The 2016-2017 iThrive wellness program had three main components:

1. Health screening (August 15 – September 16)
2. Online health assessment (September 8 – October 4)
3. Wellness activities
   (a) Fall 2016 (October 10 – December 16)
   (b) Spring 2017 (January 30 – April 25)

Steps 1 and 2 were mandatory. Participants who failed to complete them received no rewards and were not allowed to participate in subsequent wellness activities. Participants who successfully completed steps

1 and 2 were given the opportunity to participate in fall and spring wellness activities. Participation in fall activities was not required in order to participate in the spring activities.

The relationship between the different datasets employed in our study is illustrated in Appendix Figure D.1. Because most of the steps in the study were mandatory (e.g., taking the baseline survey, receiving a health screening), the datasets collected in later periods are generally only available for a strict subset of participants from previous periods. For example, health screening data are available for any participant who completed an online health assessment, but wellness activity data are not available for all participants who completed the online health assessment.

D.1.1 Online baseline survey (July 11 – July 31)

The University of Illinois provided us with a list of 12,486 active employees who met the following criteria as of June 10, 2016: (1) located at the Urbana-Champaign campus; and (2) eligible for part-time or full-time employee benefits from the Illinois Department of Central Management Services. This list included first and last names, mailing addresses, and email addresses. We dropped records that did not include a university email. We also dropped members of the research team, their family members, and other individuals heavily involved in the study. Following these exclusions, we were left with a total of 12,459 employees.

We mailed a postcard (see Appendix Figure D.2) on July 6, 2016 to each of these 12,459 employees informing them that they would receive an invitation to participate in an online survey for the Illinois Workplace Wellness Study. We included the UIUC-affiliated members of the research team in this mailing and confirmed that the postcards were delivered by July 9, 2016. The Provost of UIUC sent an email on the morning of July 11 to these employees indicating the university's support for the study (see Appendix Figure D.3).

An email invitation (see Appendix Figure D.4) containing the link to the online baseline survey was sent to each of the 12,459 employees on the morning of July 11, shortly after the email from the Provost. Reminder emails were sent on July 19, July 27, and August 1 to employees who had not yet completed the survey. The survey closed at noon on August 1, at which point 4,834 employees had successfully completed it. Participants who completed the survey immediately received a confirmation email (see Appendix Figure D.5). They also received an electronic \$30 Amazon.com gift card about one week after completing the survey (see Appendix Figure D.6).

D.1.2 Study randomization (August 1 – August 8)

We randomly assigned 3,300 of the 4,834 employees who completed the online baseline survey to one of six different iThrive treatment groups, denoted A25, A75, B25, B75, C25, and C75. Treatment groups differed only in the size of incentives offered for completing various steps of the iThrive program. Treated individuals in groups beginning with the letter A, B, or C were offered \$0, \$100, or \$200, respectively, for completing the health screening and online health assessment portions of the experiment. The second part of the treatment group name, 25 or 75, indicates the reward amount offered for each round (spring and/or fall) of wellness activities the individual completed.

For randomization, the sample was stratified by six baseline, demographic "strata" variables: (1) employee class (faculty, academic staff, or civil service); (2) sex (male or female); (3) age, as of the baseline survey launch date of July 11, 2016 (≤ 36, $37-49$, or ≥ 50); (4) above or below median annual salary; (5) quartile of annual salary; and (6) race (white or nonwhite). To create the strata, we sequentially split the sample in the order listed above for these strata variables. At each step in this sequence, we would only split a cell by the next strata variable if doing so resulted in cell sizes of at least 20. This ensured that, for every stratum, at least 2 employees could be assigned to the control and each of the 6 treatment groups (i.e., $20 \cdot p_{A,B,C} \cdot p_{25,75} > 2$, where $p_{A,B,C} \cdot p_{25,75}$ is the proportion of each stratum assigned to each treatment arm, as described below). This stratification process resulted in 69 strata, with the sample size per stratum ranging from 20 to 251.

Within each stratum, a proportion $p_{A,B,C} = 1100/4834 \approx 0.228$ of employees were randomly selected to be offered one of the three levels of incentive tied to completing the screening and health risk assessment (\$0, \$100, and \$200). This randomization was done such that exactly 1,100 employees in total would be assigned to each of these three levels of screening incentive. Next, within each stratum and screening incentive level, a proportion $p_{25,75} = 0.5$ of employees were randomly selected to be offered each of the two levels of activity incentive ($25 or $75). This resulted in six treatment groups with the following sample sizes: A25 ($N = 551$), A75 ($N = 549$), B25 ($N = 552$), B75 ($N = 548$), C25 ($N = 551$), and C75 ($N = 549$).

D.1.3 Health screening (August 15 – September 16)

We sent email invitations on August 9, 2016 to the 3,300 participants randomly selected to participate in iThrive. This email informed them of their selection and their monetary rewards for completing the different parts of the iThrive program, and explained how to sign up for a health screening (see Appendix Figure D.7). We also mailed postcards to these participants (Appendix Figure D.8) informing them of their selection. The postcards did not specify the monetary amounts and were delivered a few days after the initial email invitation. We sent reminder emails on August 12, August 23, and September 12 to participants who had not yet signed up for a health screening. Each of these participants was given login access to the iThrive website (see Appendix Figure D.9 and Appendix Figure D.10), which provided them with information about the iThrive program and reported on their progress throughout the year.

Health screenings were offered at 7 different locations on the UIUC campus, and also at Presence Covenant Medical Center, which is located about one mile away from the center of campus. A map displaying these locations is available in Appendix Figure D.11. Participants signed up for a date and time to receive their health screening using an online appointment scheduler (see Appendix Figure D.12).[1]

Appointments were available Monday through Saturday, from August 15th to September 16th, with the exception of Saturday, September 3 and Monday, September 5 (Labor Day). Appointment times were generally available from 6 AM until 10:50 AM. Only one campus location was available each day. The full schedule of appointment times and locations is displayed in Appendix Table D.1.

Participants who successfully signed up for an appointment received a confirmation email containing the date, time, and a link to a map of the location of their appointment. The online appointment scheduler sent participants an automated reminder email 24 hours prior to their appointment (see Appendix Figure D.13), and an automated text message if they had provided their cell phone number when making their

[1] A small number (<10) of participants showed up for a health screening without an appointment, but we were able to accommodate them.

appointment. We also sent participants a reminder email emphasizing that they should "not have anything to eat or drink (besides water) for 12 hours" before the health screening (see Appendix Figure D.14).

Upon showing up for their appointment, participants were asked to provide a form of identification, to sign a second informed consent form, and to complete a brief questionnaire (see Appendix Figure D.15) concerning their beliefs about their health status.[2] Participants then filled out the top half of a health screening form (Appendix Figure D.16) and were subsequently then directed to an open "station" where a clinician from Presence Covenant Medical Center measured their height, weight, waist circumference, and blood pressure. Next, they obtained blood chemistry measurements using the CardioChek Plus Analyzer, which is manufactured by PTS Diagnostics. This fingerstick measures cholesterol (total, HDL, and LDL), triglycerides, and glucose. All measures were recorded on the health screening form. At the end of the screening, a health coach reviewed the results with each individual participant in private. Depending on the measures, participants were sometimes recommended to make minor lifestyle changes or to seek medical attention. (See Appendix Figure D.17 for the guidelines employed by the health coach.) Recommendations were recorded on the health screening form. Upon departure, participants were given a carbon copy of their health screening form and a postcard reminding them to check their email for an invitation to take the online health assessment (Appendix Figure D.18). From start to finish, the entire health screening lasted on average for about 20 minutes.

D.1.4 Online health assessment (September 8 – October 4)

After completing their health screening, participants were invited over email to complete an online health assessment survey (Appendix Figure D.19). We sent reminder emails on September 21 and September 29 to participants who had not yet completed their online health assessment. After completing the survey, participants received a confirmation email from us within a few days.

The server hosting the survey became overloaded with requests on the first day of the survey (September 8), causing many participants to experience technical problems and to be unable to complete the survey.

[2]The ID was not a formal requirement, so in the small number of cases where participants did not have an ID, we allowed them to receive their health screening anyway. Fraud was not a concern because (1) participants had to make appointments online in their name prior to their arrival; and (2) all reward payments were made later in the study by direct deposit via University payroll.

This was fixed within 24 hours, although a small number of participants continued to report difficulties taking the health assessment throughout the survey period. Nevertheless, 97 percent of participants who completed the health screening managed to complete the online health assessment, so these technical glitches do not appear to have caused major difficulties for participants.

D.1.5 2016 fall wellness activities (October 10 – December 16)

We sent email invitations for the Fall 2016 wellness activities on September 27 to participants who had successfully completed their online health assessment (Appendix Figure D.20).[3] Participants were able to sign up for activities immediately, but no activities began before October 10. Signups were done via the iThrive website. Appendix Table D.2 lists the different activities that were available. Most classes were filled to capacity. Nearly 80 percent of people who registered were signed up for HealthTrails, which had unlimited capacity.

Out of 1,848 people eligible to participate, 1,306 people signed up for a wellness activity, and 903 people successfully completed them.

D.1.6 2017 spring wellness activities (January 30 – April 25)

We sent email invitations for the Spring 2017 wellness activities on January 17 to participants who had successfully completed their online health assessment (Appendix Figure D.21). Participants did not have to complete a fall activity to be eligible to participate in a spring activity. Participants were able to sign up for activities immediately, and activities began on January 25. Signups were done via the iThrive website. Appendix Table D.3 lists the different activities that were available. Most classes were filled to capacity. Over 75 percent of people who registered were signed up for Spring Into Motion, which had unlimited capacity. Out of 1,848 people eligible to participate, 1,059 people signed up for a wellness activity, and 740 people successfully completed them.

[3] We sent a separate invitation on October 3 to the small number of participants who completed their online health assessment after September 27.

D.1.7 2017 online follow-up survey (July 10 - August 9)

We mailed a postcard (see Appendix Figure D.22) on July 5, 2017 to 4,824 participants in our study.[4] We included the UIUC-affiliated members of the research team in this mailing and confirmed that the postcards were delivered by July 8, 2017.

We sent an email invitation (see Appendix Figure D.23) containing the link to the online follow-up survey to each of the 4,824 study participants on the morning of July 10. Reminder emails were sent on July 18, July 26, August 2, and August 7 to participants who had not yet completed the survey. The survey closed at 10:20 am on August 9, at which point 3,561 study participants (73.7 percent) had successfully completed it.[5] Participants who completed the survey immediately received a confirmation email. They also received an electronic $20 Amazon.com gift card about one week after completing the survey. The confirmation email and gift card were formatted similarly to the ones employed for the initial baseline survey (see Appendix Figures D.5 and D.6).

The August 2 reminder informed participants that ten people who completed the follow-up survey would be chosen at random to receive a $100 Amazon.com gift card (see Appendix Figure D.24). This new potential reward was in addition to the guaranteed $20 Amazon.com gift card. Participants who had already completed the survey prior to August 2 were included in this drawing.

D.1.8 2017 follow-up health screening (August 21 – September 22)

All study participants, including those in the control or treatment groups, were eligible to complete the one-year follow-up health screening in 2017. We randomly assigned these individuals to one of two groups, which differed only in the size of incentives ($0 or $125) offered for completing the follow-up survey.

Our method of randomization for the follow-up screening incentive combined explicit stratification plus re-randomization. Our follow-up strata were constructed by splitting the original strata by study arm. Because there were 69 original strata (see Section D.1.2) and 7 study arms (6 treatment groups plus a

[4]4,834 participants completed the 2016 baseline survey, but 10 had subsequently withdrawn from the study at the time of this invitation.

[5]The survey was accidentally reopened later that month for several weeks. Although all participants had been told that the survey would close on August 9, seven participants nevertheless completed the survey after the August 9 deadline, bringing the final number of completions up to 3,568.

control group), this resulted in $483 = 69 \times 7$ follow-up strata, with the sample size per follow-up stratum ranging from 2 to 80.

To implement the stratified re-randomization, we generated multiple potential follow-up treatment assignments T_j as follows:

1. Draw a random integer s_j and set the random-number seed to equal s_j.

2. Randomly sort all 4,834 original study participants first by follow-up strata, then within each follow-up strata. Drop the individuals ($N = 15$) who had withdrawn from the study at the time of randomization (August 4, 2017), leaving a sample of $N = 4,819$ employees to be randomized.

3. Assign alternating observations to the \$0 and \$125 follow-up screening incentive group, and let T_j denote the resulting vector of treatment assignments for each employee.

4. Test for balance between the \$0 and \$125 groups for 60 variables (pre-determined at the time of follow-up randomization) grouped into the following 8 families:

   (a) Baseline strata (6 variables).

   (b) Baseline survey (21 variables).

   (c) Salary and age (3 variables).

   (d) Employment (7 variables).

   (e) Health behavior (6 variables).

   (f) Medical spending and coverage (8 variables).

   (g) Sick days taken (2 variables).

   (h) Registration for or completion of 2016 biometric screening, HRA, or Fall 2016 or Spring 2017 wellness activities (7 variables).

   We performed joint tests for balance by family of outcomes (8 balance tests), plus individual tests for balance for each of the medical spending outcomes, with and without coverage weights for average spending outcomes (10 balance tests). In total, we performed 18 tests for balance, and we denote by p_j^{min} the minimum p-value across these tests.

After performing these steps for $j = 1$ to 10,000, we selected the treatment assignment that maximized the p-value p_j^{min} from the balance tests. Specifically, the selected treatment assignment was chosen to be T_{j^*}, where $j^* = \arg\max_j p_j^{min}$.

In total, 2,409 employees were assigned to the $0 follow-up screening incentive, while 2,410 employees were assigned to the $125 follow-up screening incentive. We sent email invitations on August 14, 2017 to these employees ($N = 4,819$) informing them of their monetary reward for completing the 2017 health screening, and explained how to sign up for it (see Appendix Figure D.25). We sent reminder emails on August 23, September 5, September 13, September 19, and September 21 to participants who had not yet signed up for a health screening.[6] The final reminder encouraged participants to walk in for a health screening even if they did not have an appointment (see Appendix Figure D.26).

The iThrive website was updated on August 14, 2017 so that treatment group participants could obtain information about the 2017 follow-up health screening and their potential rewards. For the first time, control group members were also given login access to the iThrive website. Everyone was encouraged to visit the website in the August 14 screening invitation email (Appendix Figure D.25). For control group members, the website only displayed information about the health screenings (see Appendix Figure D.27). For treatment group members, the website displayed information about the subsequent health assessment and wellness activities once the treatment group member completed a screening (see Appendix Figure D.27).

Health screenings were held in the same locations as in 2016, with the exception of the Physical Plant Services Building, which was unavailable for reservation. Unlike in 2016, people were allowed to make appointments all the way until 3:50 PM. The full schedule of appointment times and locations is reported in Appendix Table D.4.

The health screening procedure was nearly identical to the procedure employed in 2016 (see Section D.1.3 for a full description). There were only two substantive differences. First, participants were not

[6] Study participants who signed up for a screening, but later failed to show up for their appointment, were included in these reminder emails.

handed a postcard at the end of the screening reminding them to check their email for an invitation to take the online health assessment. This step was omitted in 2017 because follow-up screening participants in that year included employees from the control group, who were not eligible to take the 2017 online health assessment. Second, health screening confirmation emails were sent only to participants who had been assigned a $125 reward (see Appendix Figure D.29). Screening participants in both the control and treatment groups who were assigned a $0 reward did not receive a screening confirmation email. However, all participants could confirm their completion status by visiting the iThrive website.

D.2 Datasets

D.2.1 University administrative data

The University of Illinois provided us with an initial list of 12,459 employees who met the following criteria as of June 10, 2016: (1) located at the Urbana-Champaign campus; and (2) eligible for part-time or full-time employee benefits from the Illinois Department of Central Management Services. The university administrative datasets described below are available for all 12,459 of these employees.

Demographics

This dataset includes first and last names, mailing address, email address, exact date of birth, sex, annual salary, race (white, black, or other), employee class (faculty, academic staff, or civil service), home college (49 colleges), home organization (323 organizations), and exact hire date.

Employment history

This dataset includes employment history information up through August 15, 2017. It includes the exact hire date for all employees. Out of the initial sample of 12,459 employees, 1,537 of these employees were no longer actively employed by the university as of August 15, 2017. For these former employees, the dataset includes the exact date of employment termination and the associated reason (resigned, retired, deceased, terminated, contract ended, or other). For active employees, the dataset lists their annual salary as of

August 15, 2017.[7]

Sick leave

This dataset includes the number of sick days taken by a Civil Service employee at the monthly level, for the time period January 2015 through May 2017. For non-Civil Service employees (i.e., Academic Staff and Faculty), the dataset includes the total number of sick days taken during the two time periods August 16, 2015 through August 15, 2016, and August 16, 2016 through May 15, 2017. Sick leave for faculty (25 percent of our sample) is self-reported and exhibits little variation: more than 75 percent of the faculty in our sample reported 0 days of sick leave during the August 16, 2015 through August 15, 2016 academic year.

The vast majority of employee sick leave is noncompensable, i.e., it cannot be "cashed out" when the employee terminates employment.[8] Civil Service employees accrue sick leave at the rate of 0.0462 hours for each hour worked, which corresponds to approximately 12 days per year for a full-time employee, and this sick leave is cumulative (i.e., rolls over from one year to the next). Full-time Academic Staff and Faculty earn 12 cumulative and 13 non-cumulative sick leave days per year, and their total sick leave is recorded in the data only twice a year: on May 16 and on August 16.

Gym attendance

This dataset includes a list of the exact dates that each employee visited one of the university's campus recreational facilities during the time period January 1, 2015 through July 31, 2017. There are three recreational facilities located on the university campus: the Activities and Recreation Center (ARC), the Campus Recreation Center East (CRCE), and the Ice Arena. Membership costs $40 per month for university employees and retirees. Entering these facilities requires swiping a university identification card through a machine, which is the basis for the observations in this dataset.

[7]Civil Service, Academic Staff, and Faculty received a mid-year salary increase in the second half of February, 2017. The salary increase was explicitly merit-based, and the total salary pool was capped at 2 percent of aggregate base salaries.

[8]Prior to 1999, employees could accrue compensable sick leave. A few older employees still have positive compensable sick leave balances, but this is very rare.

D.2.2 Illinois Marathon data

The Illinois Marathon is a running event held annually in Champaign, Illinois. The races offered include a marathon, a half marathon, a 5K, and a 10K. When registering for a race, a participant must provide her name, age, sex, and hometown. That information, along with the results of the race, are published online after the races have concluded.[9]

We downloaded Illinois Marathon data for the 2014-2017 races and matched it to individuals in our study data using full name, age, sex, and hometown. An individual in our study data was counted as participating in a running event in a given year if either (a) University and Illinois Marathon records matched on full name, age (+/- 1 year), and sex; or (b) University and Illinois Marathon records matched on the first two letters of last name, age (+/- 1 year), sex, and hometown. Among University employees that match to Illinois Marathon records using *either* match measure, *both* measures generate a match in 73.7, 74.6, 84.4, and 79.6 percent of cases for the years 2014, 2015, 2016, and 2017, respectively.

D.2.3 Health insurance claims data

We obtained health insurance claims data for 8,326 university employees (anonymized for non-study participants) who were listed in our university administrative dataset and who were members of Health Alliance at any point during the period January 1, 2015 through July 31, 2017. (Note: 8,095 employees were members during the pre-period July 1, 2015 through July 31, 2016.) The dataset includes all inpatient, outpatient, and prescription drug claims with a date of service between January 1, 2015 through June 30, 2017. Each claim lists a date of service, a physician specialty code, a place of service code, and the total allowed amount, which is the sum of payments to the provider from both the insurer and the beneficiary. Health Alliance also provided an enrollment file listing start and end dates for each member.

Health Alliance, the university's most popular insurer, operates an HMO plan with a $0 medical deductible and a $100 annual pharmacy deductible. Physician visits require a $20 copay, and the plan's out-of-pocket maximum is $3,000 for the individual and $6,000 for the family.

The university offers seven different health insurance plans. One of these, Quality Care Health Plan,

[9] See http://illinoismarathon.com/resultscertificatesphotos/#results.

is a traditional indemnity insurance plan.[10] The rest are managed care plans, including four Health Maintenance Organizations (BlueAdvantage HMO, Coventry HMO, Health Alliance HMO, and HMO Illinois) and two Open Access Plans (Coventry OAP and HealthLink OAP). Beginning July 1, 2017, Coventry HMO and Coventry OAP were renamed Aetna HMO and Aetna OAP.

Employee contributions are the same for all HMO plans, and depend on income. For the 2016-2017 plan year, an employee's monthly contribution for an HMO plan ranged from $68 per month (annual salary $30,200 and below) up to $186 per month (annual salary $100,001 and above). Contributions for an employee enrolled in Quality Care Health Plan ranged from $93 per month (annual salary $30,200 and below) up to $211 per month (annual salary $100,001 and above). The seven health plans charge different contributions for dependents, with dependent contributions ranging from $96 per month (BlueAdvantage HMO) to $249 per month (Quality Care Health Plan).

D.2.4 Online survey data

2016 baseline survey

The baseline survey was administered online using survey software provided by SurveyGizmo. An email invitation containing the link to the online baseline survey was sent to 12,459 university employees. Each link was unique and pointed to a survey that could only be completed once. Survey participants navigated the survey by clicking on buttons labeled "Next" and "Back". They were allowed to skip questions and to change their answers on previous pages if so desired. In order to receive their $30 Amazon.com gift card, participants had to navigate to the end of the survey and click the "Submit" button. The software did not allow them to change their answers once the survey was submitted. Participants who exited the survey prior to completion could continue from where they left off by clicking on their invitation link again.

The software recorded that 7,468 employees clicked on the link to the survey, 4,918 employees began the survey, and 4,834 employees successfully completed the survey. Among those who completed the survey within an hour of clicking on the survey link for the first time, the average completion time was 15 minutes.

In order to assess the reliability of the survey, we compared participants' self-reported ages from the

[10] This plan was administered by Cigna up through June 30, 2017. Aetna has administered it since July 1, 2017.

survey with the ages available in the university's administrative data. Of the 4,830 participants who reported an age, only 24 (<0.5%) reported a value that differed from the university's data by more than one year.

2017 follow-up survey

The 2017 follow-up survey was administered online using survey software provided by SurveyGizmo. An email invitation containing the link to the follow-up survey was sent to 4,824 study participants.[11] The format of the invitation email and the survey were similar to the 2016 baseline survey. In order to receive their $20 Amazon.com gift card, participants had to navigate to the end of the survey and click the "Submit" button.

The software recorded that 3,642 employees clicked on the link to the survey, 3,611 employees began the survey, and 3,568 employees successfully completed the survey. Among those who completed the survey within an hour of clicking on the survey link for the first time, the average completion time was 13.3 minutes. The completion rates for the control and treatment groups were 75.4 and 73.1 percent, respectively. The difference in completion rates is marginally significant ($p = 0.079$).

In order to assess the reliability of the survey, we compared participants' self-reported ages from the survey with the ages available in the university's administrative data. Of the 3,561 participants who reported an age, only 20 (<0.006%) reported a value that differed from the university's data by more than one year.

D.2.5 Health screening data

Fall 2016 health screening

2,047 participants signed up for a health screening, and 1,900 were successfully screened. The top of each participant's screening form (see Appendix Figure D.16) contains the participant's answers to the following questions:

[11] 4,834 participants completed the 2016 baseline survey, but 10 had subsequently withdrawn from the study at the time of this invitation.

1. “Do you use tobacco of any form?”
2. “In the average week, how many times do you engage in physical activity?”
3. “If you engage in physical activity, for how long?”
4. “How often do you feel tense, anxious, or depressed?”
5. “Do you have a primary physician?”
6. “Did you fast today?”

The following biometric data were recorded on every form: height; weight; waist circumference; body mass index; systolic blood pressure; diastolic blood pressure; total cholesterol; total cholesterol ratio; HDL; LDL; triglycerides; and glucose. Finally, the form also records which (if any) of the following actions were taken by the health coach (see also Appendix Figure D.17) as a result of the patient's biometric readings:

1. Referred patient to a primary care physician
2. Advised patient to make minor lifestyle changes
3. Communicated to patient that one or more results were out of the normal range
4. Communicated to patient that the results require a medical referral
5. Communicated to patient that the results require immediate medical attention

In order to ensure accuracy, all of the data on every form was read and entered into a database twice, by two different research assistants. Any disagreements between the two entries were resolved by reexamining the original form.

D.2.6 Health questionnaire data

Fall 2016 health questionnaire

Participants were required to fill out a health questionnaire prior to receiving their health screening, so every participant who was screened (1,900 in total) is also represented in this dataset. A copy of the questionnaire is displayed in Figure D.15. As with the health screening data, these data were digitized twice in order to ensure accuracy.

D.2.7 Online health assessment and wellness activities data

Fall 2016/ Spring 2017 online health assessment and wellness activities

Out of the 1,900 participants who completed a health screening, 1,848 completed an online health assessment. These 1,848 participants constitute the set of study participants who were eligible to sign up for wellness activities in the fall and in the spring. Participants were not required to sign up for a fall activity in order to sign up for a spring activity. Out of the 1,848 people eligible to participate, 1,306 people signed up for a fall wellness activity (903 completed it) and 1,059 people signed up for a spring wellness activity (740 completed it).

The online health risk assessment (HRA) data contain the exact start dates and times that participants began their HRA, and the exact end dates and times they completed it. The wellness activity data include indicator variables for whether the participant signed up for a wellness activity, and for whether the participant completed that activity. If the participant signed up for an activity, the name of the activity was also recorded. (See Appendix Tables D.2 and D.3 for names and descriptions of the activities that were offered.) The wellness activities data also include information on how much of the activity was completed by the participant, along with the minimum threshold required to qualify for the wellness activity reward.[12]

[12]For example, the Spring 2017 "Lunchtime Walk" activity met on 8 separate occasions, and participants were required to participate in at least 6 of the walks in order to qualify for their reward. The wellness activities data contains a variable specifying how many walks each participant attended.

D.3 Online Appendix Figures

Figure D.1: Overlap among datasets

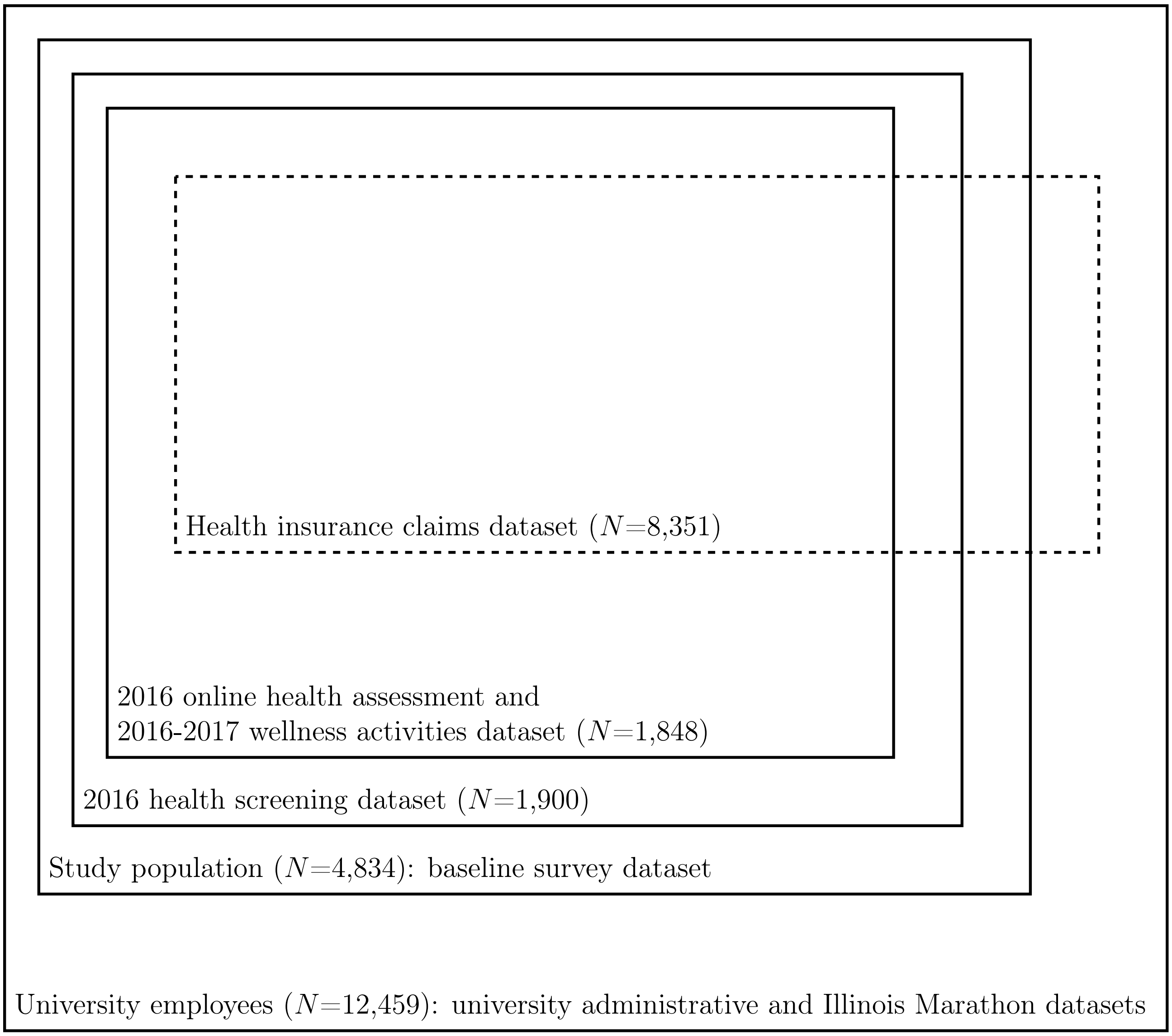
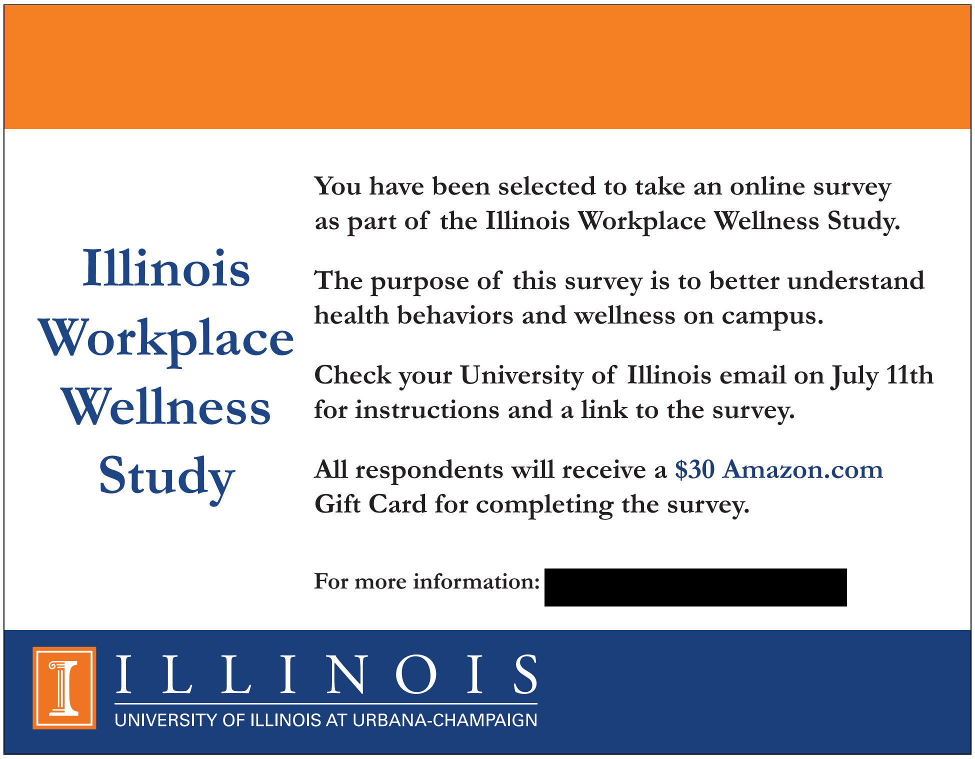


Figure D.2: Front and back sides of invitation postcard sent on July 6, 2016



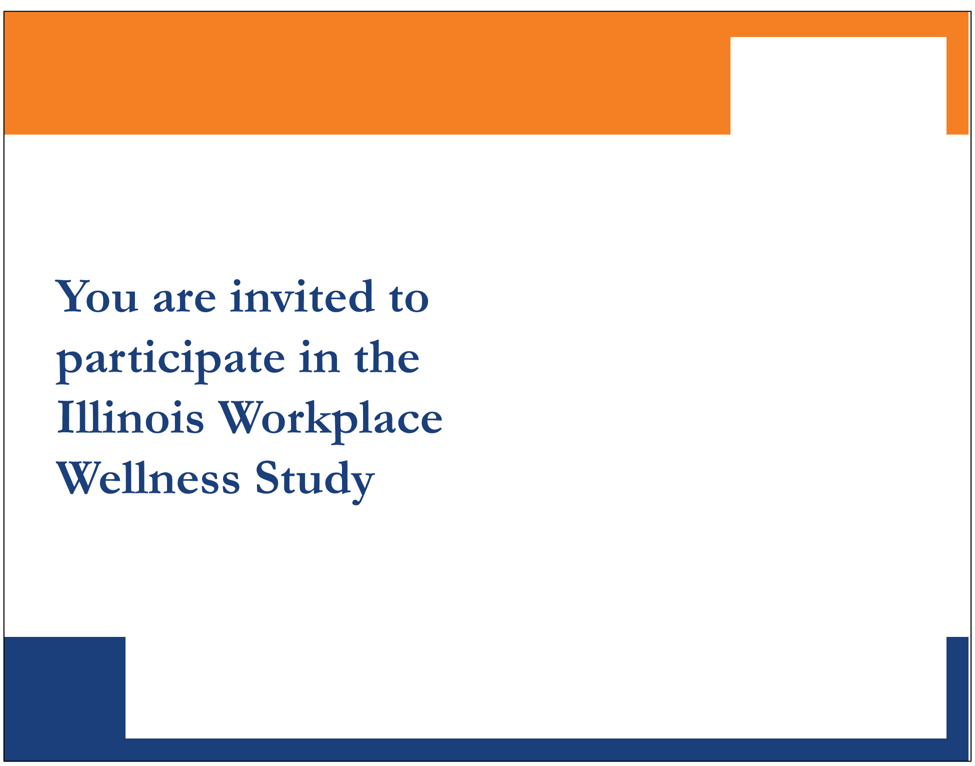


Figure D.3: Email sent from the UIUC Provost to university employees on July 11, 2016

11/16/2016 Illinois Workplace Wellness Study

Click here to see this online

OFFICE OF THE PROVOST

ACADEMIC AFFAIRS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Dear Faculty and Staff,

The Chancellor and I are pleased to announce our support of an initiative to better understand how to promote employee wellness. A research team on our campus is conducting an evaluation of worksite wellness programs over the next several months. You will soon receive an email from [redacted] asking you to participate in a brief survey. Following your participation in the survey, some of you will have the opportunity to engage in the second part of the study.

Your feedback is very important to the success of the project, and taking the survey is easy. **All data collected in this study will remain confidential**. Your individual data will never be shared with the university or your health insurer.

Please take a few minutes to complete this survey when you receive the invitation email. All respondents who complete the survey will receive a **$30 Amazon.com gift card**. Accepting this gift card is permitted under the State Officials and Employees Ethics Act.

For non-exempt civil service employees, this program is an "approved event," so that, operations permitting and with prior supervisory approval, participation that occurs during an employee's regular work schedule does not require the charging of leave benefit time.

Best regards,

Edward Feser
Interim Vice Chancellor for Academic Affairs and Provost

Office of the Provost

[redacted]

Notes: Email also available at http://illinois.edu/emailer/newsletter/100150.html.

Figure D.4: Invitation email sent to university employees on July 11, 2016

Illinois Workplace Wellness Study Invitation

From: **Illinois Workplace Wellness Study** [redacted]
Date: Monday, July 11, 2016 10:34 AM
Subject: Illinois Workplace Wellness Study Invitation
To:

Dear Colleagues,

We invite you to take part in a research study of workplace wellness programs. This study is funded by the National Institutes of Health and will help inform national health policy regarding the costs and benefits of wellness programs.

The first part of the study consists of an online survey about health behaviors and wellness on campus. The survey will take approximately 15 minutes to complete. We know that your time is valuable, so we are offering a **$30 Amazon.com gift card** to all respondents who complete the survey.

The survey is only available for a limited time, so please complete the survey promptly in order to receive your $30 gift card. To access the online survey, simply click the following URL or paste it in your browser:

http://surveys.citl.illinois.edu/go/Wellnessjx421

This survey is strictly confidential. **Your individual data will never be shared with the university or your health insurer.** Some of you who take the survey will be offered an opportunity to participate in a second part of the research study.

For non-exempt civil service employees, this program is an “approved event,” so that, operations permitting and with prior supervisory approval, participation that occurs during an employee’s regular work schedule does not require the charging of leave benefit time.

Thank you for contributing to this important research project! If you have any questions or need assistance, please contact us at [redacted]

Best regards,

Illinois Workplace Wellness Study Team

David Molitor
Assistant Professor, Department of Finance

Laura Payne
Associate Professor, Department of Recreation, Sport and Tourism

Julian Reif
Assistant Professor, Department of Finance and IGPA

Figure D.5: Text of the confirmation email sent to study participants who successfully completed the online baseline survey

From: [redacted]
Subject: Survey Confirmation: Illinois Workplace Wellness Study

Dear **[First name]**,

Congratulations! This email is confirmation that you have completed the online survey for the Illinois Workplace Wellness Study. You will soon receive an email containing your $30 Amazon.com gift card. Please allow up to one week for the gift card to be processed.

You may be selected to participate in the second part of the study. If so, we will email you within the next month.

If you have any questions or need assistance, please contact us at [redacted] or [redacted]

Regards,

Illinois Workplace Wellness Study Team

Notes: The text highlighted in yellow was appropriately customized for each participant.

Figure D.6: Electronic Amazon.com gift card sent to participants who completed the baseline survey

You've received a $30.00 Amazon.com Gift Card!

amazon.com gift card

$30.00

Claim Code:
XXXXXXXXXXXX

Redeem now

Dear [NAME],

Thank you for taking our online survey!

Regards,
Illinois Workplace Wellness Study

Redeeming your Amazon.com Gift Card

1. Visit www.Amazon.com/redeemgift.
2. Enter the Claim Code and click **Apply to Your Account**.

Gift card funds are applied automatically to eligible orders during the checkout process. Your Claim Code may also be entered during checkout. To redeem your gift card using the Amazon.com 1-Click® service, first add the gift card funds to **Your Account**.

Amazon.com Gift Cards ("GCs") may be used only for the purchase of eligible goods on Amazon.com or certain of its affiliated websites. Except as required by law, GCs cannot be transferred for value or redeemed for cash. Purchases are deducted from the GC balance. To redeem or view a GC balance, visit "Your Account" on Amazon.com. Amazon.com is not responsible if a GC is lost, stolen, destroyed or used without permission. For complete terms and conditions, see www.amazon.com/gc-legal. GCs are issued by ACI Gift Cards LLC., a Washington corporation. All Amazon ®, ™ & © are IP of Amazon.com, Inc. or its affiliates. No expiration date or service fees.

Serial Number: XXXXXXXXXXXXXXXX

Order Number: XXXXXXXXXXXXXXXXX

Notes: The text highlighted in yellow was appropriately customized for each participant.

Figure D.7: Text of invitation email sent to participants in treatment group C75 ($350 incentive) on August 9, 2016

From: [redacted]
Subject: Illinois Workplace Wellness Study: iThrive Invitation

Dear [First Name]:

Last month, you completed a health survey as part of the Illinois Workplace Wellness Study. You have been selected to participate in the second part of this research study: iThrive, a program to promote health and wellness among campus faculty and staff.

iThrive offers you the opportunity to participate in valuable health screening and wellness activities at no cost to you. In addition, **you can earn up to $350 in financial rewards**, as described below.

The opportunity to participate in iThrive is only available for a limited time. To learn more about how to get started and earn rewards, visit the iThrive website:

iThrive.illinois.edu

The iThrive website provides personalized information on your progress, links for signing up for iThrive opportunities, answers to frequently asked questions (FAQs), and a summary of your rewards. To help you get started, you will receive an invitation later today from Presence Health, in order to schedule your health screening.

The iThrive program is summarized below.

How iThrive Works

iThrive begins with a health screening and health assessment survey. Once you complete the screening and health assessment, you are eligible to enroll in wellness activities in Fall 2016 and again in Spring 2017.

Step 1: Health Screening + Health Assessment Survey ($200 reward)

The health screening is your gateway to iThrive. The purpose of a health screening is to measure physical health characteristics (e.g., height, weight, blood pressure, cholesterol) and use the information as a benchmark for health promotion and management. For your convenience, Presence Health will offer these screenings at various dates and locations across campus.

After completing the health screening, you will receive an invitation to complete an online health assessment survey. The health assessment will provide you with a detailed health summary and evaluation of health risks. Upon completion of the health screening and health assessment survey, you will receive a reward of $200.

Step 2: Wellness Activities (up to $150 reward)

After completing your health screening and health assessment survey, you will have the opportunity to participate in a wellness activity that aligns with an area of your health that you would like to improve. These areas include physical activity, weight management, stress management, chronic disease self-management, and tobacco cessation. You will have the option to participate in programs that meet in person, or you may choose to participate in one of our online, self-paced programs.

These activities will be offered in Fall 2016 and again in Spring 2017. Completing your chosen wellness activity in Fall 2016 will entitle you to a $75 reward. Completing an activity in Spring 2017 will entitle you to another $75 reward, for a total possible reward of $150 for wellness activities. If you do not complete an activity in Fall 2016, you are still eligible to participate in Spring 2017 and receive a $75 reward.

Enrolling in iThrive

You enroll in iThrive by scheduling your health screening. When scheduling your health screening, please use the email address to which this email was sent [redacted]. This email address will be referred to as your "iThrive contact email". You will receive an email from Presence Health today with a link to the online scheduler. You can also access the online scheduler now by copying and pasting the following URL into your browser:

ithrive.acuityscheduling.com

You may also visit the iThrive website at any time: iThrive.illinois.edu. This website will provide personalized information on your progress, links for signing up for iThrive opportunities, and a summary of your rewards.

For non-exempt civil service employees, iThrive is an "approved event," so that, operations permitting and with prior supervisory approval, participation that occurs during an employee's regular work schedule does not require the charging of leave benefit time.

As with every part of the Illinois Workplace Wellness Study, your participation in iThrive is completely voluntary and your **individual data will never be shared with your health insurance provider or your employer**. You can read here about the purpose of our study as well as the steps we will take to keep your information confidential. If you have any questions or need assistance, please contact us at [redacted]

Yours in good health,

Illinois Workplace Wellness Study Team

Notes: The text highlighted in yellow was appropriately customized for each participant.

Figure D.8: Front and back sides of the postcard mailed to participants selected to participate in iThrive, week of September 8, 2016

Figure D.9: Login page for the iThrive website

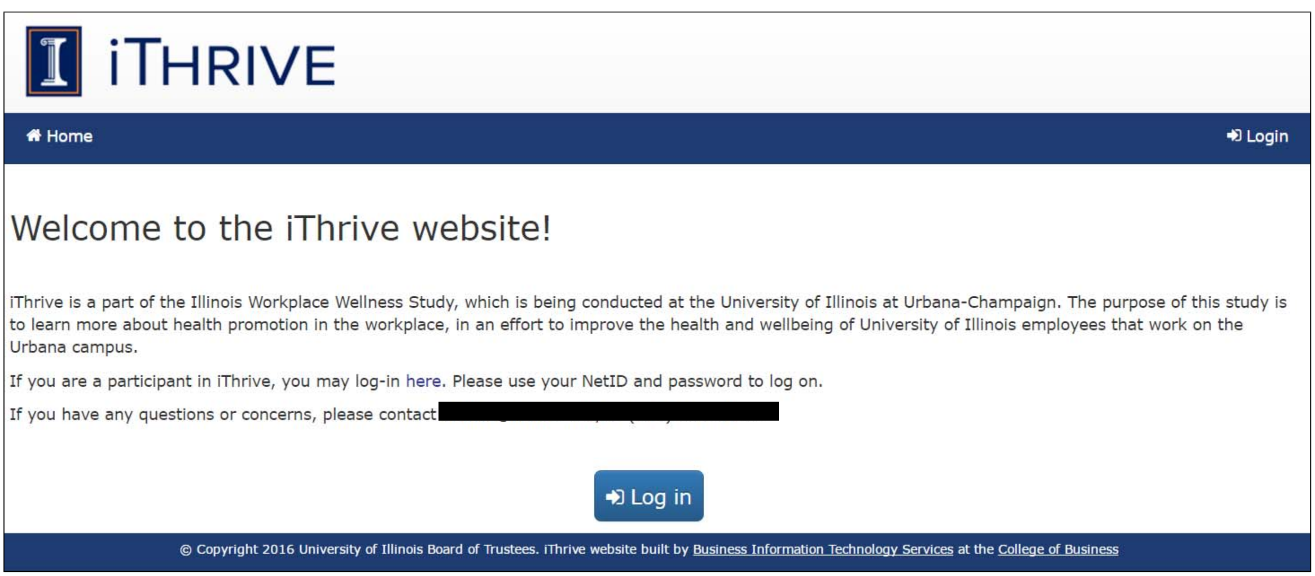
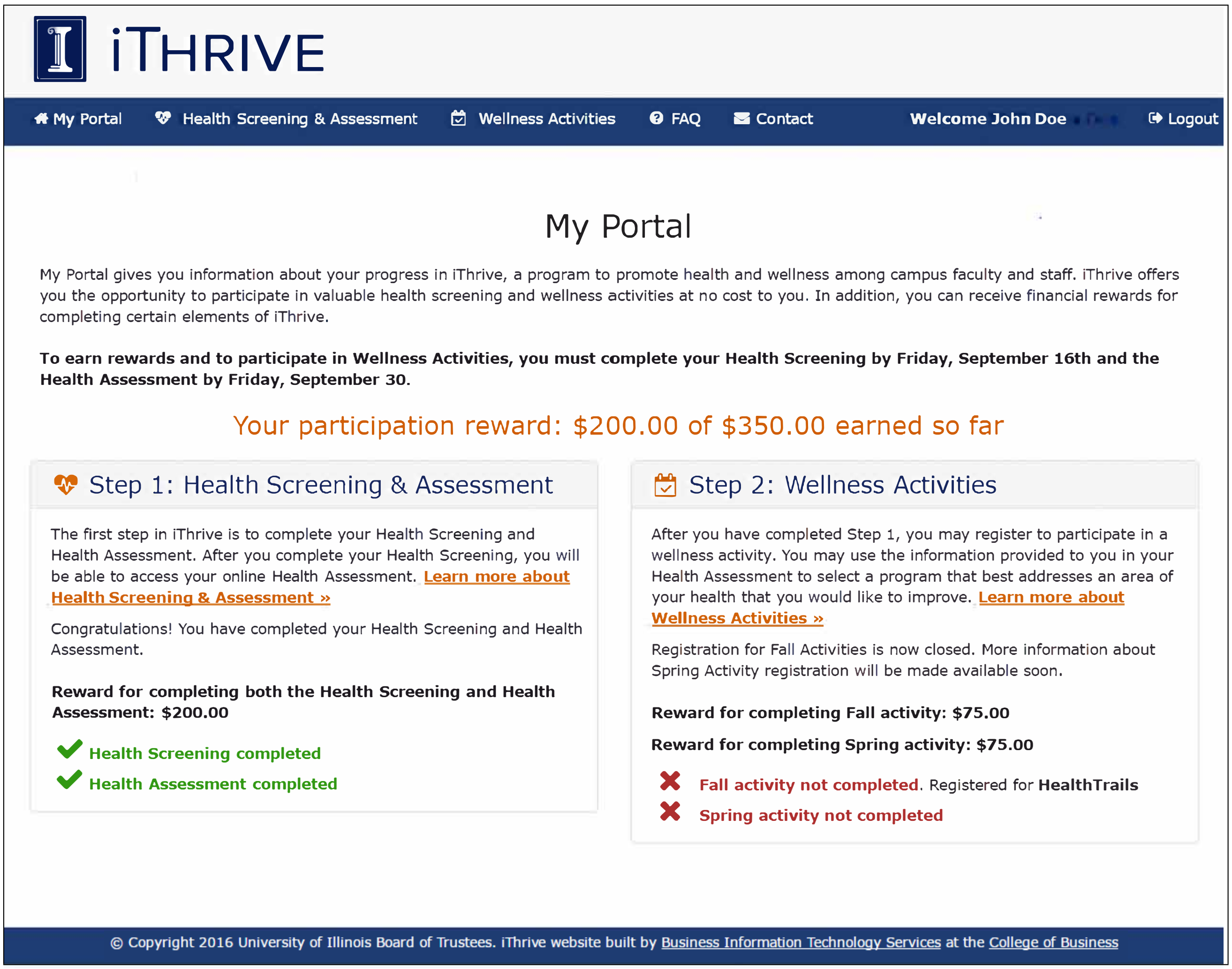
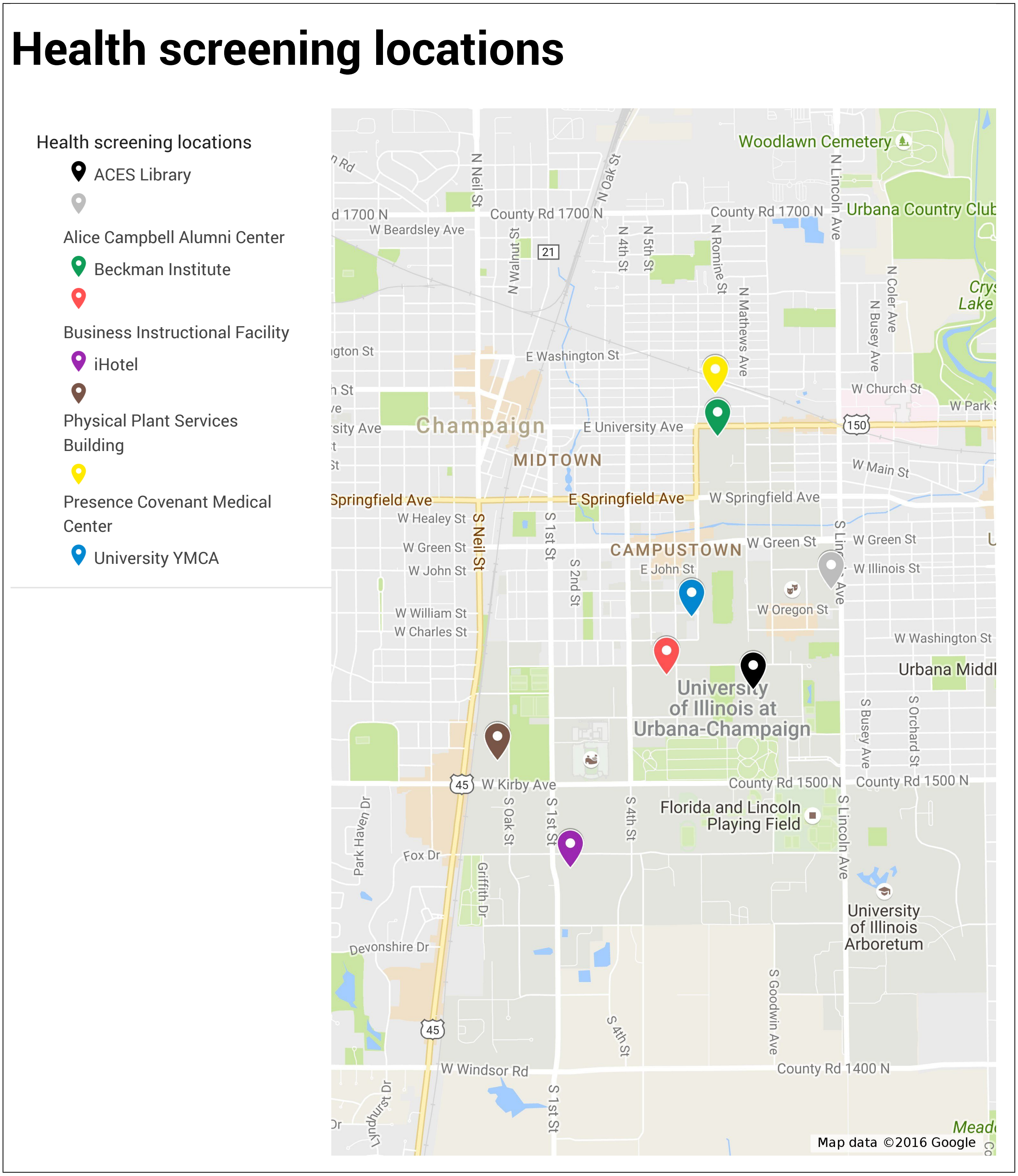


Figure D.10: Main home page for the iThrive website



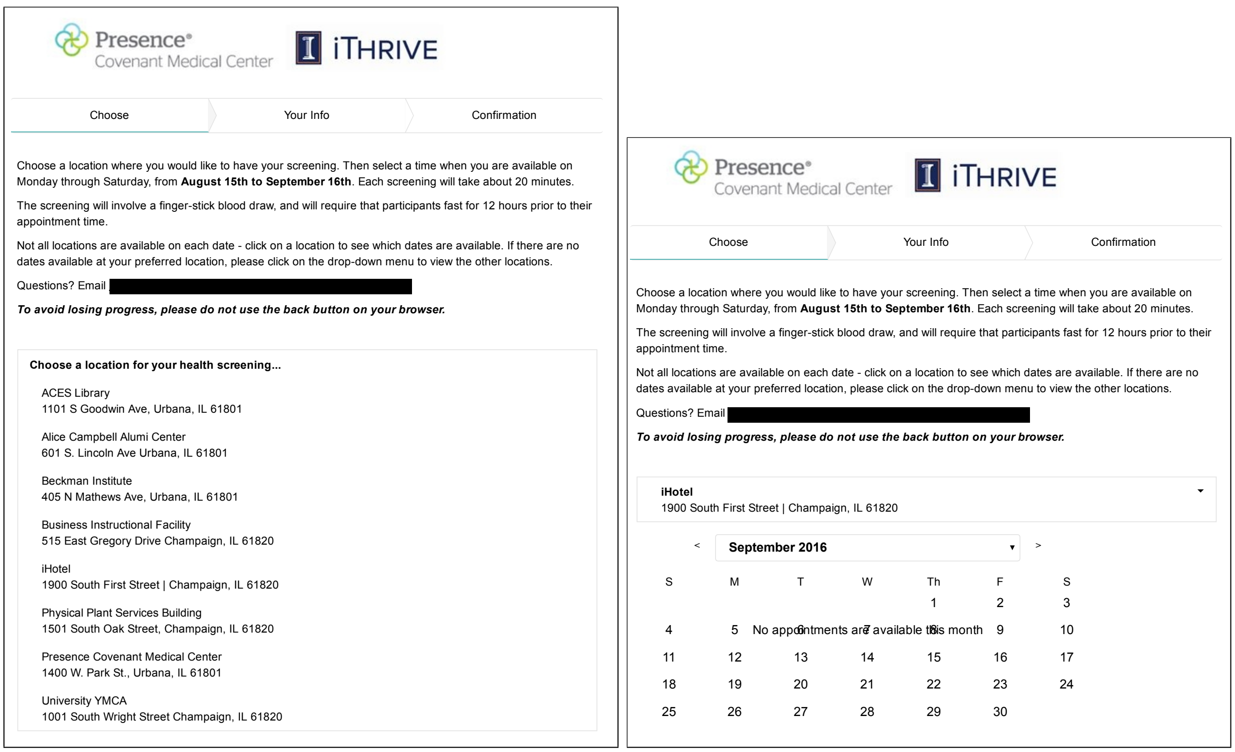
Notes: This participant was randomly assigned to treatment group C75, and thus is eligible for a total of $\$200 + 2 \times \$75 = 350$ in rewards.

Figure D.11: Screening locations



Notes: This map displays the locations of the 8 different places where health screenings were held.

Figure D.12: First and second pages of the online appointment application used to sign up for a health screening


Presence® Covenant Medical Center iTHRIVE

Choose | Your Info | Confirmation

Choose a location where you would like to have your screening. Then select a time when you are available on Monday through Saturday, from **August 15th to September 16th**. Each screening will take about 20 minutes.

The screening will involve a finger-stick blood draw, and will require that participants fast for 12 hours prior to their appointment time.

Not all locations are available on each date - click on a location to see which dates are available. If there are no dates available at your preferred location, please click on the drop-down menu to view the other locations.

Questions? Email [redacted]

To avoid losing progress, please do not use the back button on your browser.

Choose a location for your health screening...

ACES Library
1101 S Goodwin Ave, Urbana, IL 61801

Alice Campbell Alumi Center
601 S. Lincoln Ave Urbana, IL 61801

Beckman Institute
405 N Mathews Ave, Urbana, IL 61801

Business Instructional Facility
515 East Gregory Drive Champaign, IL 61820

iHotel
1900 South First Street | Champaign, IL 61820

Physical Plant Services Building
1501 South Oak Street, Champaign, IL 61820

Presence Covenant Medical Center
1400 W. Park St., Urbana, IL 61801

University YMCA
1001 South Wright Street Champaign, IL 61820

Presence® Covenant Medical Center iTHRIVE

Choose | Your Info | Confirmation

Choose a location where you would like to have your screening. Then select a time when you are available on Monday through Saturday, from **August 15th to September 16th**. Each screening will take about 20 minutes.

The screening will involve a finger-stick blood draw, and will require that participants fast for 12 hours prior to their appointment time.

Not all locations are available on each date - click on a location to see which dates are available. If there are no dates available at your preferred location, please click on the drop-down menu to view the other locations.

Questions? Email [redacted]

To avoid losing progress, please do not use the back button on your browser.

iHotel
1900 South First Street | Champaign, IL 61820

< **September 2016** >

No appointments are available this month

S	M	T	W	Th	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

Figure D.13: Example of a reminder email sent out by the online appointment scheduler

On Sep 6, 2016, at 8 20 AM, Presence Covenant Medical Center and University of Illinois iThrive 2016 [redacted]> wrote:

Appointment Reminder

for John Doe

What Beckman Institute (Beckman Institute)

When Wednesday, September 7, 2016 8:10am CDT (10 minutes)

This is a reminder your appointment for Beckman Institute is on Wednesday, September 7, 2016 8:10am CDT

REMINDER: This is a fasting health screening. Please do not have anything to drink (besides water) for 12 hours before your appointment time. Water is encouraged.

Room Locations:

ACES Library (map)	Heritage Room
Alice Campbell Alumni Center (map)	Ballroom
Beckman Institute (map)	Room 5602 on August 17 Room 1005 all other days
Business Instructional Facility (map)	Interview Rooms
iHotel (map)	Technology Room on August 19 Humanities Room all other days
Physical Plant Services Building (map)	Room 128
Presence Covenant Medical Center (map)	Auditorium A
University YMCA (map)	Wahl Room

Change/Cancel Appointment

Add to iCal/Outlook Calendar

Add to Google Calendar

Notes: These reminders were delivered one day before the participant's health screening appointment.

Figure D.14: Example of a reminder email sent by the research team to participants one day prior to their health screening

Hello,

You are receiving this email because you are scheduled for an iThrive health screening appointment **tomorrow, September 2nd, at the Funk ACES Library.** The address is as follows:

Funk ACES Library
1101 S Goodwin Ave
Urbana, IL 61801

Tomorrow's health screenings will be held in the **Heritage Room.** Enter the ACES Library from the main entrance. The Heritage Room is located on the main level of ACES, on the West Side of the atrium. Once you enter the building doors, you will continue into the Atrium where the stairs are, and you will see the Heritage Room.

Note: Please do not have anything to eat or drink (besides water) for 12 hours before your appointment time. Water is encouraged.

Please allow about 20-25 minutes for your screening appointment.

If you have any questions tonight or tomorrow morning, please email [redacted] and we will do our best to respond to your email as soon as possible.

Sincerely,

Lauren Geary

____
Lauren E. Geary
Project Manager || iThrive
University of Illinois at Urbana-Champaign
[redacted]

Figure D.15: Copy of health questionnaire given to participants prior to screening

ID________________________

We would like to ask you a few questions about your health.

1. What is your weight, in pounds? Make your best guess.

   _________ (weight in pounds)

2. What is your height, in feet and inches? Make your best guess.

   ________ft. and ________in.

Below is a drawing of a ruler with a scale from 0 to 100. For the next set of questions, please use this scale as an indicator of how confident you are in your answer.

0 10 20 30 40 50 60 70 80 90 100

Absolutely No Chance — Not Likely — Unsure — Likely — Absolutely Certain

3. Using a number from zero to one hundred, where 0 equals absolutely no chance and 100 equals absolutely certain, what do you think the chances are that you have high cholesterol today?

   __________________ (0 to 100)

4. What do you think the chances are that you have high blood pressure today?

   __________________ (0 to 100)

5. What do you think the chances are that you have impaired fasting glucose today?

   __________________ (0 to 100)

6. A body mass index that exceeds 30 indicates that a person may be obese. What do you think the chances are that your body mass index exceeds 30?

   __________________ (0 to 100)

Figure D.16: Copy of health screening form used by clinicians from Presence Covenant Medical Center to record health measures

Presence Health℠

Worksite Wellness

SCREENING REGISTRATION and CONSENT

Name: ☐ M ☐ F	Date:	
Address:	Zip Code	Date of Birth:
Telephone:	Name of primary care physician:	
Email:	If none, would you like a referral? ☐ Y ☐ N	
Insurance provider:		

I consent to the screenings listed on this page and to the collection of screening results by Presence Health. The wellness screening includes taking body measurements, vital signs, and a "finger stick" to obtain a blood sample to measure glucose, etc. I understand that my participation in the wellness screening is voluntary and that the screening results are considered preliminary and do not constitute a diagnosis of any particular disease or condition. I understand that I will be given the results of the screening and that it is my responsibility to follow up with my health care provider regarding any treatment options. I understand that my results will be kept confidential. I acknowledge that I was provided information about Presence Health's privacy practices.

Signature of patient, or, if patient is a minor, signature of parent/guardian　　Witness　　Last 4 digits of SSN

Do you use tobacco of any form?
☐ Yes ☐ No ☐ Use E-cigarette

In the average week, how many times do you engage in physical activity?
☐ None ☐ 1-2 times per week ☐ 3 or more per week

If you engage in physical activity, for how long?
☐ Do not engage ☐ 20 minutes ☐ 40 minutes

How often do you feel tense, anxious, or depressed?
☐ Rarely or Never ☐ Sometimes ☐ Often

Do you have a primary physician?
☐ Yes ☐ No

☐ Fasting ☐ Non-Fasting

Test	Results	Desirable Levels (Source-American Heart Association, Mayo Clinic)
Height		
Weight		
Waist Circumference		Ideal Range for Women - < 35 inches; Ideal Range for Men - < 40 inches
Body Mass Index		Less than 25 - Normal 25-29 - Overweight 30 or more - Obese
Blood Pressure		Less than 120/80 - Normal 120-139/80-89 - Pre-hypertension Over 140/90 - High Blood Pressure
Total Cholesterol		Less than 200 More than 240 - High
Total Cholesterol Ratio		Less than 3.5 - Optimal
HDL		More than 60 - Optimal More than 40 - Moderate
LDL		Less than 100 - Optimal primary prevention Less than 70 - Optimal for history of diagnosed cardiovascular disease
Triglycerides		Less than 150 - Optimal 151-199 - Borderline High
Glucose		Less than 100 - Normal 101-125 - Pre Diabetes
A1C		4.0 - 6.5% - Optimal

☐ PCP referral
☐ Results require medical referral
☐ Make minor lifestyle changes
☐ Results require immediate medical attention
☐ Identification of 1 or more results out of the normal range

Clinician's comments:

Revision Date: 7/16　　Form #PH-100

Notes: A carbon copy of this was given to participants upon completion of their health screening.

Figure D.17: Health coaching guidelines

Increased Blood Pressure (180/100)

1. Does the participant have a history of high blood pressure?

If yes: ask the participant if they are working with their PCP to decrease their blood pressure.

If no: make the patient aware of the damage consistently increased blood pressure has on their body.

Give educational materials.

2. Do they have a primary care provider?

If yes: tell participant to make an appointment with their provider and take the screening form with.

If no: give a list of providers and make the participant aware of the importance.

Increased Glucose (>210 Fasting)

1. Does the participant have a family history of diabetes?

If yes: ask the participant if they are working with their PCP.

If no: make the patient aware of the possibility of diabetes, and the importance of being tested. Give educational materials.

2. Do they have a primary care provider?

If yes: tell participant to make an appointment with their provider and take the screening form with.

If no: give a list of providers and make the participant aware of the importance.

Decreased Glucose (<65)

1. Ask the patient if they are feeling well.

If yes: let them know their glucose levels are low and they may want to eat something.

If no: sit them down immediately, and give them juice and a granola bar.

Increased Triglycerides (>500)

1. Ask the participant if they did indeed fast for 8-12 hours prior to health screening.

If no: then explain the test is not an accurate measurement of triglycerides.

If yes: proceed to step 2.

2. Do they have a primary care provider?

If yes: tell participant to make an appointment with their provider and take the screening form with.

If no: give a list of providers and make the participant aware of the importance.

3. Does the participant have a history of high triglycerides?

Page **1** of **2**

If yes: ask the participant if they are working with their PCP to decrease the triglycerides.

If no: make the patient aware of the damage increased triglycerides has on their body.

Give educational materials.

Increased Triglycerides (>500), Increased Total Cholesterol Ratio (>4.0)

1. Ask the participant if they did indeed fast for 8-12 hours prior to health screening.

If no: then explain the test is not an accurate measurement of triglycerides, but there is still concern with the elevated cholesterol ratio.

If yes: proceed to step 2.

2. Do they have a primary care provider?

If yes: tell participant to make an appointment with their provider and take the screening form with.

If no: give a list of providers and make the participant aware of the importance.

3. Does the participant have a family history of heart disease?

If yes: ask the participant if they are working with their PCP to prevent heart disease.

If no: make the patient aware of the damage increased triglycerides and bad cholesterol has on their body.

Give educational materials.

Increased Triglycerides (>500), Increased Total Cholesterol Ratio (>4.0), Increased Blood Pressure

1. Ask the participant if they did indeed fast for 8-12 hours prior to health screening.

If no: then explain the test is not an accurate measurement of triglycerides, but the elevated cholesterol ratio and blood pressure are cause for concern.

If yes: proceed to step 2.

2. Do they have a primary care provider?

If yes: tell participant to make an appointment with their provider and take the screening form with.

If no: give a list of providers and make the participant aware of the importance.

3. Does the participant have a family history of heart disease?

If yes: ask the participant if they are working with their PCP.

If no: make the patient aware their health screening numbers give concern for heart disease. It is essential for the participant to obtain an appointment for further assessment.

Give educational materials, and write a personal note on the screening form that states they need to see a PCP.

Page **2** of **2**

Notes: These guidelines were employed by health coaches during their private discussions with study participants immediately following the health screening.

Figure D.18: Postcard given to participants on site after they completing their health screening

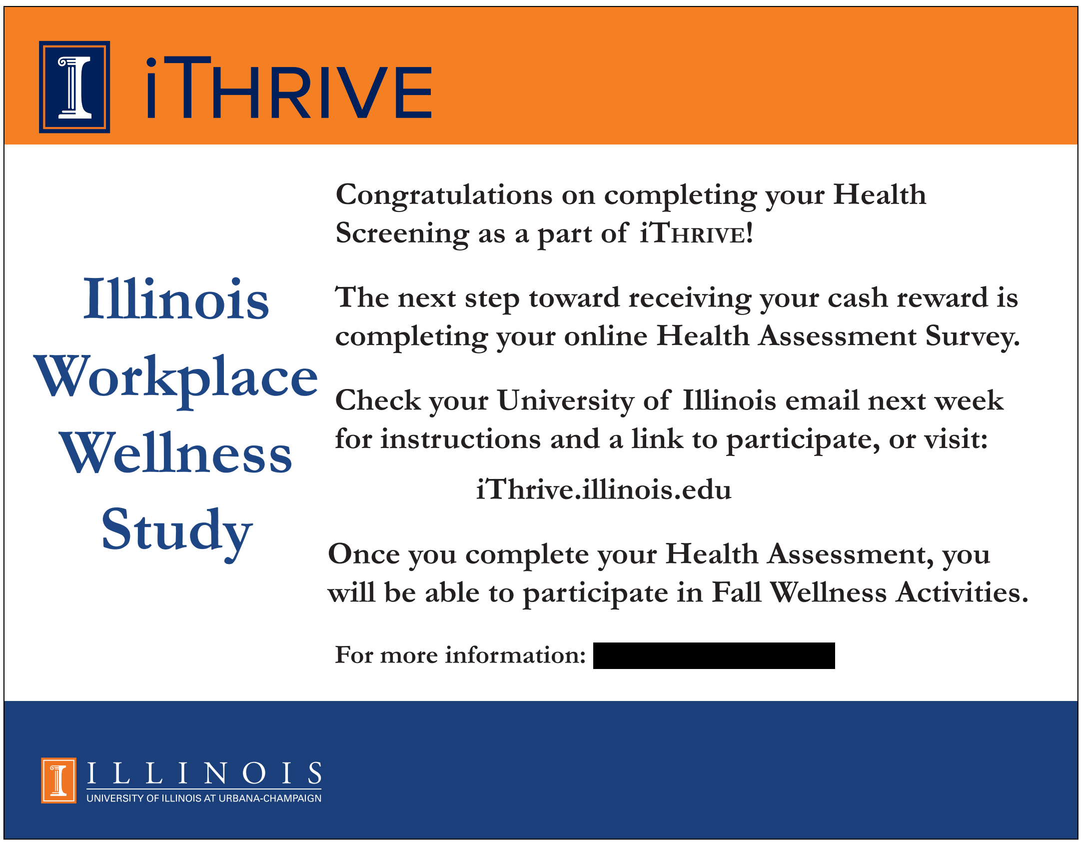


Figure D.19: Email invitation for the online health assessment

From: [redacted]
Subject: iThrive: Health Assessment Survey Invitation

Dear [First Name]:

Congratulations on completing your iThrive health screening! The next step is to complete your online Health Assessment survey, which will provide you with a personalized health summary and suggest practical ways to improve your health.

The Health Assessment survey takes about 12 minutes. After finishing this survey, you will receive a reward of $100 and will be eligible to enroll in wellness activities once registration opens.

To access the online Health Assessment survey, simply copy and paste the following URL into your browser:

ithrive.illinois.edu/healthassessment

You must log in using the following username and initial password:

Username: <username>
Password: <password>

Once you are logged in, you must accept the terms of agreement. Next, click on the "Start New Assessment" button and answer a series of questions. You must click "Finish" when you are done, in order to view your report and to become eligible to enroll in wellness activities.

Please note: Some participants have experienced technical difficulties when taking their surveys. Slow response times or error messages sometimes arise when our survey vendor's servers become overloaded. If you face any technical difficulties while taking the survey, please wait for fifteen minutes and try again later. We are sorry for any inconvenience this might cause for you.

This survey asks questions about seven dimensions of health (i.e., heart health, fitness, nutrition, mental health, diabetes risk, cancer risk, overweight/obesity risk). In order for the software to calculate a personalized wellness score for each dimension, you must answer all of the questions. Your results will give you insights you can use to make goals and plans for health improvement through iThrive programs and activities.

In the consent form you signed at the beginning of this study, you were told that you may refuse to answer any questions and withdraw at any time. This is still true with the Health Assessment, except that if you choose to skip any question in the health assessment, you cannot proceed with the survey. This software limitation only applies to the Health Assessment. If you do not wish to answer all of the survey items, you may withdraw from the study altogether.

For non-exempt civil service employees, iThrive is an "approved event," so that, operations permitting and with prior supervisory approval, participation that occurs during an employee's regular work schedule does not require the charging of leave benefit time.

As with every part of the Illinois Workplace Wellness Study, your participation in iThrive is completely voluntary and your **individual data will <u>never</u> be shared with your health insurance provider or your employer**. If you have any questions or need assistance, please contact us at [redacted] or call Lauren Geary, Project Manager, at [redacted]

Yours in good health,

Illinois Workplace Wellness Study Team

Notes: This was sent only to participants who had completed their health screening. The text highlighted in yellow was appropriately customized for each participant.

Figure D.20: Email invitation for Fall 2016 wellness activities

From: [redacted]
Subject: iThrive: Wellness Activity Registration Now Open

Dear [First Name]:

Congratulations on completing your iThrive Health Screening and online Health Assessment survey! You are now eligible to enroll in one of the iThrive Wellness Activities for Fall 2016.

You are free to choose a wellness activity, also called a "track," that best aligns with an area of your health that you would like to improve. These areas include physical activity, weight management, stress management, chronic disease management, and tobacco cessation. You will have the option to participate in classes that meet in-person, or you may choose to participate in one of our online, self-paced programs like HealthTrails.

Completing your chosen wellness track in the Fall will entitle you to a $[X] reward. If you also complete a wellness track in the Spring, you will receive an additional $[X]. You do not have to participate in an activity in the Fall in order to be eligible to participate in the Spring.

To view the set of Wellness Activities that will be offered and to enroll, log in to iThrive by copying and pasting the following URL into your browser:

https://ithrive.illinois.edu/

After you log in to iThrive, click on the "Wellness Activities" tab near the top of your home page. This page lists the different activities available to you. Below each activity is a registration link. Click on the link that corresponds to the activity in which you would like to enroll, select the option to "log in using your netid," and complete the registration form. You will receive a confirmation email when you have completed this step. Please note that you may only sign up for one fall activity.

For non-exempt civil service employees, iThrive is an "approved event," so that, operations permitting and with prior supervisory approval, participation that occurs during an employee's regular work schedule does not require the charging of leave benefit time.

As with every part of the Illinois Workplace Wellness Study, your participation in iThrive is completely voluntary and your **individual data will never be shared with your health insurance provider or your employer**. If you have any questions or need assistance, please contact us at [redacted] or call Lauren Geary, Project Manager, at [redacted]

Yours in good health,

Illinois Workplace Wellness Study Team

Notes: This was sent only to participants who had completed their online health assessment. The text highlighted in yellow was appropriately customized for each participant.

Figure D.21: Email invitation for Spring 2017 wellness activities

From: iThrive@illinois.edu
Subject: iThrive: Spring Wellness Activity Registration Now Open

Dear [First Name]:

Congratulations on all of your progress in iThrive so far. You are now eligible to enroll in one of the iThri[redacted]ies for Spring 2017.

You are free to choose a wellness activity that best aligns with an area of your health that you would like to improve. These areas include physical activity, weight management, stress management, chronic disease management, and financial wellness. You will have the option to participate in classes that meet in-person, or you may choose to participate in one of our online, self-paced programs like Spring Into Motion. Note that each activity has a limited capacity, except for Spring Into Motion. Registration will end on Friday, February 10.

Completing your chosen wellness activity in the Spring will entitle you to a $[X] reward. You are able to participate in a Wellness Activity this Spring even if you did not participate in the Fall.

To view the set of Wellness Activities that will be offered and to enroll, log in to iThrive by copying and pasting the following URL into your browser:

https://ithrive.illinois.edu/

After you log in to iThrive, click on the "Wellness Activities" tab near the top of your home page. This page lists the different activities available to you. Below each activity is a registration link. Click on the link that corresponds to the activity in which you would like to enroll, select the option to "log in using your netid," and complete the registration form. Participants with a "@uillinois.edu" email address may need to log in using the "log in using your email" option. You will receive a confirmation email when you have completed this step. Please note that you may only sign up for one Spring activity.

For non-exempt civil service employees, iThrive is an "approved event," so that, operations permitting and with prior supervisory approval, participation that occurs during an employee's regular work schedule does not require the charging of leave benefit time.

As with every part of the Illinois Workplace Wellness Study, your participation in iThrive is completely voluntary and your **individual data will never be shared with your health insurance provider or your employer**. If you have any questions or need assistance, please contact us at iThrive@illinois.edu or call Lauren Geary, Project Manager, at 217-265-8980.

Yours in good health,

Illinois Workplace Wellness Study Team

Notes: This was sent only to participants who had completed their online health assessment. The text highlighted in yellow was appropriately customized for each participant.

Figure D.22: Front and back sides of invitation postcard sent on July 6, 2017

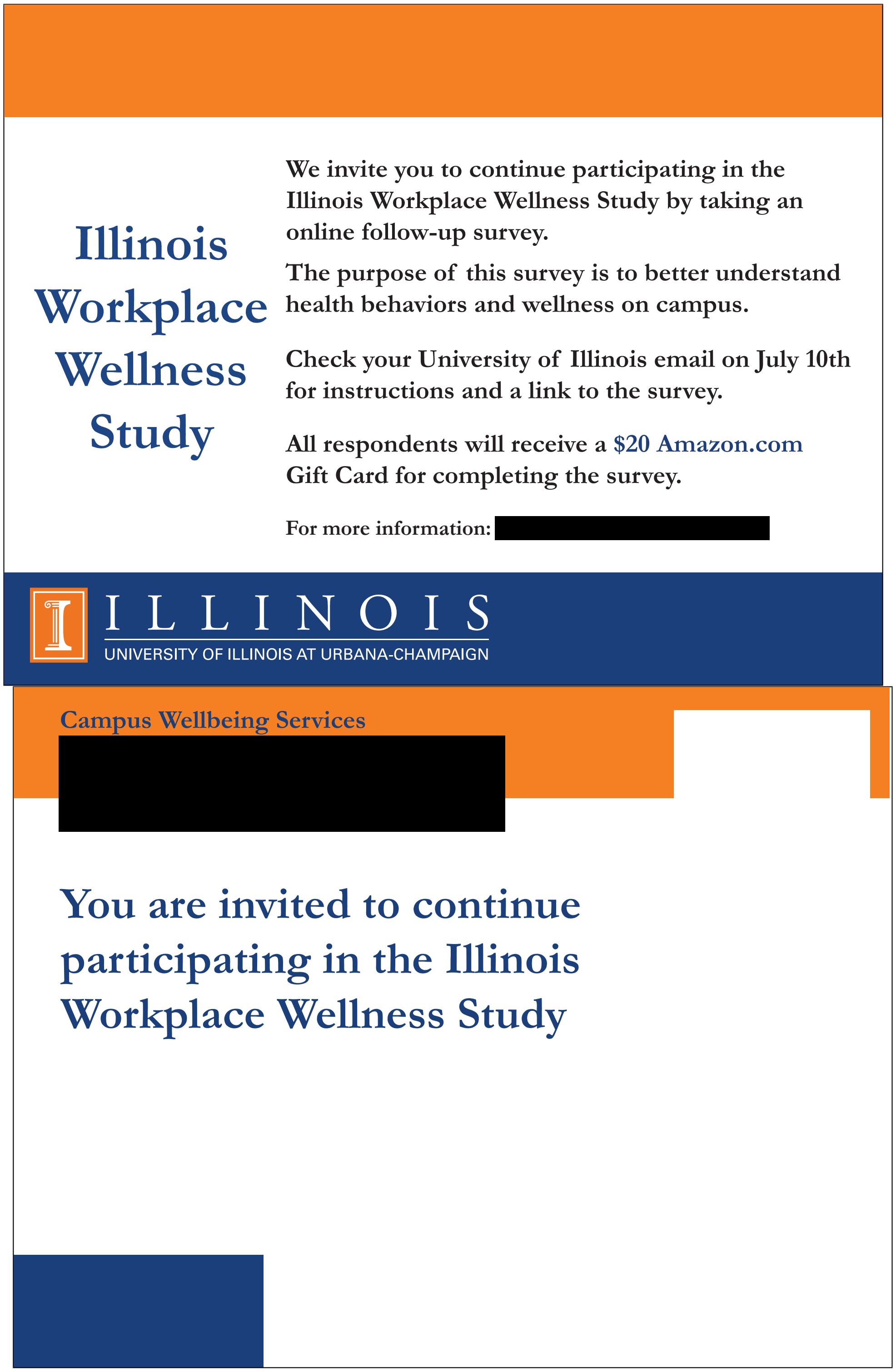


Figure D.23: One-year follow-up survey invitation sent to study participants on July 10, 2017

Dear <FirstName>,

Last summer, you participated in an online survey for the Illinois Workplace Wellness Study. Your participation has allowed the Illinois Workplace Wellness Study Team to conduct important research about workplace wellness programs on the UIUC campus.

We invite you to take part in a second survey for the Illinois Workplace Wellness Study. As before, this online survey includes questions about health behaviors and wellness on campus. The survey will take approximately 15 minutes to complete. We know that your time is valuable, so we are offering a **$20 Amazon.com gift card** to all respondents who complete the survey. This gift card is taxable.

The survey is only available for a limited time, so please complete the survey promptly in order to receive your $20 gift card. To access the online survey, simply copy and paste the following URL in your browser:

<link>

This survey is strictly confidential. **Your individual data will never be shared with the university or your health insurer.**

For non-exempt civil service employees, this program is an "approved event," so that, operations permitting and with prior supervisory approval, participation that occurs during an employee's regular work schedule does not require the charging of leave benefit time.

Thank you for contributing to this important research project! If you have any questions or need assistance, please contact us at [redacted] or [redacted]

Best regards,

Illinois Workplace Wellness Study Team

David Molitor
Assistant Professor, Department of Finance

Laura Payne
Professor, Department of Recreation, Sport and Tourism

Julian Reif
Assistant Professor, Department of Finance and IGPA

Notes: The text highlighted in yellow was appropriately customized for each participant.

Figure D.24: One-year follow-up survey reminder sent on August 2, 2017

From: [redacted]
Subject: $100 gift card drawing: Illinois Workplace Wellness Study

Dear [FIRSTNAME],

We are pleased to announce that those who complete the online survey for the Illinois Workplace Wellness Study will be entered into a drawing to win a $100 Amazon.com gift card.

Ten (10) people who complete the brief survey will be selected at random to receive a $100 Amazon.com gift card. This gift card will be <u>in addition to the $20 Amazon.com gift card</u> that all participants receive for completing the online survey. The drawing for the $100 Amazon.com gift card will occur after the survey closes. Winners will be notified by email.

To access the online survey, simply copy and paste the following URL in your browser:

<personalized study url>

If you have already completed the survey, then you will automatically be entered into the drawing.

If you have any questions or need assistance, please contact us at [redacted] or [redacted]

Best Regards,

Illinois Workplace Wellness Study Team

Notes: The text highlighted in yellow was appropriately customized for each participant. This reminder informed participants for the first time that completing the follow-up survey would enter them into a drawing for an additional $100 reward.

Figure D.25: Text of invitation email sent to study participants on August 14, 2017

Dear [First Name]:

You have been selected to participate in the 2017 iThrive Health Screenings. The iThrive Health Screenings are a component of the Illinois Workplace Wellness Study.

The iThrive program offers you the opportunity to participate in a valuable health screening at no cost to you. **In addition, you will earn $125 for completing the iThrive Health Screening.**

The opportunity to participate in the iThrive Health Screening is only available for a limited time. To learn more about iThrive and to sign up for an appointment, visit the iThrive website:

iThrive.illinois.edu

The iThrive Health Screening is summarized below.

Last month, you were invited to take the Illinois Workplace Wellness Study online survey. Even if you did not complete that survey, you are still invited to participate in the health screening. For those of you who took the survey, the random drawing has been completed and the winners have been notified.

---

iThrive Health Screening

You are invited to participate in a free health screening through the iThrive program, beginning on August 21. The purpose of a health screening is to measure physical health characteristics (e.g., height, weight, blood pressure, cholesterol) and use the information as a benchmark for health promotion and management. For your convenience, Presence Health will offer these screenings at various dates and locations across campus. Appointments typically take about 20 to 25 minutes.

Upon completion of the health screening, you will receive a reward of $125.

Scheduling your Health Screening

To schedule your health screening, copy and paste the URL below into your web browser:

https://presencehealth.acuityscheduling.com/

When scheduling your health screening, please use the email address to which this email was sent (**netid@illinois.edu**). This email address will be referred to as your "iThrive contact email".

You may also visit the iThrive website at any time: iThrive.illinois.edu. This website provides personalized information about your progress.

For non-exempt civil service employees, the iThrive Health Screening is an "approved event," so that, operations permitting and with prior supervisory approval, participation that occurs during an employee's regular work schedule does not require the charging of leave benefit time.

As with every part of the Illinois Workplace Wellness Study, your participation in the iThrive Health Screening is completely voluntary and your **individual data will never be shared with your health insurance provider or your employer**. You can read here about the purpose of our study as well as the steps we will take to keep your information confidential. If you have any questions or need assistance, please contact us at [redacted]

Yours in good health,

Illinois Workplace Wellness Study Team

Notes: The text highlighted in yellow was appropriately customized for each participant.

Figure D.26: Text of reminder email sent to study participants on September 21, 2017

Dear [First Name]:

This is your **last chance** to attend your free iThrive Health Screening. **The final day to complete your iThrive Health Screening is tomorrow, Friday September 22nd, at Beckman Institute.** To schedule a screening, copy and paste the following URL into your browser:

https://presencehealth.acuityscheduling.com/schedule.php

As a reminder, you will receive a reward of $125 after completing your iThrive Health Screening.

Walk-ins are also encouraged! Stop by Beckman Institute, Room 1005 any time between 6am and 12pm on Friday, September 22nd for an appointment.

For non-exempt civil service employees, iThrive is an "approved event," so that, operations permitting and with prior supervisory approval, participation that occurs during an employee's regular work schedule does not require the charging of leave benefit time.

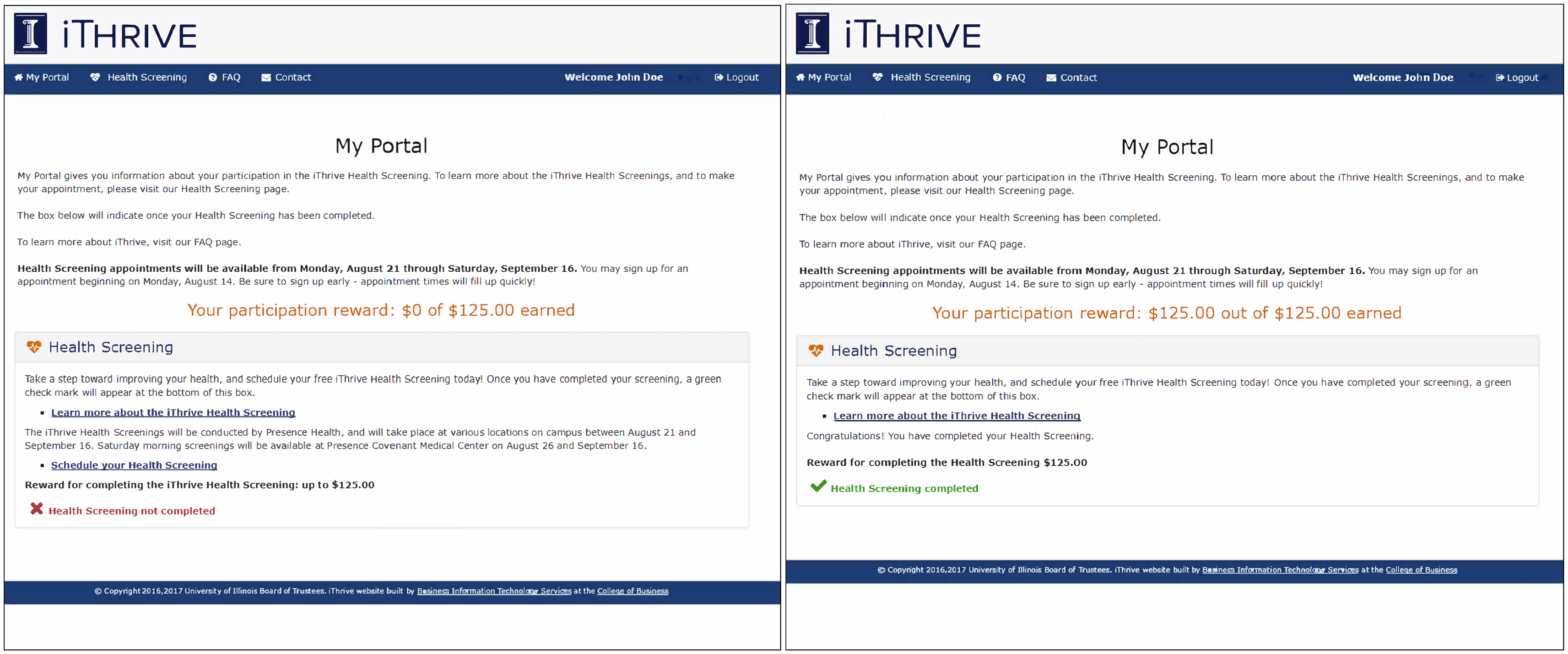
As with every part of the Illinois Workplace Wellness Study, your participation in iThrive is completely voluntary and your **individual data will never be shared with your health insurance provider or your employer**. If you have any questions or need assistance, please contact us at [redacted]

Yours in good health,

Illinois Workplace Wellness Study Team

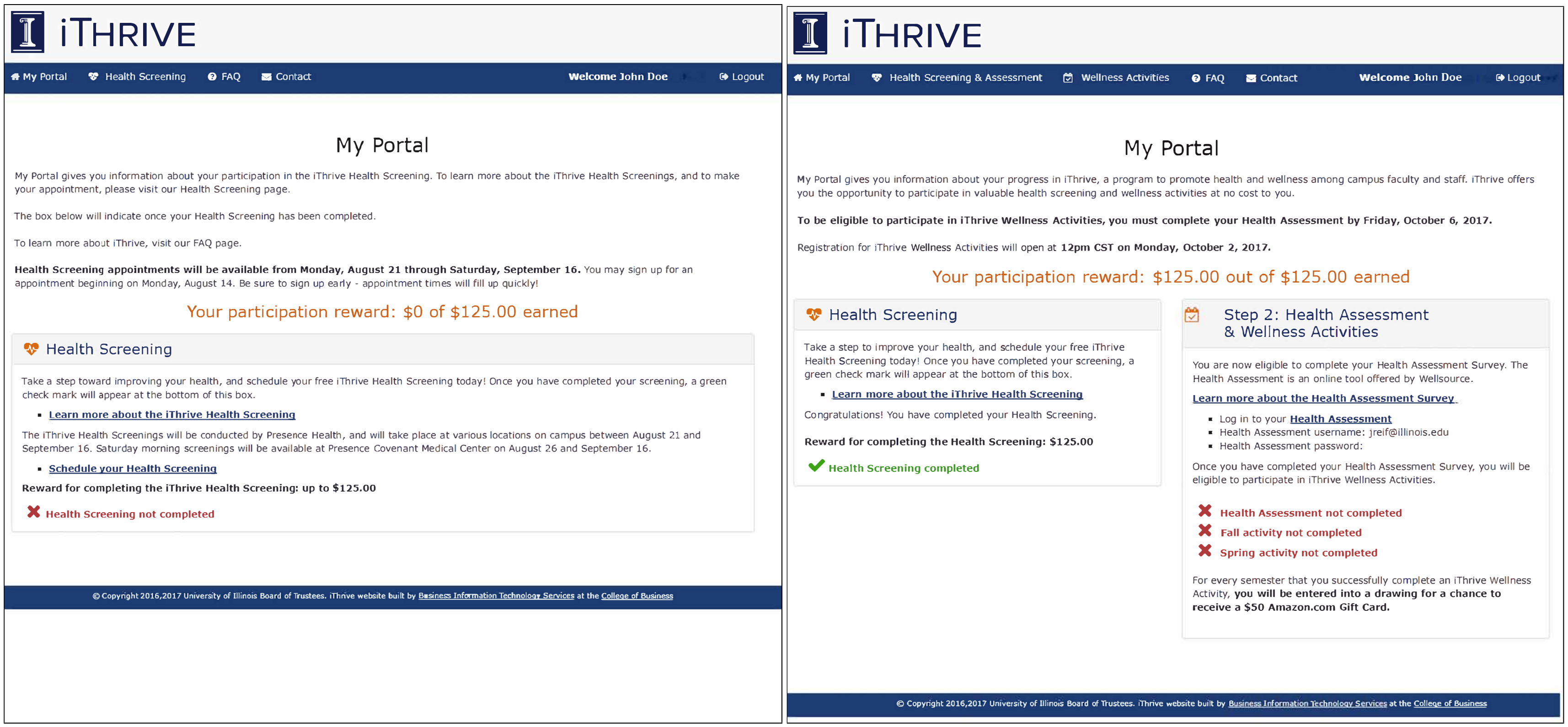
Notes: The text highlighted in yellow was appropriately customized for each participant.

Figure D.27: Main page for the 2017-2018 iThrive website for a control group member in the $125 screening reward group



Notes: Follow-up screening participants in the $0 reward group did not receive a confirmation email. However, all follow-up screening participants could confirm their completion status on the iThrive website.

Figure D.28: Main page for the 2017-2018 iThrive website for a treatment group member in the $125 screening reward group



Notes: Follow-up screening participants in the $0 reward group did not receive a confirmation email. However, all follow-up screening participants could confirm their completion status on the iThrive website.

Figure D.29: Text of the confirmation email sent to one-year follow-up screening participants in the $125 reward group

From: [redacted]
Subject: Your iThrive Health Screening Payment

Hello,

Congratulations on completing your iThrive Health Screening! Your $125 reward for completion will be processed in October, after the iThrive Health Screenings have ended. The payments will be made through direct deposit, and will be included as part of your regularly scheduled paychecks. As a reminder, these payments are taxable.

You may log in to the iThrive website at https://iThrive.illinois.edu to view your progress at any time.

Please let us know if you have any questions. We will send an email in October after all of the payments have been made.

Yours in good health,

The Illinois Workplace Wellness Study Team

Notes: Follow-up screening participants in the $0 reward group did not receive a confirmation email. However, all follow-up screening participants could confirm their completion status on the iThrive website.

D.4 Online Appendix Tables

Table D.1: Dates, locations, times, and number of health screenings performed in 2016

Date	Location	Appt Times	Capacity	Appts scheduled	Total Screened
Monday, August 15	Business Instructional Facility	6:00am - 10:20am	108	67	69
Tuesday, August 16	Business Instructional Facility	6:00am - 10:20am	108	66	65
Wednesday, August 17	Beckman Institute	6:00am - 10:20am	108	89	90
Thursday, August 18	Physical Plant Services Building	7:45am - 10:15am	64	58	57
Friday, August 19	iHotel	6:00am - 10:20am	108	91	93
Saturday, August 20	Presence Covenant Medical Center	7:00am - 10:20am	84	74	76
Monday, August 22	iHotel	6:00am - 10:20am	108	99	92
Tuesday, August 23	Business Instructional Facility	6:00am - 10:50am	120	75	75
Wednesday, August 24	Business Instructional Facility	6:00am - 10:50am	120	77	74
Thursday, August 25	Alice Campbell Alumni Center	7:45am - 10:55am	80	74	77
Friday, August 26	Beckman Institute	6:00am - 10:50am	120	100	94
Saturday, August 27	Presence Covenant Medical Center	7:00am - 9:50am	72	52	45
Monday, August 29	Beckman Institute	6:00am - 10:55am	120	97	90
Tuesday, August 30	iHotel	6:00am - 10:55am	120	109	104
Wednesday, August 31	University YMCA	6:00am - 10:50am	120	98	94
Thursday, September 1	University YMCA	6:00am - 10:50am	120	78	71
Friday, September 2	ACES Library	8:15am - 10:55am	68	66	60
Saturday, September 3	N/A	N/A	N/A	N/A	N/A
Monday, Septermber 5	N/A	N/A	N/A	N/A	N/A
Tuesday, September 6	iHotel	6:00am - 10:50am	120	117	99
Wednesday, September 7	Beckman Institute	6:00am - 10:50am	120	87	76
Thursday, September 8	University YMCA	6:00am - 10:50am	120	92	81
Friday, September 9	University YMCA	6:00am - 10:50am	120	66	55
Saturday, September 10	Presence Covenant Medical Center	7:00am - 9:50am	72	26	17
Monday, September 12	iHotel	6:00am - 10:50am	61	52	45
Tuesday, September 13	iHotel	6:00am - 10:50am	75	53	45
Wednesday, September 14	iHotel	6:00am - 10:50am	76	58	53
Thursday, September 15	iHotel	6:00am - 10:50am	76	50	42
Friday, September 16	iHotel	6:00am - 10:50am	76	76	61
Total			**2,664**	**2,047**	**1,900**

Table D.2: Description of and statistics for the Fall 2016 wellness activities

	Number of classes	Time and day of week	Start date	End date	Reward requirement	Capacity	Registered	Completed	Description
Freedom from Smoking	1	N/A	10/17/2016	12/9/2016	8 weekly calls	20	17	9	The Illinois Freedom from Smoking HelpLine is a one-on-one telephonic coaching program to help participants to quit tobacco for good. Participants are matched with a trained cessation expert. Quitline cessation specialists offer participants expert advice, an assessment of your tobacco treatment, and help you develop a customized quit-plan. Calls take place weekly, and are scheduled at your convenience.
HealthTrails	Unlimited	N/A	10/10/2016	12/4/2016	400 virtual miles	Unlimited	1027	715	HealthTrails is an eight-week self-paced, online wellness activity developed by Health Enhancement Systems – a leader in online wellness campaigns. This program allows participants to virtually travel along famous trails as they practice and record healthy lifestyle behaviors such as physical activity, nutrition, and stress management. HealthTrails is includes the option of a mobile application that allows participants to conveniently track their behaviors using their cell phone or other mobile device. The program incorporates challenging wellness goals and fun themes, as well as daily tips throughout the program. Participants who choose to register for HealthTrails can work to improve their health in the areas of: * Physical Activity * Stress Management * Healthy Eating
Live Well Be Well	1	5:15pm - 7:15pm (R)	10/13/2016	11/17/2016	Attend 5 of 6 classes	20	19	16	Live Well, Be Well is a six-week evidence-based chronic disease self-management program that was developed by Stanford University. This interactive program has been shown empowers participants through learning important lifestyle skills that enhance one's ability to effectively manage ongoing health conditions. This program is open to anyone with an ongoing health condition such as arthritis, heart disease, asthma, lung disease, diabetes, osteoporosis, cancer or any other. Caregivers may also participate. The program is taught by certified facilitators Cheri Burcham and Chelsey Byers, University of Illinois Extension community health educators.
Prudential Pathways	1	5:15pm - 6:15pm (R)	10/13/2016	11/10/2016	Attend 5 of 5 classes	25	25	20	The Prudential Pathways program offers practical, down-to-earth financial information. Participants will gain an understanding of the fundamentals of financial wellness, and personal financial planning. Prudential Pathways will be facilitated by Peggy Furlong with Prudential Financial, and will cover important topics such as: setting your financial goals, protecting your assets through risk management, investment principles, healthcare planning, retirement and asset distribution planning, tax strategies, estate planning strategies, how your employee benefits fit into your overall financial wellness, and more.
Recess for Adults	2	5:15pm - 6:00pm (W), 6:30pm - 7:15pm (W)	10/12/2016	12/7/2016	Attend 6 of 8 classes	50	49	28	Recess For Adults is an eight-week program inspired by games typically seen on a playground. This program is perfect for adults to increase their physical activity levels, and to have fun together. A typical class agenda could include, for example, "Red Light, Green Light", "Crazy Kickball", "Blob Tag", and "Group Juggle". This program meets once per week for 45 minutes, for eight weeks. The program will be led by instructor Kerri Schiller, a University of Illinois PhD student in Recreation, Sport, and Tourism.
Stress Management	1	5:15pm - 6:15pm (W)	10/19/2016	12/14/2016	Attend 6 of 8 classes	40	40	27	This eight-week program provides participants with the knowledge and skills to effectively manage stress in their lives. Participants gain an understanding of how stress affects them. They build awareness of their personal stressors and stress symptoms, of their ability to control how stress affects them, and how to address stress. The program is very interactive; in each session participants learn practical skills they can use in their daily lives. Topics include defining stress, overcoming stressful thought patterns, relaxation techniques, managing stress at work, coping with change, and more. The program is facilitated by Michele Guerra, the Director of the UI Wellness Center.
Tai Chi	3	5:15pm - 6:15pm (T), 6:30pm - 7:30pm (T, R)	10/11/2016	12/8/2016	Attend 6 of 8 classes	60	60	39	Tai Chi for Relaxation is an eight-week program that aims to improve overall health and wellness through learning basic Tai Chi movements and techniques. The class is taught by local certified Tai Chi instructor Rick Krandel, who maintains certification from the Tai Chi for Health Institute. Two sessions of Tai Chi for Relaxation are scheduled this fall. You may select either the Tuesday evening or Thursday evening sessions.
Weight Watchers at Work	2	12:00pm-12:50pm (W,R)	10/12/2016	12/8/2016	Attend 6 of 8 classes	32	32	27	Weight Watchers at Work is an eight-week weight management program, that aims to help participants to develop skills to unlock their inner strength to make healthy choices for life. Participants will learn how to see food as a fuel for a healthy life, and to find ways to move more each day. The SmartPoints plan assigns a point value to every food, and members are given a target number of points for each day. Participants can make their own choices about what foods to eat to reach their daily target number of points. Weight Watchers at Work will meet on Thursdays from 12pm to 1pm.
Well at Work	1	12:00pm-12:50pm (M)	10/10/2016	12/5/2016	Attend 6 of 8 classes	35	35	22	The Well at Work Series is an eight-week program that provides participants with practical tips on how to stay healthy at work. Each session will focus on a different aspect of workplace wellness. The brief lunch and learn format is conveniently scheduled to increase employees' ability to attend. Facilitator Michele Guerra, the Director of the UI Wellness Center, will cover a variety of workplace health-related topics, including how to: fit physical activity in at work, eat healthfullly at work, achieve work-life balance, get a good night's sleep, stay energized during the work day, relax during stressful moments, and more.
Total							**1,304**	**903**	

Table D.3: Description of and statistics for the Spring 2017 wellness activities

	Number of classes	Time and day of week	Start date	End date	Reward requirement	Capacity	Registered	Completed	Description
Active Living Every Day	1	5:15pm - 6:15pm (T)	1/31/2017	4/25/2017	Attend 9 out of 12 classes	30	12	9	Active Living Every Day (ALED) helps people become and stay physically active. ALED focuses on lifestyle physical activity into one's life and life management skills. Participants will be provided with a step-by-step process to create their own healthy lifestyle. They will learn a wide variety of life skills, including: *Setting goals *Overcoming challenges *Defusing stress *Making lasting changes, and more ALED is perfect for inactive people, or those who want to be more active, but are having difficulty doing so. Note: This is not an exercise class.
Adventures in Financial Wellness	1	5:15pm - 6:15pm (R)	2/16/2017	4/13/2017	Attend 6 out of 8 classes	36	36	21	Looking to expand or deepen your financial savvy? Sign up for Adventures in Financial Wellness. Each week, Prudential financial professionals* will provide practical information on a different financial wellness topic. Participants will gain a better working knowledge of credit, banking services, saving, investing, and funding college, taxes, life insurance and retirement planning. This program is different from the Pathways program we offered in the fall. Some information may be similar. *No Prudential financial products will be sold or promoted during this series.
Healthy Weigh	1	5:15pm - 6:15pm (W)	2/8/2017	4/5/2017	Attend 6 out of 8 classes	40	28	17	Are you looking for a safe and effective weight management program? Join the Healthy Weigh! Healthy Weigh is the UI Wellness Center's weight management program. Healthy Weigh equips participants with proper tools to lose weight safely and effectively. This program is not a diet. Participants will: *Learn how to lose and maintain a healthy weight *Attain life management skills to help them attain their weight goals *Receive group support to increase self-confidence
Live Well Be Well	1	5:15pm - 7:15pm (W)	2/22/2017	4/12/2017	Attend 5 out of 7 classes	20	9	3	Live Well, Be Well is a six-week evidence-based chronic disease self-management program that was developed by Stanford University. This interactive program has been shown empowers participants through learning important lifestyle skills that enhance one's ability to effectively manage ongoing health conditions. This program is open to anyone with an ongoing health condition such as arthritis, heart disease, asthma, lung disease, diabetes, osteoporosis, cancer or any other. Caregivers may also participate. The program is taught by certified facilitators Cheri Burcham and Chelsey Byers, University of Illinois Extension community health educators.
Lunchtime Walk	1	12:10pm - 12:55pm (M)	2/27/2017	4/24/2017	Attend 6 out of 8 sessions	35	34	21	Do you want to get more physical activity, but can't seem to find the time? It just got easier to fit in a walk during your busy day. Sign up for our Lunchtime Walk program. These walks are designed to fit into the average lunch break, allowing enough time to travel to and from the starting point, get a 30-minute walk, and return to your work area. The first three walks will be inside; once the weather warms up a bit, we will walk outside. Walkers of all abilities are welcome.
Mini Stress Relievers	1	12:10pm - 12:55pm (T)	2/14/2017	4/11/2017	Attend 6 out of 8 classes	35	35	28	Need some "me time"? Join our Mini Stress Relievers program! Each week we will feature an easy-to-do stress reduction activity. Examples of activities include: *Coloring *Practicing muscle relaxation techniques *Taking a contemplative walk *Experiencing the power of aromatherapy *And more! You will also have the opportunity to meet other campus employees in a relaxing atmosphere.
Recess for Adults	1	5:15pm - 6:00pm (W)	2/8/2017	4/5/2017	Attend 6 of 8 classes	25	25	15	Recess For Adults is an eight-week program inspired by games typically seen on a playground. This program is perfect for adults to increase their physical activity levels, and to have fun together. A typical class agenda could include, for example, "Red Light, Green Light", "Crazy Kickball", "Blob Tag", and "Group Juggle". This program meets once per week for 45 minutes, for eight weeks. The program will be led by instructor Kerri Schiller, a University of Illinois PhD student in Recreation, Sport, and Tourism.
Spring Into Motion	N/A	N/A	2/6/2017	4/2/2017	Obtain 40 "Springer Icons" (6,000 steps per day or 30 minutes of physical activity per day for 40 days)	Unlimited	808	588	Spring Into Motion is an online, self-paced wellness activity that encourages participants to be more active. The program allows participants to track either their steps or physical activity minutes each day, making progress toward a final goal. As they track their activity, participants progress through different, exciting spring events all around the world. This program is great for participants of all fitness levels. Whether you are just starting out, or have a well-established physical activity routine, Spring Into Motion will help to boost energy and improve health. For user convenience, a mobile application is also available to help with on-the-go activity tracking. Participants who own a FitBit or a Jawbone device will have the ability to sync their devices with their Spring Into Motion accounts, allowing for automatic activity tracking. Participants will strive to reach a goal of at least 6,000 steps per day or 30 minutes of physical activity per day, for at least 40 days throughout the program.
Tai Chi	3	6:30pm - 7:30pm (T), 6:30pm - 7:30pm (T, R)	2/7/2017	4/6/2017	Attend 6 of 8 classes	60	60	27	Tai Chi for is an eight-week program that aims to improve overall health and wellness through learning basic Tai Chi movements and techniques. The class is taught by local certified Tai Chi instructor Rick Krandel, who maintains certification from the Tai Chi for Health Institute. Two sessions of Tai Chi for Relaxation are scheduled this fall. You may select either the Tuesday evening or Thursday evening sessions.
Tai Chi (Advanced)	1	5:15pm - 6:15pm (T)	2/7/2017	4/4/2017	Attend 6 of 8 classes	20	12	11	Tai Chi Extension Movements is an eight-week program that aims to improve overall health and wellness through Tai Chi movements. We will be offering the Extension Movements class as an advanced section of Tai Chi, where the instructor will be teaching additional postures that were not covered in the first semester sessions. This class has a limited capacity, and is only open to participants who successfully completed an introductory Tai Chi program in the Fall (attended at least 6 of the 8 sessions).
Total							**1,059**	**740**	

Table D.4: Dates, locations, times, and number of health screenings performed in 2017

Date	Location	Appt Times	Capacity	Appts scheduled	Total Screened
Monday, August 21	Business Instructional Facility	6:00am - 11:20am, 12:40pm - 4:00pm	208	62	57
Tuesday, August 22	Beckman Institute	6:00am - 11:20am, 12:40pm - 4:00pm	208	152	138
Wednesday, August 23	Business Instructional Facility	6:00am - 11:20am, 12:40pm - 4:00pm	208	70	65
Thursday, August 24	University YMCA	6:00am - 11:20am, 12:40pm - 4:00pm	208	106	97
Friday, August 25	iHotel	6:00am - 11:20am, 12:40pm - 4:00pm	208	178	154
Saturday, August 26	Presence Covenant Medical Center	7:00am - 10:50am	96	74	67
Monday, August 28	Alice Campbell Alumni Center	7:45am - 11:15am, 12:40pm - 4:00pm	168	112	96
Tuesday, August 29	Business Instructional Facility	6:00am - 11:20am, 12:40pm - 4:00pm	208	75	63
Wednesday, August 30	ACES Library	7:45am - 11:15am, 12:40pm - 4:00pm	168	126	120
Thursday, August 31	iHotel	6:00am - 11:20am, 12:40pm - 4:00pm	208	148	138
Friday, September 1	Beckman Institute	6:00am - 11:20am, 12:40pm - 4:00pm	208	38	34
Saturday, September 2	N/A	N/A			
Monday, September 4	N/A	N/A			
Tuesday, September 5	iHotel	6:00am - 11:20am, 12:40pm - 4:00pm	208	87	75
Wednesday, September 6	Alice Campbell Alumni Center	7:45am - 11:15am, 12:40pm - 4:00pm	168	75	68
Thursday, September 7	iHotel	6:00am - 11:20am, 12:40pm - 4:00pm	208	100	85
Friday, September 8	University YMCA	6:00am - 11:20am, 12:40pm - 4:00pm	208	84	77
Saturday, September 9	N/A	N/A			
Monday, September 11	Beckman Institute	6:00am - 11:20am, 12:40pm - 4:00pm	208	101	93
Tuesday, September 12	iHotel	6:00am - 11:20am, 12:40pm - 4:00pm	208	90	82
Wednesday, September 13	University YMCA	6:00am - 11:20am, 12:40pm - 4:00pm	208	58	53
Thursday, September 14	Beckman Institute	6:00am - 11:20am, 12:40pm - 4:00pm	208	85	79
Friday, September 15	University YMCA	6:00am - 11:20am, 12:40pm - 4:00pm	208	67	58
Saturday, September 16	Presence Covenant Medical Center	7:00am - 10:50am	96	35	27
Monday, September 18	iHotel	6:00am - 11:20am	128	48	44
Tuesday, September 19	iHotel	6:00am - 11:20am	128	42	38
Wednesday, September 20	iHotel	6:00am - 11:20am	128	69	61
Thursday, September 21	University YMCA	6:00am - 11:20am	128	48	45
Friday, September 22	Beckman Institute	6:00am - 12:10pm	156	90	90
Total			**4,692**	**2,220**	**2,004**