

NBER WORKING PAPER SERIES

COMMITTING TO EXERCISE:
CONTRACT DESIGN FOR VIRTUOUS HABIT FORMATION

Jeremy D. Goldhaber-Fiebert
Erik Blumenkranz
Alan M. Garber

Working Paper 16624
<http://www.nber.org/papers/w16624>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2010

Supported in part by National Institute on Aging Grant P30 AG24957 to Stanford University, Center for Advancing Decision Making in Aging, and by the Department of Veterans Affairs. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2010 by Jeremy D. Goldhaber-Fiebert, Erik Blumenkranz, and Alan M. Garber. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Committing to Exercise: Contract Design for Virtuous Habit Formation
Jeremy D. Goldhaber-Fiebert, Erik Blumenkranz, and Alan M. Garber
NBER Working Paper No. 16624
December 2010
JEL No. D03,I1,I12

ABSTRACT

Sedentary lifestyles, obesity, and obesity-related chronic diseases have become increasingly common among U.S. adults, posing a major health policy challenge. While individuals may be interested in exercising more to reduce these health risks, they often have difficulty forming long-term exercise habits. Behavioral economic devices like commitment contracts aid individuals make repeated actions in situations where there are upfront costs and the benefits, though substantial, are delayed. It is not known whether such contracts can help individuals to sustain increased exercise. We conducted a randomized controlled trial to test whether nudges and anchoring could be used to shift the types of exercise commitment contracts people entered into using a web-based contract creation tool. At the time of contract creation, users selected a contract length (duration); number of times a week to exercise (frequency); and a financial penalty for failing to live up to the contract in a given week (stake). We randomly set the default duration shown to users (8 weeks, 12 weeks, or 16 weeks). Outcomes include: contract acceptance; chosen duration, frequency, total exercise sessions; and chosen financial stake. We analyzed the data using multivariable regressions and also developed a theoretical model of active choice in the context of nudges, fitting the model to the data using non-linear optimization. 619 users, age 18-69, were included in the study, of whom 61% accepted/signed an exercise commitment contract. Users who were shown a longer default contract durations were significantly more likely to choose a contract of longer duration. There was no difference in the likelihood of accepting contracts or in the chosen exercise frequency or financial stakes. Our model of active choice suggested that almost 50% of users were highly susceptible to default values for contract duration, with the greatest effect for users interested in exercise contracts with durations nearer to the nudged defaults. This implication of the model was confirmed by quantile regressions (greatest effect of nudges for contract durations between the 40th and 80th percentiles). With changes in default values, individuals can be nudged into longer exercise commitment contracts that obligate them to greater numbers of exercise sessions.

Jeremy D. Goldhaber-Fiebert
Stanford Health Policy
Centers for Health Policy and
Primary Care and Outcomes Research
Stanford University
117 Encina Commons
Stanford, CA 94305-6019
jeremygf@stanford.edu

Alan M. Garber
The VA Palo Alto Health Care System
and Centers for Health Policy and
Primary Care and Outcomes Research
Stanford University
117 Encina Commons
Stanford, CA 94305-6019
and also NBER
garber@stanford.edu

Erik Blumenkranz
Yale University
20 Larguita Lane
Portola Valley, CA 94028
erik.blumenkranz@yale.edu

Executive Summary

Background: Sedentary lifestyles and obesity have become increasingly common among U.S. adults. As a result, obesity-related chronic diseases are a major health policy challenge. While individuals often indicate their interest in exercising more and reducing their health risks, they often encounter difficulties in making such changes permanent – it is difficult to form long-term exercise habits. Behavioral economic devices like commitment contracts have been used in other areas such as savings where repeated actions with upfront costs lead to substantial but delayed benefits. It is not yet known whether such contracts are effective in promoting sustained increases in exercise.

Methods: We conducted a randomized controlled trial to test whether nudges and anchoring could be used to shift the types of exercise commitment contracts people entered into. The study comprised 619 users of a free, web-based tool to create exercise contracts. At the time of contract creation the users had to select a contract length (duration); number of times a week to exercise (frequency); and a financial penalty for failing to live up to the contract conditions in a given week (stake). In a full factorial design, we randomly set the default duration shown to users (8 weeks, 12 weeks, or 16 weeks) and a message noting that individuals using financial stakes were more successful. Outcomes included: 1) the likelihood of accepting a contract; 2) the duration of the contract; 3) the frequency of the contract; 4) the total number of exercise sessions chosen (frequency*duration); 5) the financial stakes chosen. We analyzed results using multivariable quantile, logistic, ordered logistic, and tobit regression models as appropriate. Additionally, we developed a theoretical model of active choice in the context of nudges, fitting the model to the data using non-linear optimization.

Results: 619 users, age 18-69, were included in the study, of whom 61% accepted/signed the exercise commitment contracts that they created. Users who were shown a longer default contract duration were significantly more likely to choose a contract of longer duration. There was no difference in the likelihood of accepting contracts, the frequency of weekly exercise chosen, or the financial stakes chosen between the randomization groups. Our model of active choice suggested that almost 50% of users were highly susceptible to default values for contract duration, with the greatest effect for users interested in exercise contracts with durations nearer to the nudged defaults. This implication of the model was confirmed by quantile regressions, which demonstrated that the largest effect on contract duration occurred in users whose contract durations were in the 40th to 80th percentiles.

Conclusions: Through the use of different default values, users of exercise commitment contracts can be nudged into longer contracts. Their total amount of contracted exercise increases since there are no significant effects on other margins of the contract (e.g., frequency, financial stakes).

I. Introduction

The high and rising prevalence of obesity among U.S. adults presents a major health policy challenge.¹ Obese adults are at increased risk for multiple diseases including hypertension, type 2 diabetes, and other cardiovascular conditions.² A recent study quantifying the loss of Quality-Adjusted Life Years (QALYS) from modifiable risk factors between 1993 and 2008 found that obesity's impact had surpassed even smoking in its detrimental effect on American health.³ The increasingly sedentary lifetime of many American adults partially explains the increase in obesity over this period. The typical American spends many hours in front of a computer or TV screen, failing to exercise despite exercise's substantial health benefits: exercise is known to significantly improve cardiovascular health, blood pressure and cholesterol and to reduce the risk of developing diseases such as diabetes and osteoporosis.⁴

For most people, the failure to exercise is unlikely to be due to a lack of knowledge about or interest in exercise. It is more likely the result of difficulties in translating initial motivation to exercise into sustained behavioral change. When asked, many American adults state preferences for improved health and claim to want to lose weight and to be more active.⁵ In addition to stated preferences, each year substantial numbers of these individuals join gyms or begin other exercise regimens, though their use declines rapidly with time. Others do not even start such exercise. Behavioral economics invokes the concept of dynamic inconsistency to explain this phenomenon. The average decision-maker tends to exhibit a changing discount function over time, specifically, exhibiting impatience in near-term trade-offs compared with long-term trade-offs. Thaler (1981) simplified this problem with two questions about apples. First, would you prefer one apple today or two apples tomorrow? Second, would you prefer one apple in one year, or two apples in one year and a day? Although the tradeoffs are fundamentally the same, many would choose an apple today over two tomorrow, while none would choose one over two in the future. Dynamic inconsistency becomes particularly relevant when the decision-maker faces a choice that presents delayed reward and immediate costs, such as exercise. While the prospect of exercising for three months and improving one's health seems attractive now, the daily decision to exercise is a much more difficult one.

Traditionally, firms have used both financial and nonfinancial incentives in order to improve employee health and performance and to reduce the costs of health insurance. Such incentives have focused on employee health and wellness, attempting to induce employees to get into better shape, lose weight, quit smoking, and/or participate in exercise programs. These programs have had mixed success. Furthermore, traditional methods can be quite costly.⁶ Their costliness stems from two main features: 1) they are offered to many people, many of whom are already engaged in the activity so need no further incentive; 2) the incentive value may be larger than necessary because the relationships between incentive size, behavior change, and improvement in health are unknown in most contexts. Furthermore, on the individual level, without an employer continually providing such an external incentive, intrinsic motivation may be insufficient to overcome dynamic inconsistency.

A more internally-focused approach to address the problem of dynamic inconsistency and failure to sustain certain behaviors over time is to adopt some form of commitment mechanism. Such commitment mechanisms would only be demanded

¹ Katherine M. Flegal; Margaret D. Carroll; Cynthia L. Ogden; Lester R. Curtin. Prevalence and Trends in Obesity Among US Adults, 1999-2008 *JAMA*. 2010;303(3):235-241.

² Sugerman HJ, Wolfe LG, Sica DA, Clore JN. "Diabetes and hypertension in severe obesity and effects of gastric bypass-induced weight loss.

³ Jia and Lubetkin. "Trends in Quality-Adjusted Life-Years Lost Contributed by Smoking and Obesity."

⁴ Myers, Jonathan. "Exercise and Cardiovascular Health"

⁵ Williamson DF, Serdula MK, Anda RF, Levy A, Byers T. "Weight loss attempts in adults: goals, duration, and rate of weight loss."

⁶ Kane *AJPM* 2004; Sutherland *MCR* 2008

by people who have recognized their dynamic inconsistency and therefore seek ways to overcome it. O'Donoghue and Rabin (1999) defined two types of consumers: sophisticated and naive. As opposed to the naive consumer, the sophisticated consumer is aware of his time-inconsistent preferences, and thus will often look to pre-commit his future actions. As far back as Strotz (1955), the study on dynamic inconsistency and commitment devices has been inextricably linked. Laibson (1997) studied commitment devices in the form of illiquid financial assets, such as retirement and pension plans, and their relationship to consumption and income. Ariely and Wertenbroch (2002) found that given the option, students would elect to self-impose deadlines for their work, implying awareness of dynamic inconsistency, and willingness to take up a commitment mechanism. Giné, Karlan and Zinman (2009) found that amongst a group of smokers in the Philippines there was substantial interest in and improved success using a commitment device in the form of a deposit contract to quit smoking.⁷

In addition to the behavioral changes that occur during the period of the actual commitment contract, long-term changes in behavior well beyond the contract period may also result. There is a substantial research literature examining the concept of habit formation. One conceptual approach is to postulate that the utility that one derives from consumption depends upon both current and past patterns of consumption. One could consider exercise to be a form of consumption, and thus a habit-forming behavior. Studies by Charness and Gneezy (2009) and Acland and Levy (2010) provide results consistent with this view, observing that student gym attendance remained higher even after they no longer received financial incentives to exercise. This type of habit-formation through short-term behavior change has had a plausible biological rationale for many years: according to physiological evidence, exercise increases pleasure sensation and decreases pain perception through a number of potential mechanisms.⁸ Sufficient numbers of repeated exposures to this positive effect should, therefore, provide biochemical reinforcement to habit formation -- a change in the utility of current consumption brought about by patterns of past consumption.

Given the theoretical basis for the value of commitment devices to address dynamic inconsistency in sustaining weight loss and exercise, the question remains: what type of commitment device would work best? The deposit contract, where money is forfeited upon failure to complete an agreement, features prominently in the recent literature on financial commitment mechanisms.⁹ One is inclined to think that the prospect of a financial loss would provide greater incentive than a financial gain based on Kahneman and Tversky's (1979) results regarding loss aversion -- there is greater disutility from the risk of an equivalently sized loss than there is a utility increase from a comparable monetary gain. The commitment contract in the context of habit formation in exercise is intended as a temporary mechanism whereby the individual manipulates the costs of not exercising in the short-term while their long-term orientation towards and experience of exercise shifts -- the commitment device is needed until the new habit is formed.

In this paper, we describe the initial steps we have taken in assessing and optimizing the use of commitment contracts to foster long-term exercise habits. To date, we have studied the demand for exercise commitment contracts as a form of commitment device that may be useful in forming longer-term exercise habits. In the context of a free, web-based commitment contract site, we describe randomized experimental results of using nudges (e.g., changes to the default value of the suggested contract length) to shift the chosen features of an exercise commitment contract -- the number of weeks the contract is for and the use of financial stakes. Specifically, we assessed whether these default values and nudges could be modified to increase the total number of exercise sessions that users contracted to undertake. This first, necessary step establishes that nudges and default

⁷ Additionally, DellaVigna and Malmendier (2004), and Gul and Pesendorfer (2004) are amongst the many others that have worked on this subject.

⁸ Carr D, Bullen BA, Skrinar GS, et al. Physical conditioning facilitates the exercise-induced secretion of beta-endorphins and beta-lipoprotein in women. *N Engl J Med* 1981;305:560-3

⁹ Giné, Karlan, and Zinman (2009) and Volpp et al (2008)

values can be used to increase the length of contracts that are selected without negatively impacting other dimensions of the contract. The next step would be to establish that individuals randomized to longer contracts engage in the same or more exercise. We address this in the final section of the paper with a description of a randomized, longitudinal study design that is currently accruing subjects that evaluates how nudge-based shifts in chosen commitment contracts relates to patterns of exercise during the contract period. In future research we plan to evaluate how these shifts in exercise in the short-term lead to exercise patterns after the commitment contracts have been completed. That is, can exercise commitment contracts be used to create long-term exercise habits?

II. Experimental Design

We evaluated a web-based tool that allows consumers to construct and undertake exercise commitment contracts aimed at achieving their personal exercise goals. Interested individuals visit the website and construct deposit contracts in which they commit to exercise over a fixed number of weeks a specified number of sessions per week with a selected amount of money at risk of forfeiture for each week in which the individual reports a failure to exercise according to the contract. Users may also designate a referee to confirm their weekly reports so as to reduce potential cheating. We term these contracts "exercise commitment contracts" or "commitment contracts" for the remainder of the paper.

In the experiment, the process of constructing a contract is as follows. The first step is to select the duration (in weeks) of the contract and the frequency of exercise per week. Each time the construction of a contract is attempted the user is randomly assigned an 8, 12, or 16 week default value displayed to them for the contract duration. The user can modify the contract duration with a simple mouse click. On the same webpage, the user also chooses a number of times per week to exercise (termed here "frequency"). No randomization is performed on the default frequency seen by the user. The next step is to designate the dollar value of a penalty for every failed week (i.e., a week where the user does not exercise at least as many times as specified by the frequency), as well as the recipient of the forfeited money. The options for the latter include charities and anti-charities (i.e., causes noxious to the user) to be named by the user or a friend or foe also to be named by the user. The anti-charity option is intended to provide an extra incentive not to fail: the thought of supporting Manchester United would send any serious Liverpool soccer fan right to the gym. Notably, the website does not suggest a specific amount of money to be staked, and no money at all is also a valid choice. However, the website randomizes whether users see a financial nudge in the form of a message informing them that "people who do [put money at stake] are twice as likely to succeed." The final steps consist of naming a referee if the user so chooses, and ultimately accepting the finished contract. After accepting a contract the user may also choose to name supporters to follow his or her progress and to be informed of weekly successes and failures.

We employed a full factorial randomization design in which the default durations (8, 12, and 16 weeks) and the presence/absence of a financial message nudge produce six possible conditions to which the user might be randomized.

For this study, data included information on all attempts to create exercise commitment contracts between February 28 and June 4, 2010. Information is available about each unique user who attempted to create an exercise contract during this period and how far each contract attempt was taken before it was completed, actively cancelled, or abandoned. We analyzed our results using an intention-to-treat approach. We merged the data for users with multiple contracts into a single entry, composed of the first condition they were randomized to and the final contract details they selected. Data on unique users included age, sex, state/country of the user, length of membership with the website and the number and types of any prior contracts they had previously undertaken.

The prospectively determined study inclusion criteria were that included individuals must have attempted to create an exercise commitment contract between February 28 and June 4, 2010 and were between 18 and 69 years of age at that time. Before analyzing the data, we removed two individuals from the dataset whose information seemed highly anomalous and either contained errors or unintentional entries (4+ exercise commitment contracts created within several minutes of each other for the same time period and with all other features largely the same). We assessed the success of our randomization based on whether the characteristics of users were largely balanced across each of the randomization conditions, implying successful randomization and also the likely balance of unobserved confounders. Successful randomization enabled us to effectively determine the causal effect of default conditions and nudges on the attributed features of the exercise commitment contract that the user ultimately selected.

Outcomes included: the duration in weeks of exercise contracts chosen; the impact of randomization on other margins of importance (e.g., the weekly frequency of exercise sessions contracted); the total number of exercise sessions contracted for; the impact of randomization on the use of and size of financial penalties for failing to successfully satisfy the exercise contract in a given week.

We assessed the impact of each randomization condition on our outcomes using multivariable quantile, order logistic, tobit, and simple logistic regression models depending on the outcome to be tested. We also developed a model to describe the activity/passivity of user choice and response to nudges (see Section III below) and fit the model to our data using non-linear optimization.

We performed all analyses using Stata/SE 11.0 (StataCorp, College Station, TX, USA) and Matlab (MathWorks, Natick, MA, USA).

III. Model of Commitment Contract Choice

We developed a simple model to distinguish between types of users considering exercise commitment contracts.

We defined two types of users: *Active Choosers* and *Suggestible Choosers*. *Active Choosers* have strong preferences for particular features of their exercise commitment contracts that they have previously determined -- in this case, they either decided on their desired contract duration before entering the website or they had a very good idea of what kind of contract they wanted to establish. In contrast, *Suggestible Choosers* have more diffuse preferences for exercise commitment contract durations and are more easily shifted by nudges and default values.

The distribution of contract durations is defined as

$$q_{i,j} = \frac{n_{i,j}}{\sum_{w=\{j\}} n_{i,w}} \quad (1)$$

The selected contract duration is split into seven categories, with $q_{i,j}$ representing the proportion of total users randomized to default value i who selected category j . The proportion $q_{i,j}$ is determined by dividing the number selecting category j by the total number of individuals randomized to i regardless of their selection from categories j in the set of all categories w .

However, randomization has an effect on the categories chosen. *Suggestible Choosers* simply choose the category that they are randomized to unlike *Active Choosers* according to the model. The challenge is that we do not know the proportion of our population who are *Active Choosers*, which we term x , nor do we know the distribution of preferences among *Active Choosers*. In this case, π_j represents the proportion of *Active Choosers* who would select category j .

If we assume that *Active Choosers* do not deviate at all from their prior preferences, then we can write the observed proportion of all users (*Active Choosers* and *Suggestible Choosers*) $\hat{q}_{i,j}$ who choose a given category j under each randomized condition i as follows:

$$\hat{q}_{i,j} = \pi_j x \text{ for } i \neq j \quad (2)$$

$$\hat{q}_{i,j} = \pi_j x + (1 - x) \text{ for } i = j \quad (3)$$

Notice that *Suggestible Choosers* only appear in equation (3) since they are entirely influenced by the randomized nudge ($i=j$).

In fact, we define a more general model for the observed proportions choosing each duration category that includes this simple model but also incorporates the idea that *Active Choosers* may be partially influenced by the randomized nudge, shown in equations (4) and (5) below. In it, we assume that *Active Choosers* whose prior preferences are close the randomized nudge are more likely to be influenced by the nudge (i.e., if one wants an 18 week contract and sees a 16 week nudge, that person is more likely to change than the person who wants a 6 week contract). We express this influence by proximity on *Active Choosers* using two variables, α and β , to express how strong the suggestibility of *Active Choosers*. The variable α can take on any value between 0 and 1, where 0 implies no suggestibility for *Active Choosers* as per our formulation in equations (2) and (3) above and where 1 implies that suggestibility operates. The variable β can take on any value greater than or equal to 1 and defines the how important the distance between a person's preferred contract duration i and the randomized nudge is for reducing the influence of susceptibility for *Active Choosers*. The larger the value of β the less influence the nudge has on people with prior preferred duration that is different than the randomized nudge.

$$\hat{q}_{i,j} = \pi_j x - \frac{\alpha \pi_j x}{|j - i|^\beta} \text{ for } i \neq j \quad (4)$$

$$\hat{q}_{i,j} = \pi_j x + \left(\sum_{w \neq i} \frac{\alpha \pi_w x}{|w - i|^\beta} \right) + (1 - x) \text{ for } i = j \quad (5)$$

The above model is identifiable in our particular experimental setup because we have three randomization conditions with respect to contract duration -- i is either 8, 12, or 16 weeks and consequently w is made up of 7 categories for j :

j=1	j=2	j=3	j=4	j=5	j=6	j=7
<8 weeks	8 weeks	9-11 weeks	12 weeks	13-15 weeks	16 weeks	>16 weeks

We therefore observe 21 combinations of $\hat{q}_{i,j}$. To identify the model, we need to solve for 7 π_j 's, for x , and for α and β -- a total of 10 unknowns.

We estimated parameters for this model with α set at zero (*Active Choosers* do not deviate from their preexisting preferences due to the presentation of a randomized default) and one to explore the importance of a "partial influence" on *Active Choosers* where α and β were allowed to vary. We repeated these analyses using both the full set of accepted contracts as well as the limited set of accepted contracts for "first time choosers" (users with no previous exercise contracts). Only accepted contracts were used since our interest was limited to the default value's effect on selection of actual exercise time.

Based on the observed results of users choosing different contract durations under each randomization condition, the model was solved to minimize the sum of squared residuals between the observed and predicted q values using Matlab.

IV. Results

Study population

Between February 28 and June 4, 2010, a total of 642 unique users attempted to create a total of 752 exercise commitment contracts. After application of the exclusion criteria, 619 users age 18-69 years who had made a total of 718 contract attempts remained eligible for analysis. The results below describe contract attempts for these 619 users, of whom 61.1% ultimately accepted a contract.¹⁰ The average age of the users was 34 years, 56% were female, and 78% attempted to create a contract from the U.S. (Table 1). When assessed using an F-test, user characteristics were not statistically different between the randomization groups, suggesting that randomization was successful.

¹⁰ 59.2% for users who had not created an exercise commitment contract before -- "first time choosers".

Table 1. Characteristics of users creating exercise contracts, those randomized to different default values, and those accepting/not accepting contracts

	All users	Users randomized to 8 weeks duration default	User randomized to 12 weeks duration default	Users randomized to 16 weeks duration default
	N=619	N=214	N=197	N=208
Age	34.2 (11.4)	34.7 (11.2)	34.0 (11.2)	33.8 (11.7)
Female (%)	56%	56%	55%	56%
Non-US (%)	22%	24%	23%	20%

Users who see longer default contract durations choose longer exercise contracts

The use of larger suggested default values to nudge users to create exercise commitment contracts with longer durations increases the total number of exercise sessions they ultimately accept without substitution on other contract margins. Specifically, higher default values for contract duration increase the chosen contract duration for users without causing them to be less likely to accept a contract, to choose fewer exercise sessions per week, to be less likely to choose a financial contract, to lower their financial contract stakes, or to be less likely to choose a referee for their contract. Table 2 shows rates of contract acceptance and other contract characteristics based upon the default values and nudges each that users were shown (see also Appendix 1). Mean contract duration increased from approximately 10 weeks among users randomized to an 8 week default value to nearly 14 weeks for those randomized to a 16 week default value. Notably, while the average financial stake for accepted contracts was consistently higher in the presence of the nudge, the significance of this value disappeared when limited to users who had never contracted before.¹¹

¹¹ Effect disappears when limited to first time choosers, see Appendix 5

Table 2. Contract Details by Randomization Group

Duration randomized default	8 Weeks		12 Weeks		16 Weeks	
Financial stake message randomized condition	Message Shown	No Message Shown	Message Shown	No Message Shown	Message Shown	No Message Shown
Accept	66	58	67	53	67	67
Accept %	60.6%	55.8%	60.9%	61.6%	62.6%	65.0%
Mean Frequency Accept	3.97	4.00	3.79	3.70	3.82	4.18
Mean Frequency Decline	3.69	3.37	3.34	3.46	3.43	3.12
Mean Duration Accept *	11.46	12.02	12.73	12.59	14.43	13.40
Mean Duration Decline *	10.16	10.52	12.79	10.33	14.98	14.50
Total Days Committed Accept	45.48	48.07	48.26	46.54	55.15	56.01
Total Days Committed Decline	37.48	35.43	42.77	35.70	51.41	45.24
% Financial Stake Accept **	38.5	38.6	41.8	43.4	37.3	38.8
% Financial Stake Decline **	24.2	18.8	18.2	26.1	30.0	36.0
Mean Financial Stake Accept (\$/wk) **	16.92	14.55	18.21	17.48	15.80	15.19
Mean Financial Stake Decline (\$/wk) **	29.88	24.17	11.67	7.50	29.44	11.44
% with a Referee Designated Accept	27.3	24.1	29.9	34.0	31.3	31.3
% with a Referee Designated Decline	65.1	54.3	58.1	51.5	70.0	41.7
% Supporters Accept	21.2	10.3	11.9	20.8	.11.9	.11.9
% Supporters Decline ***	0	0	0	0	0	0
Mean # Supporters Accept	1.14	1.33	1.25	1.18	1.00	1.25
Mean # Supporters Decline ***	0	0	0	0	0	0

* No values dropped for duration

** 3 values dropped for indicated statistics only, with 2 values of \$250 and 1 value of \$500

Note - Frequency: 70 missing values, Financial Stake: 64 missing values. Mean financial stake and mean supporters values conditional on value not equal to zero.

*** Supporters were only selected/solicited after contracts were accepted. Hence, no supporters are shown for declined contracts.

Even though contracts involve serious commitments, many users appear to be quite suggestible

Although exercise commitment contracts potentially entail upwards of 40 exercise sessions over a multi-month time period and financial stakes that can exceed several hundred dollars, fitting our model of *Active Choosers* and *Suggestible Choosers* suggest that more than 45% of users are *Suggestible Choosers* -- users whose choice of contract duration is dictated by the default value of the nudge that they are shown. This result differs by no more than 10% across models that are restricted to users choosing contracts for the first time or to models that allow *Active Choosers* to be partially influenced by the nudges. Without nudges, it appears that 6% of *Active Choosers* prefer contracts of 8 weeks duration, 20% of 12 weeks, and 8% of 16 weeks with 26% <8 weeks duration and 15% >16 weeks (see Appendix 2).

Even *Active Choosers* appear to be somewhat influenced by the default contract duration nudges. While both of our models fit the observed data well (see Appendix 2), the fit of our model in which *Active Choosers* are partially influenced is somewhat better, with lower sums of squared residuals for all users and for the subset choosing contracts for the first time (Table 3). It appears that there is more influence on the selected contract durations the nearer they are to the default. In other words,

even *Active Choosers* searching for moderate length contracts relatively close to the nudged default values are likely to shift to the default value.

Table 3. Fit of Models Considering the Effect of Nudges on Active Choosers: Sum of Squared Residuals

Model	Everybody	First Time Choosers Only
No partial influence allowed on <i>Active Choosers</i> : $\alpha=0$	0.0055	0.0047
Partial influence allowed on <i>Active Choosers</i> : $\alpha=1$	0.0029	0.0037

Nudges have their greatest influence on individuals whose preferred contract duration are near the median

Default values influence the chosen exercise contract duration most for those individuals whose contract durations are in the middle of the distribution (from the 40th percentile to the 80th percentile). The effects are robust to adjustment for the sex and age of the user choosing the contract (Table 4). Additionally, the effect becomes more pronounced after the exclusion of users with previous contracts (see Appendix 3) where the impact of the nudges extends down to the 20th percentile. The impact in the center of the distributions is consistent with our *Active Chooser* model findings regarding partial influence for individuals whose preferences are near the nudge values that we presented above. Furthermore, the wider impact on first time contractors is consistent with *Active Choosers* learning their preferred contract duration over time and therefore being less influenced by nudges after having experienced a previous contract of a given duration.

Table 4. Factors Influencing the Distribution Contract Duration*

	(1) .05	(2) .10	(3) .20	(4) .25	(5) .40	(6) .50	(7) .60	(8) .75	(9) .80	(10) .90	(11) .95
Female	-0.00 (0.91)	1.00 (0.61)	1.00 (1.38)	0.00 (0.97)	-0.00 (1.01)	0.00 (0.89)	0.00 (0.89)	0.00 (0.83)	0.00 (0.83)	0.00 (4.76)	-11.00 (8.64)
18-29	-1.00 (1.13)	-1.00 (0.80)	-1.00 (1.63)	-1.00 (1.15)	0.00 (1.21)	0.00 (1.11)	0.00 (1.12)	0.00 (1.08)	-2.00+ (1.08)	-4.00 (5.86)	-16.00 (11.11)
40-49	-1.00 (1.47)	-1.00 (0.93)	-0.00 (2.14)	-0.00 (1.42)	0.00 (1.47)	0.00 (1.29)	0.00 (1.31)	0.00 (1.22)	-2.00 (1.22)	-5.00 (6.96)	-7.00 (13.19)
50+	-0.00 (1.55)	2.00+ (1.17)	-0.00 (2.24)	-0.00 (1.52)	0.00 (1.59)	0.00 (1.44)	0.00 (1.49)	-0.000 (1.38)	-2.00 (1.38)	-5.00 (7.43)	6.00 (9.26)
8 Weeks	2.00+ (1.03)	1.00 (0.74)	-2.00 (1.63)	-2.00+ (1.14)	-4.00** (1.07)	-4.00** (1.07)	-4.00** (1.12)	-0.00 (1.06)	-0.00 (1.06)	-1.00 (5.68)	-4.00 (10.82)
16 Weeks	1.00 (1.02)	-0.00 (0.78)	-1.00 (1.71)	-0.00 (1.20)	1.00 (1.26)	4.00** (1.05)	4.00** (0.98)	4.00** (0.95)	4.00** (0.94)	0.00 (5.69)	-5.00 (10.37)
Constant	3.00** (1.08)	4.00** (0.68)	9.00** (1.88)	10.00** (1.29)	12.00** (1.23)	12.00** (1.14)	12.00** (1.14)	12.00** (1.12)	14.00** (1.06)	21.00** (5.07)	51.00** (9.80)
N	348	348	348	348	348	348	348	348	348	348	348

Quantile Regression

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

* In this table, users with previous exercise contracts are included in the regressions. Separate regressions are shown in the Appendices for users who have never had an exercise contract before.

Users who see longer default contract durations choose contracts with more total exercise sessions because they do not substitute

on other margins (weekly frequency or financial stakes)

The total number of exercise sessions contracted for is increased by nudges on default contract duration. In other words, there appears to be little substitution between contract duration and the frequency of weekly exercise sessions that users contract for. Table 5 shows that the randomization does not have a significant effect on frequency. Thus, overall, the randomization significantly alters the total amount of contracted exercise time.

Table 5. Contract Frequency

	(1)	(2)	(3)	(4)
	Everyone	Everyone	First Time Choosers	First Time Choosers
		(odds ratio)		(odds ratio)
Female	0.204 (0.198)	1.226 (0.243)	0.148 (0.208)	1.160 (0.241)
18-29	-0.001 (0.231)	0.999 (0.230)	-0.030 (0.245)	0.970 (0.238)
40-49	-0.133 (0.290)	0.875 (0.254)	-0.198 (0.307)	0.820 (0.252)
50+	-0.201 (0.330)	0.818 (0.270)	-0.221 (0.351)	0.802 (0.282)
8 Weeks	0.369 (0.240)	1.447 (0.347)	0.412 (0.254)	1.510 (0.384)
16 Weeks	0.332 (0.231)	1.394 (0.322)	0.309 (0.240)	1.362 (0.327)
cut1				
Constant	-2.559** (0.320)	0.077** (0.025)	-2.564** (0.337)	0.077** (0.026)
cut2				
Constant	-1.535** (0.267)	0.216** (0.058)	-1.573** (0.285)	0.207** (0.059)
cut3				
Constant	0.069 (0.250)	1.071 (0.267)	0.011 (0.267)	1.011 (0.271)
cut4				
Constant	0.912** (0.256)	2.489** (0.636)	0.814** (0.272)	2.256** (0.615)
cut5				
Constant	2.265** (0.283)	9.629** (2.726)	2.105** (0.297)	8.209** (2.438)
cut6				
Constant	3.029** (0.321)	20.682** (6.631)	2.853** (0.332)	17.337** (5.758)
N	348	348	313	313

Ordered Logit

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

We verified the effect on total exercise through multivariate quantile regressions of frequency multiplied by duration on the same covariates (see Appendix 4). The effect of the default value on total exercise time between the 40th and 80th percentile

remains significant, confirming our theory that the defaults exert the most influence towards the middle of the distribution. Therefore, although some users already have strong preferences, it is possible to have a large positive effect on the majority of users' total exercise time using the default value.

Randomization also has no significant effect on the use of or size of financial stakes chosen for users' exercise contracts. Having established that the randomization has a significant, positive effect on the contracted amount of exercise, we considered the effect of randomization on financial stakes. A substitution to longer contracts rather than more costly ones would not necessarily be a negative result, but it would indicate a certain level of user sophistication. We used a tobit model to account for the fact that many users choose to have no financial stake at all. Our assumption that the data is left-censored, that users would if possible select negative values, does not seem unreasonable. We found that randomization to longer or shorter suggested contract durations had no significant effect on the contract stakes selected (Table 6). Furthermore, when we examined the impact of observables on contract stakes selected (Table 7 and Appendix 5) we found that the actual contract lengths and frequencies a user selected were also not related in a significant way to the contract stakes chosen. We did find that gender and the interaction between duration and frequency both have a significant effect, although the financial message nudge did not. This interaction term signifies that although frequency and duration generally have a negative effect on financial stake, ambitious users with contracts that are high in frequency and duration often choose higher stakes. As noted above, the randomization itself did not have a strong effect on stakes selected, however.

Table 6. Contracts Stakes and Duration Randomization

	(1)	(2)	(3)	(4)
	Everyone	sigma	First Time Choosers	sigma
	(model)		(model)	
Female	-8.105*		-7.135*	
	(3.407)		(3.334)	
18-29	-3.758		-2.197	
	(4.118)		(4.045)	
40-49	1.344		2.889	
	(5.087)		(5.024)	
50+	1.240		2.281	
	(5.805)		(5.802)	
Message	1.164		0.720	
	(3.416)		(3.347)	
Prev. Exercise Ct	1.436			
	(6.084)			
Months Member	0.153		0.392	
	(0.357)		(0.395)	
8 Weeks	-1.789		0.928	
	(4.180)		(4.106)	
16 Weeks	-2.870		-1.322	
	(4.077)		(3.957)	
Constant	-3.284	26.247**	-5.354	24.428**
	(4.813)	(1.806)	(4.749)	(1.765)
N		346	346	311

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

Table 7. Contract Stake and users' observables including their chosen contract length and duration

	(1)	(2)	(3)	(4)
	Everyone	sigma	First Time Choosers	sigma
	(model)		(model)	
Female	-7.826*		-6.588*	
	(3.354)		(3.264)	
18-29	-5.294		-3.165	
	(4.055)		(3.977)	
40-49	0.130		1.862	
	(4.952)		(4.884)	
50+	0.074		1.464	
	(5.697)		(5.653)	
Message	1.386		0.556	
	(3.345)		(3.267)	
Prev. Exercise Ct	0.757			
	(5.947)			
Months Member	0.199		0.426	
	(0.354)		(0.390)	
Contract Duration	-1.264*		-1.173*	
	(0.579)		(0.566)	
Contract Frequency	-1.969		-2.746	
	(1.843)		(1.795)	
Duration*Frequency	0.264*		0.278*	
	(0.130)		(0.125)	
Constant	6.730	25.688**	7.002	23.842**
	(9.057)	(1.772)	(8.910)	(1.728)
N	346	346	311	311

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

In summary, suggesting longer durations for exercise commitment contracts through the use of default values has a strong positive effect on contract duration and overall number of exercise sessions chose. Neither the duration default value nor financial nudge had an effect on contract acceptance (see Appendix 5). Consequently, it is possible to nudge consumers into committing to longer duration contracts without sacrificing frequency of exercise, value of financial stake, or chances of acceptance.

IV. Conclusion

We often find it difficult to force ourselves to do something that we are reluctant to do, even when we know that it would enhance our well-being. Despite the obvious long-term benefits of exercise, our aversion to its immediate costs often overwhelms any internal motivation that we derive from future benefits. Consequently, many seek an external way to enforce their desire to exercise when they are faced with the immediate choice of whether or not to exercise on a given day.

Thaler and Benartzi (2001) found that default values had a strong effect on university employees' allocation of their retirement savings. The employees' allocation between different funds was highly sensitive to the menu of options they were offered, specifically, the proportion of equities in the menu. Our results are consistent with these findings. We found that except for users who desired very short or very long contract durations, the default value for an exercise contract's duration had a significant effect on the user's selection: higher default values encouraged selection of longer contracts. This result persisted and

was magnified when the data was limited to only first time users.

The goal, however, is to increase total exercise time, and not just contract duration. Fortunately, our results confirmed that contract frequency was unaffected by the randomization. Users did not substitute longer contracts in place of more frequent ones; they simply increased the length of their commitment period. We observed that the randomization did not have a significant effect on contract financial commitments or ultimate acceptance either. The results therefore indicate that it is possible to optimize default values in order to positively influence total contracted exercise time, without decreasing financial stakes or acceptance rates.

These results show that it is possible to increase users' contracted exercise time. But our study did not evaluate compliance with the exercise commitments. Ultimately, the goal is to maximize actual exercise time, rather than committed time. One could imagine a scenario in which a user over-commits to an unrealistic amount of exercise and eventually gives up, a worse result than if he or she had successfully completed a shorter contract. Further study will be required to know with more certainty whether exercise time can be increased through the optimization of default values of commitment contracts.

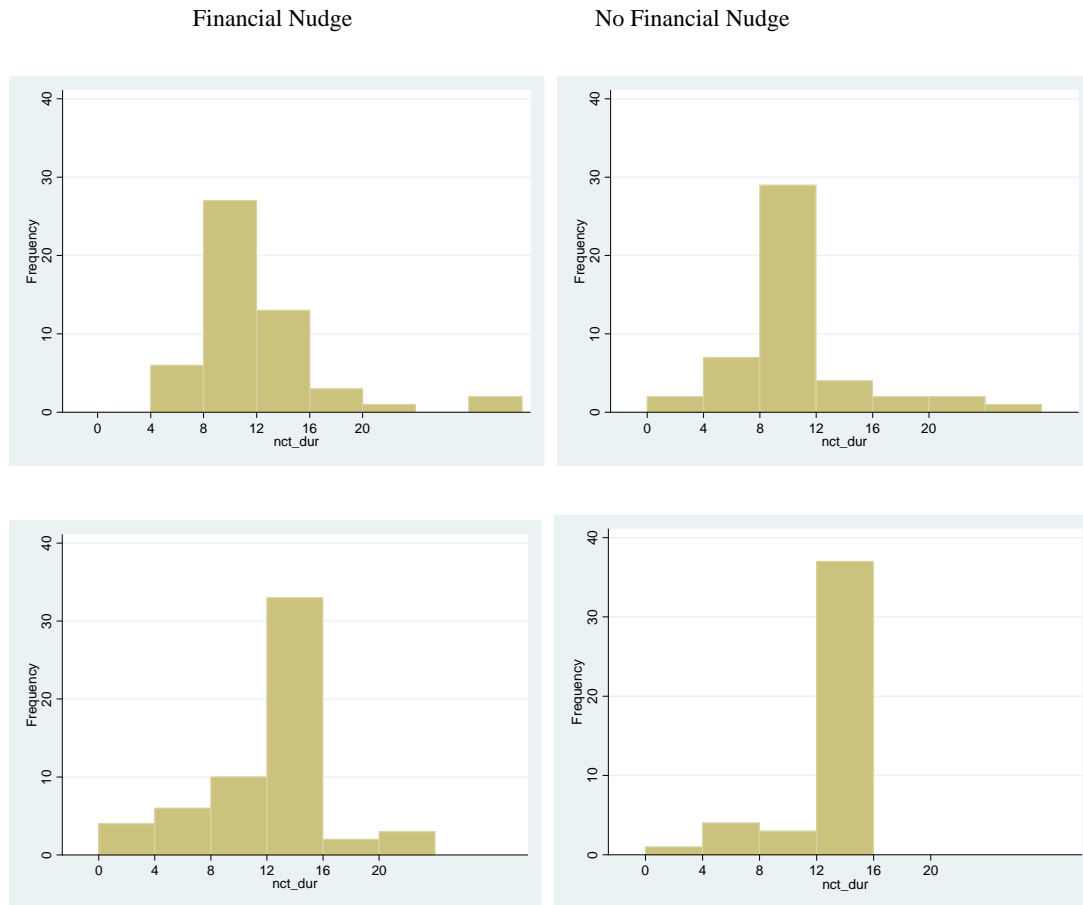
Although exercise commitment contracts may be somewhat effective at increasing short-term exercise, the real benefit of exercise is realized by sustaining it over much longer periods of time. The challenge is to find a method -- such as exercise commitment contracts -- to enable people to form long-term exercise habits. Our current findings demonstrate that nudges are quite powerful in changing the length of the exercise commitment contracts users choose without altering their other contract features (e.g., frequency of weekly exercise, size of financial stakes, etc). Current research involving a longitudinal study of over 1,000 users that assesses their exercise patterns under randomized duration nudges is underway. This research will evaluate the hypothesis that longer contracts chosen via these nudges lead to equivalent or greater engagement in exercise during the contracted period.

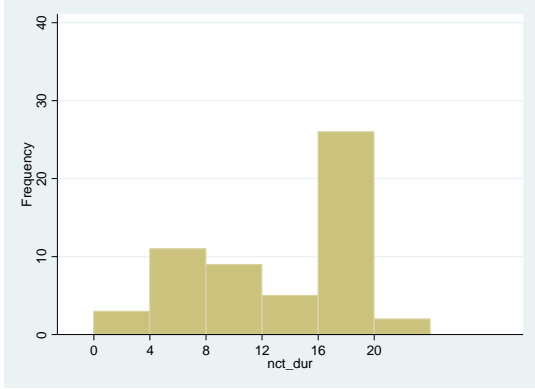
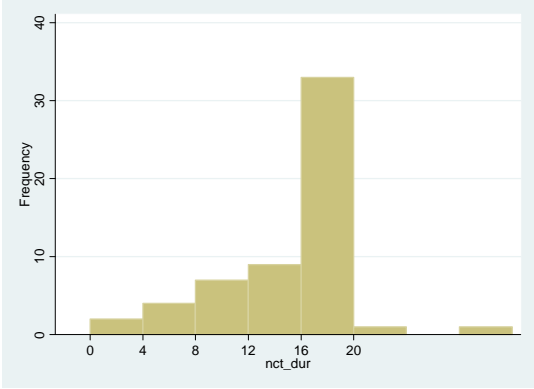
Commitment contracts are low-cost, voluntary means of linking future behavior, which would be difficult to sustain otherwise, to present intentions. Successful application of commitment contracts to exercise for individuals who find it hard to form exercise habits is an important addition to the range of tools available to improve health outcomes. To the extent that longer contracts increase long-term exercise habit formation, carefully designed default values can lead users towards selecting exercise commitment contracts with lengths optimal for addressing sedentary lifestyles. Increasing physical activity is an important element of addressing obesity and improving health among Americans.

Appendix 1.

The figures below show the distribution of chosen contract durations for individuals under each of the 6 randomization conditions (left to right, top to bottom: 1) 8 weeks / financial message nudge show; 2) 8 weeks / no financial message nudge shown; 3) 12 weeks / financial message nudge show; 4) 12 weeks / no financial message nudge shown; 5) 16 weeks / financial message nudge show; 6) 16 weeks / no financial message nudge shown. Notably, the mode is always the nudged contract duration value (e.g., 8 weeks for the 8 week randomization condition, etc.).

Figure A1.1 Duration: Accepted Contracts





Appendix 2.

The graphs below show results of the model of user choice that we fit using Matlab. We consider the effects of contract duration nudges on users making exercise commitment contracts for the first time (“first-time choosers”) and on all users in two sets of graphs. In the sets of graphs, each of the three panels shows the results for the 3 randomization conditions for contract duration: 8 weeks, 12 weeks, and 16 weeks, respectively. In each panel, 4 lines are shown. The first is “Real Values” which is the percentage of users whose contracts actually fell into each of the duration categories. The “Alpha=0” and “Alpha=1” lines refer to the fitted models (see Equations (5) and (6) above) predicting what the user choices would be under the randomization condition. “Alpha=0” refers to the simple model in which *Active Choosers* are not influenced by nudges, whereas “Alpha=1” refers to the model in which *Active Choosers* whose preferences are close to the nudge may be influenced by the nudge. The “P Values” line refers to the π 's in the model, the proportion of users choosing each contract duration category in the absence of a nudge. It can be seen that nudges shift a substantial proportion of individuals – there are many *Suggestible Choosers* among people considering exercise commitment contracts.

Figure A2.1 The effect of contract duration nudges on first-time choosers



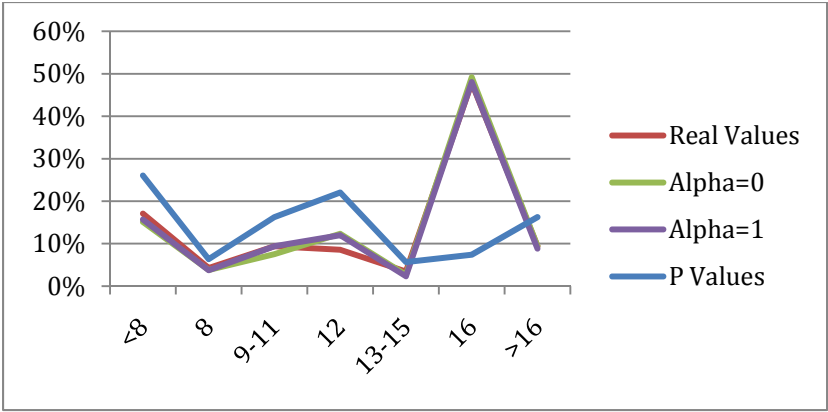
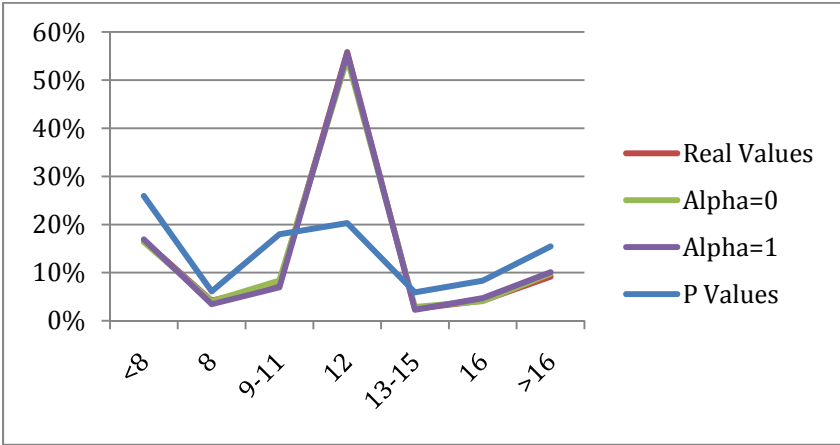
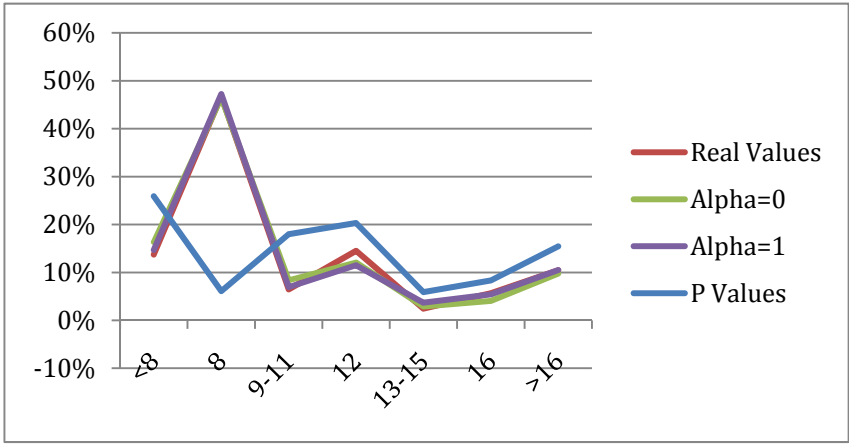
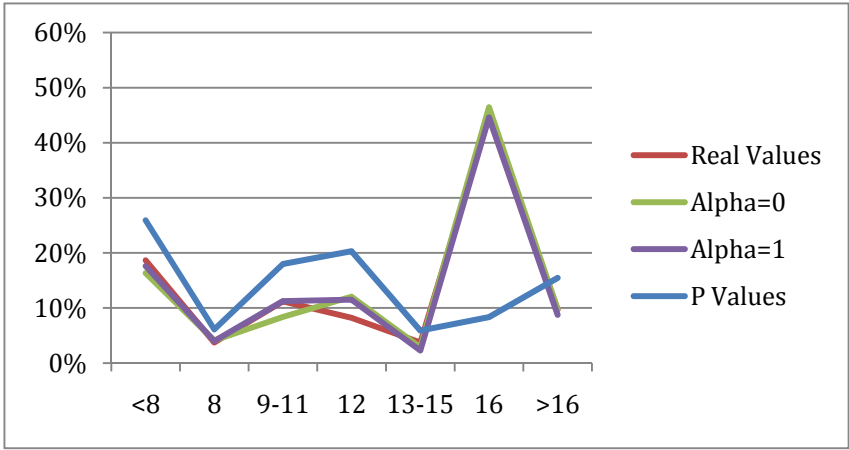


Figure A2.2 The effect of contract duration nudges on all users





Appendix 3.

Contract Duration (users with previous exercise contracts excluded)

Quantile Regression

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

	(1) .05	(2) .20	(3) .25	(4) .40	(5) .50	(6) .60	(7) .75	(8) .80	(9) .95
Female	1.000+ (0.588)	0.000 (1.377)	0.000 (0.944)	0.000 (1.001)	0.000 (0.872)	0.000** (0.000)	-0.000 (0.821)	-0.000 (0.821)	-10.000 (10.508)
18-29	-2.000** (0.746)	-2.000 (1.638)	-2.000+ (1.103)	-0.000 (1.202)	0.000 (1.087)	0.000+ (0.000)	-0.000 (1.087)	-1.000 (1.093)	-7.000 (12.596)
40-49	-2.000* (0.906)	-0.000 (2.146)	0.000 (1.452)	0.000 (1.486)	0.000 (1.277)	0.000** (0.000)	-0.000 (1.250)	-1.000 (1.253)	4.000 (15.018)
50+	2.000+ (1.172)	0.000 (2.204)	0.000 (1.401)	0.000 (1.564)	-0.000 (1.351)	-0.000** (0.000)	-0.000 (1.331)	-1.000 (1.342)	-11.000 (16.219)
8 Weeks	1.000 (0.738)	-3.000+ (1.641)	-4.000** (1.093)	-4.000** (1.051)	-4.000** (1.058)	-4.000** (0.000)	0.000 (1.064)	-0.000 (1.064)	-13.000 (12.763)
16 Weeks	-1.000 (0.653)	-1.000 (1.725)	0.000 (1.183)	1.000 (1.242)	4.000** (1.012)	4.000** (0.000)	4.000** (0.914)	4.000** (0.913)	-12.000 (12.121)
Constant	4.000** (0.472)	11.000** (1.871)	12.000** (1.210)	12.000** (1.208)	12.000** (1.106)	12.000** (0.000)	12.000** (1.147)	13.000** (1.144)	49.000** (12.623)
N	313	313	313	313	313	313	313	313	313

Appendix 4.

Contract Duration (users with previous exercise contracts included)

Quantile Regression

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

	(1) .05	(2) .20	(3) .25	(4) .40	(5) .50	(6) .60	(7) .75	(8) .80	(9) .95
Female	8.000+	5.000+	6.000	4.000	3.500	8.000*	0.000	0.000	-40.000
	(4.431)	(2.849)	(3.925)	(3.603)	(3.373)	(3.523)	(3.052)	(3.744)	(66.100)
18-29	-4.000	-7.000*	-6.000	-8.000+	-10.000*	-8.000+	-9.000*	-5.000	-20.000
	(5.138)	(3.421)	(4.714)	(4.239)	(4.015)	(4.172)	(3.640)	(4.533)	(77.982)
40-49	-3.000	-5.000	-6.000	-4.000	-3.500	-8.000	-12.000**	-15.000*	52.000
	(6.643)	(4.330)	(5.863)	(5.390)	(5.011)	(5.265)	(4.431)	(5.837)	(95.712)
50+	-5.000	-0.000	-0.000	-4.000	-6.500	-8.000	-5.000	-5.000	52.000
	(6.978)	(4.732)	(6.397)	(5.950)	(5.892)	(5.992)	(5.054)	(6.113)	(103.298)
8 Weeks	2.000	0.000	-0.000	-0.000	-2.500	-8.000+	5.000	2.000	-28.000
	(6.292)	(3.513)	(4.692)	(4.308)	(4.112)	(4.175)	(3.662)	(4.741)	(85.752)
16 Weeks	-4.000	-1.000	2.000	8.000+	9.000*	8.000+	20.000**	22.000**	-4.000
	(5.403)	(3.412)	(4.748)	(4.255)	(4.062)	(4.209)	(3.698)	(4.449)	(75.787)
Constant	8.000	24.000**	24.000**	36.000**	42.500**	48.000**	60.000**	63.000**	160.000*
	(5.808)	(3.505)	(4.763)	(4.518)	(4.335)	(4.461)	(3.627)	(4.586)	(66.054)
N	348	348	348	348	348	348	348	348	348

Appendix 5.

Contract Acceptance
Logit
Standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

	(1)	(2)	(3)	(4)
	Everyone	Everyone (odds ratio)	First Time Choosers	First Time Choosers (odds ratio)
Female	0.424*	1.528*	0.443*	1.558*
	(0.178)	(0.271)	(0.182)	(0.284)
18-29	0.039	1.040	0.068	1.070
	(0.212)	(0.221)	(0.218)	(0.233)
40-49	0.167	1.181	0.142	1.153
	(0.275)	(0.325)	(0.281)	(0.324)
50+	0.063	1.065	0.163	1.177
	(0.319)	(0.339)	(0.332)	(0.391)
8 Weeks	-0.164	0.848	-0.251	0.778
	(0.215)	(0.182)	(0.221)	(0.172)
16 Weeks	0.120	1.128	0.060	1.061
	(0.219)	(0.247)	(0.225)	(0.238)
Message	0.036	1.037	-0.010	0.990
	(0.177)	(0.184)	(0.182)	(0.180)
Prev. Exercise Ct	1.305**	3.687**		
	(0.443)	(1.634)		
Months Member	-0.030	0.971	-0.028	0.973
	(0.020)	(0.019)	(0.023)	(0.022)
Constant	0.183	1.201	0.223	1.250
	(0.248)	(0.298)	(0.256)	(0.320)

N	566	566	523	523
---	-----	-----	-----	-----