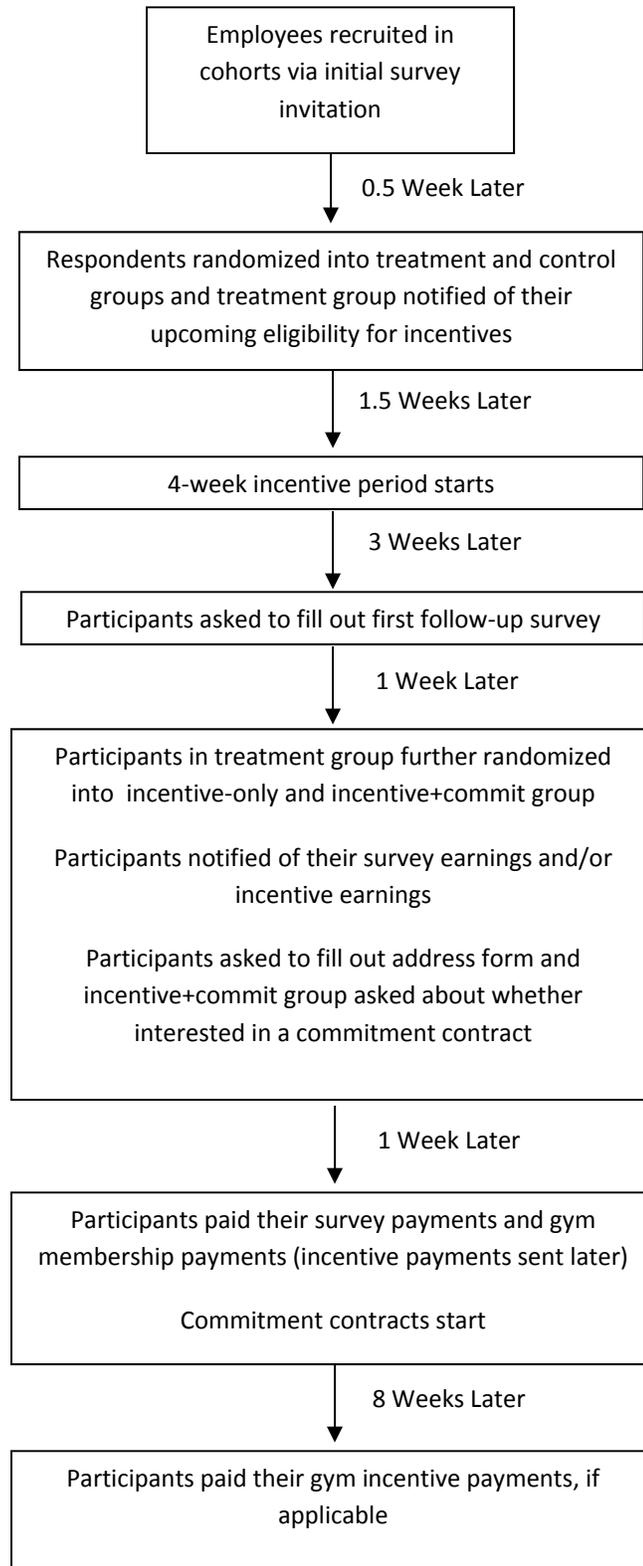


Incentives, Commitments and Habit Formation in Exercise:
Evidence from a Field Experiment with Workers at a Fortune-500 Company

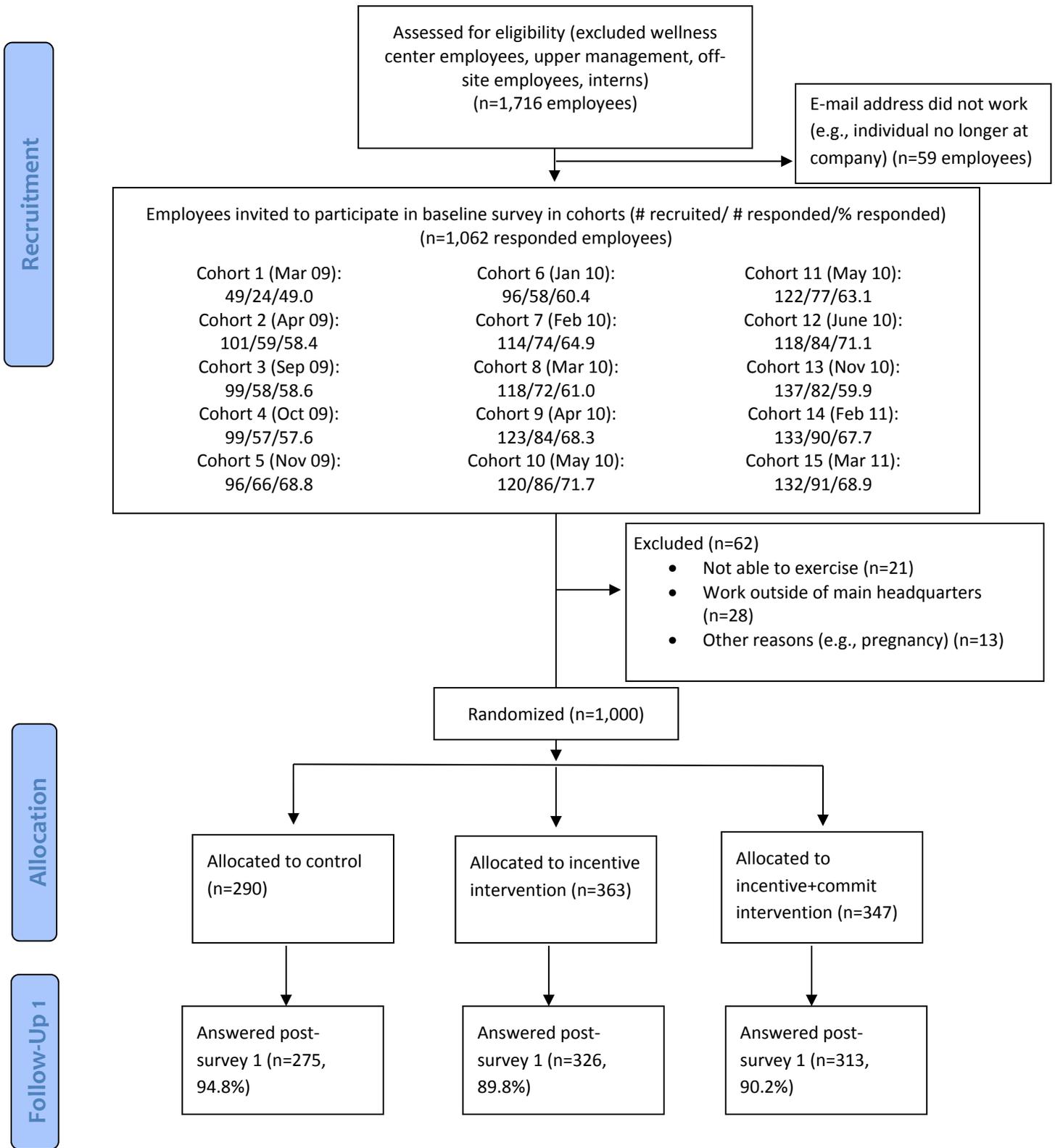
Heather Royer, Mark Stehr, and Justin Sydnor

ONLINE APPENDIX

Appendix Figure 1. Timeline



Appendix Figure 2. Flow Diagram



Appendix Table 1. Post-survey response rates as a function of treatment

Dependent variable: Indicator for whether subject responded to the post-survey

	Members			Non-Members		
	All	Below Target	Above Target	All	Below Target	Above Target
Mean for control group	0.93	0.91	0.94	0.96	0.96	0.95
Incentive-only or inc + commit	0.02 (0.03)	0.01 (0.04)	0.03 (0.04)	-0.09*** (0.02)	-0.09*** (0.02)	-0.09* (0.05)
Number of subjects	359	209	150	641	499	142

Robust standard errors in parentheses. All regression estimates control for strata. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2. OLS Regression Differences Between the Incentive+Commit Treatment and Incentive Treatment*Dependent variables: Any visit = 0/1 indicator whether individual attended gym in a given week**Weekly visits=number of visits an individual had in a given week*

	Overall		Members		Non-Members	
	Any Visit	Weekly Visits	Any Visit	Weekly Visits	Any Visit	Weekly Visits
	(1)	(2)	(3)	(4)	(5)	(6)
Control mean of dep var in pre-period	0.20	0.58	0.62	1.80	0.01	0.03
<i>Early Post-treatment (weeks 5-13)</i>						
Incentive+Commit and Incentive Difference	0.05*** (0.02)	0.07* (0.04)	0.08*** (0.03)	0.13 (0.10)	0.04** (0.02)	0.05 (0.04)
Subject-week observations	6,390	6,390	2,385	2,385	4,005	4,005
Number of subjects	710	710	265	265	445	445
<i>Late Post-treatment (weeks 14-26)</i>						
Incentive+Commit and Incentive Difference	0.03* (0.02)	0.05 (0.05)	0.06 (0.04)	0.09 (0.11)	0.02 (0.02)	0.01 (0.04)
Subject-week observations	9,230	9,230	3,445	3,445	5,785	5,785
Number of subjects	710	710	265	265	445	445

Notes: Robust standard errors clustered by individual in parentheses. All regressions only include the sample of those in the incentive or the incentive+commit treatment group. Estimates presented are the difference between the incentive+commit and incentive group during the specified time period. All regressions include strata fixed effects (i.e., study cohort x exercise above or below target fixed effects x member fixed effects (in pooled regressions)), separate week fixed effects, and either any visit fixed effects (any visit regressions) or weekly visits fixed effects (weekly visits regressions) for each week of the incentive period. For the overall regressions, separate week fixed effects for members and non-members are included. *** p<0.01, ** p<0.05, * p<0.1

Appendix A: The Substitution Effect after Accounting for Non-Response Bias

First, it is imperative to define the substitution effect. We define the substitution effect as the ratio of the effect of the incentives on overall days of weekly exercise to the effect of the incentives on days of weekly company gym exercise. A ratio of 1 indicates that there is no substitution whereas a ratio of 0 reflects complete substitution. Mathematically, we define the true substitution effect as the following:

$$\text{True substitution} = \frac{\bar{Y}_{treated}^{overall} - \bar{Y}_{control}^{overall}}{\bar{Y}_{treated}^{gym} - \bar{Y}_{control}^{gym}}$$

where $\bar{Y}_{treated}^{overall}$ is the average overall weekly days of exercise for the treated group (i.e., the incentive-only and incentive+commit group combined) during the incentive period, $\bar{Y}_{control}^{overall}$ is the average overall weekly days of exercise for the control group, $\bar{Y}_{treated}^{gym}$ is the average weekly days of company gym exercise for the treated group during the incentive period, and $\bar{Y}_{control}^{gym}$ is the average weekly days of company gym exercise for the control group. In contrast, the measured substitution effect from the post-survey data:

$$\text{Measured substitution} = \frac{\bar{Y}_{treated,r}^{overall} - \bar{Y}_{control,r}^{overall}}{\bar{Y}_{treated,r}^{gym} - \bar{Y}_{control,r}^{gym}}$$

where the means pertain to the group of survey responders (i.e., $\bar{Y}_{treated,r}^{overall}$ is the average weekly days of overall exercise for the treated group survey responders).

The measure of true substitution above can be expanded as follows:

$$\text{True Substitution} = \frac{p_{treated,r} \bar{Y}_{treated,r}^{overall} + (1 - p_{treated,r}) \bar{Y}_{treated,nr}^{overall} - (p_{control,r} \bar{Y}_{control,r}^{overall} + (1 - p_{control,r}) \bar{Y}_{control,nr}^{overall})}{p_{treated,r} \bar{Y}_{treated,r}^{gym} + (1 - p_{treated,r}) \bar{Y}_{treated,nr}^{gym} - (p_{control,r} \bar{Y}_{control,r}^{gym} + (1 - p_{control,r}) \bar{Y}_{control,nr}^{gym})}$$

where $p_{treated,r}$ is the probability of a treated individual responding to the survey and $p_{control,r}$ is the probability of a control individual responding to the survey. The nr subscripts on the averages denote non-responders (e.g., $\bar{Y}_{control,nr}^{overall}$ is the average weekly days of overall exercise for the control group non-responders). We can simplify this expression as $\bar{Y}_{treated,nr}^{gym} = 0$ and $\bar{Y}_{control,nr}^{gym} = 0$ because in our data, we observe that all non-responders did not attend the gym during the incentive period. Therefore, the true substitution effect simplifies to

$$\begin{aligned} \text{True Substitution} &= \frac{p_{treated,r} \bar{Y}_{treated,r}^{overall} + (1 - p_{treated,r}) \bar{Y}_{treated,nr}^{overall} - (p_{control,r} \bar{Y}_{control,r}^{overall} + (1 - p_{control,r}) \bar{Y}_{control,nr}^{overall})}{p_{treated,r} \bar{Y}_{treated,r}^{gym} - p_{control,r} \bar{Y}_{control,r}^{gym}} \\ &= \frac{\bar{Y}_{treated,nr}^{overall} - \bar{Y}_{control,nr}^{overall} + p_{treated,r} (\bar{Y}_{treated,r}^{overall} - \bar{Y}_{treated,nr}^{overall}) - p_{control,r} (\bar{Y}_{control,r}^{overall} - \bar{Y}_{control,nr}^{overall})}{p_{treated,r} \bar{Y}_{treated,r}^{gym} - p_{control,r} \bar{Y}_{control,r}^{gym}} \end{aligned}$$

We have data on all objects in this formula except for $\bar{Y}_{treated,nr}^{overall}$ and $\bar{Y}_{control,nr}^{overall}$. One might consider a reasonable approximation for these objects to be their pre-incentive levels from the initial survey. Using these values, we can show that our estimates of true substitution and measured substitution are close to one another; they differ by at most 10%.