Web Appendix for

Does Access to Human Coaches Lead to More Weight Loss than With AI Coaches Alone?

Anuj Kapoor
University of Missouri

Sridhar Narayanan
Stanford University

Puneet Manchanda
University of Michigan

Web Appendix

A Web Appendices

A.1 Web Appendix: App Screens

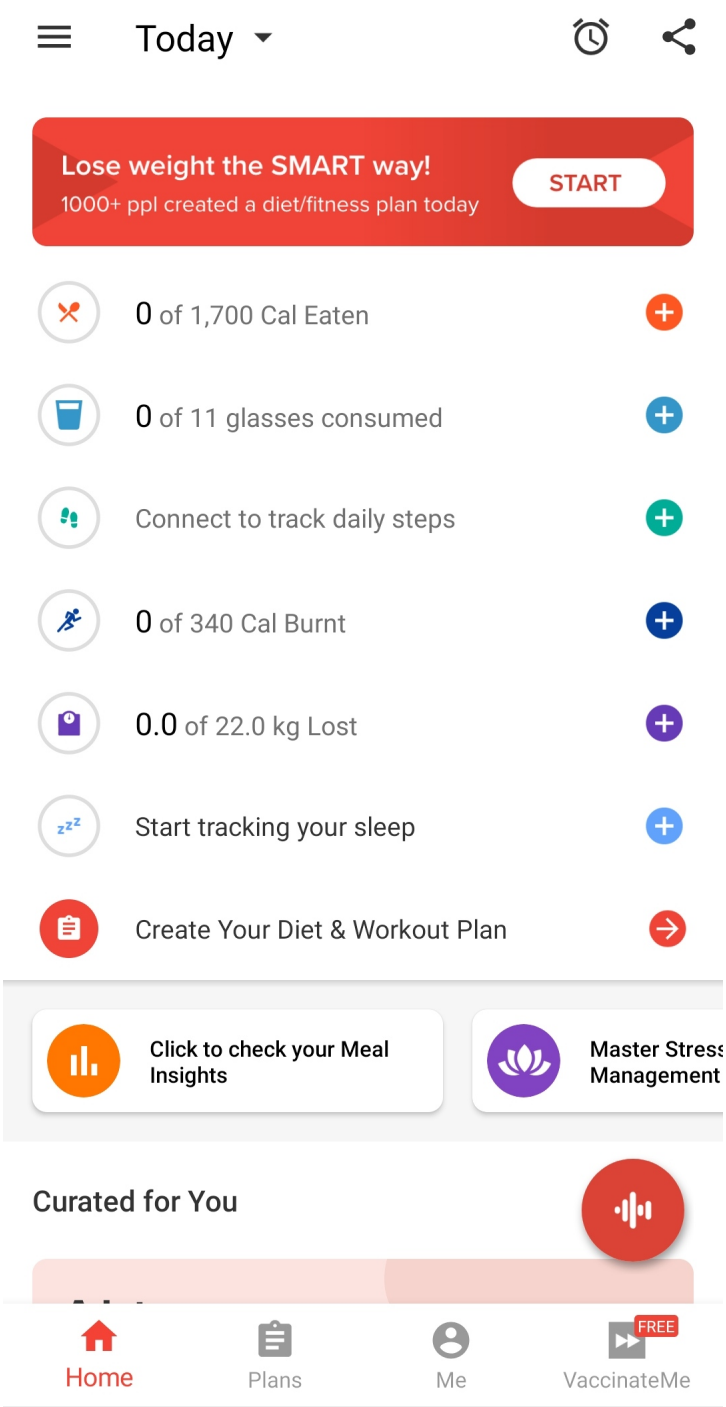

Figure A.1: FREE Plan app screen

Notes: The figure above shows the interface of the application under the 'Free' plan. The plan has tracking features but no access to the personalized insights, coaching and motivational features of paid plans

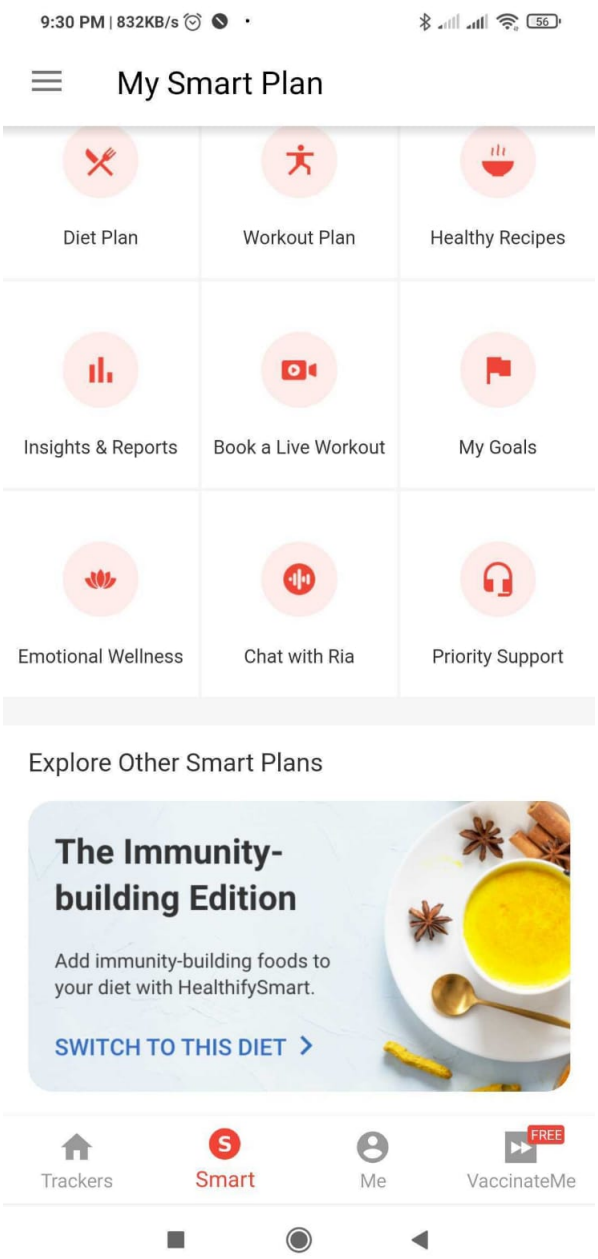

Figure A.2: Smart Plan app screen

Notes: The Figure shows a snapshot of the interface of a user enrolled for the `Smart Plan` on the HealthifyMe app. This is a paid plan with an AI-only coach.

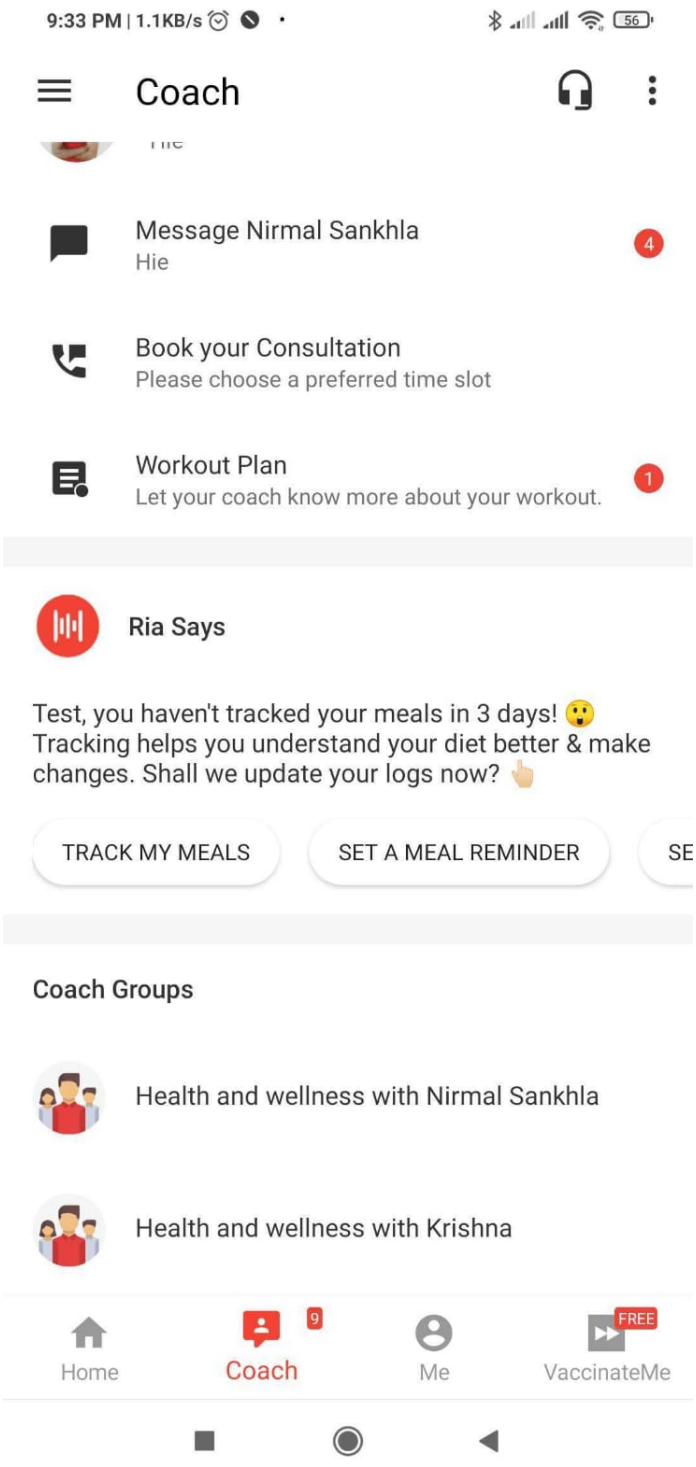

Figure A.3: Coach Plan app screen

Notes: The Figure shows a snapshot of the interface of a user enrolled for the Paid Plan involving a Human coach on the HealthifyMe app.

A.2 Web Appendix: Performance of the Classification Algorithm

The first part of the matching procedure is to build a predictive model for whether the consumer selects the AI+Human or AI plan. For this purpose, we used a Random Forest classification algorithm due to its superior performance in our context. We used 60% data for training, 20% for validation, and the final 20% for testing. We also attempted the predictive model using other approaches, including a binomial logit model but selected the Random Forest classifier for its superior predictive performance.[1]

Table A.1 presents the impurity-based feature importance of different variables used in the Random Forest Classifier model. The length of the user's free journey in months, their age, the number of weight loss spells they had in their free journey, their weight logging frequency, and food logging frequency are variables that are of high importance for the classification task.

The training accuracy of the model on the data it was trained on (60% of the sample) is 96.28 i.e., this is a measure of in-sample 'fit'. The Out of Bag (OOB) score, calculated on a 20% validation sample is 94.32. This sample is used during the training of the model. The accuracy on an additional 20% test sample, which was not used during the training of the model at all, was 92.45.

Table A.1: Importance of variables in Random Forest Classifier

	Importance
months	0.231
age	0.213
avg_spells_per_month	0.172
avg_weight_logs_per_day	0.139
avg_goal_achievement	0.052
avg_food_calories	0.041
avg_food_logs	0.040
bmi	0.037
avg_goal_loss	0.021
height	0.018
avg_actual_loss	0.017
weight	0.016
gender	0.002

Notes: The table presents the impurity-based feature importance of different variables used in the Random-Forest Classifier model. The method calculates the mean decrease in the impurity measure (as measured by the GINI-impurity), with the average taken across all the trees in the forest. The table presents the importance of the variables in descending order, implying that the variable "months" have the highest importance and "gender" the least.

[1] More details on the predictive model and comparisons with other approaches are in Table A.2.

Table A.2: Comparison of Alternative Classification Approaches

	Random Forest	**Logit**
Validation accuracy	0.9494	0.8318
Test accuracy	0.9493	0.8341
Train accuracy	0.9628	–
OOB estimate	0.9432	–

A.3 Web Appendix: Distribution Matched and Unmatched Sample

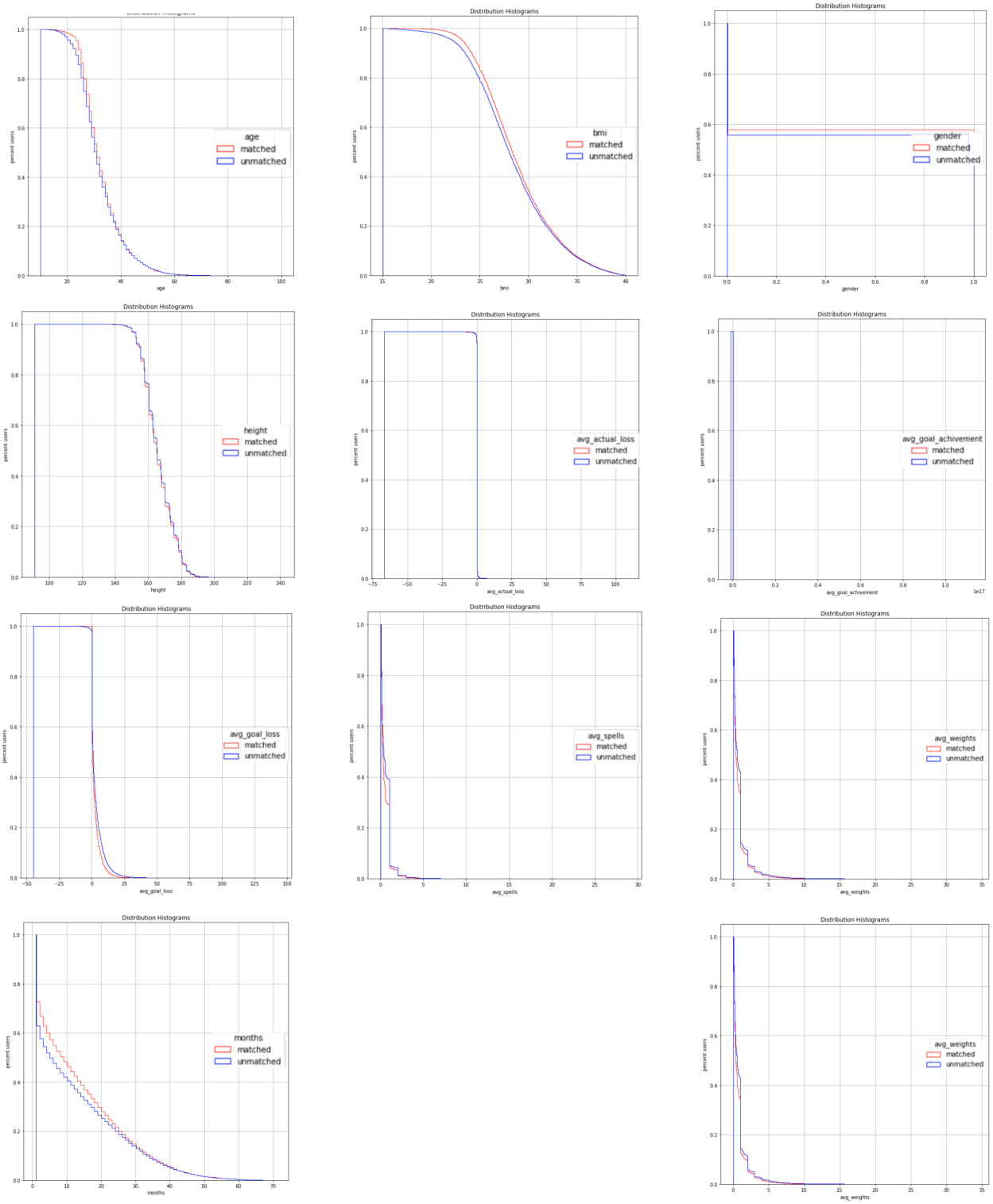

Figure A.4: Distribution Matched and Unmatched Sample

A.4 Web Appendix: Distributions of Variables Used for Matching, Before and After Matching

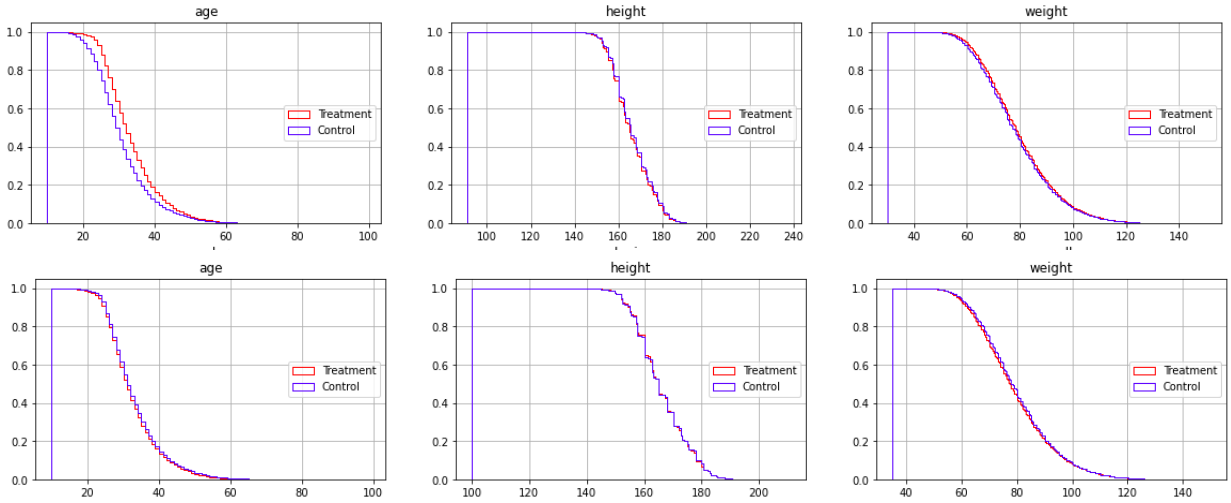

Figure A.5: Distribution Before (upper panel) and After (lower panel) Matching (A)

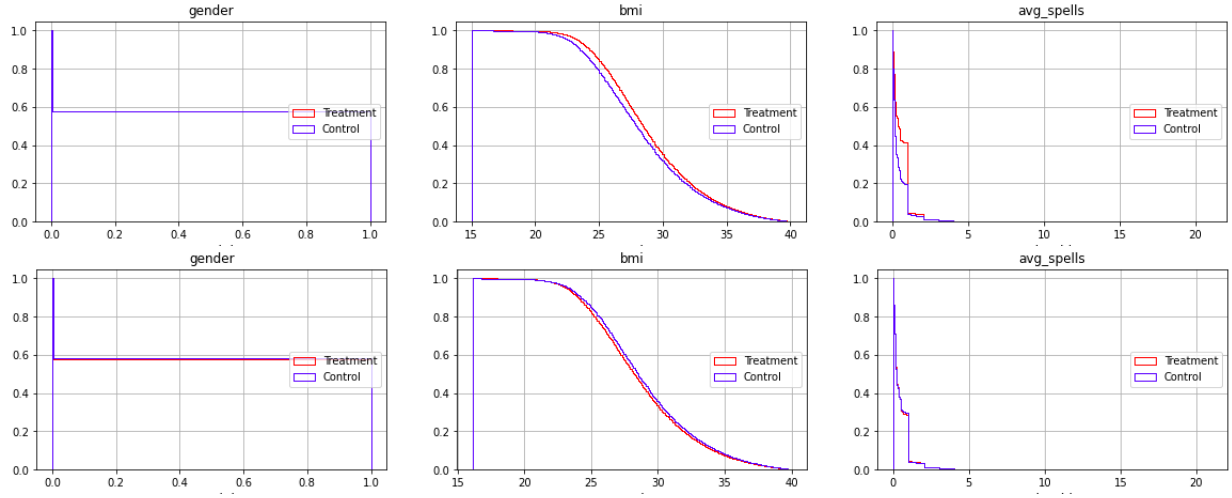

Figure A.6: Distribution Before (upper panel) and After (lower panel) Matching (B)

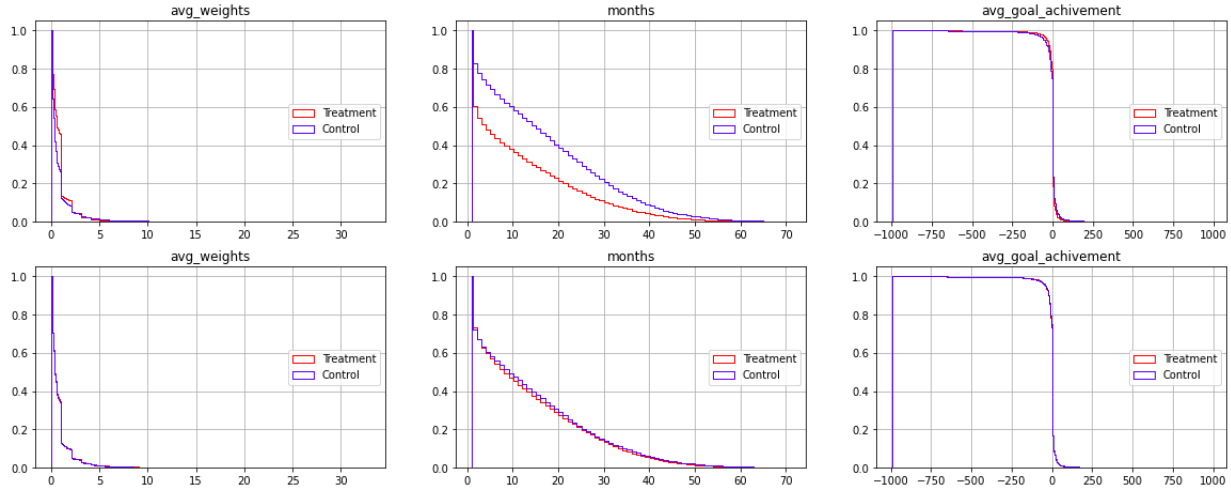

Figure A.7: Distribution Before (upper panel) and After (lower panel) Matching (C)

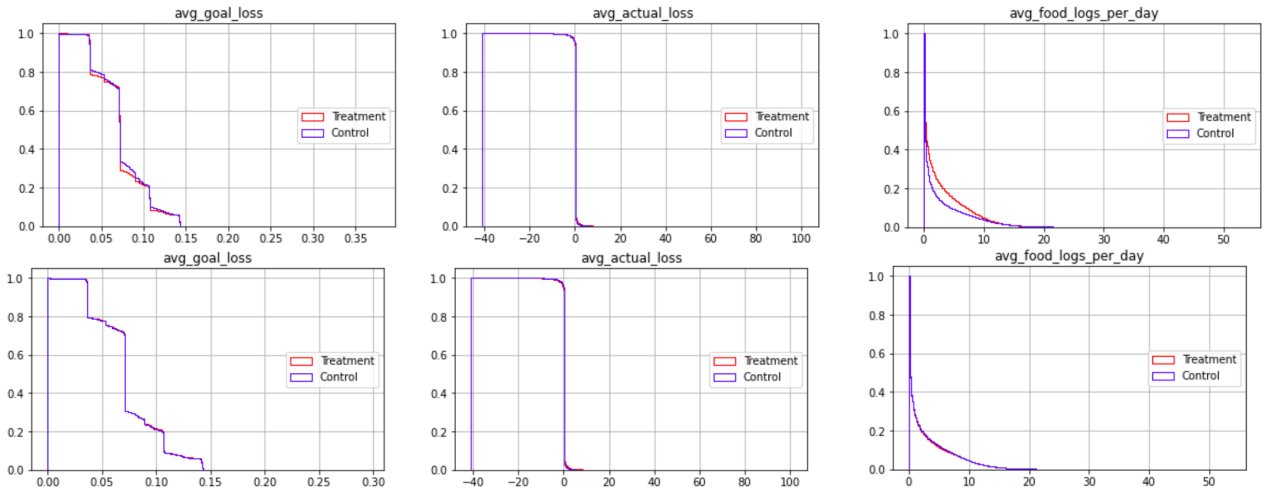

Figure A.8: Distribution Before (upper panel) and After (lower panel) Matching (D)

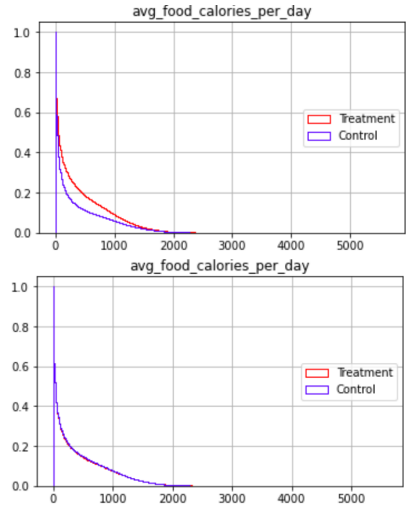

Figure A.9: Distribution Before (upper panel) and After (lower panel) Matching (D)

Table A.3: Kolmogorov-Smirnov Test Results - Before and After Matching

	Statistic (Before Matching)	Statistic (After Matching)
age	0.164***	0.090***
height	0.032***	0.017
weight	0.051***	0.030
gender	0.001	0.001
BMI	0.084***	0.068***
avg_spells	0.270***	0.009
avg_weight_logs_per_day	0.267***	0.004
months	0.285***	0.140***
avg_goal_achievement	0.089***	0.007
avg_goal_loss	0.211***	0.003
avg_actual_loss	0.06***	0.004
avg_food_logs_per_day	0.350***	0.003
avg_food_calories_per_day	0.333***	0.003

Notes: The table shows the results of the Kolmogorov-Smirnov Test for the empirical cumulative distribution function (cdf) of the key variables of interest, before and after the matching procedure. The KS test statistic is reported in the column along with the p-values. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

A.5 Web Appendix: Rosenbaum Bounds

The PSM method is based on observables. In order to deal with potential selection on unobservables, we used the Rosenbaum bounds, as it is common in the literature (Rosenbaum 2002). Rosenbaum bounds provide information about the level of uncertainty contained in matching estimators by showing how significant the influence of a confounding variable must be to undermine the conclusions of a matching analysis. As a result, they have been used in many studies (Manchanda et al. 2015).

We report the results from the Rosenbaum bounds analysis in Table A.4. As can be seen from the table, the effect of treatment (i.e., AI+Human vs AI) on weight loss is positive and significant at the 5% level for the Rosenbaum bounds test statistic $\Gamma < 1.3$ (the relevant metric here is sig+, since the hypothesized direction of the treatment effect is positive) and at the 10% level at $\Gamma < 1.5$. The rough interpretation of Γ is that it represents the odds ratio of selection into treatment due to unobserved selection of two units matched on observables. Thus, a unit from our sample has to be at least 1.3 to 1.5 times more likely (depending on the desired significance level of 5 or 10%) to select into treatment based on unobservable factors to nullify our reported treatment effect. Our reported values of Γ compares well with those reported in the prior literature (e.g., Manchanda et al. (2015) report a value of 1.6, Altonji et al. (2005), Oster (2019) report a value of 1 or above as appropriate). Thus, we conclude that selection on unobservables is not a major concern in our case.

Table A.4: Rosenbaum Sensitivity Test for Hodges-Lehmann Point Estimate

Gamma	sig+	sig-	Lower bound (t-hat+)	Upper bound (t-hat-)
1.0	.022	.022	27.75000	27.75
1.1	.027	.021	22.95000	30.05
1.2	.039	.020	15.95000	31.75
1.3	.056	.019	11.15000	33.15
1.4	.076	.018	8.25000	34.35
1.5	.104	.018	6.05000	35.35
1.6	.137	.017	4.55000	36.35
1.7	.184	.017	3.35000	37.15
1.8	.256	.017	2.35000	37.95
1.9	.364	.016	1.55000	38.55
2.0	.516	.016	0.94999	39.35

Notes: Gamma is the Odds of Differential Assignment to Treatment Due to Unobserved Factors. Upper Bound and Lower Bound are the t-hat- and t-hat+ respectively. We see that results stop being statistically significant at the 0.05 level from $\Gamma = 1.3$ and the 0.10 level from $\Gamma = 1.5$.

A.6 Web Appendix: Detailed Summary Statistics

Table A.5: Comparison of Variables for Treatment (AI+Human Coach) and Control (AI) Users (Before Matching)

(a) Control Group

	Age	Height	Weight	Gender	BMI	Avg_ spells_ per_mth	Avg_ weight_ logs	Months	Avg_goal _achiev -ement	Avg_goal _loss	Avg _actual _loss	Avg_food _logs _per_day	Avg_food _calories _per_day
mean	30.813	166.161	78.566	0.574	28.369	0.388	0.651	17.393	-5.349	0.075	-0.025	1.383	172.510
std	8.176	9.371	14.314	0.494	4.045	0.591	1.166	14.857	46.444	3.311	0.599	3.150	373.480
min	11.000	91.000	30.000	0.000	15.039	0.014	0.014	1.000	-995.148	0.000	-21.333	0.000	0.000
25%	25.000	160.000	68.000	0.000	25.384	0.667	0.100	3.000	-5.887	0.042	-0.003	0.021	3.070
50%	29.000	165.100	77.000	1.000	27.915	0.152	0.250	15.000	0.500	0.220	0.000	0.143	20.469
75%	35.000	172.720	87.500	1.000	30.950	0.500	1.000	28.000	1.899	2.000	0.000	0.858	114.397
max	99.000	236.220	150.000	1.000	39.989	16.000	29.000	71.000	929.248	56.550	50.006	53.500	5649.853

(b) Treatment Group

	Age	Height	Weight	Gender	BMI	Avg_ spells_ per_mth	Avg_ weight_ logs	Months	Avg_goal _achiev -ement	Avg_goal _loss	Avg _actual _loss	Avg_food _logs _per_day	Avg_food _calories _per_day
mean	33.398	165.549	79.451	0.577	28.889	0.585	0.815	10.740	-3.589	0.074	-0.008	1.951	257.807
std	7.904	9.249	14.122	0.494	3.878	0.607	0.027	12.942	36.168	3.632	0.627	3.454	437.245
min	10.000	127.000	37.600	0.000	16.937	0.014	0.014	1.000	-953.435	0.000	-22.002	0.000	0.000
25%	28.000	158.000	69.000	0.000	26.050	0.111	0.167	1.000	-2.094	0.103	-0.001	0.038	5.520
50%	32.000	165.000	78.000	1.000	28.405	0.363	0.560	4.000	0.497	0.932	0.000	0.297	41.356
75%	38.000	172.720	88.000	1.000	31.280	1.000	1.000	18.000	1.163	3.555	0.000	2.087	290.528
max	99.000	208.000	149.000	1.000	39.995	21.000	33.000	70.000	983.778	59.375	52.500	52.667	5586.651

Note: All the values are rounded up to three decimal points.

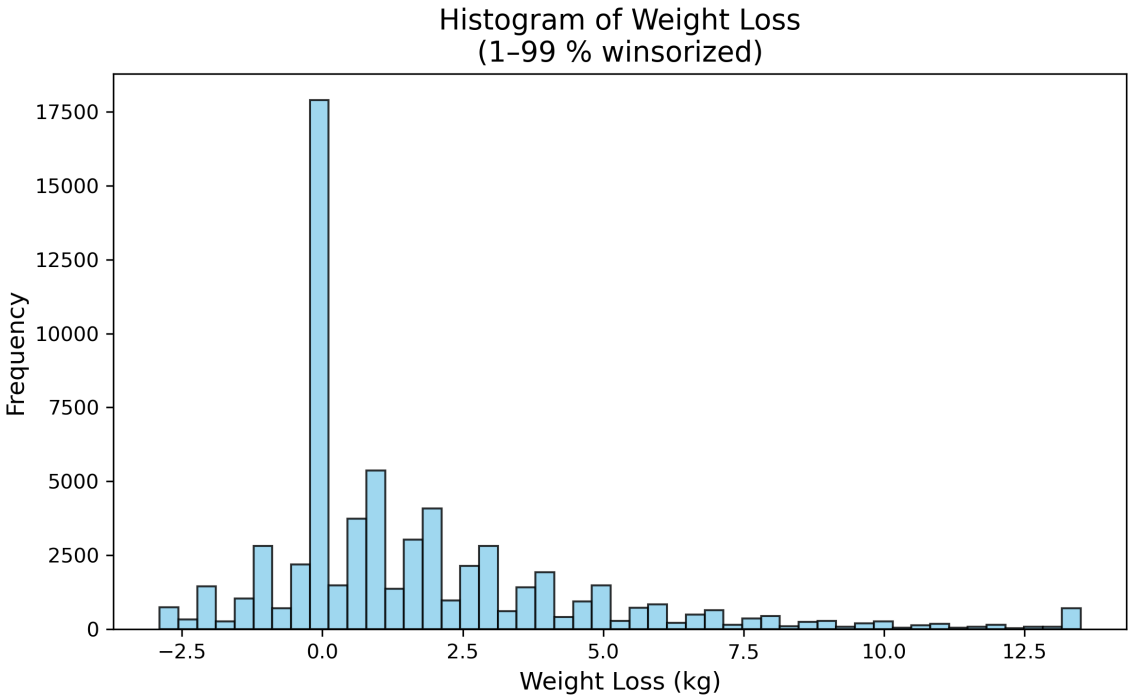

Figure A.10: Histogram of Weight Loss (1st–99th Percentile Winsorized)

A.7 Web Appendix: Prior Studies in the Weight Loss Domain

Table A.6: Weight Loss Papers

Paper	Nature of comparison	Type of estimate	Estimate of weight loss
Chin et al. (2016)	Male-app versus Female-app	Retrospective cohort study	Baseline BMI was 30.2±0.1kg/m2 for males and 28.0±0.0kg/m2 for females. Among the participants (35,921 Noom users), 77.9% reported a decrease in body weight while they were using the app with 22.7% experiencing more than 10% weight loss compared with baseline with a higher weight loss success rate in males (83.9 versus 76.1%), resulting in final BMIs of 28.1±0.1kg/m2 for males and 26.5±0.0kg/m2 for females.
Uetake and Yang (2020)	Weight loss program with In-person group meetings as an important component	Observational Study	The weight loss for the average peer leads to individual weight gain, as a 1 kg increase in the average peer performerís weight loss is associated with an individual's decrease in weight loss by about 0.02 kg. In contrast, weight loss by the top performer leads to increased individual weight loss, as a 1 kg increase in the top performerís weight loss is associated with an individual's increase in weight loss by about 0.01 kg. A large number of observations (about 41%) involve participants gaining weight.
Ingels et al. (2017)	DPM program (West Virginia) (2015-16)	Observational Study	The study divided participants into three tracking groups: rare (¡33% of days tracked, n = 25), inconsistent (33–66%, n = 5), and consistent (¿66%, n = 15). The average total weight loss for consistent trackers was 5.6 pounds (SD = 12 0) for a 49-week period.
Knäuper et al. (2018)	The NIH-developed Diabetes Prevention Program (DPP) across two groups i.e., standard or enhanced DPP (by integrating habit formation tools (i.e., if-then plans))	Field Experiment	The McGill Healthy Weight Program was administered in 22 sessions over 12 months. Participants were randomly assigned by computer-generated 1:1 sequence either to the standard or the enhanced DPP. The active control group received the standard group-based DPP delivered over 1 year (12 weekly core sessions, 4 transitional sessions over 3 months, and 6 monthly support sessions). The enhanced DPP group followed the same program as the standard DPP group, but instructions for if-then planning were integrated into it. On average, participants lost 9.98% of their initial body weight in the program. At baseline, the standard DPP (Diabetes Prevention Program) group had a slightly higher mean weight than the enhanced DPP group. Controlling for this difference, weight loss did not differ between the groups over the course of the intervention. Both groups displayed significant reductions in weight from baseline to 3 months and 12 months, losing, on average of 20.36 pounds over the course of the program.

Table A.7: Weight Loss Papers (Continued ..)

Paper	Nature of comparison	Type of estimate	Estimate of weight loss
Liu et al. (2015)	Mobile phone app interventions compared with control group	Meta-analysis	Compared with the control group, mobile phone app interventions resulted in significant decreases in body weight, with the pooled estimates of the net change in body weight being -1.04 kg (95% CI -1.75 to -0.34; I2 = 41%).
Volpp et al. (2008)	Fifty-seven healthy participants aged 30-70 years with a body mass index of 30-40 were randomized to 3 weight loss plans: monthly weigh-ins, a lottery incentive program, or a deposit contract that allowed for participant matching, with a weight loss goal of 1 lb (0.45 kg) a week for 16 weeks.	A Randomized Trial	The incentive groups lost significantly more weight than the control group (mean, 3.9 lb). Compared with the control group, the lottery group lost a mean of 13.1 lb (95% confidence interval [CI] of the difference in means, 1.95-16.40; P=.02) and the deposit contract group lost a mean of 14.0 lb (95% CI of the difference in means, 3.69-16.43; P =.006).
Wharton et al. (2014)	Diet tracking and weight loss were compared across participants during an 8-week weight loss trial. Participants tracked intake using 1 of 3 methods: the mobile app "Lose It!", the memo feature on a smartphone, or a traditional paper-and-pencil method.	RCT	No difference in weight loss was noted between groups.
Yancy et al. (2019)	All participants were advised to weigh themselves daily, with a goal of 6 or more days per week, and received text messaging feedback on their performance. Incentive group participants were eligible for a lottery-based incentive worth an expected value of $3.98 in week 1 that escalated by 0.43 each week they achieved their self-weighing goal during months 1 to 6 (phase 1), followed by no incentives during months 7 to 12 (phase 2).	RCT	Mean weight changes at the end of phase 1 were 1.1 (95% CI, 2.1 to 0.1) kg in the incentive group and 1.9 (95% CI, 2.9 to 0.8) kg in the control group, with a mean difference of 0.7 (95% CI, 0.7 to 2.2) kg (P = .30 for comparison). At the end of phase 2, mean weight changes were 0.2 (95% CI, 1.2 to 1.7) kg in the incentive group and 0.6 (95% CI, 2.0 to 0.8) kg in the control group, with a mean difference of 0.8 (95% CI, 1.2 to 2.8) kg (P = .41 for comparison).

Table A.8: Weight Loss Papers (Continued..)

Paper	Nature of comparison	Type of estimate	Estimate of weight loss
Laing et al. (2014)	evaluate the effect of introducing primary care patients to a free smartphone app for weight loss. 2 academic primary care clinics. 6 months of usual care without (n = 107) or with (n = 105) assistance in downloading the MyFitnessPal app (MyFitnessPal).	RCT	After 6 months, weight change was minimal, with no difference between groups (mean between-group difference, -0.30 kg [95% CI, -1.50 to 0.95 kg]; P = 0.63).
Patel et al. (2019)	Participants were randomized to a 12-week stand-alone weight loss intervention using the MyFitnessPal smartphone app for daily self-monitoring of either (1) both weight and diet, with weekly lessons, action plans, and feedback (Simultaneous); (2) weight through week 4, then added diet, with the same behavioral components (Sequential); or (3) only diet (App-Only).	RCT	There was no difference in weight change at 3 months between the Sequential arm (mean -2.7 kg, 95% CI -3.9 to -1.5) and either the App-Only arm (-2.4 kg, -3.7 to -1.2; P=.78) or the Simultaneous arm (-2.8 kg, -4.0 to -1.5; P=.72).

A.8 Web Appendix: Heterogeneous Treatment Effects Results

Table A.9: Impact of AI+Human Coaches Based on Gender of Consumers

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
Male						
AI + Human Coach	0.902***	0.047***	1.337***	0.014***	43.502***	0.201***
	(0.051)	(0.003)	(0.143)	(0.001)	(1.624)	(0.023)
Age	-0.004	0.001***	-0.093***	-0.001***	2.202***	0.014***
	(0.003)	(0.000)	(0.009)	(0.000)	(0.136)	(0.002)
BMI	-0.117***	-0.028***	2.060***	0.016***	-4.727***	-0.024***
	(0.007)	(0.000)	(0.022)	(0.000)	(0.229)	(0.003)
Spell Duration in Months	0.196***	0.005***	0.345***	0.003***	-2.422***	-0.112***
	(0.007)	(0.000)	(0.016)	(0.000)	(0.148)	(0.002)
Baseline (Mean)	1.468	0.182	16.472	0.178	56.849	1.125
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26748	26748	26748	26748	26748	26748
R-squared Adj.	0.072	0.227	0.538	0.425	0.113	0.101
Female						
AI + Human Coach	0.901***	0.038***	1.525***	0.018***	58.563***	0.255***
	(0.042)	(0.002)	(0.112)	(0.001)	(1.161)	(0.018)
Age	0.00	0.002***	-0.132***	-0.002***	1.466***	0.005***
	(0.005)	(0.000)	(0.012)	(0.000)	(0.090)	(0.001)
BMI	-0.0835***	-0.0207***	1.7522***	0.0159***	-2.6182***	-0.0049**
	(0.011)	(0.000)	(0.020)	(0.000)	(0.148)	(0.002)
Spell Duration in Months	0.1673***	0.0055***	0.3270***	0.0038***	-1.8557***	-0.1148***
	(0.006)	(0.0003)	(0.0142)	(0.0002)	(0.112)	(0.002)
Baseline(Mean)	1.030	0.155	15.471	0.199	46.528	1.057
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39386	39386	39386	39386	39386	39386
R-squared Adj.	0.085	0.205	0.541	0.447	0.130	0.111

Notes: The table shows a regression with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

Table A.10: Impact of AI+Human Coaches Based on the Age of Consumers

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No.of Food Logs	Weekly No. of Weight Logs
Younger						
AI + Human Coach	0.877***	0.044***	1.226***	0.014***	51.450***	0.188***
	(0.044)	(0.003)	(0.115)	(0.001)	(1.162)	(0.019)
Age	-0.015**	0.001**	-0.159***	-0.002***	1.133***	0.007**
	(0.007)	(0.000)	(0.020)	(0.000)	(0.193)	(0.003)
Female	-0.477***	-0.041***	-0.016	0.033***	4.008***	0.019
	(0.043)	(0.003)	(0.113)	(0.001)	(1.203)	(0.018)
BMI	-0.101***	-0.023***	1.885***	0.016***	-2.942***	-0.009***
	(0.011)	(0.000)	(0.020)	(0.000)	(0.148)	(0.002)
Spell Duration in Months	0.179***	0.006***	0.330***	0.004***	-2.133***	-0.115***
	(0.007)	(0.000)	(0.014)	(0.000)	(0.117)	(0.002)
Baseline(Mean)	1.189	0.158	16.251	0.197	41.853	1.071
Time Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33084	33084	33084	33084	33084	33084
R-squared Adj.	0.063	0.215	0.565	0.461	0.115	0.111
Older						
AI + Human Coach	0.937***	0.039***	1.686***	0.018***	53.593***	0.278***
	(0.050)	(0.003)	(0.136)	(0.002)	(1.516)	(0.021)
Age	0.001	0.002***	-0.099***	-0.001***	1.781***	0.004*
	(0.006)	(0.000)	(0.015)	(0.000)	(0.146)	(0.002)
Female	-0.513***	-0.040***	-0.456***	0.026***	-5.671***	-0.106***
	(0.048)	(0.003)	(0.130)	(0.001)	(1.586)	(0.021)
BMI	-0.091***	-0.024***	1.8486***	0.016***	-3.927***	-0.016***
	(0.011)	(0.001)	(0.024)	(0.000)	(0.207)	(0.003)
Spell Duration in Months	0.182***	0.005***	0.346***	0.004***	-2.158***	-0.113***
	(0.007)	(0.000)	(0.016)	(0.000)	(0.141)	(0.002)
Baseline(Mean)	1.218	0.173	15.513	0.185	58.830	1.097
Time Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33047	33047	33047	33047	33047	33047
R-squared Adj.	0.099	0.215	0.511	0.426	0.110	0.102

Notes: The table shows a regression with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

Table A.11: Impact of AI+Human Coaches on Based on Different BMIs of Consumers

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No.of Food Logs	Weekly No. of Weight Logs
Below Median						
AI + Human Coach	1.085***	0.068***	1.139***	0.015***	60.412***	0.271***
	(0.046)	(0.003)	(0.083)	(0.001)	(1.496)	(0.019)
Age	-0.003	0.002***	-0.097***	-0.001***	2.142***	0.011***
	(0.004)	(0.000)	(0.006)	(0.000)	(0.129)	(0.002)
Female	-0.713***	-0.082***	0.421***	0.031***	-8.772***	-0.106***
	(0.049)	(0.003)	(0.084)	(0.001)	(1.605)	(0.021)
BMI	-0.231***	-0.054***	1.483***	0.017***	-5.592***	-0.025***
	(0.054)	(0.001)	(0.051)	(0.000)	(0.452)	(0.005)
Spell Duration in Months	0.213***	0.008***	0.317***	0.004***	-2.199***	-0.118***
	(0.008)	(0.000)	(0.011)	(0.000)	(0.159)	(0.002)
Baseline(Mean)	1.355	0.227	10.176	0.140	58.276	1.137
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33075	33075	33075	33075	33075	33075
R-squared Adj.	0.114	0.213	0.263	0.278	0.119	0.091
Above Median						
AI + Human Coach	0.739***	0.020***	1.783***	0.018***	44.896***	0.193***
	(0.045)	(0.002)	(0.152)	(0.002)	(1.136)	(0.012)
Age	0.002	0.001***	-0.132***	-0.001***	1.512***	0.007***
	(0.004)	(0.000)	(0.014)	(0.000)	(0.088)	(0.001)
Female	-0.342***	-0.014***	-1.024***	0.030***	5.258***	0.004
	(0.045)	(0.002)	(0.144)	(0.002)	(1.168)	(0.018)
BMI	-0.054***	-0.010***	2.050***	0.014***	-1.853***	0.001
	(0.008)	(0.000)	(0.032)	(0.000)	(0.193)	(0.003)
Spell Duration in Months	0.158***	0.004***	0.347***	0.003***	-2.083***	-0.110***
	(0.007)	(0.000)	(0.016)	(0.000)	(0.112)	(0.002)
Baseline(Mean)	1.060	0.107	21.321	0.239	43.312	1.033
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	32613	32613	32613	32613	32613	32613
R-squared Adj.	0.050	0.065	0.355	0.241	0.097	0.117

Notes: The table shows a regression with robust standard errors clustered at the user level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.12: Impact of AI+Human Coaches Based on Start Date of Consumer

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No.of Food Logs	Weekly No. of Weight Logs
			Below Start Date Split			
Earlier Subscriber						
AI + Human Coach	0.798***	0.043***	1.282***	0.014***	49.189***	0.286***
	(0.047)	(0.002)	(0.116)	(0.001)	(1.246)	(0.018)
Age	-0.000	0.001***	-0.115***	-0.001***	1.882***	0.010***
	(0.004)	(0.000)	(0.009)	(0.000)	(0.104)	(0.001)
Female	-0.479***	-0.039***	-0.388***	0.028***	-3.555***	-0.036**
	(0.045)	(0.003)	(0.112)	(0.0012)	(1.292)	(0.018)
BMI	-0.115***	-0.024***	1.882***	0.016***	-3.2402***	-0.012***
	(0.010)	(0.000)	(0.012)	(0.000)	(0.165)	(0.002)
Spell Duration in Months	0.210***	0.006***	0.386***	0.004***	-1.085***	-0.112***
	(0.008)	(0.000)	(0.016)	(0.000)	(0.151)	(0.002)
Baseline(Mean)	1.120	0.155	16.091	0.193	45.231	1.040
Time Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33149	33149	33149	33149	33149	33149
R-squared Adj.	0.062	0.217	0.543	0.449	0.108	0.094
Later Subscriber						
AI + Human Coach	0.982***	0.040***	1.595***	0.018***	55.092***	0.181***
	(0.045)	(0.003)	(0.122)	(0.001)	(1.306)	(0.020)
Age	-0.001	0.002***	-0.123***	-0.001***	1.721***	0.008***
	(0.004)	(0.000)	(0.011)	(0.000)	(0.099)	(0.002)
Female	-0.511***	-0.042***	-0.095	0.032***	1.178	-0.056***
	(0.047)	(0.003)	(0.121)	(0.001)	(1.381)	(0.021)
BMI	-0.078***	-0.023***	1.849***	0.016***	-3.708***	-0.013***
	(0.011)	(0.000)	(0.022)	(0.000)	(0.171)	(0.003)
Spell Duration in Months	0.161***	0.005***	0.306***	0.003***	-2.871***	-0.115***
	(0.007)	(0.000)	(0.013)	(0.000)	(0.115)	(0.002)
Baseline(Mean)	1.284	0.177	15.658	0.188	55.786	1.125
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	32984	32984	32984	32984	32984	32984
R-squared Adj.	0.103	0.219	0.532	0.438	0.123	0.114

Notes: The table shows a regression with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

A.9 Web Appendix: Heterogeneity Results with Matching for each Subsample

In this Appendix, we present results of the analysis of heterogeneity of AI+Human effects across different observable characteristics of consumers such as gender, age, etc. with a different matching procedure. In our main results, we first found matched samples of AI+Human and AI consumers and then subsampled them based on gender, age, etc. In this section, we first subsample the entire data based on age, gender, etc. and then conduct a separate matching procedure within each subsample to find matched sets of AI+Human versus AI consumers within each subsample. We find that our findings here are largely consistent with the results presented in Section 4.3 of the main paper.

Table A.13: Male and Female Consumers - with separate matching for each subsample

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
Males						
AI + Human Coach	0.813***	0.048***	1.100***	0.011***	41.912***	0.202***
	(0.084)	(0.005)	(0.225)	(0.002)	(2.483)	(0.034)
Age	-0.006	0.001***	-0.086***	-0.001***	2.204***	0.014***
	(0.004)	(0.000)	(0.011)	(0.000)	(0.174)	(0.002)
BMI	-0.136***	-0.030***	2.103***	0.016***	-5.218***	-0.026***
	(0.009)	(0.000)	(0.027)	(0.002)	(0.284)	(0.003)
Spell duration in Months	0.234***	0.006***	0.384***	0.004***	-1.993***	-0.108***
	(0.009)	(0.000)	(0.019)	(0.000)	(0.189)	(0.002)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	19625	19625	19625	19625	19625	19625
R-squared Adj.	0.081	0.245	0.550	0.437	0.096	0.010
Females						
AI + Human Coach	0.894***	0.040***	1.382***	0.016***	58.164***	0.242***
	(0.049)	(0.003)	(0.172)	(0.000)	(1.484)	(0.027)
Age	0.002	0.002***	-0.133***	-0.002***	1.584***	0.007***
	(0.004)	(0.000)	(0.013)	(0.000)	(0.105)	(0.001)
BMI	-0.082***	-0.220***	1.785***	0.016***	-3.003***	-0.007***
	(0.005)	(0.000)	(0.021)	(0.000)	(0.179)	(0.003)
Spell duration in Months	0.201***	0.007***	0.349***	0.004***	-1.291***	-0.110***
	(0.008)	(0.000)	(0.017)	(0.000)	(0.149)	(0.002)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	30229	30229	30229	30229	30229	30229
R-squared Adj.	0.058	0.219	0.553	0.456	0.010	0.010

Notes: The table shows regressions with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

Table A.14: Younger and Older Consumers - with separate matching for each subsample

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
Younger						
AI + Human Coach	0.803***	0.044***	1.065***	0.013***	50.683***	0.184***
	(0.061)	(0.003)	(0.150)	(0.001)	(1.466)	(0.024)
Age	-0.012	0.001**	-0.166***	-0.002***	1.100***	0.008***
	(0.007)	(0.000)	(0.019)	(0.000)	(0.215)	(0.002)
BMI	-0.110***	-0.024***	1.908***	0.016***	-3.318***	-0.011***
	(0.013)	(0.000)	(0.022)	(0.000)	(0.177)	(0.002)
Spell duration in Months	0.212***	0.007***	0.362***	0.004***	-1.866***	-0.112***
	(0.008)	(0.000)	(0.015)	(0.000)	(0.144)	(0.002)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25879	25879	25879	25879	25879	25879
R-squared Adj.	0.069	0.225	0.577	0.474	0.092	0.105
Older						
AI + Human Coach	0.893***	0.420***	1.903***	0.010***	53.047***	0.331***
	(0.077)	(0.004)	(0.220)	(0.002)	(1.891)	(0.029)
Age	0.003	0.001***	-0.099***	-0.001***	1.761***	0.003***
	(0.007)	(0.000)	(0.019)	(0.000)	(0.168)	(0.002)
BMI	-0.105***	-0.025***	1.870***	0.016***	-4.421***	-0.016***
	(0.014)	(0.000)	(0.029)	(0.000)	(0.240)	(0.003)
Spell duration in Months	0.213***	0.007***	0.375***	0.004***	-1.523***	-0.106***
	(0.009)	(0.000)	(0.020)	(0.000)	(0.170)	(0.002)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24961	24961	24961	24961	24961	24961
R-squared Adj.	0.105	0.225	0.507	0.423	0.088	0.097

Notes: The table shows regressions with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

Table A.15: Consumers with BMI below and above median - with separate matching for each subsample

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
BMI Below Median						
AI + Human Coach	1.121***	0.070***	1.095***	0.014***	57.637***	0.223***
	(0.061)	(0.004)	(0.117)	(0.001)	(1.945)	(0.029)
Age	-0.002***	0.002***	-0.093***	-0.001***	2.239***	0.013***
	(0.005)	(0.000)	(0.007)	(0.000)	(0.150)	(0.002)
Female	-0.700***	-0.086***	0.533***	0.032***	-6.494***	0.116***
	(0.054)	(0.004)	(0.093)	(0.001)	(1.885)	(0.025)
BMI	-0.217***	-0.056***	1.525***	0.017***	-6.110***	-0.028***
	(0.050)	(0.001)	(0.051)	(0.000)	(0.550)	(0.006)
Spell duration in Months	0.264***	0.010***	0.365***	0.004***	-1.471***	-0.117***
	(0.010)	(0.000)	(0.014)	(0.000)	(0.205)	(0.003)
Baseline(Mean)	1.997	0.271	10.776	0.148	93.195	1.311
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24951	24951	24951	24951	24951	24951
R-squared Adj.	0.141	0.220	0.280	0.286	0.089	0.081
BMI Above Median						
AI + Human Coach	0.732***	0.019***	1.731***	0.017***	45.110***	0.240***
	(0.055)	(0.003)	(0.198)	(0.002)	(1.479)	(0.023)
Age	0.002	0.001***	-0.120***	-0.001***	1.560***	0.006***
	(0.005)	(0.000)	(0.011)	(0.000)	(0.105)	(0.001)
Female	-0.320***	-0.014***	-0.950***	0.030***	8.082***	0.012***
	(0.052)	(0.002)	(0.154)	(0.002)	(1.391)	(0.018)
BMI	-0.059***	-0.010***	2.097***	0.014***	-2.088***	-0.000***
	(0.009)	(0.000)	(0.030)	(0.000)	(0.235)	(0.003)
Spell duration in Months	0.177***	0.004***	0.359***	0.003***	-1.820***	-0.103***
	(0.008)	(0.000)	(0.017)	(0.000)	(0.136)	(0.002)
Baseline(Mean)	1.563	0.121	22.494	0.250	70.653	1.140
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25279	25279	25279	25279	25279	25279
R-squared Adj.	0.051	0.070	0.368	0.251	0.076	0.119

Notes: The table shows regressions with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

Table A.16: Earlier and Later Subscribers - with separate matching for each subsample

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
Earlier Subscriber						
AI + Human Coach	0.771***	0.042***	1.331***	0.013***	49.757***	0.287***
	(0.053)	(0.003)	(0.142)	(0.001)	(1.430)	(0.022)
Age	-0.002	0.001***	-0.120***	-0.001***	2.085***	0.011***
	(0.003)	(0.000)	(0.009)	(0.000)	(0.116)	(0.001)
Female	-0.439***	-0.038***	0.314***	0.029***	-2.234***	-0.033***
	(0.048)	(0.003)	(0.121)	(0.001)	(1.447)	(0.019)
BMI	-0.109***	-0.024***	1.910***	0.016***	-3.516***	-0.013***
	(0.007)	(0.000)	(0.020)	(0.000)	(0.185)	(0.002)
Spell duration in Months	0.236***	0.007***	0.414***	0.004***	-0.555***	-0.109***
	(0.009)	(0.000)	(0.018)	(0.000)	(0.174)	(0.002)
Baseline(Mean)	1.689	0.192	16.456	0.198	77.377	1.226
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26752	26752	26752	26752	26752	26752
R-squared Adj.	0.069	0.222	0.553	0.455	0.098	0.087
Later Subscriber						
AI + Human Coach	0.896**	0.038***	1.311***	0.015***	51.425***	0.168***
	(0.060)	(0.004)	(0.169)	(0.002)	(1.613)	(0.026)
Age	-0.000	0.001***	-0.121***	-0.001***	1.690***	0.008***
	(0.005)	(0.000)	(0.013)	(0.000)	(0.115)	(0.002)
Female	-0.529***	-0.044***	-0.168***	0.031***	-4.157***	0.034***
	(0.055)	(0.003)	(0.134)	(0.001)	(1.584)	(0.022)
BMI	-0.090***	-0.024***	1.870***	0.015***	-4.155***	-0.016***
	(0.013)	(0.000)	(0.025)	(0.000)	(0.201)	(0.003)
Spell duration in Months	0.182***	0.005***	0.315***	0.003***	-2.635***	-0.110***
	(0.008)	(0.000)	(0.016)	(0.000)	(0.137)	(0.002)
Baseline(Mean)	1.883	0.205	16.378	0.197	87.062	1.232
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26461	26461	26461	26461	26461	26461
R-squared Adj.	0.107	0.226	0.533	0.439	0.099	0.114

Notes: The table shows regressions with robust standard errors clustered at the user level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

A.10 Web Appendix: Regression results for Weight Loss (Dependent Variable) at a Fixed Time Duration: Non-Overlapping Sub-Samples

The main results remain the same as in Section 4.4 ("Treatment Effects for Fixed Time Periods") of the main paper. The impact of AI+Human coaches increases with time, but at diminishing rates. On average, there is an approximately 67-79% (with variation based on which sub-sample we are looking at) increase in the weight loss impact of AI+Human coaches at the 6-month mark relative to the 3-month mark. This is quite comparable to the approximately 75% impact we saw in Section 4.4.

Table A.17: Regression results for Weight Loss (Dependent Variable) at a Fixed Time Duration: Non-Overlapping Sub-Samples

T2 - T1	Y	AI+Human Coach
6-3 months	3 months	1.154***
		(0.157)
9-6 months	3 months	0.78***
		(0.159)
9-6 months	6 months	1.397***
		(0.202)
12-9 months	3 months	1.206***
		(0.329)
12-9 months	6 months	2.145***
		(0.395)
12-9 months	9 months	2.553***
		(0.438)
12+ months	3 months	2.112***
		(0.290)
12+ months	6 months	3.538***
		(0.352)
12+ months	9 months	3.87***
		(0.402)
12+ months	12 months	3.833***
		(0.441)

Notes: The table reports the regression results for Weight Loss at a Fixed Time Duration (i.e Weight Loss after Y months for a sample of users who stay in the app for time T1-T2) as the dependent variable. The coefficients for independent variables are reported in the table. The controls in these regressions are the same as in the main regressions reported in Table 5 in the paper

A.11 Web Appendix: Other Robustness Checks

Table A.18: Comparing Different Matching Procedures

	Weight Loss	Weight Loss Fraction	Weight Loss Goal	Weight Loss Goal Fraction	Weekly no. of food logs	Weekly no. of weight logs
Nearest Neighbor (NNR)	0.948***	0.041***	1.567***	0.018***	51.910***	0.222***
	(0.035)	(0.002)	(0.098)	(0.001)	(0.977)	(0.015)
Mahalanobis without replacement	0.894***	0.042***	1.445***	0.016***	52.386***	0.231***
	(0.034)	(0.002)	(0.089)	(0.001)	(0.957)	(0.014)
Mahalanobis (outlier removal)	1.049***	0.050***	1.600***	0.018***	53.217***	0.259***
	(0.039)	(0.002)	(0.105)	(0.001)	(1.144)	(0.017)
Mahalanobis (0.1)	0.901***	0.045***	1.237***	0.014***	51.175***	0.205***
	(0.047)	(0.002)	(0.108)	(0.001)	(1.088)	(0.017)

Notes: The table shows a regression with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1. Estimates are calculated on the matched sample of adopters of the Coach plan and the AI plan using different matching methods. Month Fixed effects are used, and the unit of analysis is the user spell. The dependent variable in Column-1 is Weight Loss, Column-2 is the fraction of Weight Loss Fraction, Column-3 is goal loss, Column-4 is the fraction of goal loss, Column-5 is the Average Number of Food Logs, and Column-6 is the Avg No. of Weight Logs. The independent variables are indicators for the user's adoption of a paid plan, age of the user, BMI of the user, and the number of months since the user joined the corresponding plan. Mahalanobis (0.02) - Mahalanobis matching without replacement with 0.02 threshold on the propensity score. Mahalanobis (0.1) - Mahalanobis matching without replacement with 0.1 thresholds on the propensity score. NNR - nearest neighbor matching with replacement. Mahalanobis (outlier removal) - Mahalanobis matching without replacement with 0.02 threshold after outlier removal in the pre-period.

Table A.19: Robustness to Excluding the Control Variables from the Main Specification

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
AI + Human Coach	0.925***	0.045***	1.309***	0.015***	51.656***	0.208***
	(0.029)	(0.002)	(0.081)	(0.001)	(0.761)	(0.011)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared Adj.	0.015	0.011	0.004	0.005	0.065	0.005

Notes: The table shows regression results with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

Table A.20: Robustness to Including Spell Count Fixed Effects

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No.of Food Logs	Weekly No. of Weight Logs
AI + Human Coach	0.838***	0.046***	1.379***	0.015***	52.393***	0.246***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Age	0.003	0.002***	-0.112***	-0.001***	1.767***	0.009***
	(0.145)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Female	-0.496***	-0.04***	-0.271***	0.03***	-1.0	-0.049***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.178)	(0.0)
BMI	-0.103***	-0.023***	1.847***	0.016***	-3.484***	-0.011***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Spell Duration in Months	0.174***	0.006***	0.303***	0.003***	-2.136***	-0.111***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Spell Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared Adj.	0.059	0.231	0.072	0.840	0.100	0.100

Notes: This table reports results from regressions that include fixed effects for the number of spells (i.e., spell count) alongside time fixed effects. The goal is to account for heterogeneity in users' engagement levels, as reflected in the number of active goal-setting periods. The estimated effect of AI + Human Coach remains stable in magnitude and statistically significant across all outcomes, indicating that the main results are not driven by differences in spell frequency. Standard errors are clustered at the user level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.21: Joint interaction model (all moderators simultaneously)

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
AI + Human Coach	0.908***	0.042***	1.444***	0.016***	52.382***	0.231***
	(0.032)	(0.002)	(0.087)	(0.001)	(0.945)	(0.014)
Age	-0.002	0.002***	-0.140***	-0.002***	1.633***	0.006***
	(0.003)	(0.000)	(0.015)	(0.000)	(0.111)	(0.002)
Female	-0.464***	-0.032***	-0.411***	0.027***	-9.528***	-0.072***
	(0.047)	(0.003)	(0.157)	(0.002)	(1.403)	(0.024)
BMI	-0.056***	-0.019***	1.784***	0.015***	-2.254***	-0.006*
	(0.006)	(0.001)	(0.026)	(0.000)	(0.176)	(0.003)
Spell Duration in Months	0.181***	0.005***	0.336***	0.004***	-2.116***	-0.113***
	(0.005)	(0.000)	(0.011)	(0.000)	(0.093)	(0.001)
AI+Human × Female	-0.055	-0.013***	0.281	0.005***	13.747***	0.040
	(0.065)	(0.004)	(0.186)	(0.002)	(1.958)	(0.030)
AI+Human × Age	0.000	-0.000	0.035**	0.000**	0.278*	0.004*
	(0.005)	(0.000)	(0.017)	(0.000)	(0.154)	(0.002)
AI+Human × BMI	-0.051***	-0.007***	0.153***	0.001***	-2.030***	-0.011***
	(0.008)	(0.001)	(0.030)	(0.000)	(0.250)	(0.004)
Baseline (Mean)	1.187	0.166	15.860	0.191	50.636	1.084
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	66,131	66,131	66,131	66,131	66,131	66,131
R-squared Adj.	0.066	0.219	0.543	0.447	0.121	0.105

Notes. The standard errors reported here are robust to multiple comparisons, using the Bonferroni correction.

Table A.22: Analysis for the first spell alone

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
AI + Human Coach	0.799***	0.035***	1.338***	0.015***	51.582***	0.290***
	(0.039)	(0.002)	(0.074)	(0.001)	(0.885)	(0.013)
Age	0.004	0.001***	-0.104***	-0.001***	1.799***	0.010***
	(0.003)	(0.000)	(0.005)	(0.000)	(0.070)	(0.001)
Female	-0.628***	-0.041***	-0.553***	0.028***	-0.818	-0.044***
	(0.040)	(0.002)	(0.074)	(0.001)	(0.939)	(0.013)
BMI	-0.107***	-0.023***	1.889***	0.016***	-3.252***	-0.011***
	(0.005)	(0.000)	(0.012)	(0.000)	(0.115)	(0.002)
Spell Duration in Months	0.151***	0.005***	0.299***	0.003***	-2.716***	-0.108***
	(0.006)	(0.000)	(0.011)	(0.000)	(0.103)	(0.001)
Baseline Mean	1.453	0.163	17.131	0.205	47.279	0.906
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39,108	39,108	39,108	39,108	39,108	39,108
Adj. R-squared	0.049	0.205	0.538	0.437	0.130	0.122

Table A.23: Analysis with Dependent Variables at the Weekly Level (Normalized to Spell Duration)

	Weekly Weight Loss (Kgs)	Weekly Weight Loss Fraction
AI + Human Coach	0.0329***	0.0043***
	(0.009)	(0.002)
Age	0.0007	0.002***
	(0.001)	(0.000)
Female	-0.022*	-0.0019***
	(0.002)	(0.000)
BMI	-0.0278***	-0.0029***
	(0.000)	(0.000)
Baseline (mean)	0.0318	0.0026
Time Fixed Effects	Yes	Yes
R-squared Adj.	0.011	0.143

Notes: The table shows regression results with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

Table A.24: Results: Alternative method of estimating weight at the end of the time window

	3-month window	6-month window	9-month window	12-month window
AI + Human Coach	0.677***	1.965***	4.506***	4.638**
	(0.148)	(0.325)	(0.717)	(1.923)
Age	-0.015**	0.016	-0.019	-0.087
	(0.007)	(0.020)	(0.044)	(0.101)
Female	-0.988***	-0.563**	-0.707	-5.142***
	(0.112)	(0.283)	(0.724)	(1.967)
BMI	-0.126***	-0.205***	-0.373***	-0.325
	(0.016)	(0.039)	(0.093)	(0.284)
Spell Duration in Months	0.074***	0.175***	0.286	0.100
	(0.017)	(0.051)	(0.192)	(1.013)
Observations	15,214	3,047	461	48
Adj. R-squared	0.032	0.046	0.153	0.145

Notes.: In this table, we report the results where the ending weight was interpolated using weights reported 3, 6, 9, and 12 months before and after the date when the time window ended. In cases where the weight on that date was available, that was used for analysis directly.

A.12 Web Appendix: PSM matching Free versus Paid Plans

In this section, we report the analysis of paid (i.e., both AI-only and AI+Human plans) versus free plans. A PSM-matching procedure was conducted like for our main analysis. We report the histograms of propensity scores before and after matching, histograms for the variables used for matching before and after matching, and the regression results for the matched samples.

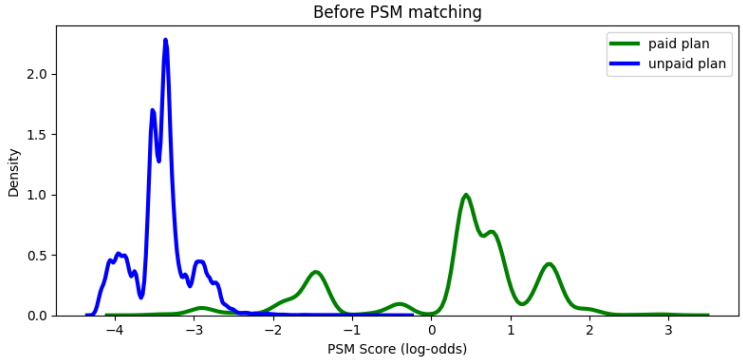

Figure A.11: Propensity Scores Before Matching

Notes: The figure above shows the propensity scores *before* matching. The blue line represents the propensity scores for the Free plans users, and the green line for the Paid plans users.

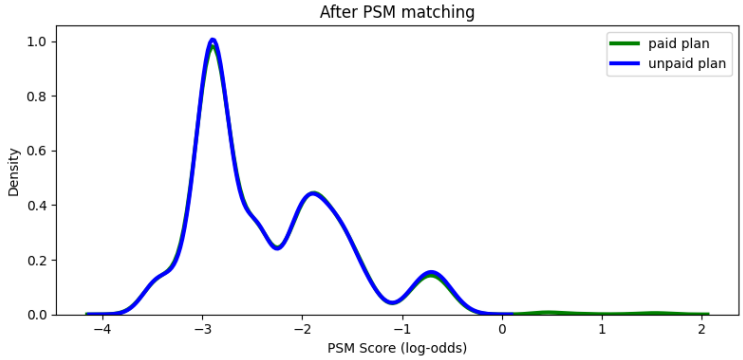

Figure A.12: Propensity Scores After Matching

Notes: The Figure above shows the propensity scores *after* matching. The blue line represents the propensity scores for the Free plans users, and the green line for the Paid plans users.

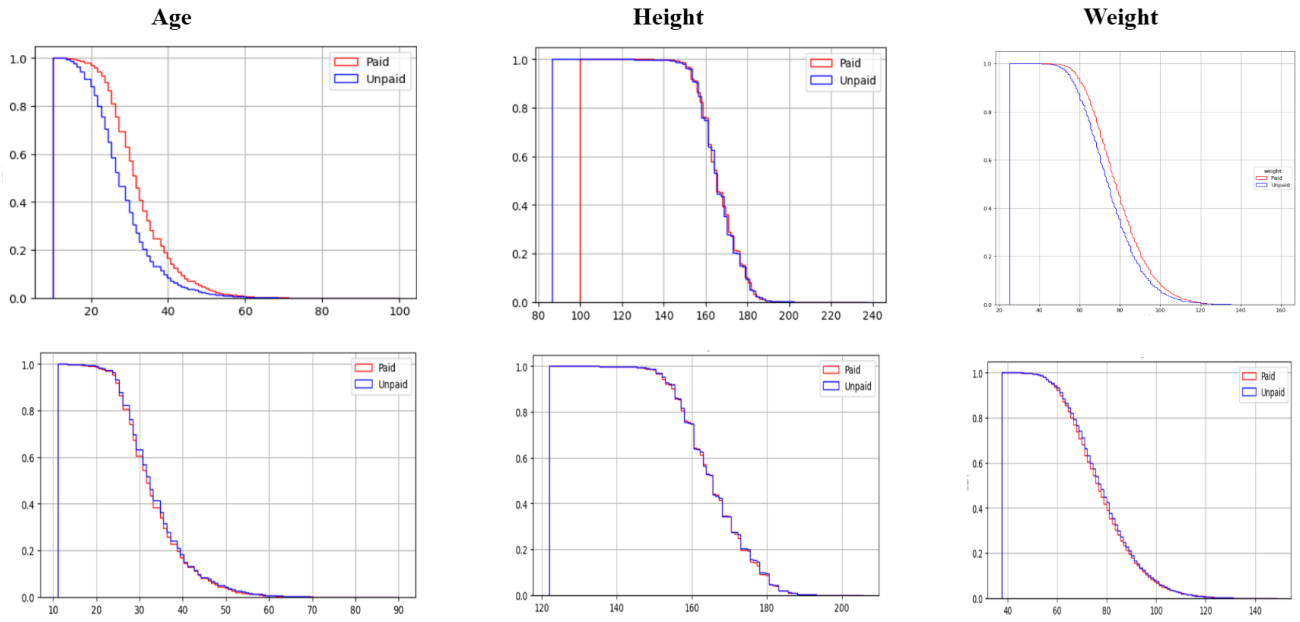

Figure A.13: Distribution of Free vs Paid plans Before (upper panel) and After (lower panel) Matching

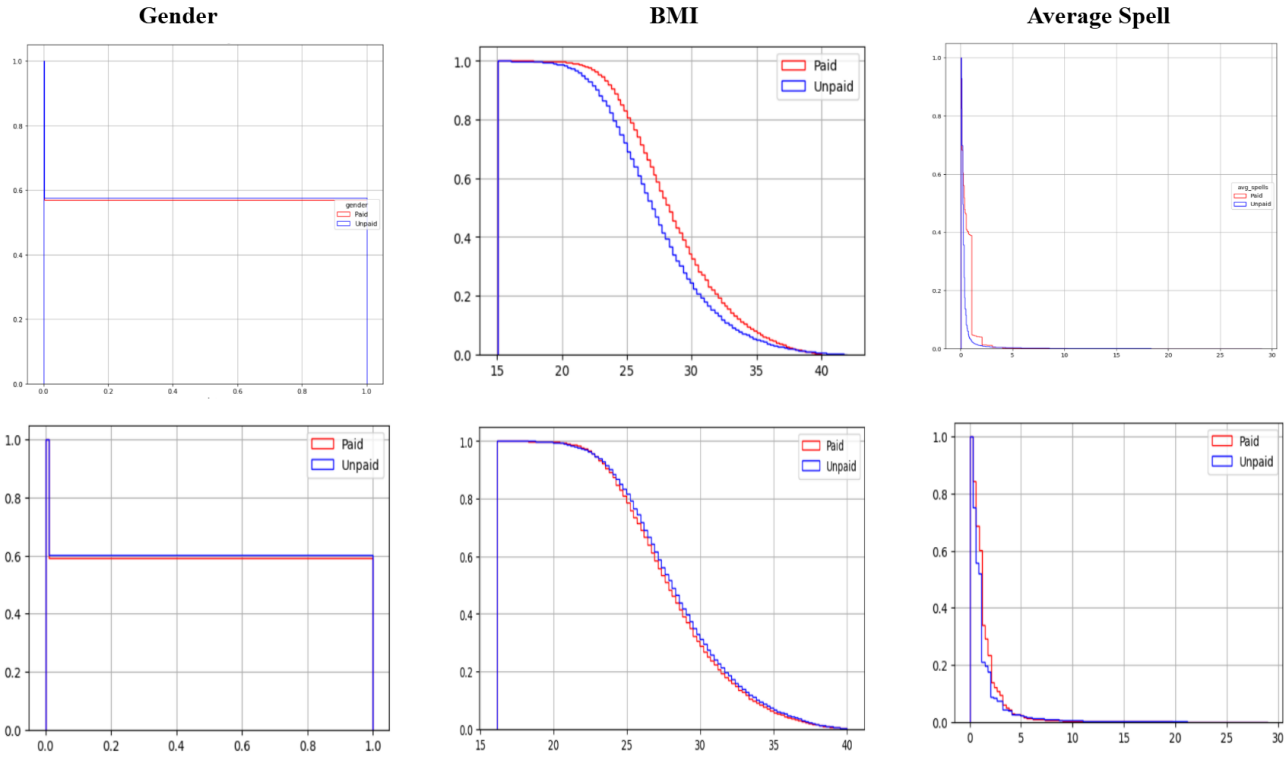

Figure A.14: Distribution of Free vs Paid plans Before (upper panel) and After (lower panel) Matching

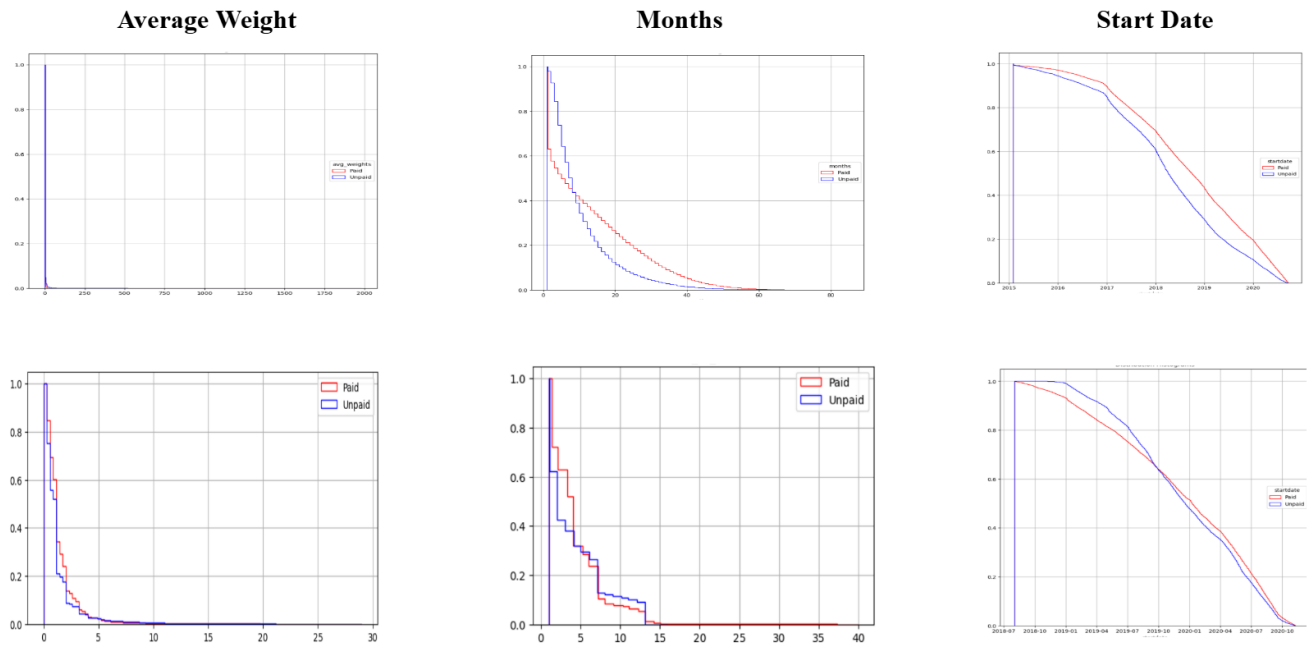

Figure A.15: Distribution of Free vs Paid plans Before (upper panel) and After (lower panel) Matching

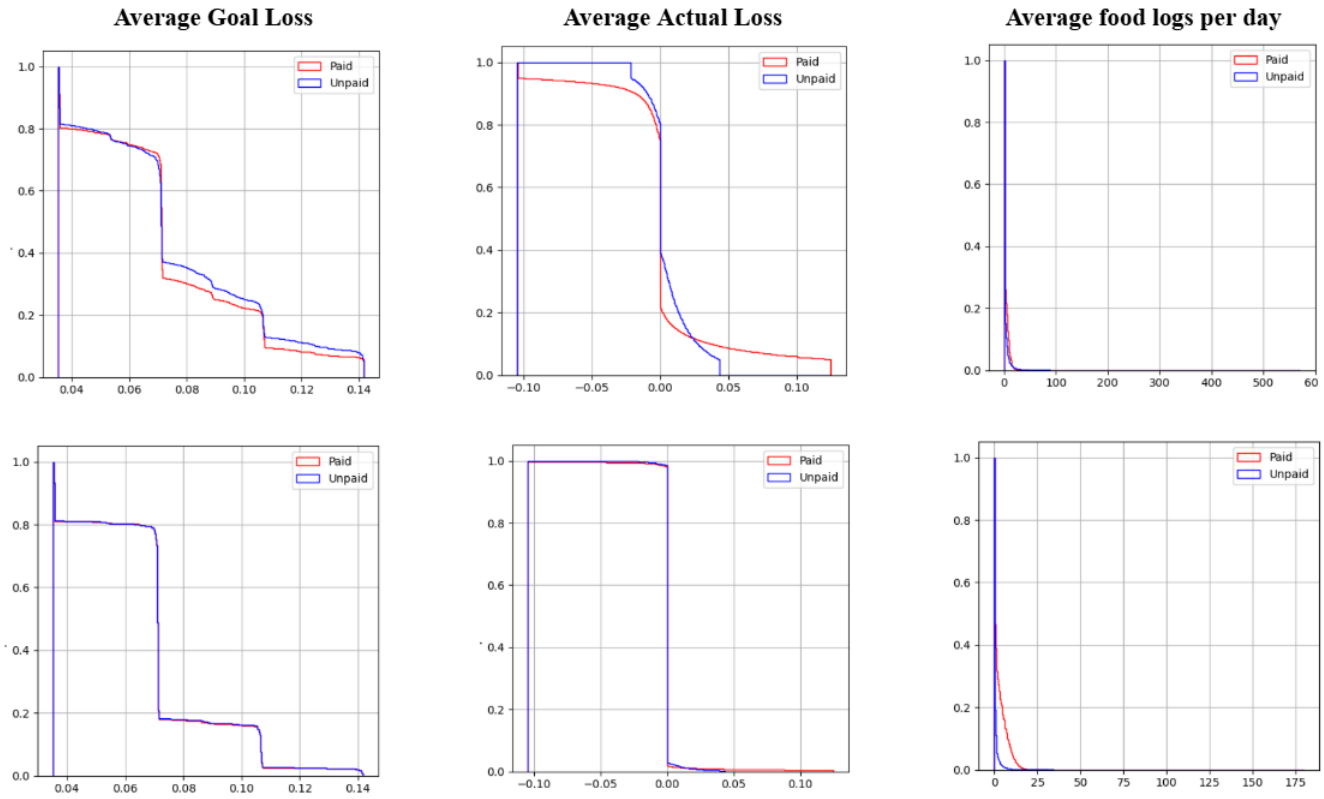

Figure A.16: Distribution of Free vs Paid plan Before (upper panel) and After (lower panel) Matching

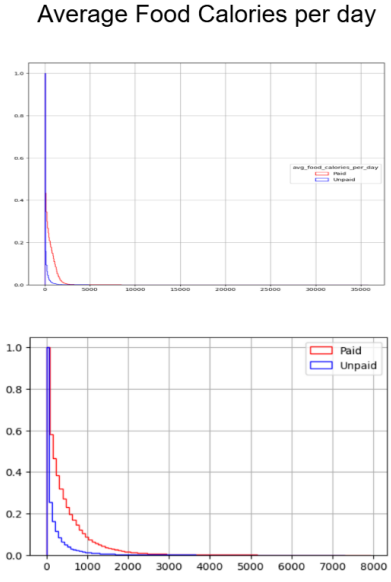

Figure A.17: Distribution of Free vs Paid plan Before (upper panel) and After (lower panel) Matching

Table A.25: Comparison of Free versus Paid Plans

	Weight Loss (Kgs)	Weight Loss Fraction	Weight Loss Goal (Kgs)	Weight Loss Goal Fraction	Weekly No. of Food Logs	Weekly No. of Weight Logs
Paid Plans	1.562***	0.096***	0.527***	0.002***	52.334***	0.206***
	(0.011)	(0.001)	(0.028)	(0.000)	(0.323)	(0.005)
Age	0.004***	0.001***	-0.081***	-0.001***	1.266***	0.007***
	(0.001)	(0.000)	(0.002)	(0.000)	(0.019)	(0.000)
Female	-0.238***	-0.018***	-0.691***	0.025***	-0.753**	-0.019***
	(0.011)	(0.001)	(0.027)	(0.000)	(0.317)	(0.005)
BMI	-0.056***	-0.017***	2.052***	0.018***	-1.881***	-0.011***
	(0.001)	(0.000)	(0.004)	(0.000)	(0.043)	(0.001)
Spell Duration in Months	0.021***	0.001***	0.029***	0.000***	-1.037***	-0.054***
	(0.001)	(0.000)	(0.002)	(0.000)	(0.024)	(0.000)
Baseline Mean	0.082	0.079	17.300	0.210	19.941	0.818
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	182288	182288	182288	182288	182288	182288
Adj. R-squared	0.109	0.253	0.616	0.005	0.065	0.005

Notes: The table shows regression results with robust standard errors clustered at the user level in parentheses. *** p <0.01, ** p <0.05, * p <0.1

References

Altonji JG, Elder TE, Taber CR (2005) Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy* 113:151–184.

Chin SO, Keum C, Woo J, Park J, Choi HJ, Woo Jt, Rhee SY (2016) Successful weight reduction and maintenance by using a smartphone application in those with overweight and obesity. *Scientific reports* 6(1):1–8.

Ingels JS, Misra R, Stewart J, Lucke-Wold B, Shawley-Brzoska S (2017) The effect of adherence to dietary tracking on weight loss: using hlm to model weight loss over time. *Journal of diabetes research* 2017.

Knäuper B, Carriere K, Frayn M, Ivanova E, Xu Z, Ames-Bull A, Islam F, Lowensteyn I, Sadikaj G, Luszczynska A, et al. (2018) The effects of if-then plans on weight loss: Results of the mcgill chip healthy weight program randomized controlled trial. *Obesity* 26(8):1285–1295.

Laing BY, Mangione CM, Tseng CH, Leng M, Vaisberg E, Mahida M, Bholat M, Glazier E, Morisky DE, Bell DS (2014) Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. *Annals of internal medicine* 161(10_Supplement):S5–S12.

Liu F, Kong X, Cao J, Chen S, Li C, Huang J, Gu D, Kelly TN (2015) Mobile phone intervention and weight loss among overweight and obese adults: a meta-analysis of randomized controlled trials. *American journal of epidemiology* 181(5):337–348.

Manchanda P, Packard G, Pattabhiramaiah A (2015) Social dollars: The economic impact of customer participation in a firm-sponsored online customer community. *Marketing Science* 34(3):367–387.

Oster E (2019) Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business and Economic Statistics* 37(2):187–204.

Patel ML, Hopkins CM, Brooks TL, Bennett GG (2019) Comparing self-monitoring strategies for weight loss in a smartphone app: randomized controlled trial. *JMIR mHealth and uHealth* 7(2):e12209.

Rosenbaum PR (2002) *Observational Studies* (Springer).

Uetake K, Yang N (2020) Inspiration from the "biggest loser": social interactions in a weight loss program. *Marketing Science* 39(3):487–499.

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

Wharton CM, Johnston CS, Cunningham BK, Sterner D (2014) Dietary self-monitoring, but not dietary quality, improves with use of smartphone app technology in an 8-week weight loss trial. *Journal of nutrition education and behavior* 46(5):440–444.

Yancy WS, Shaw PA, Reale C, Hilbert V, Yan J, Zhu J, Troxel AB, Foster GD, Volpp KG (2019) Effect of escalating financial incentive rewards on maintenance of weight loss: a randomized clinical trial. *JAMA network open* 2(11):e1914393–e1914393.