Online Appendix for:

Designing Incentives for Impatient People:
An RCT Promoting Exercise to Manage Diabetes

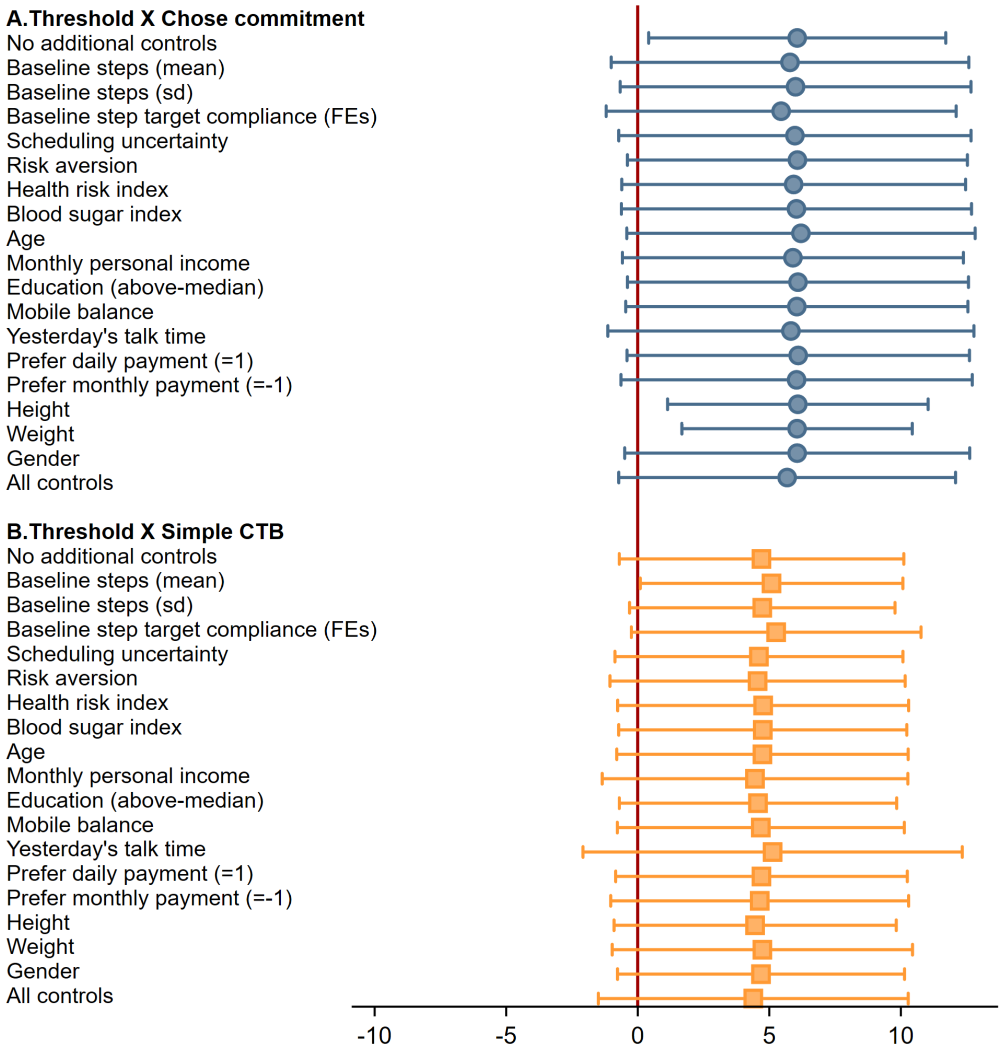
Shilpa Aggarwal
Indian School of Business

Rebecca Dizon-Ross
University of Chicago

Ariel Zucker
UC Santa Cruz

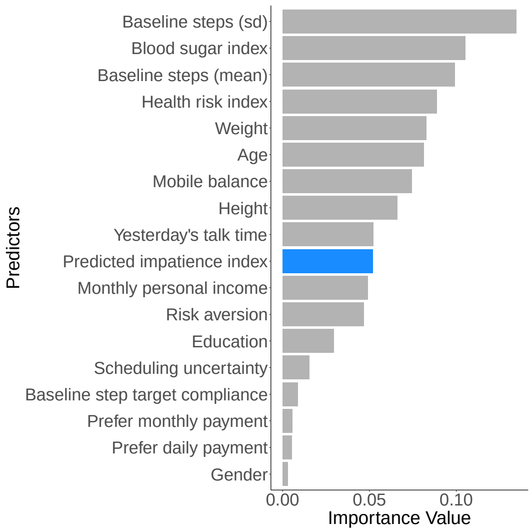
Table of Contents

F Additional Tables and Figures

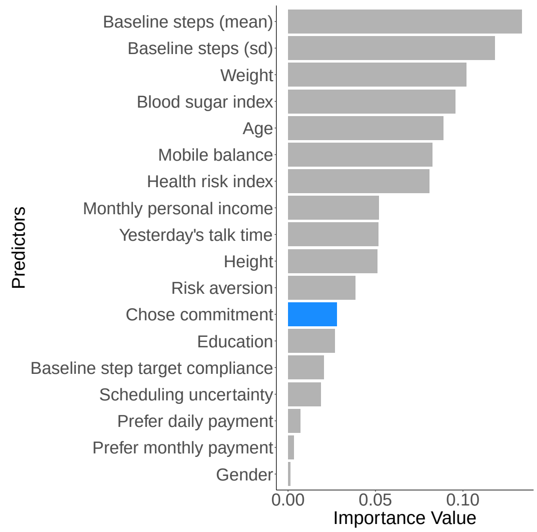


Appendix Figure F.1: Threshold Heterogeneity in Choosing Threshold and Choosing to Walk Later Is Robust to a Variety of Controls

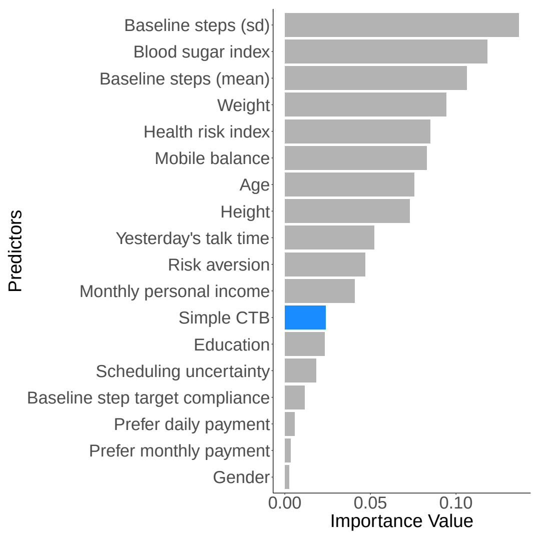
Notes: This figure replicates Figure 3 using different impatience measures. Panel A uses demand for commitment and Panel B uses simple CTB. See the notes to Table 3 for more detail on these impatience measures. All other details are the same as in Figure 3; see Figure 3 notes for more details.



(a) Predicted Impatience



(b) Chose Commitment

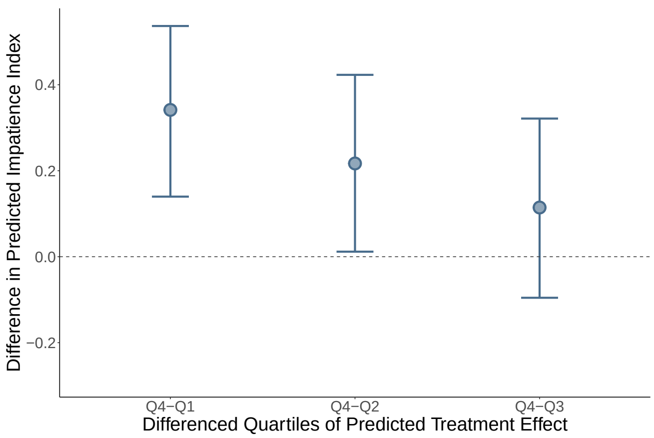


(c) Simple CTB

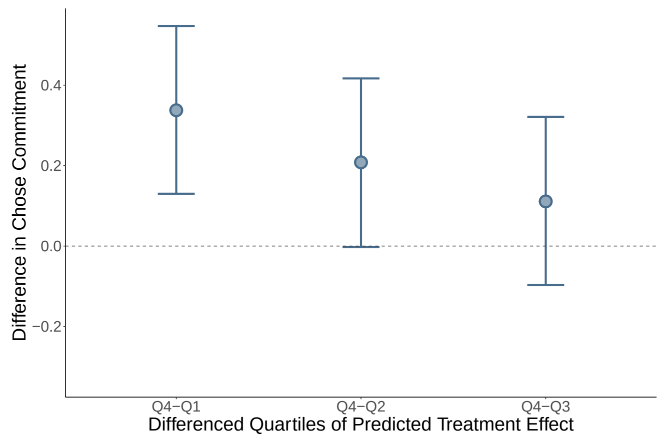
Appendix Figure F.2: Importance Results for Other Impatience Measures

Notes: This figure is analogous to Figure A.2. It displays the importance of each predictor included in a causal forest prediction of the Threshold treatment effect on average compliance at the individual level. Variable importance is a weighted sum of the number of splits on the variable of the causal forest at each depth. Predictors include the controls shown in Panel A of Figure 3, except that this analysis uses continuous versions of the baseline compliance and education variables (because the importance analysis more naturally handles continuous variables). Missing values of predictor variables are imputed with the treatment-group mean; we also include an indicator for whether each variable is missing (each of which the analysis assigned importance values of 0, and hence which we do not depict for brevity). We implement the Causal Forest using the GRF package in R (Tibshirani et al., 2023).

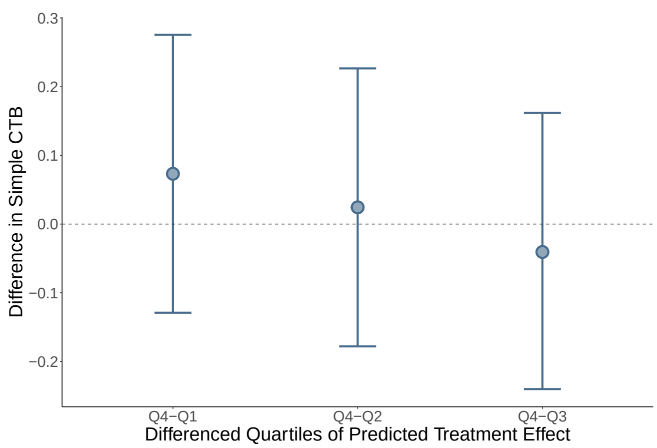
Appendix Figure F.3: Classification Analysis Results for Other Impatience Measures



(a) Difference in Predicted Impatience Across Quartiles of the Predicted Threshold Treatment Effect



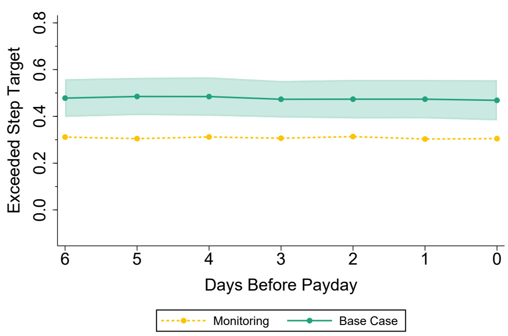
(b) Difference in Chose Commitment Across Quartiles of the Predicted Threshold Treatment Effect



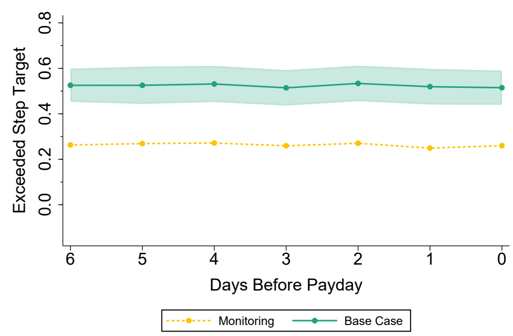
(c) Difference in Simple CTB Across Quartiles of the Predicted Threshold Treatment Effect

Notes: This figure replicates Panel (b) of Figure A.3 using the predicted impatience index, chose commitment and Simple CTB instead of the actual impatience index.

Appendix Figure F.4: No Heterogeneity by Impatience in Compliance Pattern Across the Pay-cycle



(a) Above Median Impatience Index



(b) Below Median Impatience Index

Notes: The figures show the probability of exceeding the daily 10,000-step target for the base case relative to the monitoring group, according to days remaining until payday. Each Panel is limited to above/below-median values of the impatience index. Effects control for payday day-of-week fixed effects, day-of-week fixed effects, day-of-week relative to survey day-of-week fixed effects, and the same controls as in Table 2. The shaded area represents a collection of confidence intervals from tests of equality within each daily period between the incentive and monitoring groups from regressions with the same controls as in Table 2. p-values for the test that the payday spikes are equal across above/below-median samples for each impatience measure are: Impatience index: 0.462; Predicted impatience index: 0.803; Chose commitment: 0.647; Simple CTB: 0.100.

Appendix Table F.1: Participants Understood Their Assigned Contracts

Contract Type	Question	% Correct		
		At Contract Launch	First Call	Any Call
Base Case n=902	How many recharges would you receive on (*payment day of week*) if you walked 10,000 steps on exactly 1 day over the period (*payment day of week*) to (*payment day of week*–1)?	0.99	0.99	1.00
	How many times over the course of this week would you receive recharges if you walked 10,000 steps on exactly 5 days over the period (*payment day of week*) to (*payment day of week*–1)?	1.00	1.00	1.00
4-Day Threshold n=794	What is the minimum number of days that you need to walk to get a recharge?	-	0.90	0.95
	How many recharges would you receive at the end of this week if you walked 10,000 steps on exactly 1 day this week?	0.93	0.98	1.00
	How many recharges would you receive at the end of this week if you walked 10,000 steps on exactly 4 days this week?	0.99	0.99	1.00
	How many recharges would you receive at the end of this week if you walked 10,000 steps on exactly 6 days this week?	0.98	1.00	1.00
5-Day Threshold n=312	What is the minimum number of days that you need to walk to get a recharge?	-	0.88	0.93
	How many recharges would you receive at the end of this week if you walked 10,000 steps on exactly 1 day this week?	0.91	0.96	0.99
	How many recharges would you receive at the end of this week if you walked 10,000 steps on exactly 5 days this week?	0.98	0.99	0.99
	How many recharges would you receive at the end of this week if you walked 10,000 steps on exactly 6 days this week?	0.99	0.99	0.99
Daily n=166	How many times over the course of this week would you receive recharges if you walked 10,000 steps on exactly 1 day ?	1.00	0.98	1.00
	How many times over the course of this week would you receive recharges if you walked 10,000 steps on exactly 5 days ?	-	0.99	0.99
Monthly n=164	How many recharges would you receive on (*payment day of week*) if you walked 10,000 steps on exactly 1 day over this week ?	0.99	0.99	1.00
	How many recharges would you receive on (*payment day of week*) if you walked 10,000 steps on exactly 5 days in this week ?	1.00	1.00	1.00
Monitoring n=203	How do you report your steps to us?	1.00	0.99	1.00
	How large is the Fitbit wearing bonus?	-	0.78	0.99

Notes: This table shows the share of participants who correctly answered questions about their contract. Participants were initially asked these questions when contracts were first explained ("At Contract Launch"). Questions were asked again over the phone at a later date ("First Call"). Those who answered questions incorrectly were asked again in two subsequent follow-up calls. The "Any Call" column represents the proportion of participants who got the questions right at any of these phone calls. Some questions were not asked at the initial contract launch phase. Each participant in the monthly, base case, and threshold groups was always paid on the same day of the week, which is labeled *"payment day of week"*.

Appendix Table F.2: Threshold Heterogeneity Results are Robust to Ways of Constructing the "Chose Commitment" and "Simple CTB" Measures

Dependent variable:	Exceeded step target (×100)							
Impatience measure:	Chose commitment					Simple CTB		
	Average	Either	Both	4-Day	5-Day	Average	Either	Both
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Impatience × Threshold	6.06**	5.31*	5.74*	6.32**	4.75	4.70*	5.16*	3.82
	[0.42, 11.71]	[-0.49,11.11]	[-0.19,11.68]	[0.56, 12.08]	[-3.26, 12.76]	[-0.70, 10.11]	[-0.02, 10.34]	[-1.60, 9.25]
Base Case mean	49.9	49.8	49.9	50	49.8	50.2	50.2	50.2
# Individuals	1,798	1,809	1,798	1,523	1,097	1,967	1,967	1,967
# Observations	144,099	145,005	144,099	122,277	87,990	157,799	157,799	157,799

Notes: This table shows robustness of results in columns 5 and 6 of Table 3 to different ways of constructing the Chose Commitment and Simple CTB variables. For Chose Commitment, "average" is the main specification in Table 3 and is the average of preference for 4-day and 5-day threshold contracts versus the linear contract. "Either" means preferring either 4-day or 5-day threshold, and "Both" means preferring both threshold contracts. "4-day" and "5-day" only look at the preference for 4-day and 5-day threshold respectively. For "Simple CTB", "Average" is the main specification and is the average between choosing to walk more earlier in two CTB-style walking choices, "Either" means choosing to walk earlier in either choice and "Both" means choosing to walk earlier in both choices. Controls are the same as in Table 2. The sample includes the base case and threshold groups. Data are at the individual × day level. Bootstrapped 95% confidence are in brackets. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.3: Lee Bounds on the Impacts of Incentives on Exercise

Definition of missing:	No steps data	Did not wear Fitbit	No data from Fitbit	Lost data entire period	Withdrew immediately	Mid-period withdrawal	Other reasons
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Daily steps							
Regression estimate	1269	1269	1338	1338	1338	1338	1338
(conditional on nonmissing data)	[245]	[245]	[261]	[261]	[261]	[261]	[261]
Lee lower bound	1053	882	1230	1315	1297	1226	1303
	[276]	[209]	[306]	[292]	[270]	[230]	[289]
Lee upper bound	1426	1571	1572	1351	1430	1581	1358
	[328]	[349]	[333]	[277]	[269]	[250]	[279]
B. Met 10k step target							
Regression estimate	0.223	0.223	0.205	0.205	0.205	0.205	0.205
(conditional on nonmissing data)	[0.024]	[0.024]	[0.022]	[0.022]	[0.022]	[0.022]	[0.022]
Lee lower bound	0.215	0.208	0.200	0.204	0.203	0.200	0.204
	[0.030]	[0.029]	[0.022]	[0.024]	[0.023]	[0.020]	[0.022]
Lee upper bound	0.232	0.242	0.216	0.206	0.209	0.217	0.206
	[0.030]	[0.031]	[0.022]	[0.024]	[0.024]	[0.021]	[0.022]
# Individuals	2,607	2,559	2,607	2,568	2,598	2,561	2,566
# Observations	218,988	205,732	218,988	206,488	209,008	211,551	206,320

Notes: This table reports regression estimates and Lee bounds estimates (accounting for different types of missing pedometer data) of the effect of Incentives relative to Monitoring on exercise during the intervention period. Standard errors in parentheses. The regression estimates and Lee bounds condition on data not being missing, using different definitions of missing data in each column. Regression estimates are not comparable to those reported in Table 2 because each column conditions on the "type of missing" indicator in the first row being equal to 0 and does not include controls. Data are at the individual × day level.

Appendix Table F.4: Summaries From Minute-Level Pedometer Data

	Incentives	Monitoring	I - M	p-value: I=M
	(1)	(2)	(3)	(4)
A. Activity (by minute)				
Average daily activity	213	197	17	0.001
Average steps per minute	41	38	3	0.001
B. Time of day				
Average start time	07:11	07:16	5	0.441
Average end time	20:49	20:50	1	0.742
C. High step counts per minute (share)				
Steps > 242	0	0	0	.
Steps > 150	0	0	0	0.322
# Individuals	2,368	201		

Notes: This table presents various statistics at the respondent × minute level in the incentive and monitoring groups for the days on which minute-by-minute data were available (typically 10 days of minute-wise data prior to each sync). "Average daily activity" is the average number of minutes in which a step was recorded each day. "Average steps per minute" is the average steps per minute in which at least one step was recorded. Average start/end time is the average time the first/last step was recorded by the fitbit on that day. The "High step counts per minute (share)" variables are the share of days on which we recorded steps-per-minute over the stated thresholds. High step count thresholds (242 and 150) were determined based on the average number of steps an individual takes when running at 5 mph and 8 mph, respectively. Only one individual's minute-by-minute data coincide with jogging at a pace greater than 5 miles per hour, and only for a total of 15 minutes over one day in the intervention period.

Appendix Table F.5: HbA1c and RBS are Predictive of Each Other

Dependent variable:	Endline HbA1c	Endline RBS
	(1)	(2)
Baseline HbA1c (SDs)	0.60***	0.33***
	[0.045]	[0.057]
Baseline RBS (SDs)	0.25***	0.37***
	[0.044]	[0.055]
# Individuals	560	561

Notes: This table reports estimates from regressing standardized HbA1c (column 1) and standardized RBS (column 2) at endline on standardized HbA1c and standardized RBS at baseline. Standard errors in parentheses. The sample is the control group only. Data are at the individual level. No additional controls are included. Significance levels: * 10%, ** 5%, *** 1%

Appendix Table F.6: Table 2 Results Robust to Different Controls

	No controls				Stratum fixed effects				Lasso-selected controls			
Dependent variable:	Exceeded step target	Daily steps	Daily steps (if > 0)	Earned payment when target met	Exceeded step target	Daily steps	Daily steps (if > 0)	Earned payment when target met	Exceeded step target	Daily steps	Daily steps (if > 0)	Earned payment when target met
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Pooled incentives												
Incentives	0.205***	1337.6***	1271.4***	0.950***	0.200***	1263.7***	1158.0***	0.952***	0.196***	1287.1***	1144.2***	0.952***
	[0.0224]	[261.1]	[246.1]	[0.00231]	[0.0185]	[208.7]	[188.1]	[0.00309]	[0.0180]	[211.4]	[190.3]	[0.00282]
B. Unpooled incentives												
Base Case	0.208***	1356.6***	1208.8***	1.000***	0.210***	1386.2***	1199.1***	1.006***	0.207***	1411.4***	1197.0***	1.005***
	[0.0241]	[277.0]	[258.6]	[1.62e-13]	[0.0201]	[222.0]	[199.4]	[0.00267]	[0.0196]	[225.0]	[201.8]	[0.00223]
Threshold	0.207***	1337.9***	1315.2***	0.890***	0.198***	1214.7***	1139.8***	0.892***	0.194***	1238.0***	1125.3***	0.892***
	[0.0240]	[277.1]	[259.3]	[0.00505]	[0.0199]	[220.8]	[198.0]	[0.00547]	[0.0194]	[223.2]	[200.3]	[0.00533]
Daily	0.207***	1202.7***	1363.9***	1.000***	0.200***	1120.7***	1279.2***	1.003***	0.199***	1126.7***	1245.0***	1.003***
	[0.0345]	[389.5]	[346.0]	[2.02e-13]	[0.0303]	[331.0]	[277.3]	[0.00365]	[0.0302]	[332.2]	[279.2]	[0.00296]
Monthly	0.198***	1568.6***	1482.3***	1.000***	0.177***	1265.7***	1174.2***	1.002***	0.179***	1302.6***	1152.4***	1.000***
	[0.0348]	[393.8]	[365.4]	[3.52e-13]	[0.0288]	[307.4]	[270.2]	[0.00335]	[0.0281]	[311.0]	[272.3]	[0.00271]
Small Payment	0.147***	820.5	658.5	1.000***	0.137***	728.1*	549.8	1.000***	0.128***	740.7*	510.6	0.999***
	[0.0485]	[524.0]	[477.9]	[5.58e-14]	[0.0383]	[386.1]	[334.9]	[0.00499]	[0.0382]	[381.0]	[331.3]	[0.00417]
p-value for Base Case vs												
Daily	0.980	0.630	0.570	.	0.710	0.350	0.730	0.340	0.790	0.310	0.830	0.560
Monthly	0.760	0.520	0.360	.	0.180	0.630	0.910	0.160	0.270	0.670	0.840	0.020
Threshold	0.980	0.910	0.480	<0.001	0.360	0.210	0.620	<0.001	0.350	0.210	0.550	<0.001
Small Payment	0.180	0.260	0.200	.	0.040	0.060	0.030	0.200	0.030	0.050	0.020	0.150
Monitoring mean	0.294	6,774	7,986	0	0.294	6,774	7,986	0	0.294	6,774	7,986	0
# Individuals	2,559	2,559	2,557	2,394	2,559	2,559	2,557	2,394	2,559	2,559	2,557	2,394
# Observations	205,732	205,732	180,018	99,406	205,732	205,732	180,018	99,406	205,732	205,732	180,018	99,406

Notes: This table replicates the Table 2 estimates with different sets of controls. Columns 1–4 do not use controls, columns 5–8 use the same controls as in 2 along with stratum fixed effects, and columns 9–12 use controls selected by double-Lasso. We allow lasso to select from the following list of controls: female, age, age squared, weight, weight squared, indicator for missing weight, height, height squared, indicator for missing height, yearmonth and day of week fixed effects. In addition, column 9 controls for the number of days in phase-in the target was met, its square, and an SMS treatment indicator. Columns 10–12 control for average baseline steps, average baseline steps squared, an indicator for missing baseline steps, and an SMS treatment indicator. See the notes for Table 2 for more information. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.7: Quantile Regression Estimates Show That the Linear and Threshold Contracts Similarly Impact the Distribution of Individual-Level and Weekly Compliance

Dependent variable:	Share of days met step target in intervention period			Share of days met step target in week		
Percentile:	25	50	75	25	50	75
	(1)	(2)	(3)	(4)	(5)	(6)
5-Day Threshold	0.108***	0.238***	0.353***	0.105***	0.228***	0.351***
	[0.024]	[0.031]	[0.047]	[0.020]	[0.035]	[0.053]
4-Day Threshold	0.093***	0.206***	0.323***	0.092***	0.214***	0.327***
	[0.021]	[0.026]	[0.044]	[0.017]	[0.025]	[0.051]
Base Case	0.116***	0.245***	0.336***	0.111***	0.246***	0.328***
	[0.020]	[0.025]	[0.042]	[0.016]	[0.024]	[0.051]
p-value: 5-Day vs 4-Day	.48	.28	.29	.46	.69	.25
p-value: 5-Day vs Base Case	0.705	0.792	0.523	0.712	0.575	0.266
p-value: 4-Day vs Base Case	0.201	0.107	0.560	0.152	0.157	0.948
Monitoring mean	0.292	0.469	0.723	0.294	0.523	0.801
# Individuals	2,133	2,133	2,133	2,168	2,168	2,168
# Observations	2,133	2,133	2,133	24,864	24,864	24,864

Notes: This table shows quantile regressions where the dependent variable is the share of days a participant met their step target in a given week (columns 1–3) or during the intervention period (columns 4–6). Data in columns 1–3 are at the individual × week level; in columns 4–6 they are at the individual level. The sample includes the base case, threshold, and monitoring groups. Controls are the same as in Table 2, except that, because the data are not at the individual × day level, we do not include day-of-week fixed effects. Also, in columns 1–3 we include year-month fixed effects for the first year-month of the intervention period, and in columns 4–6, we include year-month fixed effects for the first year-month of the week. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.8: Threshold Heterogeneity in Chose Commitment Is Robust to Ways of Handling "No Preference" Responses

Dependent variable:	Exceeded step target (×100)			
	Excluding no preference	No preference as Threshold	No preference as Base Case	No preference as separate
	(1)	(2)	(3)	(4)
Impatience × Threshold	6.06**	5.67*	5.52*	6.13*
	[0.42, 11.71]	[-0.49, 11.84]	[-0.23, 11.28]	[-0.27, 12.53]
Base Case mean	49.9	50.2	50.2	50.2
# Individuals	1,798	1,969	1,969	1,969
# Observations	144,099	157,946	157,946	157,946

Notes: This table shows robustness of results in column 5 of Table 3 to different ways of handling participants with no preference between the 4- or 5-day threshold and base case contract. Column 1 uses the same specification as in column 5 of Table 3 by counting no preference as missing. Column 2 counts no preference as choosing Threshold and column 3 counts no preference as choosing Base Case. Column 4 counts no preference as a separate group by adding a dummy and its interaction with the indicator for threshold treatment. Controls are as in Table 2. Bootstrapped 95% confidence are in brackets. Data are at the individual × day level. The sample includes the threshold and base case groups. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.9: Threshold Heterogeneity Results Similar with Steps as Outcome or When Analyze Threshold Groups Separately

Impatience measure:	Impatience index	Above-median impatience index	Predicted impatience index	Above-median predicted index	Chose commitment	Simple CTB
	(1)	(2)	(3)	(4)	(5)	(6)
A. Dependent variable = steps						
Impatience × Threshold	289*	576	238**	521**	580**	157
	[-41, 619]	[-254,1405]	[16, 444]	[8, 925]	[84, 1075]	[-427, 741]
Threshold	-143	-401	-166	-360**	-457***	-253
	[-442, 157]	[-939, 138]	[-367, 41]	[-615, -64]	[-697, -216]	[-671, 164]
Impatience	-209	-444	-229***	-549**	-248	-41
	[-474, 56]	[-1001, 113]	[-379, -72]	[-823, -101]	[-700, 205]	[-461, 380]
Base Case mean	8,098	8,098	8,131	8,131	8,091	8,131
B. Dependent variable = exceeded step target (×100)						
Impatience × 5-Day Threshold	3.52*	6.00	3.66***	7.29**	6.10	3.76
	[-0.05, 7.08]	[-1.73,13.73]	[1.32, 5.94]	[0.88, 11.26]	[-1.26, 13.45]	[-2.22, 9.74]
5-Day Threshold	-1.72	-4.31	-1.71	-4.42***	-4.92***	-3.79
	[-5.16, 1.73]	[-10.03,1.41]	[-4.01, 0.56]	[-7.09, -1.22]	[-8.02, -1.82]	[-8.31, 0.73]
Impatience × 4-Day Threshold	5.00*	7.95	1.76	2.51	6.20	7.10**
	[-0.94, 10.94]	[-4.13,20.02]	[-1.58, 4.58]	[-4.44, 8.64]	[-1.44, 13.83]	[0.20, 14.00]
4-Day Threshold	-0.14	-3.71	0.17	-0.84	-2.81	-3.76
	[-4.53, 4.26]	[-11.18,3.77]	[-2.96, 3.38]	[-4.54, 3.15]	[-7.86, 2.25]	[-9.01, 1.49]
Impatience	-2.97**	-5.03**	-2.39***	-5.32**	-2.37	-2.68
	[-5.36, -0.58]	[-9.96,-0.10]	[-3.90, -0.85]	[-8.15, -0.69]	[-7.23, 2.48]	[-7.31, 1.96]
Base Case mean	50.4	50.4	50.2	50.2	49.9	50.2
# Individuals	1,075	1,075	1,969	1,969	1,798	1,967
# Observations	86,215	86,215	157,946	157,946	144,099	157,799

Notes: Panel A shows that the Threshold heterogeneity reported in Table 3 is robust to using daily steps as the outcome. Panel B shows heterogeneity in the 4- and 5-day threshold treatments by impatience. The impatience measure changes across columns; its units in columns 1 and 3 are standard deviations. The sample includes the base case and threshold groups only. Specifications in columns 1 and 2 include only participants who were enrolled after we started measuring the impatience index; columns 3–6 include everyone. Threshold pools the 4- and 5-day threshold groups. Bootstrap draws were done at the individual level, and bootstrapped 95% confidence intervals are in brackets. See the notes to Table 3 for a detailed description of the bootstrap procedure. For Panel A: The Gaussian standard errors and p-values for the column 1 $Impatience \times Threshold$ coefficient are 192.76 and 0.134, respectively; for column 2 the corresponding values are 379.15 and 0.129; for column 5 the corresponding values are 300.19 and 0.054; for column 6 the values are 284.4 and 0.580 . For Panel B: The Gaussian standard errors and p-values for the column 1 $Impatience \times 5$ - day $Threshold$ coefficient are 4.16 and 0.088, respectively; for column 2 the corresponding values are 4.16 and 0.223; for column 5 the corresponding values are 4.16 and 0.088; for column 6 the values are 4.16 and 0.088 . The Gaussian standard errors and p-values for the column 1 $Impatience \times 4$ - day $Threshold$ coefficient are 3.08 and 0.223, respectively; for column 2 the corresponding values are 3.08 and 0.223; for column 5 the corresponding values are 3.08 and 0.223; for column 6 the values are 3.08 and 0.223 . Controls are the same as in Table 2. Data are at the individual × day level. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.10: No Significant Heterogeneity in Post-Intervention Persistence by Impatience

Dependent variable:	Post-endline exceeded step target (× 100)			
Impatience measure:	Impatience index		Predicted index	
Sample:	Late		Full	
	(1)	(2)	(3)	(4)
A. Exceeded step target (× 100)				
Impatience × Incentives	0.716	-0.551	-0.023	-0.893
	[-1.162,2.593]	[-1.982,0.881]	[-1.95,2.22]	[-5.88,3.56]
Impatience	-0.901	-0.241	-0.778	-0.112
	[-2.335,0.534]	[-1.533,1.052]	[-2.48,1.40]	[-3.77,5.39]
Incentives	-51.286	8.238	-50.743	8.615
	[-58.814,-43.758]	[5.311,11.166]	[5.75,11.58]	[1.25,11.13]
Baseline steps	21.030	28.474	21.017	28.516
	[14.715,27.345]	[20.958,35.991]	[20.55,34.51]	[5.77,17.34]
Intervention steps	0.392		0.390	
	[0.339,0.444]		[0.00,0.00]	
Base mean: exceeded step target	23	23	23	23
B. Average daily steps				
Impatience × Incentives	179.047	-17.391	55.594	-150.602
	[-172.712,530.806]	[-344.876,310.094]	[-376.99,279.88]	[-799.65,287.12]
Impatience	-219.229	-96.560	-128.002	15.522
	[-494.472,56.014]	[-347.841,154.720]	[-324.84,264.30]	[-366.29,694.08]
Incentives	-4.7e+03	608.318	-4.6e+03	694.355
	[-5.4e+03,-4.0e+03]	[79.120,1137.516]	[262.17,1161.73]	[-706.15,1013.48]
Baseline steps	0.191	0.475	0.189	0.472
	[0.128,0.255]	[0.421,0.528]	[0.43,0.52]	[0.45,0.54]
Intervention steps	0.630		0.628	
	[0.569,0.691]		[0.00,0.00]	
Base mean: steps	5,113	5,113	5,113	5,113
# Individuals	1,122	1,122	1,122	1,122
# Observations	91,756	91,756	91,756	91,756

Notes: This table shows heterogeneity by time preferences in persistence of treatment effects. The sample includes everyone who walked in the post-intervention period. Controls are the same as in Table 2. Base Case is the omitted group, and individual group level dummies are not reported. Because we have no intervention step data for the control group, regressions that include intervention steps only include incentive and monitoring groups. We add a missing intervention period dummy to prevent Control from dropping out of the sample. All units are standard deviations on the indexes. Data are at the individual × day level. 95% confidence intervals bootstrapped at the person level are in brackets; see the notes to 3 for more detail on the bootstrap procedure. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.11: Commitment Device Estimates Robust To Different Ways of Defining Who Chose the Threshold Contract

Dependent variable:	Exceeded step target (×100)				Paid when exceeding step target (×100)			
Definition of preferring threshold:	Both	Either	5-Day	4-Day	Both	Either	5-Day	4-Day
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Heterogeneity method								
Commitment	2.12**	1.85*	1.09	2.46**	6.78***	5.71***	7.99***	5.52***
	[1.08]	[0.99]	[1.65]	[1.10]	[0.45]	[0.41]	[0.94]	[0.47]
Threshold group mean	49.38	49.43	50.34	49.00	88.87	88.90	85.58	90.20
# Individuals	1,798	1,809	1,097	1,523	1,681	1,692	1,034	1,419
# Observations	144,099	145,005	87,990	122,277	71,525	71,944	43,929	60,564
B. Synthetic group method								
Commitment	2.13	1.82	1.03	2.35	6.65***	5.74***	8.12***	5.37***
	[1.41]	[1.40]	[2.28]	[1.59]	[0.64]	[0.65]	[1.41]	[0.67]
Threshold group mean	49.38	49.43	50.34	49.00	88.87	88.90	85.58	90.20
# Individuals	1,931	1,954	879	1,517	1,809	1,833	831	1,415
# Observations	154,336	156,334	70,202	121,438	78,031	78,736	35,331	61,307
Fraction preferring threshold	0.46	0.54	0.45	0.53	0.46	0.54	0.45	0.53

Notes: This table shows the robustness of the estimated effect of a hypothetical commitment device to different definitions of who selected each contract. The outcomes are compliance (columns 1–4) and the fraction of paid compliance (columns 5–8). We use incentivized choices between Base Case and 5- or 4-Day Thresholds to identify who prefers the threshold contract. Preferring the threshold is defined in columns 1 and 5 (2 and 6) as choosing 4-Day and (or) 5-Day Threshold over Base Case and in columns 3 and 7 (4 and 8) as preferring 5-Day (4-Day) Threshold to Base Case. Columns 1 and 5 correspond to the estimates from row 1 of Figure H.1. Panel A estimates the commitment effect relative to Threshold using Method 1 from Figure H.1, which estimates the commitment effect as the fraction of participants preferring the base case contract times the treatment effect of moving from Threshold to Base Case among participants who prefer the Base Case. The sample includes Base Case and the 4- and 5-Day Threshold in columns 1, 2, 5, and 6; Base Case and the 5-Day Threshold in columns 3 and 7; and Base Case and the 4-Day Threshold in columns 4 and 8. Panel B estimates the commitment effect relative to Threshold using Method 2 from Figure H.1, i.e., using a synthetic personalized group. See the notes for Figure H.1 for details for how the synthetic group is constructed. The sample includes the threshold and synthetic commitment groups; for columns 1, 2, and 5, 6 we use the full groups, while in columns 3 and 7 (4 and 8) we exclude the members of both groups that were randomly assigned to 4-Day (5-Day) Threshold. Controls are the same as Table 2. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.12: Walking Does Not Vary Significantly Across the Pay Cycle

Dependent variable:	Exceeded step target (×100)				
Payment frequency:	Weekly		Monthly		
	(1)	(2)	(3)	(4)	(5)
Days before payday	0.11		0.11		
	[0.09]		[0.09]		
Payday		-0.63		-0.63	
		[0.55]		[0.55]	
Payweek					-0.12
					[1.02]
Sample mean	50.19	50.19	50.19	50.19	49.28
# Individuals	890	890	890	890	163
# Observations	71,672	71,672	71,672	71,672	13,333

Notes: The columns show the effect of days until payday on the probability of meeting the step target in the base case and monthly groups; the sample in columns 1 and 2 is restricted to the base case group, and the sample in columns 3–5 is restricted to the monthly group. We control for payday day-of-week fixed effects, day-of-week fixed effects, day-of-week relative to launch survey day-of-week fixed effects, a day-of-contract-period time trend, and the same controls as in Table 2. Data are at the individual × day level. Standard errors, in brackets, are clustered at the individual level. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.13: Effect of Incentives on BMI, Blood Pressure, and Waist Circumference

Sample:	Full sample effects			Above-median baseline blood sugar effects		
	Body mass index	Mean arterial BP	Waist circumference	Body mass index	Mean arterial BP	Waist circumference
	(1)	(2)	(3)	(4)	(5)	(6)
Incentives	-0.0525	0.0884	-0.211	0.0195	-0.0811	-0.275
	[0.0409]	[0.426]	[0.284]	[0.0570]	[0.605]	[0.396]
Monitoring	0.0657	1.121	-0.0352	0.00127	-0.478	0.345
	[0.0838]	[0.739]	[0.438]	[0.0830]	[1.083]	[0.590]
Sample	Full	Full	Full	Above-median blood sugar	Above-median blood sugar	Above-median blood sugar
Control mean	26.45	103.02	94.44	26.09	103.96	94.57
# Individuals	3,058	3,056	3,059	1,527	1,529	1,525

Notes: This table shows the effect of incentives on the endline components of the health risk index not included in Table 4. Columns 4–6 restricts to the above-median blood sugar index sample. The blood sugar index is constructed as in Table 4. Controls are as described in Table 4 notes. The sample includes the incentive, monitoring, and control groups. Data are at the individual level. Standard errors are in brackets. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.14: Impacts of Incentives on Health, Robustness to Different Controls

	Full sample effects				Above-median baseline blood sugar sample effects			
	Blood sugar index	HbA1c	Random blood sugar	Health risk index	Blood sugar index	HbA1c	Random blood sugar	Health risk index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. No controls								
Incentives	-0.044	-0.068	-5.53	-0.055	-0.092*	-0.15	-11.4*	-0.13**
	[0.043]	[0.11]	[4.37]	[0.047]	[0.053]	[0.14]	[6.12]	[0.060]
Monitoring	0.0073	-0.078	4.62	0.058	-0.070	-0.30	-1.35	-0.11
	[0.074]	[0.19]	[7.94]	[0.078]	[0.088]	[0.22]	[10.8]	[0.10]
p-value: I = M	0.435	0.952	0.153	0.102	0.770	0.463	0.294	0.803
Control mean	0.00	8.44	193.83	0.00	0.64	10.09	248.26	0.45
# Individuals	3,067	3,066	3,067	3,068	1,530	1,529	1,530	1,531
Panel B. Stratum fixed effects								
Incentives	-0.05*	-0.07	-6.70*	-0.05*	-0.10**	-0.13	-14.01**	-0.09**
	[0.03]	[0.07]	[3.44]	[0.03]	[0.05]	[0.12]	[5.85]	[0.04]
Monitoring	-0.02	-0.14	2.10	0.02	-0.06	-0.31	-0.24	-0.05
	[0.05]	[0.12]	[6.36]	[0.04]	[0.08]	[0.19]	[10.37]	[0.07]
p-value: I = M	0.492	0.515	0.124	0.120	0.576	0.278	0.138	0.546
Control mean	0.00	8.44	193.83	0.00	0.64	10.09	248.26	0.45
# Individuals	3,067	3,066	3,067	3,068	1,530	1,529	1,530	1,531
Panel C. Lasso-selected controls								
Incentives	-0.05**	-0.08	-6.04*	-0.05*	-0.10**	-0.15	-11.95**	-0.08**
	[0.03]	[0.07]	[3.52]	[0.02]	[0.05]	[0.12]	[5.90]	[0.04]
Monitoring	-0.03	-0.14	1.29	0.01	-0.07	-0.33*	0.85	-0.05
	[0.05]	[0.12]	[6.61]	[0.04]	[0.08]	[0.20]	[10.48]	[0.07]
p-value: I = M	0.517	0.573	0.220	0.129	0.631	0.306	0.170	0.553
Control mean	0.00	8.44	193.83	0.00	0.64	10.09	248.26	0.45
# Individuals	3,067	3,066	3,067	3,068	1,530	1,529	1,530	1,531

Notes: This table reports the results of the specifications displayed in Table 4 with different controls. Panel A include no controls, Panel B include the same controls as 4 along with stratum fixed effects, Panel C include controls selected by double-Lasso. We allow lasso to select from the following list of controls: female, age, age squared, weight, weight squared, weight missing indicator, height, height squared, height missing indicator, completed endline survey indicator, and date and hour of endline completion fixed effects. Panel C also control for the baseline value of the outcome (or index components for indices), along with an SMS treatment indicator. Standard errors are in brackets. Data are at the individual level. The sample includes the incentive, monitoring, and control groups. *p*-value: I = M is the *p*-value for incentives vs monitoring. See Table 4 for more information on outcome variables and controls. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.15: Impact of Incentives on Fitness and Mental Health

A. Mental Health	Mental health index	Felt happy	Less nervous	Peaceful	Energy	Less blue	Less worn	Less harm to social life
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Incentives	0.095**	0.088*	0.026	0.054	0.062	0.016	0.090**	0.053
	[0.045]	[0.045]	[0.044]	[0.047]	[0.048]	[0.044]	[0.042]	[0.032]
Monitoring	0.16**	0.074	0.13	0.095	0.032	0.13*	0.17***	0.049
	[0.073]	[0.075]	[0.077]	[0.083]	[0.082]	[0.075]	[0.066]	[0.053]
p-value: M = I	0.34	0.82	0.14	0.59	0.68	0.09	0.14	0.93
Control mean	0.00	3.06	3.48	3.35	3.30	3.86	4.40	4.71
# Individuals	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068

B. Fitness	Fitness time trial index	Seconds to walk 4m	Seconds for 5 sit-stands
	(1)	(2)	(3)
Incentives	0.024	0.042	-0.10
	[0.045]	[0.043]	[0.12]
Monitoring	0.069	0.080	-0.088
	[0.077]	[0.076]	[0.19]
p-value: M = I	0.50	0.57	0.94
Control mean	0.00	3.88	13.18
# Individuals	2,890	2,825	2,793

Notes: The Mental health index averages the values of seven questions adapted from RAND's 36-Item Short Form Survey. A large value of the Fitness time trial index indicates low fitness. The sample includes the incentive, monitoring, and control groups. Controls are the same as described in the Table 4 notes, along with the same set of additional controls described in the Table F.16 notes. Robust standard errors are in brackets. Data are at the individual level. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.16: Impacts of Incentives on Diet and Addictive Consumption

A. Healthy diet

	Healthy diet index	Wheat meals	Meals with vegetables	Servings of fruit	Negative of rice meals	Negative of junkfood pieces	Negative of spoons sugar in coffee	Negative of sweets yesterday)	Avoid unhealthy food
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Incentives	0.052	0.028	0.060**	0.038	0.029	-0.020	-0.019	-0.028	0.0037
	[0.044]	[0.029]	[0.030]	[0.035]	[0.033]	[0.066]	[0.047]	[0.038]	[0.018]
Monitoring	0.023	0.019	0.082	0.062	-0.0068	0.13	-0.026	-0.048	-0.040
	[0.085]	[0.053]	[0.054]	[0.062]	[0.060]	[0.10]	[0.081]	[0.082]	[0.033]
p-value: M = I	0.71	0.85	0.66	0.68	0.51	0.08	0.92	0.80	0.14
Control mean	0.00	0.49	0.58	0.53	-2.34	-0.91	-1.12	-0.35	0.83
# Individuals	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068	3,068

B. Addictive consumption

	Addictive good consumption index	Average daily areca	Average daily alcohol	Average daily cigarettes
	(1)	(2)	(3)	(4)
Incentives	-0.014	0.034	-0.036	-0.056
	[0.037]	[0.037]	[0.028]	[0.095]
Monitoring	-0.0036	0.015	-0.016	-0.018
	[0.060]	[0.068]	[0.038]	[0.14]
p-value: M = I	0.85	0.76	0.46	0.77
Control mean	0.00	0.13	0.11	1.02
# Individuals	3,068	3,068	3,068	3,068

Notes: The Healthy Diet Index is composed of the average values of eight diet questions, standardized by their average and standard deviation in the control group; a larger value indicates a healthier diet. The Addictive Good Consumption Index is an index created by the average self-reported daily consumption of areca, alcoholic drinks, and cigarettes, standardized by their average and standard deviation in the control group; a larger value indicates higher consumption. The omitted category is Control. All specifications control for the baseline value of the dependent variable (or index components for indices), the baseline value of the dependent variable squared (or index components squared for indices), an SMS treatment indicator, and the following controls: age, weight, height, gender, and their second-order polynomials, as well as endline completion date, hour of endline completion, and dummy for late completion. Standard errors, in brackets, are clustered at the individual level. The sample includes the incentive, monitoring, and control groups. Data are at the individual level. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.17: Impact of the Base Case and Threshold Contracts on the Histogram of Weekly Compliance

Dependent variable:	Met step target exactly X times in the week							
Days	0	1	2	3	4	5	6	7
Base Case	-18.85***	-4.78***	-2.08**	1.42**	0.99	3.62***	6.86***	12.82***
	[2.37]	[1.08]	[0.88]	[0.70]	[0.79]	[0.88]	[0.81]	[1.65]
5-Day Threshold	-14.11***	-5.04***	-3.66***	-1.24	-2.09**	0.93	7.45***	17.75***
	[2.72]	[1.20]	[0.97]	[0.78]	[0.83]	[1.02]	[1.14]	[2.29]
4-Day Threshold	-13.28***	-5.27***	-3.35***	-1.25*	-0.11	2.31**	6.57***	14.38***
	[2.43]	[1.09]	[0.88]	[0.67]	[0.81]	[0.90]	[0.86]	[1.74]
p-value: Base Case = 5-Day Threshold	0.015	0.734	0.012	0.000	0.000	0.001	0.580	0.017
p-value: Base Case = 4-Day Threshold	0.000	0.385	0.007	0.000	0.031	0.038	0.692	0.276
Monitoring Mean	43.70	13.56	8.92	5.76	6.58	6.71	5.72	9.05
# Individuals	2,167	2,167	2,167	2,167	2,167	2,167	2,167	2,167
# Observations	24,721	24,721	24,721	24,721	24,721	24,721	24,721	24,721

Notes: This table shows the results from Figure D.3. The table shows regressions of an indicator for meeting the step target exactly X times in the week in Base Case, Threshold, and Monitoring. Data are at the individual × week level. Controls are the same as in Table 2, except that, because the data are at the individual × week (not individual × day) level, we exclude day-of-week fixed effects. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table F.18: Main Sample and Validation Sample Have Similar Characteristics

	Main sample	Validation sample	p-value	Norm. Diff.
	(1)	(2)	(3)	(4)
A. Demographics				
Age	49.56	50.55	0.331	-0.120
	(8.51)	(7.98)		
Female (=1)	0.42	0.41	0.815	0.028
	(0.49)	(0.50)		
Labor force participation (=1)	0.74	0.75	0.963	-0.006
	(0.44)	(0.44)		
Household size	3.91	3.73	0.370	0.126
	(1.62)	(1.08)		
B. Health				
Overweight (=1)	0.61	0.65	0.513	-0.079
	(0.49)	(0.48)		
BMI	26.42	27.06	0.226	-0.131
	(4.34)	(5.31)		
Systolic BP (mmHg)	133.38	135.83	0.290	-0.134
	(19.16)	(17.21)		
Diastolic BP (mmHg)	88.48	91.17	0.045	-0.246
	(11.10)	(10.76)		
C. Walking - phase-in				
Exceeded step target (=1)	0.25	0.21	0.321	0.116
	(0.32)	(0.34)		
Average daily steps	7004.04	6539.98	0.331	0.119
	(3981.43)	(3837.98)		
F-test for joint orthogonality				
p-value			0.48	
Sample size				
Number of individuals	3232	71		

Notes: Means are reported for each variable and standard deviations are in parentheses. Main sample is our primary experimental sample. Validation sample is the sample used to validate our impatience index as described in Appendix C. Norm. Diff. is normalized differences. All variables are as in Table 1. The number of individuals with pedometer data differs from the total number of individuals because a few participants withdrew immediately. The F-statistic is obtained by running regressions with all characteristics. Data are at the individual level.

G Misreporting Steps, Confusion, and Suspensions

Procedures to Curb Misreporting Because incentive payments were determined by self-reported data and not pedometer data, we implemented a number of checks to ensure integrity of step reporting. Within each 28-day sync period, respondents who incorrectly over-reported meeting a 10k step target on more than 25% of days were flagged for cheating and suspended from receiving recharges for 7 days, and those who over-reported on 10–25% of days were flagged for cheating but only given a warning. Those who were flagged for cheating more than once were terminated from the program. Fewer than 5% of Incentive participants were suspended for cheating and only 1 was terminated (Table G.1)

During the intervention, we also attempted to flag participants who appeared to be confused about how to read their pedometers or report properly. We flagged those whose reported steps were either more than 10% higher than their pedometer steps or more than 15% lower than their pedometer steps on 40% of days as "confused" (unless their misreporting was indicative of cheating). Those who were flagged received a refresher from the surveyors on how to use the step-reporting system. We did not require pedometer and reported steps to match exactly because our pedometers record daily steps until midnight, but respondents typically reported their daily steps before midnight. As a result, we expected pedometer and reported steps to diverge slightly, either because respondents continued to walk after reporting their steps or because respondents (incorrectly) estimated the number of additional steps they would take post-reporting, and reported that amount instead.

We also took measures to encourage regular reporting for all groups. We offered a 50 INR "pedometer wearing and reporting bonus" to participants during the pre-intervention period if they wore the pedometer and reported steps on 80% of days to ensure that all participants were familiar with the step reporting system. At contract launch, we also briefly encouraged all but Control participants to report steps regularly during the intervention period, and offered a larger 200 INR pedometer wearing and reporting bonus for wearing and reporting during the intervention period. Finally, if participants did not report for a number of consecutive days, we would send them a text message reminder to report.

Rates of Misreporting and Confusion Our analysis only uses pedometer data (not reported data), so misreporting would not bias our conclusions. However, it is still interesting to examine the prevalence of misreporting. The prevalence of misreporting, defined as reporting steps above 10,000 when the pedometer itself records fewer than 10,000 steps, is less than 5% and, interestingly, balanced across incentive and monitoring groups (column 1 of Table G.2). The balance with the monitoring group, who had no incentives to over-report, suggests that over-reporting was mainly unintentional participant mistakes. The incentive group also appeared to put more effort into making correct step reports, with fewer divergences in either the positive or the negative directions (columns 2-4 of Table G.2).

Appendix Table G.1: Summary Statistics on Audits and Suspensions

	Count		Share	
	Incentives	Monitoring	Incentives	Monitoring
	(1)	(2)	(3)	(4)
Shared Fitbit ever	3	0	0.004	0.000
Suspended for cheating	100	N/A	0.042	N/A
Terminated for cheating	1	N/A	0.000	N/A
Total:	2,404	203	0.92	0.08

Notes: We randomly audited around 1,000 individuals from both the incentive and monitoring groups to look for evidence of pedometer sharing. The first row in columns 3 and 4 is conditional on being audited.

Appendix Table G.2: Misreporting, Confusion and Cheating by Treatment Group

Variable type:	Reporting	Confusion		
Dependent variable:	Incorrectly reported over 10k steps	Over-reported or under-reported	Over-reported by at least 10%	Under-reported by at least 15%
	(1)	(2)	(3)	(4)
Incentives	0.0079	-0.081***	-0.059***	-0.022**
	[0.01]	[0.02]	[0.02]	[0.01]
Monitoring mean	0.049	0.272	0.167	0.104
# Individuals	2,542	2,542	2,542	2,542
# Observations	173,131	173,131	173,131	173,131

Notes: Each observation is a respondent × day. Column 2 shows whether a respondent over-reported by at least 10% or under-reported by at least 15%. The omitted group is the monitoring group. Controls are the same as Table 2. Standard errors, in brackets, are clustered at the individual level. The sample includes the incentive and monitoring groups. Significance levels: * 10%, ** 5%, *** 1%.

H Personalizing Time-Bundled Thresholds

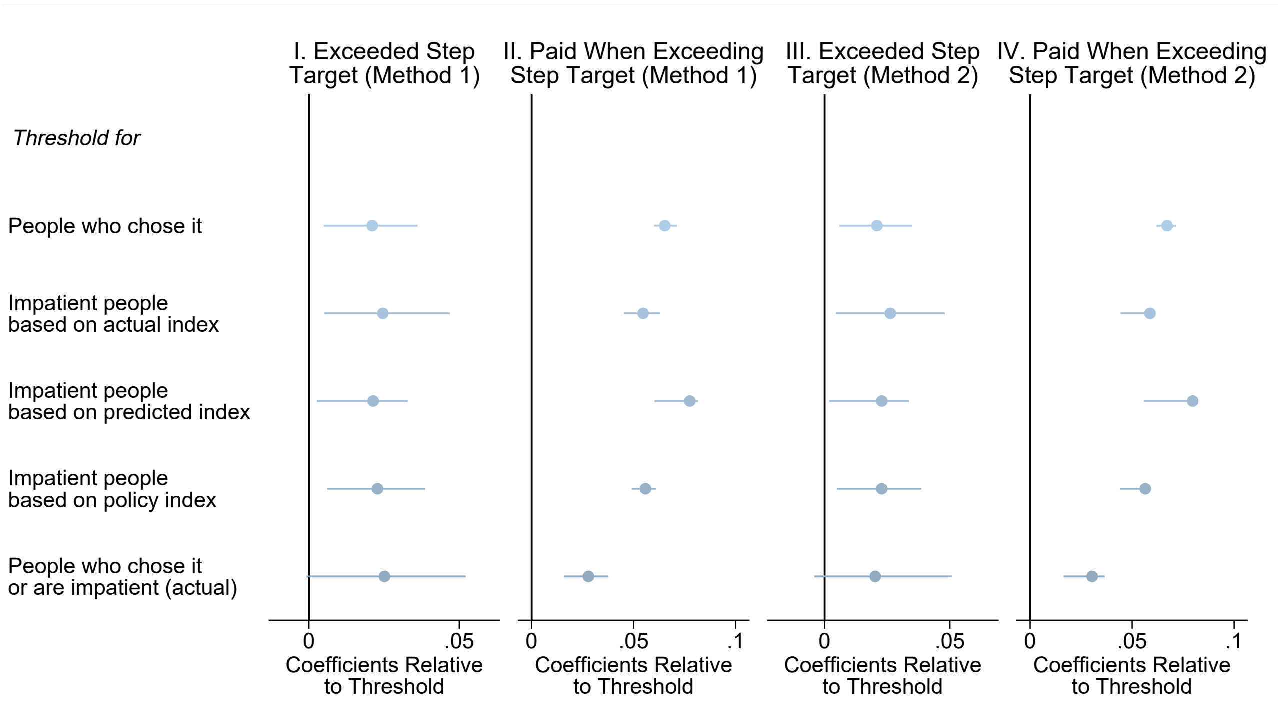
We now compare the performance of personalizing the assignment of time-bundled thresholds to assigning the threshold to everyone. While our experiment did not include personalized assignment mechanisms, we gathered impatience measures prior to randomization which, paired with random assignment, allow us to estimate how personalization would have performed. For example, relative to the threshold treatment, a commitment device that allowed participants to choose their treatments would differ in assigning the linear contract to those who preferred it. The estimated treatment effect of the commitment device relative to the threshold would thus be $p^L \times \tau^L_{BC-TH}$ where p^L is the proportion of participants who preferred the linear contract offered in the Base Case, and τ^L_{BC-TH} is the estimated treatment effect of Base Case relative to Threshold among participants who preferred the linear contract.[71]

Column I of Figure H.1 displays the estimated treatment effects of personalization (relative to assigning everyone to Threshold) on compliance.[72] The first 4 rows show the effects of personalizing with choice, the actual impatience index, the predicted impatience index, and the policy prediction (described in Section Appendix E), respectively. Because all of the impatience measures predict higher compliance in Threshold (Table 3), personalizing based on each of them significantly increase compliance by roughly 2 pp.[73] However, as shown in column II of Figure H.1, each method also significantly decreases cost-effectiveness, with effects ranging from 5-7 pp. Thus, no personalized approach unambiguously outperforms assigning everyone to the threshold. This may in part reflect the imperfection of each impatience measure. Indeed, using multiple measures to decrease exclusion errors is a promising approach – Row 5 shows that assigning the threshold to those who have above-median impatience index *or* who chose it improves cost-effectiveness without decreasing compliance relative to either measure alone.

[71] For expositional simplicity, we describe this estimate as though we implemented a single threshold contract, while in fact we implemented two. We measured preferences for each threshold contract (4- and 5-day) relative to the linear contract. Over 90% of participants either always preferred linear or always preferred threshold payment, so our main specification uses an indicator that the participant preferred *both* threshold contracts as the measure of preferring the threshold (and 1 minus that indicator as the measure of preferring linear) and uses the pooled threshold groups to calculate τ^L_{BC-TH}. The estimates are robust to two other methods: (1) using an indicator that the participant preferred *either* threshold contract, and (2) only using the 4- (or 5-) day contract preference and threshold treatment effects.

[72] Column III of Figure H.1 shows that the results are robust to a different estimation method: for each measure of impatience, we construct a "synthetic personalized group" consisting of participants who were randomly assigned the contract that they "should" have been according to that impatience measure (e.g., people with below-median actual impatience randomly assigned to the base case group). We then compare this synthetic personalized group to the threshold groups, as described in the Figure H.1 notes.

[73] The success of choice-based personalization in our setting compared to, for example, Bai et al. 2021 may reflect the relatively high demand for commitment (around 50%). This may in turn reflect that people are relatively sophisticated in the domain of walking (Dizon-Ross and Zucker, 2023).



Appendix Figure H.1: Personalizing Thresholds Increases Compliance but Decreases Cost-Effectiveness

Notes: This figure compares the effect of personalizing the assignment of linear and threshold contracts relative to assigning all participants to the threshold contract. The first row assigns the threshold to the people who prefer both threshold contracts to the linear contract, and the linear contract to those who do not. Rows 2-4 assign the threshold contract to those with above-median values of the respective index and the linear contract to those with below-median. Row 5 assigns the threshold contract to those who are assigned it in rows 1 *or* 2. Columns I and II show estimates from the method described in the main text of Section H ("Method 1"), while Columns III and IV show estimates from the "synthetic personalized group" method ("Method 2"). In Method 1, the effect of each assignment mechanism is calculated by multiplying the fraction of people assigned to the linear contract by the treatment effect of Base Case relative to Threshold in that subgroup. For Method 2, the synthetic personalized groups are constructed by duplicating the Threshold and Base Case groups and keeping only participants who preferred and were randomly assigned to the Base Case or likewise for the Threshold (Row 1), have above-median actual impatience and assigned to Threshold or below-median actual impatience and assigned to Base Case (Row 2), etc. To compare the outcomes of each synthetic personalized group with the threshold groups, each observation is weighted by the inverse of the probability of assignment to their treatment group (Base Case or Threshold) to account for over-representation of people who prefer the contract that is more frequently assigned. Confidence intervals are bootstrapped to account for the randomness in the fraction of people assigned threshold in each subgroup. All comparisons come from regressions with the same controls as Table 2. The sample sizes for rows 1-5 are 1798, 1075, 1969, 1746 and 953respectively for Method 1, and 1836, 1010, 1811, 1685 and 1021for Method 2.

I Theoretical Predictions: Additional Proofs

I.1 Proofs of Section B.2 Propositions

We begin by proving Proposition 1 for $T \geq 2$. We then prove Propositions 3, 4, and 5.

Proposition 1 ($T = K$, Threshold Compliance and Impatience Over Effort)**.** *Let $T > 1$. Fix all parameters other than $\delta^{(t)}$. Take any threshold contract with threshold level $K = T$; denote the threshold payment M. Compliance in the threshold contract will be weakly decreasing in $\delta^{(t)}$ for all $t \leq T - 1$.*

Proof. Let $V_{t,j}^{(1)}$ be the value of being on day t having complied on all previous days 1 through $t-1$, where the value is evaluated from the perspective of the agent on day $j \leq t$. Let $V_{t,j}^{(0)}$ be the value of being on day t having *not* complied on at least one of the previous days 1 through $t-1$, again evaluated from the day j perspective. And let $V_{t,j}^{(1-0)} = V_{t,j}^{(1)} - V_{t,j}^{(0)}$. Correspondingly, let $w_t(e_t, 1)$ be the compliance decision on day t if the person has effort cost e_t and has complied on all prior days, and let $w_t(e_t, 0)$ be the compliance decision on day t if the person has effort cost e_t and has not complied on all prior days. If the person has complied on all previous days, we thus have that day t compliance is as follows:

$$w_t(e_t, 1) = \begin{cases} 1 & \text{if } e_t < V_{t+1,t}^{(1-0)} \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

and as follows if the person has not complied on all previous days

$$w_t(e_t, 0) = \begin{cases} 1 & \text{if } e_t < 0 \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

We look at naifs first and then sophisticates. For both types, we begin by examining day T and then use the day T result to show results for days $t < T$. On day j, naifs think that, on day T, conditional on complying on days 1 through $T-1$, their day-T self will comply if $\delta^{(T-j)} e_T < d^{(T-j)} M$, or equivalently if $d^{(T-j)} M - \delta^{(T-j)} e_T > 0$. Their value if they comply is the discounted payment net of discounted effort costs, $d^{(T-j)} M - \delta^{(T-j)} e_T$. Hence, we have

$$\begin{aligned} V_{T,j}^{(1)} &= \mathbb{E}\left[\left(d^{(T-j)} M - \delta^{(T-j)} e_T\right) \mathbb{1}\{d^{(T-j)} M - \delta^{(T-j)} e_T > 0\} \middle| e_1, \ldots, e_j\right] \ , \ j = 1, \ldots, T \\ &= \mathbb{E}\left[\max\{d^{(T-j)} M - \delta^{(T-j)} e_T, 0\} | e_1, \ldots, e_j\right] \ , \ j = 1, \ldots, T \end{aligned} \tag{26}$$

They also think that, on any day t including T, if they haven't complied on all days through $t-1$, they will comply if $\delta^{(t-j)} e_t < 0$, which is equivalent to $e_t < 0$, which yields

$$V_{t,j}^{(0)} = \mathbb{E}\left[-\delta^{(t-j)} e_t \mathbb{1}\{e_t < 0\} | e_1, \ldots, e_j\right] \ , \ j = 1, \ldots, t \tag{27}$$

As a result, we have that:

$$V_{T,j}^{(1-0)} = \mathbb{E}\left[\max\{d^{(T-j)} M - \delta^{(T-j)} e_T, 0\} + \delta^{(T-j)} e_T \mathbb{1}\{e_T < 0\} | e_1, \ldots, e_j\right] \tag{28}$$

To show that this expectation is decreasing in $\delta^{(T-j)}$, we show that the argument, $\max\{d^{(T-j)} M - \delta^{(T-j)} e_T, 0\} + \delta^{(T-j)} e_T \mathbb{1}\{e_T < 0\}$, is decreasing in δ for all values of e_T. Consider two cases:

1. Case 1: $e_T > 0$. In this case,

$$\max\{d^{(T-j)}M - \delta^{(T-j)}e_T, 0\} + \delta^{(T-j)}e_T\mathbb{1}\{e_T < 0\} = \max\{d^{(T-j)}M - \delta^{(T-j)}e_T, 0\},$$

   which is decreasing in $\delta^{(T-j)}$.

2. Case 2: $e_T \leq 0$ In this case, letting $u = -e_T \geq 0$, we have

$$\begin{aligned}&\max\{d^{(T-j)}M - \delta^{(T-j)}e_T, 0\} + \delta^{(T-j)}e_T\mathbb{1}\{e_T < 0\}\\ &= \begin{cases}\max\{d^{(T-j)}M + \delta^{(T-j)}u, 0\} - \delta^{(T-j)}u & \text{if } \ e_T \neq 0\\ d^{(T-j)}M & \text{if } \ e_T = 0\end{cases}\\ &= d^{(T-j)}M,\end{aligned}$$

   which is invariant to $\delta^{(T-j)}$.

Thus, $\max\{d^{(T-j)}M - \delta^{(T-j)}e_T, 0\} + \delta^{(T-j)}e_T\mathbb{1}\{e_T < 0\}$ is weakly decreasing in $\delta^{(T-j)}$ for all e_t, and so, by taking expectations, equation (28) must also be decreasing in $\delta^{(T-j)}$.

In addition, on day j, naifs think that, conditional on having complied on days 1 through $t-1$, they will comply on day $t \geq j$, if $\delta^{(t-j)}e_t < V^{(1-0)}_{t+1,j}$. So, for $t \leq T-1$ we have:

$$\begin{aligned}V^{(1)}_{t,j} &= \mathbb{E}\left[\left(V^{(1-0)}_{t+1,j} - \delta^{(t-j)}e_t\right)\mathbb{1}\left\{\delta^{(t-j)}e_t < V^{(1-0)}_{t+1,j}\right\}|e_1,\ldots,e_j\right]\ ,\ j = 1,\ldots,t\\ &= \mathbb{E}\left[\max\{V^{(1-0)}_{t+1,j} - \delta^{(t-j)}e_t, 0\}|e_1,\ldots,e_j\right]\ ,\ j = 1,\ldots,t\end{aligned}$$

Combined with equation (27) this yields:

$$V^{(1-0)}_{t,j} = \mathbb{E}\left[\max\{V^{(1-0)}_{t+1,j} - \delta^{(t-j)}e_t, 0\} + \delta^{(t-j)}e_t\mathbb{1}\{\delta^{(t-j)}e_t < 0\}\middle|e_1,\ldots,e_j\right]\ ,\ j = 1,\ldots,t \quad (29)$$

Equations (28) and (29) thus recursively define all of the $V^{(1-0)}_{t,j}$ for any $t \leq T$ and $j \leq t$. Since we already showed that $V^{(1-0)}_{T,j}$ is weakly decreasing in $\delta^{(T-j)}$ for all $j \leq T$ (equation (28)), we can then use reverse induction from $t = T,\ldots,j$ using equations (28) and (29) to see that $V^{(1-0)}_{t,j}$ is decreasing in all $\delta^{(T-j)},\ldots,\delta^{(t-j)}$ for any $t \leq T$ and $j \leq t$.[74]

[74] We make the induction hypothesis that $V^{(1-0)}_{t+1,j}$ is weakly decreasing in $\delta^{(1)},\ldots,\delta^{(t)}$ and show that, under this hypothesis, $V^{(1-0)}_{t,j}$ is also weakly decreasing in $\delta^{(1)},\ldots,\delta^{(t)}$. Since we have already shown that $V^{(1-0)}_{T,j}$ is decreasing in all $\delta^{(1)},\ldots,\delta^{(T-1)}$, the result then follows. To show that $V^{(1-0)}_{t,j}$ is weakly decreasing in all $\delta^{(1)},\ldots,\delta^{(t+1)}$, we show that the argument of equation (29), $\max\{V^{(1-0)}_{t+1,j} - \delta^{(t-j)}e_t, 0\} + \delta^{(t-j)}e_t\mathbb{1}\{\delta^{(t-j)}e_t < 0\}$, is decreasing in $\delta^{(t-j)}$ for all e_t. Again there are two cases:

1. Case 1: $e_t > 0$. In this case,

$$\max\{V^{(1-0)}_{t+1,j} - \delta^{(t-j)}e_T, 0\} + \delta^{(t-j)}e_t\mathbb{1}\{e_t < 0\} = \max\{V^{(1-0)}_{t+1,j} - \delta^{(t-j)}e_T, 0\},$$

   which is weakly decreasing in $\delta^{(t-j)}$ under the induction hypothesis.

2. Case 2: $e_t \leq 0$ In this case, letting $u = -e_t \geq 0$, we have

$$\max\{V^{(1-0)}_{t+1,j} - \delta^{(t-j)}e_T, 0\} + \delta^{(t-j)}e_t\mathbb{1}\{e_t < 0\} = \max\{V^{(1-0)}_{t+1,j} + \delta^{(t-j)}u, 0\} - \delta^{(t-j)}u = V^{(1-0)}_{t+1,j},$$

   which is again weakly decreasing in $\delta^{(t-j)}$ under the induction hypothesis.

The fact that $V_{t,j}^{(1-0)}$ is decreasing in all $\delta^{(T-j)}, \ldots, \delta^{(t-j)}$ for any $t \leq T$ and $j \leq t$ shows that day t compliance is also weakly decreasing in all $\delta^{(T-t)}, \ldots, \delta^{(t-t)}$, since one complies on day t if $e_t < V_{t+1,t}^{(1-0)}$ (equation (24)). Hence, overall compliance C from days $1, \ldots, T$, is weakly decreasing in $\delta^{(1)}, \ldots, \delta^{(T-1)}$ for naifs.

Sophisticates know that, conditional on complying on all prior days, on day T they will comply if $e_T < M$. Thus, equation (28) becomes:

$$V_{T,j}^{(1-0)} = \mathbb{E}\left[\left(d^{(T-j)}M - \delta^{(T-j)}e_T\right)\mathbb{1}\left\{e_T < M\right\} + \delta^{(T-j)}e_T\mathbb{1}\{e_T < 0\}\middle|e_1, \ldots, e_j\right] \quad j = 1, \ldots, T \tag{30}$$

This is weakly decreasing in $\delta^{(T-j)}$ since the argument is weakly decreasing in $\delta^{(T-j)}$ for all e_T:

1. $e_T > 0$: In this case, $\left(d^{(T-j)}M - \delta^{(T-j)}e_T\right)\mathbb{1}\left\{e_T < M\right\} + \delta^{(T-j)}e_T\mathbb{1}\{e_T < 0\} = \left(d^{(T-j)}M - \delta^{(T-j)}e_T\right)\mathbb{1}\left\{e_T < M\right\}$, which is weakly decreasing in $\delta^{(T-j)}$.
2. $e_T \leq 0$: In this case, $\left(d^{(T-j)}M - \delta^{(T-j)}e_T\right)\mathbb{1}\left\{e_T < M\right\} + \delta^{(T-j)}e_T\mathbb{1}\{e_T < 0\} = \left(d^{(T-j)}M - \delta^{(T-j)}e_T\right) + \delta^{(T-j)}e_T = d^{(T-j)}M$, which is invariant to $\delta^{(T-j)}$.

Thus, $V_{T,j}^{(1-0)}$ is weakly decreasing in $\delta^{(T-j)}$.

Sophisticates also know that, on day $t \leq T-1$, if they have complied on all previous days, they will comply if $e_t < V_{t+1,t}^{(1-0)}$ and so equation (29) becomes:

$$V_{t,j}^{(1-0)} = \mathbb{E}\left[\left(V_{t+1,j}^{(1-0)} - \delta^{(t-j)}e_t\right)\mathbb{1}\left\{e_t < V_{t+1,t}^{(1-0)}\right\} + \delta^{(t-j)}e_t\mathbb{1}\{e_t < 0\}\middle|e_1, \ldots, e_j\right] \quad j = 1, \ldots, t \tag{31}$$

Since we showed above that $V_{T,j}^{(1-0)}$ is weakly decreasing in $\delta^{(T-j)}$ for all $j \leq T$, one can thus use equation (31) and the same reverse induction argument as for naifs to show this implies that $V_{t,j}^{(1-0)}$ is decreasing in all $\delta^{(T-j)}, \ldots, \delta^{(t-j)}$ for all $j \leq t \leq T$.[75] By the same argument used for naifs, this then implies overall compliance C is weakly decreasing in $\delta^{(1)}, \ldots, \delta^{(T-1)}$ for sophisticates. □

Proposition 3 (Perfect Correlation, Threshold Effectiveness and Impatience Over Effort)**.** *Let there be perfect correlation in costs across periods ($e_t = e_{t'} \equiv e$ for all t, t'). For simplicity, let $\delta^{(t)} < 1$ for all $t > 0$ if $\delta^{(t)} < 1$ for any t. Fix all parameters other than $\delta^{(t)}$ for some $t \leq T-1$.*

---

[75] Again the induction hypothesis is that $V_{t+1,j}^{(1-0)}$ is weakly decreasing in $\delta^{(1)}, \ldots, \delta^{(t)}$. One can then use equation (31) to show that this implies that $V_{t,j}^{(1-0)}$ is weakly decreasing in $\delta^{(1)}, \ldots, \delta^{(t)}$ because the argument, $\left(V_{t+1,j}^{(1-0)} - \delta^{(t-j)}e_t\right)\mathbb{1}\left\{e_t < V_{t+1,j}^{(1-0)}\right\} + \delta^{(t-j)}e_t\mathbb{1}\{e_t < 0\}$, is weakly decreasing in $\delta^{(1)}, \ldots, \delta^{(t)}$ for all e_t. There are two cases::

1. $e_t > 0$: In this case, $\left(V_{t+1,j}^{(1-0)} - \delta^{(t-j)}e_t\right)\mathbb{1}\left\{e_t < V_{t+1,j}^{(1-0)}\right\} + \delta^{(t-j)}e_t\mathbb{1}\{e_t < 0\} = \left(V_{t+1,j}^{(1-0)} - \delta^{(t-j)}e_t\right)\mathbb{1}\left\{e_t < V_{t+1,j}^{(1-0)}\right\}$, which is weakly decreasing in $\delta^{(t-j)}$ under the induction hypothesis.
2. $e_t \leq 0$: In this case, $\left(V_{t+1,j}^{(1-0)} - \delta^{(t-j)}e_t\right)\mathbb{1}\left\{e_t < V_{t+1,j}^{(1-0)}\right\} + \delta^{(t-j)}e_t\mathbb{1}\{e_T < 0\} = V_{t+1,j}^{(1-0)}$, which is weakly decreasing in $\delta^{(t-j)}$ under the induction hypothesis.

Since we have already shown that $V_{T,j}^{(1-0)}$ is weakly decreasing in $\delta^{(1)}, \ldots, \delta^{(T-1)}$ the result is thus shown.

Take any threshold contract with threshold level $K \leq T$. Compliance and effectiveness in the threshold contract will be weakly decreasing in $\delta^{(t)}$.

Proof. We first examine compliance and then examine effectiveness.

To gain intuition for the compliance result, first think about a person who is fully patient over both effort and payment: $\delta^{(t)} = 1$ and $d^{(t)} = 1$ for all t. That person will comply on all days if $e < m'$ (with m' the per-day reward in the threshold contract) and on no days if $e \geq m'$. In contrast, we now show that when people are impatient over effort, they often will comply even when $e > m'$.

When people are impatient, there are two cases. The first (less interesting) case is where it would be worthwhile for the agent to comply on at least K days in a separable contract paying m': $e < d^{(T-K+1)}m'$. In that case, the threshold does not "bind" and the person just complies on all days t for which $e < d^{(T-t)}m'$. Compliance is just like in the separable contract paying m' and is invariant to $\delta^{(t)}$.

The second (interesting) case is where the agent would not comply on at least K days in a separable contract paying m' ($e \geq d^{(T-K+1)}m'$) and so the threshold "binds." In this case, note that agents will never comply more than K days total.[76]

A naif who is impatient over effort (i.e., for whom $\delta^{(t)} < 1$ for all $t > 0$) will never comply before day $T-K+1$ (i.e., before the last K days). In period $T-K+1$, the naif will comply if on day $T-K+1$:

$$\sum_{t=T-K+1}^{T} \delta^{(t-(T-K+1))} e \leq d^{(K-1)} K m' \tag{32}$$

Compliance on day $T-K+1$ is thus decreasing in $\delta^{(t)}$ for all t from 1 to K. If the naif complies on day $T-K+1$, the naif will then comply on all future days. Hence, compliance is decreasing in $\delta^{(t)}$ for all t from 1 to K.

A sophisticate who is impatient over effort will always comply when a naif with the same discount rates would. In addition, the sophisticate may comply before the last K days as well.[77]

To formalize the sophisticate's conditions for compliance, consider all combinations of size K taken from the days 1 through T. There will be $\binom{T}{K}$ such combinations.[78] Order each combination chronologically and index the ordered days as days $j = 1, ..., K$ with values t_1 through t_k (e.g., if the combination is day 1 and day 3, then $t_1 = 1$ and $t_2 = 3$). A sophisticate will comply exactly K times if, for *any* of the $\binom{T}{K}$ combinations, *all* of the following K constraints

[76] Once people have reached the threshold, they will only comply on the other days if they would have complied on those days for a piece rate of m' and, since the agent would not have complied K days in a separable contract pay m', there will be no additional days that satisfy that criterion after they have reached the threshold.

[77] For example, take the case where $T = 3$ and $K = 2$. There may be cases where the individual would not find it worthwhile to comply on day 2, since $(1 + \delta^{(1)})e > 2dm'$, but would find it worthwhile to comply on day 1, since $(1 + \delta^{(2)})e < 2dm'$. In that case, the sophisticate would comply on days 1 and 3.

[78] In our example with $T = 3$ and $K = 2$, the combinations would be $1, 3$ and $2, 3$.

hold:

$$\sum_{j=1}^{K} \delta^{(t_j - t_1)} e \leq d^{(T-t_1)} K m'$$
$$\sum_{j=2}^{K} \delta^{(t_j - t_2)} e \leq d^{(T-t_2)} K m'$$
$$.....$$
$$\sum_{j=K}^{K} \delta^{(t_j - t_K)} e \leq d^{(T-t_K)} K m' \tag{33}$$

Since any of these constraints is weakly more likely to hold the lower any $\delta^{(t_j - t_1)}$, the result is thus shown for sophisticates as well.

Having shown that compliance in the threshold contract is weakly decreasing in $\delta^{(t)}$, we now just need to show that cost-effectiveness is not increasing in $\delta^{(t)}$ and the effectiveness result follows. To show this, we note that, in the perfect correlation case, regardless of $\delta^{(t)}$, any agent who complies on at least one day will always follow through to reach the threshold and achieve payment. Payments will thus be $m'C$ and cost-effectiveness will thus be $\frac{1}{m'}$ regardless of the discount factors. This is invariant to $\delta^{(t)}$. □

Proposition 4. *Let $T = 3$. Let the cost of effort on each day be binary, taking on either a "high value" (e_H) or a "low value" (e_L), with $e_H \geq e_L$. Let agents observe the full sequence of costs e_1, e_2, e_3 on day 1. Let $\delta^{(t)} = \delta^t$ (i.e., let the discount factor over effort be exponential) and let $d^{(t)} = 1$. Fix all parameters other than δ. Consider a threshold contract with $K = 2$, where the agent must thus comply on at least 2 days in order to receive payment. Compliance and effectiveness in the threshold contract are weakly higher for someone with a discount factor $\delta < 1$ than for someone with discount factor $\delta = 1$.*

Proof. We first consider different values of e_H and e_L. First, if $e_H < m'$, then $\sum_{t=1}^{3} w_t = 3$ for all δ and so the prediction trivially goes through. Second, if $e_L \geq m'$, then $\sum_{t=1}^{3} w_t = 0$ for $\delta = 1$. However, some people with $\delta < 1$ may walk in at least one period due to the standard cost-bundling effect (e.g., if they have costs of e_L every period and if $e_L + \delta e_L < 2m'$, then they would walk twice). Thus, the prediction goes through in that case as well. We thus have proved the prediction in the cases where $e_H < m'$ and $e_L \geq m'$ and so we next consider the cases where $e_H \geq m'$ and $e_L < m'$.

To prove the prediction, we examine all 8 potential sequences of costs and prove it separately for each case. Note that we only consider the cases where $e_H \geq m'$ and $e_L < m'$.

1. Cases 1 and 2: e_L, e_L, e_L and e_H, e_H, e_H. Since in these cases, costs are constant across periods, the prediction goes through by using the same arguments as in the proof for the case when costs are perfectly correlated across periods (Proposition 7b).
2. Case 3: e_H, e_H, e_L: Again, neither sophisticates nor naifs walk in period 1 but both walk in period 2 and period 3 if $e_H + \delta e_L < 2m'$ (note that by the assumptions above, since

$e_L < m'$, they will always follow-through so there is no follow-through constraint). Thus total compliance is decreasing in δ.

3. Case 4: e_H, e_L, e_H. Again, nobody walks in period 1. Sophisticates walk in periods 2 and 3 if $e_L + \delta e_H < 2m'$ and $e_H < 2m'$. Naifs walk in period 2 if $e_L + \delta e_H < 2m'$ and in period 3 if they've walked in period 2 and $e_H < 2m'$. Again, total compliance is decreasing in δ.
4. Case 5: e_L, e_H, e_H. Sophisticates walk in period 1 if $e_L + \delta^2 e_H < 2m'$ and they know they will follow through ($e_H < 2m'$). Naifs walk in period 1 if $e_L + \delta^2 e_H < 2m'$. Neither type walks in period 2 since $e_H \geq m'$. Both types walk in period 3 if they walked in period 1 and $e_H < 2m'$. Again total compliance is thus decreasing in δ.
5. Cases 6, 7, and 8: e_L, e_H, e_L; e_L, e_L, e_H; and e_H, e_L, e_L. All people, regardless of δ, walk in the two periods where the cost is e_L, since $e_L + e_L < 2m'$. Nobody walks in the period where the cost is e_H since they know they will walk in the other periods and $e_H \geq m'$. Thus, the prediction (trivially) holds.

To prove the effectiveness part of the result, we examine sophisticates first and then naifs and show that cost-effectiveness is non-increasing in δ for both types. Sophisticates will always get paid for every day they comply. Thus, regardless of δ, if compliance is non-0, cost-effectiveness will be $\frac{1}{m'}$, and hence non-increasing in δ. In contrast with sophisticates, naifs can sometimes not receive payment for a day on which they comply. In case 4, naifs will walk on day 2 if $e_L + \delta e_H < 2m'$ but not walk on day 3—and hence not be paid—if $e_H > 2m'$. Those two conditions are more likely to hold in conjunction the lower is δ. Similarly in case 5, naifs will walk on day 1 if $e_L + \delta^2 < 2m'$ but not receive payment if $e_H > 2m'$, which is again more likely to occur the lower is δ. Since having days of compliance that the principal does not have to pay for increases cost-effectiveness, this means that the lower is δ, the weakly higher cost-effectiveness is for naifs.

Hence, since we have shown that compliance is decreasing in δ whereas cost-effectiveness is non-increasing (and in particular, flat for sophisticates and weakly decreasing for naifs), then we have shown that *effectiveness* is also weakly decreasing in δ.

□

For sophisticates, we can also show a stronger result. In simulations with most realistic cost distributions, this stronger result goes through for naifs as well.

Proposition 5. *Let $T = 3$. Let costs be weakly positive and let agents observe the full sequence of costs e_1, e_2, e_3 on day 1. Let $\delta^{(t)} = \delta^t$ (i.e., let the discount factor over effort be exponential) and let $d^{(t)} = 1$. Fix all parameters other than δ. Consider a threshold contract with $K = 2$, where the agent must thus comply on at least 2 days in order to receive payment. For sophisticates, compliance and effectiveness in the threshold contract are weakly decreasing in the discount factor δ.*

Proof. We begin by examining compliance and then turn to effectiveness. For the compliance result, we first define some useful notation. Let X_t be the "walking stock" coming into period

t (i.e., sum from period 1 to period $t-1$ of whether the person complied $X_t = \sum_{i=1}^{t-1} w_i$). Let $w_t(X_t)$ be a dummy for whether the person complies in period t as a function of the walking stock coming into period t.

To examine compliance, we work backward. In period 3, behavior will depend on the walking stock X_3:

$$\begin{aligned} w_3(2) &= \mathbb{1}\{e_3 < m'\} \\ w_3(1) &= \mathbb{1}\{e_3 < 2m'\} \\ w_3(0) &= \mathbb{1}\{e_3 < 0\}. \end{aligned}$$

We show that the prediction holds by showing that it holds under all potential cases for e_3.

Case 1: $m' \leq e_3 < 2m'$ In this case, walking in period 3 is

$$\begin{aligned} w_3(2) &= 0 \\ w_3(1) &= 1 \\ w_3(0) &= 0. \end{aligned}$$

Note that this implies the person will never walk three times. Walking in period 2 is

$$\begin{aligned} w_2(1) &= \mathbb{1}\{e_2 \leq \delta e_3\} \\ w_2(0) &= \mathbb{1}\{e_2 + \delta e_3 < 2m'\}. \end{aligned}$$

In period 1, consider two cases:

1. $e_2 + \delta e_3 < 2m'$: she knows she will walk at least twice, and the only question is whether to walk now or later. If $e_1 < min\{\delta e_2, \delta^2 e_3\}$, then she will walk in period 1; if not, then she will wait and walk in periods 2 and 3. Either way, she walks twice.
2. $e_2 + \delta e_3 \geq 2m'$: she knows she will not walk later, so she will walk if $e_1 + min\{\delta e_2, \delta^2 e_3\} < 2m'$.

Thus we can see that when $m' \leq e_3 < 2m'$, overall compliance is as follows:

$$\text{Days walked} = \begin{cases} 2 & \text{if } e_2 + \delta e_3 \leq 2m' \text{ OR } e_1 + \delta min\{e_2, \delta e_3\} \leq 2m' \\ 0 & \text{otherwise.} \end{cases}$$

Thus, compliance is obviously decreasing in δ.

Case 2: $e_3 \geq 2m'$ In this case, the person will never walk in period 3 regardless of the walking stock. Thus, overall compliance is as follows:

$$\text{Days walked} = \begin{cases} 2 & \text{if } e_1 + \delta e_2 < 2m' \text{ AND } e_2 < 2m' \\ 0 & \text{otherwise.} \end{cases}$$

This is again decreasing in δ.

Case 3: $e_3 < m'$ In this case, walking in period 3 is

$$w_3(2) = 1$$
$$w_3(1) = 1$$
$$w_3(0) = 0.$$

There are two cases to consider for e_2:

1. $e_2 < m'$: in this case (for $\delta \leq 1$), discount rates do not matter since the person will walk regardless in periods 2 and 3. Then they walk in period 1 if $e_1 < m'$.
2. $e_2 \geq m'$: in this case, the person will not walk in period 2 with walking stock 1. Thus, the maximum the person will ever walk is two periods (the first or the second and then the third).

$$\text{Days walked} = \begin{cases} 2 & \text{if } (e_1 + \delta^2 e_3 < 2m' \ \& \ e_3 < 2m') \text{ or } (e_2 + \delta e_3 < 2m' \ \& \ e_3 < 2m') \\ 0 & \text{otherwise.} \end{cases}$$

Thus days walked is again weakly decreasing in δ.

Thus, we have shown the compliance portion of the result, as we have shown that compliance is weakly decreasing in δ for all potential values of e_3.

To prove the effectiveness part of the result, note that sophisticates will always get paid for every day they comply. Thus, regardless of δ, if compliance is non-0, cost-effectiveness will be $\frac{1}{m'}$. Hence, since compliance is decreasing in δ whereas cost-effectiveness is non-increasing, then effectiveness is also decreasing in δ.

□

I.2 Proofs of Section B.3 Propositions

We now provide the proofs for Propositions 6 - 8b.

Proposition 6. *Let $d = 1$ and $T = 2$. Fix all parameters other than δ, and take a linear contract that induces compliance $C > 0$.*
(a) If agents are naive and e_2 is weakly increasing in e_1, in a first order stochastic dominance sense,[79] *then for sufficiently small δ, there exists a threshold contract with $K = 2$ that has at least two times higher cost-effectiveness (and $1 + \frac{1}{C}$ times higher cost-effectiveness if costs are IID) and that generates compliance $\frac{1+C}{2}$ of the linear contract.*
(b) If agents are sophisticated and costs are IID, then for sufficiently small δ, there exists a threshold contract with $K = 2$ that has at least $1+C$ times higher cost-effectiveness and that generates compliance at least $\frac{1+C}{2}$ of the linear contract.

Proof. Take a linear contract with payment m that induces compliance $C > 0$. Equation (3) implies that compliance in a linear contract is $C = \frac{1}{T}\sum_{t=1}^{T} F(d^{(T-t)}m)$, which simplifies to $C = F(m)$when $d^{(T-t)} = 1$. Recall that the cost-effectiveness of a linear contract is $\frac{1}{m}$ (see Section 2.2).

[79]Note that this assumption flexibly accommodates the range from IID to perfect positive correlation, just ruling out negative correlation.

(a) Naifs: Consider a threshold contract that pays $M = m + \varepsilon$. On day 1, the naive agent thinks that, conditional on complying on day 1, she will comply on day 2 if $\delta e_2 < M$. The perceived probability of day 2 compliance conditional on day 1 compliance is $F_{e_2|e_1}(\frac{m+\varepsilon}{\delta})$. For $\delta \simeq 0$, $F_{e_2|e_1}(\frac{m+\varepsilon}{\delta}) \simeq 1$. Hence, for $\delta \simeq 0$, on day 1, the naive agent will comply if $e_1 + \delta E[e_2|e_1] < m + \varepsilon$; the probability of effort on day 1 thus approaches $F(m)$ as $\delta \to 0, \varepsilon \to 0$. Conditional on complying on day 1, the probability of compliance on day 2 then approaches $F_{e_2|e_1<m}(m)$. This is equal to $F(m)$ if costs are IID and is weakly greater than $F(m)$ under our more general assumption that e_2 is weakly increasing in e_1. Overall compliance is thus equal to $0.5(F(m) + F(m)F_{e_2|e_1<m}(m)) = 0.5(C + CF_{e_2|e_1<m}(m)) \geq 0.5C(1 + C)$. Expected payment per period then approaches $0.5mF(m)F_{e_2|e_1<m}(m) = 0.5mCF_{e_2|e_1<m}(m)$. Cost-effectiveness thus approaches $\frac{1}{m}\left(1 + \frac{1}{F_{e_2|e_1<m}(m)}\right) \geq 2/m$. This means the contract generates compliance of at least $(1 + C)/2$ times that of the linear contract and has at least 2 times higher cost-effectiveness. If costs are IID, $F_{e_2|e_1<m}(m) = F(m) = C$, and so cost-effectiveness approaches $\frac{1}{m}\left(1 + \frac{1}{C}\right)$, which is $1 + 1/C$ times larger than the cost-effectiveness of the linear contract.

(b) Sophisticates with IID costs: Now consider a threshold contract that pays $M = m/p' + \varepsilon$ for p' defined as a fixed point to $F(m/p') = p'$. The intermediate value theorem tells us that such a solution exists for $p' \in [C, 1]$ because F is continuous, $F(m/1) \leq 1$, and $F(m/C) \geq F(m) = C$.

Under this threshold contract, conditional on working in the first period, the probability of working in the second period is $F(M) = F(m/p' + \varepsilon) \geq F(m/p') = p'$, with $F(M) \simeq p'$ for $\varepsilon \simeq 0$. Hence, the expected payment conditional on working in the first period is $MF(M) \geq \frac{m}{p'}p' = m$, with this payment approximately m for $\varepsilon \simeq 0$. Therefore, for $\delta \simeq 0$, the probability of effort in the first period is at least $C = F(m)$, and approaches $F(m)$ for $\varepsilon \to 0, \delta \to 0$.

Taking $\varepsilon \to 0$ and then $\delta \to 0$: Total compliance in this contract is approximately $\frac{1}{2}(F(m) + F(m)F(M)) = \frac{1}{2}C(1 + p')$, with $\frac{1}{2}C(1 + p') \geq \frac{1}{2}C(1 + C)$ since $p' \geq C$. Payment per period is approximately $\frac{1}{2}MCp'$, with C the probability of working in the first period and p' the probability of working in the second period conditional on working in the first period; we have $\frac{1}{2}MCp' \simeq \frac{1}{2}\frac{m}{p'}Cp' = \frac{1}{2}mC$. Hence, cost-effectiveness is approximately $(\frac{1}{2}C(1+p'))/(\frac{1}{2}mC) = (1+p')/m \geq (1+C)/m$. □

Proposition 7a (Perfect Correlation, $M = 2m$)**.** *Let $T = 2$. Fix all parameters other than δ. Consider a linear contract with payment m and a threshold contract with payment $2m$. Then, regardless of agent type, the threshold contract is more effective than the linear contract if $\delta < 2d - 1$. If $\delta \geq 2d - 1$, then the linear contract may be more effective.*

Proof. As before, with perfect correlation, the agent takes effort either in both periods or in neither of a threshold contract. Thus the cost-effectiveness of the threshold contract will be $1/m$ and is thus the same as the cost-effectiveness of the linear contract. Therefore, whichever contract has higher compliance will be more effective. On day 1 of the linear, the agent complies if $e_1 < dm$, and on day 2 if $e_2 < m$, and so compliance in the linear contract is $\frac{1}{2}\left(F(dm) + F(m)\right) \leq F(m)$. In the threshold contract, on day 1 (and consequently day 2) the agent complies if $e_1(1 + \delta)d2m$, and so compliance is $F\left(\frac{2d}{1+\delta}m\right)$. Thus, if $\frac{2d}{1+\delta}m > m$ (i.e., if $\delta < 2d - 1$), the threshold contract has higher compliance (and hence effectiveness) than the linear. If that is not true, then the linear could have higher effectiveness. □

Proposition 7b (Perfect Correlation)**.** *Let $T = 2$. Fix all parameters other than δ, and take any linear contract that induces compliance $C > 0$. Let there be perfect correlation in costs across days*

($e_1 = e_2$). Then, regardless of agent type, there exists a threshold contract that induces compliance of at least C and that has approximately $2\frac{d}{1+\delta}$ times greater cost-effectiveness than the linear contract. Hence, if $\delta < 2d - 1$, the most effective contract will always be a threshold contract.

Proof. With perfect correlation, the agent takes effort either in both periods or in neither of a threshold contract. Therefore, as long as the agent ever exerts effort, the cost-effectiveness is equal to 2 divided by the threshold payment.

Suppose a linear contract paying m induces $C > 0$ and has cost-effectiveness $\frac{1}{m}$. Note that, because $C = \frac{1}{2}(F(dm) + F(m))$, this implies that $F(m) \geq C$.

Consider a threshold contract with payment $M = m\frac{1+\delta}{d}$. Note that this contract will have cost effectiveness of $2\frac{d}{(1+\delta)m}$, which is $2\frac{d}{(1+\delta)}$ times the cost-effectiveness of the linear contract. On day 1 (and consequently day 2), the agent complies under the threshold contract if $e_1(1+\delta) < dM$ (where the left side comes from the fact that $e_1 = e_2$). With payment $M = m\frac{1+\delta}{d}$, the agent thus complies if $e_1 < m$. Thus, the threshold contract achieves compliance of $F(m) \geq C$. □

Proposition 8a (IID Uniform, $M = 2m$)**.** *Let $d = 1$. Fix all parameters other than δ. Let costs be independently drawn each day from a uniform[0,1] distribution. Take any threshold contract paying $M < 2$ and compare it with the linear contract paying $m = \frac{M}{2}$.*
(a) If $M < 1$, the threshold contract is always more cost-effective, but whether it has higher compliance (and hence whether it is more effective) depends on δ. There is a type-specific "cutoff value" such that if δ is less than the cutoff value for a given type, then the threshold contract is more effective, as it generates greater compliance.
(b) If $1 \leq M < 2$,[80] then the threshold contract is more effective.

Proof. Note that we take the general solution for compliance and payments for threshold contracts from the proof for Proposition 8b.

For a linear contract with payment level $\frac{M}{2}$, we have:

$$C = \frac{M}{2}$$
$$P = \frac{M^2}{4}$$
$$\frac{C}{P} = \frac{2}{M}$$
$$E = \lambda\frac{M}{2} - \frac{M^2}{4}$$

Now we consider multiple cases for what the threshold contract compliance and payments would be depending on the parameters.

(a) $0 < M < 1$ We begin with naifs and then move to sophisticates. For naifs, there are two cases:

Case 1: $M < \delta$ for Naifs In this case, $E[e_2|e_2 < M/\delta] = \frac{M}{2\delta}$, giving that

$$e_1^* = (M - \delta\frac{M}{2\delta})\frac{M}{\delta} = \frac{M^2}{2\delta}$$

[80] Note that the principal would never pay $M > 2$ since $M = 2$ achieves 100% compliance regardless of δ.

Thus,

$$\begin{aligned} C &= .5\left[\frac{M^2}{2\delta}+\frac{M^3}{2\delta}\right] \\ P &= .5\frac{M^4}{2\delta} \end{aligned}$$

Thus, cost-effectiveness is:

$$\frac{C}{P}=\frac{1+M}{M^2}$$

and effectiveness is:

$$E=.5\lambda\left[\frac{M^2}{2\delta}+\frac{M^3}{2\delta}\right]-.5\frac{M^4}{2\delta}$$

The threshold has higher cost-effectiveness if:

$$\frac{2}{M}<\frac{1+M}{M^2}.$$

This holds if $2M < 1+M$ which is always true for $M < 1$. Thus, the threshold is always more cost-effective in this case.

The threshold has higher compliance if:

$$\frac{M}{2}<.5\left[\frac{M^2}{2\delta}+\frac{M^3}{2\delta}\right]$$

which simplifies to

$$\delta<\left[\frac{M}{2}+\frac{M^2}{2}\right].$$

This expression is not satisfied because $M < \delta$. Therefore, in this case, the threshold has lower compliance, and may have lower effectiveness. In fact, for $M < \delta$, whether the threshold has higher effectiveness depends on λ, the principal's marginal return to compliance: the higher λ, the more likely the threshold is to have higher effectiveness. Thus, in this range of relatively large δ we are above the cutoff value for naif types, and it is possible that the threshold will have either higher or lower effectiveness.

Case 2: $\boldsymbol{\delta < M}$ for Naifs Because $M > \delta e_2$,

$$e_1^*=E[(M-\delta e_2)\mathbb{1}\{M-\delta e_2>0\}]=E[M-\delta e_2]=M-\delta/2$$

Thus,

$$\begin{aligned} C &= .5(M-\delta/2)(1+M) \\ P &= .5(M-\delta/2)M^2 \end{aligned}$$

giving cost-effectiveness of

$$\frac{C}{P}=\frac{1+M}{M^2}$$

and effectiveness of

$$E=.5\lambda(M-\delta/2)(1+M)-.5(M-\delta/2)M^2.$$

The cost-effectiveness of the threshold contract is the same as in case 1, and so the threshold contract is again always more cost-effective.

Compliance of the threshold contract is higher than in the linear if:

$$.5M < .5(M - \frac{\delta}{2})(1 + M)$$

which simplifies to:

$$\begin{aligned}
M &< (M - \frac{\delta}{2})(1 + M) \\
M &< M(1 + M) - \frac{\delta}{2}(1 + M) \\
M &< M + M^2 - \frac{\delta}{2}(1 + M) \\
0 &< M^2 - \frac{\delta}{2}(1 + M) \\
\frac{\delta}{2}(1 + M) &< M^2 \\
\delta &< \frac{2M^2}{1 + M}
\end{aligned}$$

Note that, for $M < 1$, it is always true that $\frac{2M^2}{1+M} < M$.

Hence we can see that $\frac{2M^2}{1+M}$ is the cutoff value for naifs. For naifs, if $\delta < \frac{2M^2}{1+M}$ and $M < 1$, the threshold will always be more effective than the linear contract.

For sophisticates, there is just one case:

Case 3: $M < 1$ for Sophisticates In this case,

$$e_1^* = \left(M - \delta\frac{M}{2}\right)M = M^2(1 - \delta/2)$$

Thus,

$$\begin{aligned}
C &= .5(M^2 + M^3)(1 - \delta/2) \\
P &= .5(M^4)(1 - \delta/2)
\end{aligned}$$

Thus, cost-effectiveness is:

$$\frac{C}{P} = \frac{1 + M}{M^2}$$

The cost-effectiveness of the threshold contract is the same as in cases 1 and 2, and so, again, the threshold is always more cost-effective.

The compliance of the threshold contract is higher if:

$$.5M < .5(M^2 + M^3)(1 - \delta/2)$$

which holds if all of the following hold:

$$\begin{aligned}1 &< (M+M^2)(1-\delta/2)\\ \frac{1}{M+M^2} &< 1-\delta/2\\ \delta &< 2-\frac{2}{M+M^2}\end{aligned}$$

Thus, the cutoff value for sophisticates is $2-\frac{2}{M+M^2}$. If $\delta < 2-\frac{2}{M+M^2}$, the threshold contract is more effective. For larger δ, the linear contract may be more effective.

(b) $1 \leq M < 2$ Here naifs and sophisticates behave the same and there are two cases.

Case 4: $1 < M < 1+\delta/2$ In this case, because $M > \delta e_2$ and $M > e_2$

$$e_1^* = M - \delta/2$$

Because $M-\delta/2 < 1$,

$$\begin{aligned}C &= (M-\delta/2)\\ P &= .5M(M-\delta/2)\end{aligned}$$

giving

$$\frac{C}{P} = \frac{2}{M}.$$

This is the same cost-effectiveness as the linear contract. Hence, whichever contract has higher compliance will have higher effectiveness. Threshold compliance will be higher if:

$$\begin{aligned}M/2 &< (M-\delta/2)\\ \delta/2 &< M/2\\ \delta &< M\end{aligned}$$

which is always true assuming that $\delta \leq 1$, since $M > 1$. Hence the threshold is always more effective.

Case 5: $1+\delta/2 < M < 2$ Again, because $M > \delta e_2$ and $M > e_2$

$$e_1^* = M - \delta/2$$

Because $M-\delta/2 > 1$,

$$\begin{aligned}C &= 1\\ P &= .5M\end{aligned}$$

giving

$$\frac{C}{P} = \frac{2}{M},$$

which is again the same as the cost-effectiveness of the linear contract. Hence, the threshold will have higher effectiveness if it has higher compliance, which is true if

$$M/2 < 1,$$

which will always be the case for $M < 2$. Hence, the threshold is always more effective. $\square$

Proposition 8b (IID Uniform, Optimal Contracts). *Let $d = 1$. Fix all parameters other than δ. Let costs be independently drawn each day from a uniform[0,1] distribution. Whether the most effective threshold contract is more effective than the most effective linear contract depends on δ as well as λ, the principal's marginal return to compliance. For a wide and plausible range of values of λ,[81] there exists a "cutoff" value of δ such that the threshold contract is more effective when δ is below the cutoff, and the linear contract is more effective when δ is above the cutoff. For the remaining values of λ, either the threshold contract is always more effective, or the linear contract is always more effective, but in either case the effectiveness of the threshold relative to linear is decreasing in δ.*

Proof. We begin with a more precise statement of the result, before proceeding to prove the result. Specifically, the following describes how the effectiveness of optimal threshold contract relative to the optimal linear one depends on the value of δ in different ranges of λ values:

(a) Naifs for $0 < \lambda < 0.225$, and naifs and sophisticates for $0.225 \leq \lambda < 1$ and $3 \leq \lambda \leq 2 + \sqrt{2}$. In these cases, there is a "cutoff" value of δ such that the threshold contract is more effective when δ is below the cutoff, and the linear contract is more effective when δ is above the cutoff.

(b) Naifs and Sophisticates for $1 \leq \lambda < 3$. In this case, the threshold contract is more effective than the linear contract for all δ, with the gap decreasing in δ.

(c) Sophisticates for $\lambda < 0.225$ and naifs and sophisticates for $\lambda > 2 + \sqrt{2}$. In this case, the linear contract is always more effective, with the gap increasing in δ.

To prove the result, we begin by calculating the optimal linear and threshold contracts. For both, we proceed in two steps: we first solve for the compliance, effectiveness, and cost-effectiveness of any given linear or threshold contract, and then we solve for the optimal contract. Finally, we compare the optimal linear and threshold contracts within different ranges of λ.

Linear Contract Compliance and Effectiveness: Consider a linear contract with payment level $\frac{M}{2}$. Substituting this into the formulas from Section 2, we have the following values for compliance, daily payment, cost-effectiveness, and effectiveness, respectively:

$$C = \frac{M}{2}$$
$$P = \frac{M^2}{4}$$
$$\frac{C}{P} = \frac{2}{M}$$
$$E = \lambda\frac{M}{2} - \frac{M^2}{4}$$

[81] See the beginning of the proof for specific ranges for both naifs and sophisticates.

Optimal Linear Contract: We want to choose the payment level to maximize contract effectiveness. The first-order condition for maximizing effectiveness is:

$$\frac{\partial E}{\partial M} = \frac{\lambda}{2} - \frac{M}{2} = 0$$

Denoting the arg max as M^{L*}, the payment level in the optimal linear contract is thus:

$$M^{L*} = \lambda$$

and the effectiveness of the optimal linear contract (which we will denote as E^{L*}) is:

$$\begin{aligned} E^{L*} &= \lambda\frac{M^{L*}}{2} - \frac{M^{L*2}}{4} \\ &= \lambda\frac{\lambda}{2} - \frac{\lambda^2}{4} \\ &= \frac{\lambda^2}{4} \end{aligned}$$

Threshold Contract Compliance and Effectiveness: We begin by solving for compliance, payments, and effectiveness in a two period threshold contract with payment level M. In the two-period IID threshold case, the agent complies in period 2 if they complied in period 1 and $e_2 < M$. Moreover, equation (15) implies that the agent will comply in period 1 if:

$$e_1 < E[(M - \delta e_2)w_{2,1}|w_1 = 1]. \tag{34}$$

Let $e_1^* = E[(M - \delta e_2)w_{2,1}|w_1 = 1]$ be the maximum effort cost that results in compliance. For naifs, for whom $w_{2,1}|^{(w_1=1)} = \mathbb{1}\{M - \delta e_2 > 0\}$,

$$\begin{aligned} e_1^* &= E[(M - \delta e_2)\mathbb{1}\{M - \delta e_2 > 0\}] \\ &= E[M - \delta e_2|\delta e_2 < M] \times Prob(\delta e_2 < M) \\ &= (M - \delta E[e_2|e_2 < M/\delta])F(M/\delta) \end{aligned}$$

For sophisticates, for whom $w_{2,1}|^{(w_1=1)} = \mathbb{1}\{M - e_2 > 0\}$,

$$\begin{aligned} e_1^* &= E[(M - \delta e_2)\mathbb{1}\{M - e_2 > 0\}] \\ &= E[M - \delta e_2|e_2 < M] \times Prob(e_2 < M) \\ &= (M - \delta E[e_2|e_2 < M])F(M) \end{aligned}$$

Compliance and payments are functions of e_1^*:

$$\begin{aligned} C &= .5[F(e_1^*) + F(e_1^*)F(M)] \\ P &= .5MF(e_1^*)F(M) \end{aligned}$$

Effectiveness depends on the size of M and δ. When $\mathbf{0 < M < 1}$, we explore two cases for naifs and a single case for sophisticates based on the relative size of δ.:

Case 1: $0 < M < \delta < 1$ for Naifs In this case, $E[e_2|e_2 < M/\delta] = \frac{M}{2\delta}$, giving the following values for e_1^*, C, and P :

$$\begin{aligned}
e_1^* &= (M - \delta\frac{M}{2\delta})\frac{M}{\delta} \\
&= \frac{M^2}{2\delta} \\
C &= .5\left[\frac{M^2}{2\delta} + \frac{M^3}{2\delta}\right] \\
P &= .5\frac{M^4}{2\delta}
\end{aligned}$$

Thus, cost-effectiveness and effectiveness, respectively, are:

$$\frac{C}{P} = \frac{1+M}{M^2}$$

and

$$E = .5\lambda\left[\frac{M^2}{2\delta} + \frac{M^3}{2\delta}\right] - .5\frac{M^4}{2\delta}$$

Case 2: $0 < \delta < M < 1$ for Naifs In this case, because $M > \delta e_2$, the value e_1^* is:

$$e_1^* = E[(M - \delta e_2)\mathbb{1}\{M - \delta e_2 > 0\}] = E[M - \delta e_2] = M - \delta/2$$

This yields compliance and payments of:

$$\begin{aligned}
C &= .5(M - \delta/2)(1 + M) \\
P &= .5(M - \delta/2)M^2
\end{aligned}$$

This gives cost-effectiveness and effectiveness, respectively, of:

$$\frac{C}{P} = \frac{1+M}{M^2}$$

and

$$E = .5\lambda(M - \delta/2)(1 + M) - .5(M - \delta/2)M^2.$$

Case 3: $0 < M < 1$ for Sophisticates In this case, the value e_1^* is:

$$e_1^* = \left(M - \delta\frac{M}{2}\right)M = M^2(1 - \delta/2)$$

So compliance and payments are:

$$\begin{aligned}
C &= .5(M^2 + M^3)(1 - \delta/2) \\
P &= .5(M^4)(1 - \delta/2)
\end{aligned}$$

and cost-effectiveness and effectiveness, respectively, are:

$$\frac{C}{P} = \frac{1+M}{M^2}$$

and

$$E = .5\lambda(M^2 + M^3)(1 - \delta/2) - .5(M^4)(1 - \delta/2).$$

For larger values of M, such that $\mathbf{1 \leq M < 2}$, naifs and sophisticates behave the same way. We consider two more cases.

Case 4: $\mathbf{1 < M < 1 + \delta/2}$ for Naifs and Sophisticates In this case, because $M > \delta e_2$ and $M > e_2$, the value e_1^* is:

$$e_1^* = M - \delta/2$$

Furthermore, because $M - \delta/2 < 1$, compliance and payments are:

$$\begin{aligned} C &= (M - \delta/2) \\ P &= .5M(M - \delta/2) \end{aligned}$$

giving cost-effectiveness and effectiveness, respectively, of

$$\frac{C}{P} = \frac{2}{M}$$

and

$$E = \lambda(M - \delta/2) - .5M(M - \delta/2).$$

Case 5: $\mathbf{1 + \delta/2 < M < 2}$ for Naifs and Sophisticates Again, because $M > \delta e_2$ and $M > e_2$, the value e_1^* is:

$$e_1^* = M - \delta/2$$

Because in this case $M - \delta/2 > 1$, compliance and payments are:

$$\begin{aligned} C &= 1 \\ P &= .5M \end{aligned}$$

giving cost-effectiveness and effectiveness, respectively, of

$$\frac{C}{P} = \frac{2}{M}$$

and

$$E = \lambda - .5M.$$

Having solved for compliance, payments, and effectiveness for naifs and sophisticates and for all M between 0 and 2, we now derive the payment level of the optimal threshold contract, which we denote as M^{T*}, and its effectiveness, which we denote as E^{T*}. We first consider sophisticates and then naifs.

Optimal threshold contract for sophisticates:
Aggregating cases 3-5 above, we have that effectiveness for sophisticates is as follows:

$$E = \begin{cases} .5(1-\delta/2)\left(\lambda(M^2+M^3)-M^4\right) & \text{if } M<1 \\ \lambda(M-\delta/2)-M^2/2+\delta M/4 & \text{if } 1\le M<1+\delta/2 \\ \lambda - M/2 & \text{if } 1+\delta/2 \le M \end{cases}$$

The derivative of effectiveness with respect to the payment level M is:

$$\frac{\partial E}{\partial M} = \begin{cases} .5(1-\delta/2)\left(\lambda(2M+3M^2)-4M^3\right) & \text{if } M<1 \\ \lambda - M+\delta/4 & \text{if } 1\le M<1+\delta/2 \\ -1/2 & \text{if } 1+\delta/2 \le M \end{cases}$$

The payment level of the optimal threshold contract, M^{T*}, will set this derivative equal to zero. Note that if $1+\delta/2 \le M$, it follows that $\frac{\partial E}{\partial M} < 0$ (since $M = 1+\delta/2$ achieves full compliance). Hence, M^{T*} is always smaller than $1+\delta/2$. However, the exact value of M^{T*} depends on the value of λ. We consider three cases, (A) - (C).

Case A: $\lambda \ge 1+\delta/4$
In this case, we have that $\frac{\partial E}{\partial M}|^{1\le M<1+\delta/2} = \lambda - M + \delta/4 > 0$ for $1 \le M < 1+\delta/2$. In addition, $\frac{\partial E}{\partial M}|^{M<1} = .5(1-\delta/2)\left(\lambda(2M+3M^2)-4M^3\right)$ is always positive.[82] Combined with the fact that $\frac{\partial E}{\partial M}|^{M>1+\delta/2} < 0$, the optimal payment is:

$$M^{T*}|^{\lambda>1+\delta/4} = 1+\delta/2.$$

and the effectiveness of the optimal threshold contract is

$$\begin{aligned} E^{T*}|^{\lambda>1+\delta/4} &= \lambda - M^*/2 \\ &= \lambda - .5 - \delta/4 \end{aligned}$$

Case B: $\lambda < 1-\delta/4$
In this case, $\frac{\partial E}{\partial M}|^{1\le M<1+\delta/2} = \lambda - M + \delta/4 < 0$ for all $1 \le M < 1+\delta/2$. Recall that $\frac{\partial E}{\partial M}|^{M>1+\delta/2} < 0$ in all cases. Hence $\frac{\partial E}{\partial M}|^{M>1} < 0$, which implies that the optimum must have $M \le 1$.

We hence set the $\frac{\partial E}{\partial M}|^{M<1} = 0$, which yields:

$$\frac{\partial E}{\partial M}|^{M<1} = .5(1-\delta/2)\left(\lambda(2M+3M^2)-4M^3\right) = 0$$

which implies

$$\lambda(2M+3M^2)-4M^3 = 0$$

[82] This is because, given $\lambda \ge 1$, the function $\lambda(2M+3M^2)-4M^3$ increases at $M=0$ and is never 0 in (0,1].

or that

$$\lambda(2 + 3M) - 4M^2 = 0$$

The solution to this quadratic is:

$$M = \lambda \left(\frac{3}{8} + \sqrt{\frac{9}{64} + \frac{1}{2\lambda}} \right)$$

This M falls in the region $M < 1$ whenever $\lambda < \frac{4}{5}$. When $\lambda \geq \frac{4}{5}$, $\frac{\partial E}{\partial M}|^{M<1} > 0$ for all $M < 1$, which (combined with the fact that $\frac{\partial E}{\partial M}|^{M>1} < 0$) implies that the optimal M must be at the "kink point" where M=1:

$$M^{T*}|^{\lambda<1-\delta/4 \text{ \& } \lambda>4/5} = 1$$

Note that having $\lambda \geq \frac{4}{5}$ while $\lambda < 1 - \delta/4$ implies a relatively low δ.

Thus we have:

$$M^{T*}|^{\lambda<1-\delta/4} = \begin{cases} \lambda \left(\frac{3}{8} + \sqrt{\frac{9}{64} + \frac{1}{2\lambda}} \right) & \text{if } \lambda < 4/5 \& \lambda < 1 - \delta/4 \\ 1 & \text{if } \lambda \geq 4/5 \& \lambda < 1 - \delta/4 \end{cases}$$

This implies that maximized effectiveness when $\lambda < 4/5$ is:

$$\begin{aligned} E^{T*}|^{\lambda<1-\delta/4 \text{ \& } \lambda<4/5} &= .5(1-\delta/2)\left(\lambda(M^2+M^3) - M^4\right)|^{M=\lambda\left(\frac{3}{8}+\sqrt{\frac{9}{64}+\frac{1}{2\lambda}}\right)} \\ &= \frac{1}{16}(2-\delta)\left(\frac{3}{8}+\frac{1}{8}\sqrt{9+\frac{32}{\lambda}}\right)^2 \lambda^3 \left(4+\lambda\left(\frac{15}{16}+\frac{1}{16}\left(-9-\frac{32}{\lambda}\right)+\frac{1}{8}\sqrt{9+\frac{32}{\lambda}}\right)\right) \end{aligned}$$

When $\lambda \geq 4/5$, maximized effectiveness is:

$$\begin{aligned} E^{T*}|^{\lambda<1-\delta/4 \text{ \& } \lambda\geq4/5} &= \lambda(M-\delta/2) - M^2/2 + \delta M/4|^{M=1} \\ &= \lambda(1-\delta/2) - 1/2 + \delta/4 \\ &= \lambda - 1/2 - \delta(\lambda/2 - 1/4) \end{aligned}$$

Note that both of these are decreasing in δ (where the latter holds because $\lambda/2 - 1/4 > 0$ when $\lambda > 4/5$.

Case C: $1 - \delta/4 \leq \lambda < 1 + \delta/4$ In this case, we have that $\frac{\partial E}{\partial M}|^{1\leq M<1+\delta/2} = \lambda - M + \delta/4 = 0$ somewhere in the region of $1 \leq M < 1 + \delta/2$ — that is, there is a local max in this region.

There are two subcases.

Subcase C(i): $1 - \delta/4 \leq \lambda < 1 + \delta/4$ **and** $\lambda \geq 4/5$
If $\lambda \geq 4/5$, then $\frac{\partial E}{\partial M}|^{M<1} > 0$, which means that the optimum must be the local max in the region of $1 \leq M < 1 + \delta/2$.

We thus solve for this local maximum by finding the M at which $\frac{\partial E}{\partial M}|^{1\leq M<1+\delta/2}$ is 0:

$$\frac{\partial E}{\partial M}|^{1\leq M<1+\delta/2} = \lambda - M^* + \delta/4 = 0$$

which implies that

$$M^* = \lambda + \delta/4$$

which means that

$$\begin{aligned} E^{T*}\Big|^{1-\delta/4<\lambda<1+\delta/4 \ \& \ \lambda>4/5} &= \lambda(M^* - \delta/2) - M^{*2}/2 + \delta M^*/4 \\ &= \lambda(\lambda + \delta/4 - \delta/2) - (\lambda + \delta/4)^2/2 + \delta(\lambda + \delta/4)/4 \\ &= \lambda^2 - \lambda\delta/4 - \lambda^2/2 - \lambda\delta/4 - \delta^2/32 + \lambda\delta/4 + \delta^2/16 \\ &= \lambda^2/2 - \lambda\delta/4 + \delta^2/32 \end{aligned}$$

Note again that this is decreasing in δ for all $\lambda > 4/5$ and $\delta \leq 1$.[83]

Subcase C(ii): $1 - \delta/4 \leq \lambda < 1 + \delta/4$ **and** $\lambda < 4/5$

In this case, there are two local maxima: one when $M < 1$ and one when $1 \leq M < 1+\delta/2$. The global maximum thus is the larger of those two values:

$$E^{T*} = \max\Bigg\{ \lambda^2/2 - \lambda\delta/4 + \delta^2/32, \\ \frac{1}{16}(2-\delta)\left(\frac{3}{8} + \frac{1}{8}\sqrt{9 + \frac{32}{\lambda}}\right)^2 \lambda^3 \left(4 + \lambda\left(\frac{15}{16} + \frac{1}{16}\left(-9 - \frac{32}{\lambda}\right) + \frac{1}{8}\sqrt{9 + \frac{32}{\lambda}}\right)\right)\Bigg\}$$

We next aggregate the cases into a single solution for the effectiveness of the most effective threshold for sophisticates as a function of δ. We then compare the most effective threshold and linear contracts as δ changes. However, the solution function depends on λ.

Threshold vs. Linear Effectiveness with $\lambda \geq 4/5$.

When $\lambda \geq 4/5$, we aggregate the effectiveness function of the optimal threshold contract from cases A-C as:

$$E^{T*}|^{\lambda \geq 4/5} = \begin{cases} \lambda - 1/2 - \delta(\lambda/2 - 1/4) & \text{if } \lambda < 1 - \delta/4 \\ \lambda^2/2 - \lambda\delta/4 + \delta^2/32 & \text{if } 1 - \delta/4 \leq \lambda < 1 + \delta/4 \\ \lambda - .5 - \delta/4 & \text{if } \lambda > 1 + \delta/4. \end{cases}$$

We can rewrite effectiveness more transparently as a function of δ. If $4/5 \leq \lambda < 1$, we have:

$$E^{T*} = \begin{cases} \lambda - 1/2 - \delta(\lambda/2 - 1/4) & \text{if } \delta < 4(1 - \lambda) \\ \lambda^2/2 - \lambda\delta/4 + \delta^2/32 & \text{if } \delta \geq 4(1 - \lambda) \end{cases}$$

and if $1 \leq \lambda$, we have

$$E^{T*} = \begin{cases} \lambda^2/2 - \lambda\delta/4 + \delta^2/32 & \text{if } \delta > 4(\lambda - 1) \\ \lambda - .5 - \delta/4 & \text{if } \delta \leq 4(\lambda - 1) \end{cases}$$

[83] This is because the function $-\lambda/4 + \delta/16$ is negative for all $\lambda \geq 4/5$ as long as $\delta < 16/5$.

Note that each of these functions is continuous in δ. Moreover, because each segment is decreasing in δ, we achieve the important result: $\frac{\partial E^{T*}}{\partial \delta} < 0$. That is, the effectiveness of the most effective threshold contract is decreasing in δ.

Now we compare the effectiveness of the optimal threshold and linear contracts in the region $\lambda \geq 4/5$. First consider the case where $4/5 \leq \lambda < 1$. For $\delta < 4(1-\lambda)$, $E^{T*} > E^{L*}$ would require $\delta > \frac{\lambda^2/4-\lambda+1/2}{1/4-\lambda/2}$, but this value is greater than $4(1-\lambda)$ for $4/5 \leq \lambda < 1$. So the linear contract is always more effective if $\delta < 4(1-\lambda)$. For $\delta \geq 4(1-\lambda)$, in order for $E^{T*} > E^{L*} = \lambda^2/4$, it would require that $\lambda^2/2 - \lambda\delta/4 + \delta^2/32 > \lambda^2/4$ or $\delta < (4-2\sqrt{2})\lambda$. Since $(4-2\sqrt{2})\lambda > 1$ if $\lambda > \frac{1}{4-2\sqrt{2}} \approx 0.85$, the threshold contract will always be more effective for $\lambda > 0.85$ and $\delta \geq 4(1-\lambda)$. And then for $\lambda \leq 0.85$, which contract is more effective depends on the exact value of δ.

In case where $\lambda \geq 1$, if $\delta > 4(\lambda-1)$, $E^{T*} > E^{L*}$ would require $\delta < (4-2\sqrt{2})\lambda$, which is always true for $\lambda \geq 1$. If $\delta \leq 4(\lambda-1)$, $E^{T*} > E^{L*}$ would require $\delta < -\lambda^2+4\lambda-2$. This holds for all $\delta \in [0,1]$ if $\lambda < 3$, for some δ if $3 \leq \lambda < 2+\sqrt{2}$, and no δ if $\lambda \geq 2+\sqrt{2}$.

Threshold vs. Linear Effectiveness with $\lambda < 4/5$ Now, we write the effectiveness of the optimal threshold contract as a function of λ and δ when $\lambda < 4/5$.

Let $\xi(\lambda) = \frac{1}{16}\left(\frac{3}{8}+\frac{1}{8}\sqrt{9+\frac{32}{\lambda}}\right)^2 \lambda^3 \left(4+\lambda\left(\frac{15}{16}+\frac{1}{16}\left(-9-\frac{32}{\lambda}\right)+\frac{1}{8}\sqrt{9+\frac{32}{\lambda}}\right)\right)$. Then we have

$$E^{T*}|^{\lambda<4/5.} = \begin{cases} (2-\delta)\,\xi(\lambda) & \text{if } \lambda < 1-\delta/4 \\ \max\Big\{(2-\delta)\,\xi(\lambda), \lambda^2/2 - \lambda\delta/4 + \delta^2/32\Big\} & \text{if } 1-\delta/4 \leq \lambda \end{cases}$$

or equivalently:

$$E^{T*}|^{\lambda<4/5.} = \begin{cases} (2-\delta)\,\xi(\lambda) & \text{if } \delta < 4(1-\lambda) \\ \max\Big\{(2-\delta)\,\xi(\lambda), \lambda^2/2 - \lambda\delta/4 + \delta^2/32\Big\} & \text{if } 4(1-\lambda) \leq \delta \end{cases}$$

If $\lambda < 0.75$, then $\delta < 4(1-\lambda)$ and so we have the $E^{T*} = (2-\delta)\,\xi(\lambda)$. This function is continuous and decreasing in δ.

Threshold effectiveness will in this case be higher than linear if

$$(2-\delta)\,\xi(\lambda) > \lambda^2/4.$$

This implies that threshold effectiveness is higher if

$$\delta < 2 - \frac{\lambda^2}{4\xi(\lambda)}.$$

Since the function $2 - \frac{\lambda^2}{4\xi(\lambda)}$ is negative for $\lambda \leq 0.225$, the linear contract is always more effective for this range of λ. For $\lambda > 0.225$, there is a cutoff value for δ where the optimal threshold contract is more effective for δ below the threshold.

If $0.75 \leq \lambda < 0.8$, we need some additional analysis on the function E^{T*}. Both $(2-\delta)\,\xi(\lambda)$ and $\lambda^2/2 - \lambda\delta/4 + \delta^2/32$ are continuous for $\delta \in [0,1]$, and E^{T*} is continuous at $\delta = 4(1-\lambda)$ since $(2-\delta)\,\xi(\lambda) > \lambda^2/2 - \lambda\delta/4 + \delta^2/32$ at $\delta = 4(1-\lambda)$. Also the maximum of two continuous functions is continuous, so E^{T*} is continuous in δ. Then if $E^{T*} > E^{L*}$ when $\delta = 0$ and $E^{T*} < E^{L*}$ when $\delta = 1$, there is some threshold value of δ for which the linear and threshold contracts will have the same effectiveness, and above that the threshold will have higher effectiveness and below that the linear will. This is true as long as $\lambda > 0.225$, which holds for all λ in this interval. So again there is a cutoff value for δ below which the threshold contract is more effective.

Optimal threshold contract for naifs: Again using the formulas from the proof of Proposition 8a, we have that effectiveness is as follows:

$$E = \begin{cases} .5\lambda\left[\frac{M^2}{2\delta} + \frac{M^3}{2\delta}\right] - .5\frac{M^4}{2\delta} & \text{if } M \leq \delta \\ .5\lambda(M-\delta/2)(1+M) - .5(M-\delta/2)M^2 & \text{if } \delta < M < 1 \\ \lambda(M-\delta/2) - M^2/2 + \delta M/4 & \text{if } 1 \leq M < 1+\delta/2 \\ \lambda - M/2 & \text{if } 1+\delta/2 \leq M \end{cases}$$

The derivative of effectiveness w.r.t. M is hence:

$$\frac{\partial E}{\partial M} = \begin{cases} 0.5\lambda\left[\frac{M}{\delta} + \frac{3M^2}{2\delta}\right] - \frac{M^3}{\delta} & \text{if } M \leq \delta \\ -0.5M^2 + 0.5\lambda(1+M) + 0.5\lambda(-(\delta/2)+M) - M(-(\delta/2)+M) & \text{if } \delta < M < 1 \\ \lambda - M + \delta/4 & \text{if } 1 \leq M < 1+\delta/2 \\ -1/2 & \text{if } 1+\delta/2 \leq M \end{cases}$$

Note that this is the same as sophisticates when $M \geq 1$.

Again we derive the payment and effectiveness of the optimal contract based on the value of λ. We consider two cases, (D) and (E).

Case D: $\lambda \geq 4/5$

When $\lambda > 4/5$, $\frac{\partial E}{\partial M} > 0$ for all $M \leq \delta$, and we have the following cases:

- If $\lambda \geq \frac{1.5-\delta/2}{1.5-\delta/4}$, $\frac{\partial E}{\partial M} > 0$ for all $M < 1$ and the sophisticate results go through. Note that $\lambda \geq \frac{1.5-\delta/2}{1.5-\delta/4}$ implies $\lambda \geq 1 - \delta/4$.
- If $\lambda < 1 - \frac{\delta}{4}$, $\lambda < \frac{1.5-\delta/2}{1.5-\delta/4}$. $\frac{\partial E}{\partial M} < 0$ for $M \geq 1$ and also for some $M \in (\delta, 1)$, so there is an optimum in $(\delta, 1)$ and it is global, so the sophisticate results go through as well.
- If $1 - \frac{\delta}{4} \leq \lambda < \frac{1.5-\delta/2}{1.5-\delta/4}$, there are two local optima, one in $(\delta, 1)$ and another in $[1, 1+\frac{\delta}{2})$, so the global optimum is the maximum between the two. Also threshold efficiency is decreasing in δ for a given λ, and increasing in λ for a given δ. Also, at $\lambda = 4/5$, there is a cutoff value of δ when linear contract becomes more effective. So we can let $\delta = \frac{1.5-1.5\lambda}{1/2-\lambda/4}$, and solve for the λ value such that $E^{T*} = E^{L*}$, and the solution is $\lambda = 0.81$. So there is a cutoff value of δ for when linear contract becomes more effective if $\lambda < 0.81$; otherwise the threshold contract is always more effective.

Case E: $\lambda < 4/5$

From the discussion of sophisticates, we know in this case that if $\lambda < 1 - \delta/4$, the optimum will have $M < 1$; if $1 - \delta/4 \leq \lambda < 4/5$, there will be another local optimum in $[1, 1+\delta/2)$, and the global optimum will be the maximum between the two. Explicitly, in case $\lambda < 1 - \delta/4$, we have

$$M^* = \begin{cases} \frac{\frac{3}{4}\lambda + \sqrt{\frac{9}{16}\lambda^2 + 2\lambda}}{2} & \text{if } M^* \leq \delta \\ \frac{\lambda + \delta/2 + \sqrt{\lambda^2 - \frac{1}{2}\delta\lambda + 3\lambda + \frac{\delta^2}{4}}}{3} & \text{if } M^* > \delta \end{cases}$$

Let $\delta^* = \frac{\frac{3}{4}\lambda + \sqrt{\frac{9}{16}\lambda^2 + 2\lambda}}{2}$. This turns out to be the solution for $\delta = \frac{\lambda + \delta/2 + \sqrt{\lambda^2 - \frac{1}{2}\delta\lambda + 3\lambda + \frac{\delta^2}{4}}}{3}$, so we have

$$M^* = \begin{cases} \frac{\frac{3}{4}\lambda + \sqrt{\frac{9}{16}\lambda^2 + 2\lambda}}{2} & \text{if } \delta \geq \delta^* \\ \frac{\lambda + \delta/2 + \sqrt{\lambda^2 - \frac{1}{2}\delta\lambda + 3\lambda + \frac{\delta^2}{4}}}{3} & \text{if } \delta < \delta^* \end{cases}$$

So

$$E^{T*} = \begin{cases} .5\lambda \left[\frac{M^{*2}}{2\delta} + \frac{M^{*3}}{2\delta}\right] - .5\frac{M^{*4}}{2\delta} & \text{if } \delta \geq \delta^* \\ .5\lambda(M^* - \delta/2)(1 + M^*) - .5(M^* - \delta/2)M^{*2} & \text{if } \delta < \delta^* \end{cases}$$

When $1 - \delta/4 \leq \lambda < 4/5$, there is another optimum at $M = \lambda + \delta/4 \in [1, 1+\delta/2)$. For simplicity, let $E_1 = .5\lambda \left[\frac{M^{*2}}{2\delta} + \frac{M^{*3}}{2\delta}\right] - .5\frac{M^{*4}}{2\delta}$, $E_2 = .5\lambda(M^* - \delta/2)(1+M^*) - .5(M^* - \delta/2)M^{*2}$, and $E_3 = \frac{1}{2}\lambda - \frac{\delta\lambda}{4} - \frac{\delta^2}{16}$ which is $\lambda(M - \delta/2) - M^2/2 + \delta M/4$ evaluated at $\lambda + \delta/4$. We have

$$E^{T*} = \begin{cases} \max\{E_1, E_3\} & \text{if } \delta \geq \delta^* \\ \max\{E_2, E_3\} & \text{if } \delta < \delta^* \end{cases}$$

We know that E^{T*} is continuous in δ since the maximum of a function is continuous in the parameter if it is maximized on a compact domain, and in this case we are considering $M \in [0, 1 + \delta/2]$. So we can compare E^{T*} and E^{L*} by analyzing their values at $\delta = 0$ and $\delta = 1$. If $E^{T*} < E^{L*}$ at one endpoint and $E^{T*} > E^{L*}$ at another, we can conclude that there is a cutoff δ where threshold contract becomes more effective beyond.

If $\delta = 0$, then $E = 0.5\lambda(M - \delta/2)(1 + M) - 0.5(M - \delta/2)M^2$. This function is maximized on the region from $0 < M < 1$ (i.e., $\frac{\partial E}{\partial M} = 0$) when $M = \frac{1}{6}\left(\delta + 2\lambda + \sqrt{\delta^2 + 12\lambda - 2\delta\lambda + 4\lambda^2}\right)$. The corresponding maximized value of effectiveness is greater than the effectiveness of the optimal linear contract, $\lambda^2/4$, when $\delta = 0$, for all $\lambda > 0$.

If $\delta = 1$, then $E = \max\{E_1, E_3\}$, which is less than the effectiveness of the optimal linear contract, $\lambda^2/4$ for all λ.

Hence, we have that maximized effectiveness from the threshold is greater than maximized effectiveness from the linear, $E^{T*} > E^{L*}$, when $\delta = 0$, while the opposite is true, $E^{T*} < E^{L*}$, when $\delta = 1$. Since maximized effectiveness is continuous in δ,[84] this implies that there is a cutoff δ for which the effectiveness of the optimal threshold is the same as the effectiveness of the optimal linear, and that the effectiveness of the optimal threshold is above the linear for δ below the threshold (and vice verse for δ above the threshold).

□

[84] This follows because $.5\lambda \left[\frac{M^2}{2\delta} + \frac{M^3}{2\delta}\right] - .5\frac{M^4}{2\delta} = .5\lambda(M - \delta/2)(1 + M) - .5(M - \delta/2)M^2$ when $M = \delta$.

J CTB Time Preference Measurement

We adapted the convex time budget (CTB) methodology of Andreoni and Sprenger (2012a) to try to measure time preferences in two domains, walking and mobile recharges. Unfortunately, it did not work for either domain. As a result, we do not use the full CTB measures for analysis and instead use the simple versions of CTB described in Section 4.2. In Section J.1 we summarize why we believe our full CTB measurement was not a reliable measure of time preferences in this setting. In Section J.2 we briefly summarize evidence that the Simple CTB measures worked better. Section J.3 further expands upon section J.1 and provides additional evidence.

J.1 Performance of the Full CTB

We believe our implementation of the full CTB methodology of Andreoni and Sprenger (2012a) was unsuccessful because respondents did not understand it. The complex methodology was difficult to explain to our participants, who had limited familiarity with screens, sliders, or complicated exercises. Due to survey length constraints, we also included fewer questions (and gave less practice) than previous laboratory studies.

A number of patterns in the data suggest that participant understanding was limited. First, law of demand violations are far more common than in previous studies.[85] As shown in Table J.1, 57% of the sample violated the law of demand at least once. For reference, participants in Augenblick et al. (2015) had 16 opportunities to violate monotonicity, while ours had just 2. If understanding were similar in both contexts one would expect a higher share of the Augenblick et al. (2015) sample to ever violate the law of demand, but the share in their sample was only 16%.

Second, the CTB estimates do not correlate with any of the behaviors one would expect them to. The CTB estimates in the steps/effort domain do not correlate with exercise and health, and the estimates in the recharge domain do not correlate consistently with our proxies for impatience over recharges (e.g., balances).

Finally, there are a number of other problems with the full CTB data, such as low follow-through on the incentivized activity and low convergence of the parameters. We describe these issues in more depth in Section J.3.

For all of these reasons, we do not think our CTB estimates are a reliable measure of discount rates in this setting and do not use them for analysis.

J.2 Performance of the Simple CTB

The Simple CTB measures seem to have performed better than the full CTB exercise. For example, only 18% of the participants had any law-of-demand violations in these simpler questions, much lower than the 57% in the full CTB, even though participants had the same number of opportunities for violations in both question sets. The 18% estimate is much closer to the 16% found in Augenblick et al. (2015). The percent of future-biased choices (19%) is

[85] We can only examine law of demand violations in the effort domain because we did not include exchange rate variation in the recharge domain, so cannot estimate the demand curve.

Appendix Table J.1: Law of Demand Violations in Effort Allocations

	# of violators	% of sample
	(1)	(2)
Violates 0/7	1,318	41.3
Violates 7/14	1,493	46.8
Violates at least once	1,805	56.6
Violates both	1,006	31.5
Total:	3,232	100

Notes: This table summarizes law of demand violations in the full CTB in the recharge domain. Violators allocate more steps to the future date when we increase the interest rate from 1 to 1.25. We varied the exchange rate for two questions: today versus 7 days from now, and 7 versus 14 days; rows 1 and 2 show violations for these two questions separately and row 3 and 4 show percentages of people who violated at least once or both.

also closer to what is found in Augenblick et al. (2015) (which finds 17%) than to the higher estimates from the full CTB (26%).

Note that these estimates come from the performance of the simple CTB over recharges but not over effort; given the specific questions we asked in the effort domain, we cannot calculate law of demand violations nor future bias, so we cannot compare the measures on that front. However, as shown in Table A.1, the simple CTB over effort correlates in the expected direction with exercise (i.e., people who look more impatient under the simple CTB have lower steps). In contrast, the full CTB estimates do not correlate in the expected direction with any behaviors. Hence, the simple CTB still appears to be the better measure for our context.

J.3 Implementation of the Full CTB

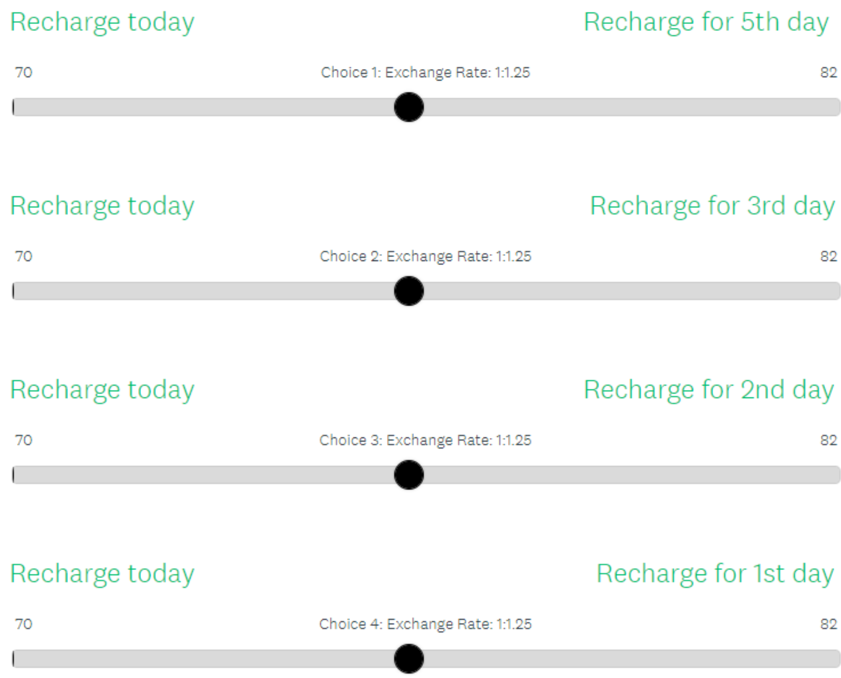
We first discuss the methodology used for the full CTB. We then show that the full CTB measures do not correlate with the behaviors that we would expect. Finally, we describe additional problems with the full CTB implementation.

J.3.1 Estimation Methodology

Our full CTB uses the full CTB methodology of Augenblick et al. (2015). In each CTB choice in our full CTB module, the participant is asked to allocate a fixed budget of either steps or mobile recharges between a "sooner" and a "later" date using a slider bar. In particular, each choice allows the respondent to choose an allocation of consumption on the sooner and later dates, c_t, c_{t+k} that satisfies the budget constraint

$$c_t + \frac{1}{r}c_{t+k} = m \tag{35}$$

where the sooner date t, the later date $t + k$, the interest rate r, and the budget m change between each choice. A sample slider screen allowing for such choices is shown in Figure J.1.



Appendix Figure J.1: Sample Decision Screen for Mobile Recharges

Notes: In this example, the interest rate, r, is 1.25; the total budget, m, is 140; the "sooner" date is Today; and the "later" date decreases from 5 days from today in the first choice to 1 day from today in the final choice. The sliders are shown positioned at the choice $(c_t = 70, c_{t+k} = 82)$.

We asked participants to make six allocations in the recharge domain, and eight allocations in the step domain, as summarized in Table J.2. We assume a time-separable and good-separable CRRA utility function with quasihyperbolic discounting[86]. In the domain of recharges, individuals will then seek to maximize utility,

$$U\left(c_t, c_{t+k}\right) = \frac{1}{\alpha}\left(c_t - \omega\right)^{\alpha} + \beta\delta^k \frac{1}{\alpha}\left(c_{t+k} - \omega\right)^{\alpha} \tag{36}$$

and in the step domain, individuals will seek to minimize costs of effort

$$C\left(c_t, c_{t+k}\right) = \frac{1}{\alpha}\left(c_t + \omega\right)^{\alpha} + \beta\delta^k \frac{1}{\alpha}\left(c_{t+k} + \omega\right)^{\alpha} \tag{37}$$

The variation in consumption choices as the budget constraint varies identify the time preference parameters—in particular, the daily discount factor δ and the present-bias parameter β—as well as the concavity or convexity of preferences α. Due to budget and time constraints, we had to keep the module short and so did not implement interest rate variation for the recharge tradeoffs, only for the step tradeoffs. Thus α is identified for the effort estimation only, not the recharge one; for the recharge estimation, we calibrate α using the estimate of α from Augenblick et al. (2015) in the financial payment domain.

[86]Unlike in Appendix C.2 where the quasihyperbolic discounting model we used only has one parameter δ_{QH} or d_{QH}, here we use both β and δ since we estimated them simultaneously.

We recover individual-level structural estimates of time preference and concavity parameters from the allocations (c_t, c_{t+k}) using a two-limit Tobit specification of the intertemporal Euler condition following Augenblick et al. (2015).

$$\log\left(\frac{c_t+\omega}{c_{t+k}+\omega}\right) = \frac{\log(\beta)}{\alpha-1}1_{t=0} + \frac{\log(\delta)}{\alpha-1}k - \frac{1}{\alpha-1}\log(r) \tag{38}$$

Details on the estimation strategy can be found in the Online Appendix of Augenblick et al. (2015). Because our predictions concern overall impatience, not whether an individual is time-consistent, on the time preference side, we want one single summary measure capturing impatience. To do so, we estimate two different variants. In one, we set $\beta = 1$ for everyone at the estimation stage and simply estimate δ at the individual level. In the second, we estimate the equation as above, allowing both β and δ to vary at the individual level, and use $\beta \times \delta$ as our measure of individual-level impatience. In both estimation procedures, we allow α to vary at the individual-level in the steps domain, since we considered individual-level convexity of the step function to be an important potential confound.[87] However, the results we describe next are similar if we do not allow α to vary at the individual-level for steps.

Appendix Table J.2: CTB Allocation Parameters

Summary of convex time budget allocations					
Question no.	t	k	r	Recharge domain	Step domain
1	7	7	1	X	X
2	0	7	1	X	X
3	0	5	1	X	X
4	0	3	1	X	X
5	0	2	1	X	X
6	0	1	1	X	X
7	7	7	1.25		X
8	0	7	1.25		X

Notes: This table summarizes the parameters of the six CTB allocations made over recharges, and the eight CTB allocations made over steps.

Our CTB environment builds on a number of features from previous studies. First, the choices are made after the one-week phase-in period in which all participants have pedometers and report their daily steps, ensuring that participants are familiar with the costs of walking. This allows for meaningful allocations of steps between sooner and later dates. Second, the responses are designed to be incentive compatible; all respondents were informed that we would

[87] Indeed, when we estimate impatience (e.g., δ) but do not allow α to vary, that estimated δ correlates as strongly with α as it does with the δ estimated allowing α to vary, suggesting that convexity is an important confound indeed.

implement their choice from a randomly selected survey question. We set the probabilities such that for most respondents the randomly selected survey question was a multiple price list of lotteries over money (which measures risk preferences), but for a few a CTB allocation was selected. Because the allocations might have interfered with any walking program offered, we excluded the 40 respondents who were randomly selected to receive one of their allocations from the experimental sample.[88] To try to ensure that participants complete the allocated steps, we offer a large cash completion bonus of 500 INR in the step domain if the allocation is selected to be implemented, and the steps are completed as allocated, with the bonus to be delivered 15 days from the date of the survey (which is 1 day after the latest "later" day used).

We also take a number of precautions to avoid various potential confounds, including confounds reflecting fixed costs or benefits of taking an action, or confounds due to the time of day of measurement.[89] However, we were not able to address one potential confound to our estimates of time-preferences across individuals fully: variation across people in the cost of walking over time, or in the benefit of receiving a recharge over time. For example, an individual with a particularly busy week after the time-preference survey, and therefore relatively high costs to steps in the near-term relative to the distant future, will appear to be particularly impatient over steps in our data (he will wish to put off walking). An individual with a relatively free week just after the time-preference survey will instead appear particularly forward-looking (he will not wish to put off walking). The same concerns can also arise with recharges.

J.3.2 CTB Estimates: Problems with Convergence and Lack of Correlation

Table J.3 displays the summary statistics as well as the convergence statistics. The CTB parameter estimates themselves are not robust and are inconsistent with typical priors. First,

[88] This means we have CTB data from a total of 3,232 people: the 3,192 in the experimental sample plus the 40 selected to receive "real-stakes" allocations. For completeness, we summarize in this section the CTB data for all 3,232 but the results are the same if we restrict to the experimental sample.

[89] To avoid confounds related to fixed costs or benefits, such as the effort of wearing a pedometer or the psychological benefit of receiving a free recharge, we include minimum allocations on both sooner and later days in each domain. The minimum allocations were chosen to be high enough that any fixed costs would be included (e.g. one could not easily achieve the minimums by simply shaking the pedometer) but low enough to avoid corner solutions. In the step domain, this required a modification of the CTB methodology: individual-specific minimum allocations. Our step allocations also featured individual-specific total step budgets m, which were chosen to be large enough that achieving them would require some effort beyond simply wearing the pedometer but small enough that participants would certainly achieve them in exchange for the completion bonus. Specifically, minimum steps on each day are calculated as $\frac{X}{10}$, and the total step budget m is $X + 2\frac{X}{10}$, respectively, where $X \in \{3000, 4000, 5000\}$ is the element closest to the participant's average daily walking during the phase-in period. That is, minimum steps are one of 300, 400, or 500 on each day, and the total step budget is one of 3,600, 4,800, or 6,000. To avoid confounding impatience with the time of day that the baseline time-preference survey was administered (which could influence the desirability of walking and/or recharges delivered in the next 24 hours), as well as to capture heterogeneity in time preferences including any present-bias for very short beta-windows, we required that all walking on any date be conducted within a 2 hour period, which was chosen to start at the time immediately after the time-preference survey would end (e.g., if the survey ended at 4pm, the time period for any day's walking would be 5-7pm). The short window could potentially bias our overall measures of impatience downwards, as uncertainty about future schedules in a short time window could lead participants to want to get their walking done early when they had more certainty over their schedule. However, our primary purpose was to capture heterogeneity in time-preferences, and we considered the potential loss in validity of aggregate time preference estimates to be worth the ability to capture heterogeneity in time preferences in the time frames near to the present.

we do not have estimates for a large, endogenous share of the sample. The estimates do not converge (i.e., we are unable to estimate discount rate parameters) for 38 to 44% of the sample in the recharge domain, and 23 to 44% of the sample in the steps domain. Moreover, many of the participants with estimates that converge in the effort domain have an estimated $\alpha < 1$, which violates the first order conditions for estimation and is often associated with non-sensible δ and β estimates. When we exclude these estimates, we are left with estimates for only 34 to 38% of the sample in the effort domain. Second, we have a high rate of negative estimated discount rates: 43% for steps and 61% for recharges. This is more than the usual rate of negative individual-level estimates.

Appendix Table J.3: Summary Statistics For CTB Parameters

	Full sample		$\alpha > 1$	
Parameters estimated:	$\beta\delta$	δ	$\beta\delta$	δ
	(1)	(2)	(3)	(4)
A. Effort				
Beta	2.066	–	1.573	–
Delta	0.883	0.997	1.015	0.999
Alpha	0.244	0.723	1.673	1.576
% of sample:	77.2	56.3	34	38
# Individuals:	2,494	1,821	1,092	1,225
B. Recharges				
Beta	0.972	–	–	–
Delta	0.989	0.996	–	–
% of sample:	55.9	62.2	–	–
# Individuals:	1,808	2,011	–	–

Notes: This table displays means and convergence rates of individual-level CTB parameters in both the effort and recharge domains. Columns 1 - 2 display average values for the parameters from the full sample of individuals with parameters that converged. In the effort domain, in columns 3 - 4, we ignore all individuals whose estimated *alpha* was below 1, as handled similarly in Andreoni and Sprenger (2012a), as that is inconsistent with the first order conditions. We winsorize all parameters at the top and bottom 1 percentiles. We allow α to vary at the individual level in the effort domain, and in the recharge domain, we calibrate α to be 0.975, which is the estimated value in Augenblick et al. (2015). Delta is estimated by allowing δ to vary at the individual level and setting β to 1. Beta-delta is estimating by allowing both δ and β to vary. We derive these two parameters from an estimation that allows δ and β to vary at the individual level. Significance levels: * 10%, ** 5%, *** 1%.

Tables J.4 and J.5 show that the estimated CTB parameters do not correlate in the expected direction with measured behaviors. In particular, Table J.4 shows that the CTB estimates in the steps/effort domain do not correlate with exercise and health,[90] and Table J.5 shows that the estimates in the recharge domain do not correlate with recharge balances, usage, or credit constraint proxies. The CTB measures do correlate at the 1% level with our measure of marginal propensity to consume recharges, but the correlations go in opposite directions for the two CTB measures (δ from an estimation setting $\beta = 1$ vs. $\beta\delta$ estimated allowing both parameters to vary) so is likely noise.

Appendix Table J.4: CTB Estimates of Discount Factors Over Steps Do Not Correlate With Measured Behaviors

Covariate type:	Exercise		Baseline indices			
Dependent variable:	Daily steps	Daily exercise (min)	Health index	Negative vices index	Healthy diet index	# Individuals
Delta	-0.018	0.009	-0.038	0.010	0.025	1,342
Beta-delta	0.016	0.018	0.014	0.010	0.027	1,086

Notes: This table displays the correlations between CTB parameters in the effort domain and a few baseline health covariates. We normalize impatience variables so that a higher value corresponds to greater impatience, and we normalize health outcomes so that higher values correspond to healthier outcomes. All CTB parameters have been winsorized at the top and bottom 1 percentile to remove outliers. Delta is measured from an estimation that allows δ and α to vary at the individual level, while excluding β. Beta-delta is a measure of beta times the average delta over one week. We estimate the two parameters by allowing β, δ, and α to vary at the individual level. Significance levels: * 10%, ** 5%, *** 1%.

Appendix Table J.5: CTB Estimates of Discount Factors Over Recharges Do Not Correlate With Other Proxies for Impatience Over Recharges

Covariate type:	Recharge variables			Credit constraint proxies			
Dependent variable:	Negative mobile balance	Negative yesterday's talk time	Prefers daily (=1)	Prefers monthly (=-1)	Negative wealth index	Negative monthly household income	# Individuals
Delta	0.027	0.012	-0.141***	0.045	-0.010	-0.008	1,837
Beta-delta	-0.002	-0.022	0.145***	-0.019	-0.015	0.033	1,652

Notes: This table displays the correlations between CTB parameters in the recharge domain and baseline measures that should be related to credit constraints and discount rates over recharges. We normalize impatience variables so that a higher value corresponds to greater impatience, and we normalize the proxies so that higher values correspond to higher expected discount rates; hence, the prediction is that coefficients should be positive. All CTB parameters have been winsorized at the top and bottom 1 percentile to remove outliers. We use two main estimation specifications, and to identify parameters, we calibrate α to be 0.975, the value of α estimated in Augenblick et al. (2015). Delta is estimated by allowing δ to vary at the individual level and excluding β. Beta-delta is a measure of the average delta over one week multiplied by beta. We derive these two parameters from an estimation that allows δ and β to vary at the individual level. Significance levels: * 10%, ** 5%, *** 1%.

[90] Table J.4 shows the correlations when we exclude the effort estimates from participants with estimated $\alpha < 1$, but the results are similar when we include all estimates together.

J.3.3 Additional Problems With the Full CTB Data

Finally, we provide more detail on other problems with the Full CTB, in addition to the law of demand violations, the lack of convergence, and the lack of correlation described earlier.

First, in the effort task, there was low follow-through on the incentivized activity: fewer than 50% of participants selected to complete the step task did so despite large rewards (500 INR) for completion. While this partly reflects a logistical glitch (we failed to give respondents intended reminder calls the day before their activity), the lack of follow-through may also indicate a lack of respondent understanding. Regardless, the poor follow-through is problematic methodologically: identification requires that, when participants make their allocation decisions, they think they will follow-through with certainty, which seems unrealistic given how few followed through in practice.

Second, respondents on average allocated more steps to today than the future even when the interest rate was 1:1. Although they could be future-biased, the following other potential explanations are concerning for interpretation: respondents were confused; they saw steps as consumption instead of a cost (violating the first order conditions underlying estimation); or uncertainty over future walking costs and schedules led participants to want to finish steps sooner, which would confound discount rate estimates with risk aversion and uncertainty.

Third, day-specific shocks appear to be important in practice. 19% of respondents' allocations of steps to the sooner date are neither monotonically weakly increasing nor monotonically weakly decreasing across questions which feature the same sooner date (today) but a monotonically decreasing later date (questions 2-6). These allocations cannot be rationalized with a discount rate that is either weakly decreasing *or* increasing with lag length without day-specific utility shocks. The same holds for 24% of respondents in the recharge domain. These types of shocks would also confound estimation.

K Monitoring Treatment Impacts on Walking

The health results suggest that the monitoring treatment had limited impact, although the results are somewhat imprecise. Did the monitoring treatment not affect exercise, or were the exercise impacts too small to translate into measurable health impacts? We now present an analysis of the effects of monitoring on exercise. Because we do not have pedometer walking data from the control group, we use a before-after design. We find that monitoring alone has limited impact on overall steps. Monitoring does however change the distribution of steps, increasing the share of days on which participants met the 10,000-step target but decreasing the steps taken on other days for a null effect on total exercise.

Our before-after design compares pedometer-measured walking in the monitoring group during the phase-in period (during which we had not given participants a walking goal and just told them to walk the same as they normally do) to their behavior during the intervention period. This strategy will be biased either in the presence of within-person time trends in walking, or if the phase-in period directly affects walking behavior. We control for year-month fixed effects to help address time trends, but the latter concern is more difficult, as the phase-in period likely did increase walking above normal, either because of Hawthorne effects or because participants received a pedometer and a step-reporting system, which are two of the elements of the monitoring treatment itself (the other three remaining that we can still evaluate are (a) a daily 10,000 step goal, (b) positive feedback for meeting the step goal through SMS messages and the step-reporting system, and (c) periodic walking summaries). Thus, we consider a pre-post comparison of walking in the monitoring group to be a lower bound of the monitoring program treatment effect.

One can visualize the variation used for our pre-post estimate in Figure A.4, Panels (a) and (b). Walking increases immediately during the intervention period for the monitoring group, although the effects decay over time.

We next estimate the pre-post monitoring effect controlling for date effects. In order to increase the precision of our estimated year-month fixed effects, we include the incentive group in the regression as well since that group is much larger. We thus estimate the following difference-in-differences regression using data from both the intervention and phase-in periods for the incentive and monitoring groups:

$$\begin{aligned} y_{it} = & \alpha + \beta_1 Intervention\,Period_{it} + \beta_2 incentives_i + \beta_3 (Intervention\,Period_{it} \times incentives_i) \\ & + \boldsymbol{X}_i' \gamma + \boldsymbol{\mu}_m + \varepsilon_{it}, \end{aligned} \tag{39}$$

where y_{it} are daily pedometer outcomes measured during both the phase-in and the intervention period, $Intervention\,Period_{it}$ is an indicator for whether individual i has been randomized into their contract at time t, $incentives_i$ is an indicator for whether i is in an incentive treatment group, $\boldsymbol{X}_i$ is a vector of individual-specific controls, and $\boldsymbol{\mu}_m$ is a vector of month fixed effects. The coefficient β_1 — the coefficient of interest — is the pre-post difference in pedometer outcomes within the monitoring group (controlling for aggregate time effects).

Table K.1 presents the results. Column 2 shows that the monitoring group achieves the 10,000-step target on approximately 7% more days in the intervention period than in the phase-in period, an effect significant at the 1% level and equal to roughly 36% of the estimated impact of incentives. In contrast, the estimated effect on steps is very small in magnitude, varies across specifications, and is in fact sometimes negative (columns 4-6). Thus, the monitoring treatment, if anything, appears to do more to make walking consistent across days than it does to increase total steps.

Appendix Table K.1: Impacts of Monitoring (Pre-Post) and Incentives (Difference-In-Differences) on Exercise Outcomes

	Achieved 10K steps			Daily steps		
	(1)	(2)	(3)	(4)	(5)	(6)
Incentives	0.012	0.013	0.012	66.7	66.4	48.9
	[0.024]	[0.024]	[0.014]	[268.1]	[266.9]	[112.3]
Intervention period	0.057***	0.073***	0.064***	-130.4	108.0	-18.5
	[0.020]	[0.020]	[0.020]	[237.8]	[240.8]	[234.1]
Intervention period X Incentives	0.19***	0.19***	0.19***	1270.9***	1258.9***	1212.7***
	[0.021]	[0.021]	[0.021]	[248.6]	[249.2]	[243.4]
Year-month FEs	No	Yes	Yes	No	Yes	Yes
Individual controls	No	No	Yes	No	No	Yes
Monitoring phase-in mean	.24	.24	.24	6,904.8	6,904.8	6,904.8
# Individuals	2,604	2,604	2,604	2,604	2,604	2,604
Observations	221,214	221,214	221,214	221,214	221,214	221,214

Notes: This table shows coefficient estimates from regressions of the form specified in equation (39). The outcomes are from daily panel data from the pedometers. Standard errors, in brackets, are clustered at the individual level. Individual controls are the same as Table 2. The omitted category is Monitoring in the phase-in period. The coefficient in the second row, on $Intervention\,Period_{it}$, corresponds to the pre-post estimate of the Monitoring effect. Significance levels: * 10%, ** 5%, *** 1%.