When Commitment Fails – Evidence from a Field Experiment*

ANETT JOHN†

May 2018

Commitment products can remedy self-control problems. However, imperfect knowledge about their preferences may (discontinuously) lead individuals to select into incentive-incompatible commitments, which reduce their welfare. I conduct a field experiment where low-income individuals were randomly offered a new installment-savings commitment account. Individuals chose a personalized savings plan and a default penalty themselves. A majority appears to choose a harmful contract: While the average effect on bank savings is large, 55 percent of clients default, and incur monetary losses. A possible explanation is that the chosen penalties were too low (the commitment was too weak) to overcome clients' self-control problems. Measures of sophisticated hyperbolic discounting correlate negatively with take-up and default, and positively with penalty choices – consistent with theoretical predictions that partial sophisticates adopt weak commitments and then default, while full sophisticates are more cautious about committing, but better able to choose incentive-compatible contracts.

Keywords: commitment, hyperbolic discounting, partial sophistication
JEL classification: C93, D03, D14, O12

*I would like to extend my gratitude to Oriana Bandiera, Maitreesh Ghatak, and Gharad Bryan, for their invaluable support and advice. I am deeply indebted to Dean Karlan, Ann Mayuga, Faith McCollister, Megan McGuire, Yoeri Suykerbuyk and Eva Ghirmai of IPA Philippines, without whom this project would not have been possible. I am grateful to Nageeb Ali, Douglas Bernheim, Andrea Canidio, Kristina Czura, Stefano DellaVigna, Jonathan de Quidt, Paul Heidhues, Alex Imas, Supreet Kaur, David Laibson, Matthew Levy, Lance Lochner, George Loewenstein, Johannes Spinnewijn, Betty Wilkinson, Erina Ytsma, and various seminar audiences for helpful comments and discussions. I also thank 1st Valley Bank of Cagayan de Oro, Philippines, for a productive collaboration. I gratefully acknowledge the financial support of the Yale Savings and Payments Research Fund at Innovations for Poverty Action, sponsored by a grant from the Bill & Melinda Gates Foundation. This research further received generous support from the Royal Economic Society. All errors and omissions are my own.

†CREST, email: anett.john@ensae.fr. This paper was previously circulated under the name Anett Hofmann.

1. Introduction

Commitment is ubiquitous. Contrary to predictions of the standard neoclassical model, the last decade has seen a surge of evidence documenting a demand for (self-)commitment contracts - roughly understood as a voluntary restriction of one's future choice set, in order to overcome intrapersonal conflicts.[1] Applications are as broad as the scope of human ambition, and range from gym memberships, diet clubs and pension savings to self-imposed binding deadlines for academic papers.[2] More informal arrangements include taking only a fixed amount of cash (and no credit cards) when going shopping, or not keeping chocolate in the house. In developing countries, documented demand for commitment devices goes back to the literature on rotating savings and credit organisations (ROSCAs), the wandering deposit collectors of South Asia and Africa, and more recent studies on newly introduced commitment savings products.[3]

Why do people self-commit? Commitment entails the voluntary imposition of constraints on future choices, thereby putting a cost on flexibility, which is weakly welfare-reducing from a neoclassical perspective. Among the most frequently cited models to rationalize the observed demand for commitment are those of quasi-hyperbolic discounting (Laibson (1997) and O'Donoghue and Rabin (1999)).[4] They suggest that agents are more impatient over current trade-offs (now vs. tomorrow) than over future trade-offs (one year vs. one year plus one day). As a result, agents procrastinate activities that involve immediate costs and later rewards (saving for a new TV, going to the gym), and do too much of activities that involve immediate gratification but later costs (using high-interest credit cards, buying temptation goods). If individuals with such preferences realize their own time-inconsistency, they will have a positive willingness to pay for commitment devices which eliminate tempting options from their future choice sets (or make them more expensive), thus allowing them to follow through with their plans (to save, to eat healthily, to exercise). In theory, this will increase their welfare from a long-run perspective. Empirically, commitment devices have been shown to increase savings levels (Ashraf, Karlan and Yin (2006)), agricultural input use (Brune et al. (2016)), pension contributions (Benartzi and Thaler (2004)), microenterprise investment (Dupas and Robinson (2013)), and chances of successful smoking cessation (Giné, Karlan and Zinman (2010)), as well as to reduce heavy daytime drinking (Schilbach (2015)).

But are people good at choosing the 'right' commitment contract? By construction, correctly choosing a welfare-improving contract requires knowledge about one's future preferences: To determine whether a contract will enable her to follow through with a plan, the agent needs to anticipate how her future selves will behave under the contract. Telling her social network that she plans on running a marathon, for instance, requires the agent to assess whether the threat of embarrassment upon withdrawal is sufficient to overcome her laziness. If the agent is overconfident, the contract may result in undesirable behaviour (she does not run the marathon), and she may

[1] This paper focuses purely on self-commitment. It does not address commitment contracts adopted with strategic motives with respect to others, or commitments entered into for convenience or other immediate benefits. As an example, the purchase of Christmas gifts in October qualifies as self-commitment if the agent fears not having enough money left in December, but not if the agent's motivation is purely to avoid the Christmas rush.

[2] See DellaVigna and Malmendier (2006) for gym memberships, Benartzi and Thaler (2004) for 401(k) pension savings, and Ariely and Wertenbroch (2002) for academic assignments. For an overview of commitment devices, see Bryan, Karlan and Nelson (2010).

[3] On ROSCAs, see Anderson and Baland (2002), or Gugerty (2007). See e.g. Besley (1995) on West Africa's susu collectors. Ashraf, Karlan and Yin (2006), Brune et al. (2016) and Dupas and Robinson (2013) study withdrawal-restriction savings accounts. Duflo, Kremer and Robinson (2011) study commitments to use fertilizer.

[4] Models of temptation and self-control (Gul and Pesendorfer (2001)) and dual-self models (Fudenberg and Levine (2006)) also predict a demand for commitment, with similar implications for observed behaviour.

be harmed rather than helped (she gains no health benefits, but pays the cost of embarrassment). The very nature of most commitment contracts is to impose monetary or social penalties for undesirable behaviour.[5] Thus, adopting a commitment device that is ill-suited to one's preferences may backfire and become a threat to welfare.

This paper argues that commitment can be harmful if agents select into the wrong commitment contract - and presents evidence that they frequently do. In a framework where a penalty is conditional on a binary savings decision, I first outline theoretically why offering commitment improves the welfare of fully sophisticated agents, but is likely to harm partially sophisticated agents. Given a continuous penalty, this effect is discontinuous at full sophistication, as any amount of naiveté leads agents to select into contracts which violate their incentive constraints. Furthermore, commitment adoption *increases* in naiveté,[6] as naive agents underestimate the cost of effective commitment. Second, I conduct a randomized experiment in the Philippines where individuals could sign up for a new commitment savings account with fixed regular installments. Adopters choose the stakes of the contract (in form of a default penalty) themselves. I find that the *average* effect on bank savings is large and significant: The Intent-to-Treat (ITT) effect on bank savings is roughly three times that of a conventional withdrawal-restriction product that was offered as a control treatment. However, the *median* client appears to choose a 'harmful' contract: 55 percent of clients default on their savings contract, and incur the associated penalty. The magnitude and timing of defaults is difficult to reconcile with rational expectations and idiosyncratic shocks (a 'bad luck' scenario). Instead, it is suggestive of individuals making mistakes in contract choice. A possible explanation that is supported by the data is that the chosen stakes were too low (the commitment was too weak) to overcome clients' self-control problems. In addition, both take-up and default are *negatively* predicted by measures of sophisticated hyperbolic discounting. This is consistent with the notion that those who are fully aware of their bias realize the commitment is too weak for them, and stay away. The results from a repeat marketing stage with the offer of 'pre-ordering' the product for a second round support the hypothesis that a significant share of clients took up the commitment contract by mistake. Alternative explanations for default that find some empirical support are income optimism and household conflict. A pure stochastic shock explanation appears unlikely.

I partnered with 1st Valley Bank, a rural bank based in Mindanao, Philippines. The sample population of 913 individuals was obtained by conducting a door-to-door baseline survey in low-income areas in proximity to two selected bank branches. The baseline survey elicited time preferences (described in Section 3.3), risk aversion, financial claims from others, cognitive ability, financial literacy, intrahousehold bargaining power, household demographics, and measures of saving, borrowing, and household expenditures. After the baseline survey, all individuals were provided with a marketing treatment, which included a personalized savings plan for an upcoming expenditure and a free non-commitment savings account with 100 pesos (U.S. $2.50) opening balance.[7] Personal savings plans featured a self-chosen goal date, goal amount, and a weekly or bi-weekly installment plan (see Figure A1). The idea was to encourage individuals to save for their lump-sum expenses (such as school fees, business capital, or house repairs), rather than following the common practice of borrowing at high informal moneylender rates. At the end of the marketing visit, a randomly chosen 50 percent (the 'Installment Savings' group) were

[5] Examples of monetary penalties include any type of commitment contract with front-loaded fees, such as retirement savings products with acquisition or management costs. Such fees generate a 'J curve' in the asset value, resulting in high negative returns if the contract is cancelled during the initial years. A similar argument can be made for front-loaded gym membership costs.

[6] Conditional on an awareness that one cannot save without commitment. The latter rules out perfect naiveté.

[7] At the time of marketing (October 2012), the exchange rate was 42 Philippine pesos per U.S. dollar.

offered a new installment-savings commitment account ('IS account'). This account committed clients to make fixed weekly or bi-weekly deposits and pay a penalty upon default, which effectively made all features of the personal savings plan binding. The default penalty was chosen by the client upon contract signing, and framed as a charity donation.[8] The interest rate was equal to the standard market interest rate.[9]

As a benchmark treatment, 25 percent of the sample (the 'Withdrawal Restriction' group) were offered the commitment savings account studied in Ashraf, Karlan and Yin (2006), Giné, Karlan and Zinman (2010), and Brune et al. (2016). This withdrawal-restriction account ('WR account') allowed individuals to restrict withdrawals before either the goal date or the goal amount from their savings plan had been reached. The account did not include any obligation to make further deposits after the opening balance. The remaining 25 percent of the sample received no further intervention after the marketing treatment, and constitute the control group. For the control group, none of the savings plan features were binding. Since individuals' expenditures were due at different times, the outcome of interest are individuals' savings at the time of their goal date. The study concluded with a comprehensive endline survey, as well as a repeat marketing stage where IS clients could opt to 'pre-order' the product for a second round.

Demand for commitment is high, even in a general low-income population with little previous bank exposure: Take-up rates were 27 percent for the installment-savings account and 42 percent for the withdrawal-restriction account, despite the prior universal provision of free standard savings accounts. Offering an installment commitment was more effective at increasing savings: By the time individuals reached their goal date (an average of 130 days later), bank savings in the Installment Savings group had increased by 429 pesos (U.S. $10.20, ITT) relative to the control group. The corresponding effect for the Withdrawal Restriction group was 148 pesos (U.S. $3.50).[10] The control group saved an average of 27 pesos. The scale of effects suggests that a commitment product with fixed regular installments is highly effective at increasing savings on *average*. However, these averages hide a lot of heterogeneity in the case of both products: 55 percent of IS clients defaulted on their savings contract, incurring penalties between 150 and 300 pesos - the equivalent of a day's wage. More than half of defaults occur immediately after account opening, which makes it unlikely that defaulters benefitted from the savings contract via higher savings. Similarly, 79 percent of Withdrawal Restriction clients made no further deposits after the opening balance. While default is not formally defined for the WR product, abandoning the contract is costly: For those who had chosen a binding goal amount (45 percent), their initial savings were tied up indefinitely, or until the bank would exhaust their account with dormancy fees.[11]

I show that installment-savings completion is strongly bi-modal, in that most clients either (i) stop depositing immediately after the opening balance or (ii) complete their savings plan in full. I interpret this as evidence against a shock explanation, where individuals rationally default following large random shocks to their income or expenditures. Using measures of time-inconsistency and sophistication from the baseline survey, the data suggest that present-biased preferences by themselves do not predict take-up of a commitment product, but they do predict default. In contrast, sophistication drives both take-up and default: For a given level of time-inconsistency, more sophisticated agents are *less* likely to adopt commitment. Conditional on take-up, sophisticates choose stronger

[8] The concept is similar to the Stickk initiative (www.stickk.com), applied to a development context.

[9] As of September 2012, the interest rate was 1.5 percent p.a. for all offered accounts. It decreased to 1 percent in November 2012. The inflation rate for 2012 was 3.1 percent.

[10] The ITT measures the effect of being *offered* the product, and already accounts for any penalty charges. An increase of 429 pesos (148 pesos) corresponds to 20 percent (7 percent) of median weekly household income, or 18 percent (5 percent) of their median savings goal.

[11] Dormancy fees are very common with Philippine banks, and commonly start after two years of inactivity.

commitments (higher penalties), and are less likely to default. This is consistent with the theory: Sophistication increases both the perceived cost of commitment and the agent's ability to choose an incentive-compatible contract.

This paper builds and expands on the literature in several ways. To the author's knowledge, it is the first study to link heterogeneous effects (and possible welfare risks) of commitment contracts to a proxy measure of (partially) sophisticated time-inconsistency. This makes it closest in spirit to DellaVigna and Malmendier (2006), who show that U.S. consumers choose gym contracts which are cost-inefficient given their attendance frequency. The study further relates to Heidhues and Kőszegi (2009), who show theoretically that commitment is likely to reduce welfare when agents are partially sophisticated. Their model differs from the one presented in this paper in some key assumptions and welfare implications: Heidhues and Kőszegi (2009) assume that commitment is desirable at all levels of sophistication, and derive that welfare is lowest just below full sophistication, as agents commit to high penalties which remain insufficient to deter consumption. In my model, commitment adoption is endogenous, and decreases in the amount of sophistication. The cost of commitment endogenously results from the presence of uncertainty, as unexpected shocks force the agent to default. The negative welfare effect of large but insufficient penalties is compensated by the probability that these large penalties are prohibitive for adopting commitment. Consequently, the group most at risk of negative welfare effects are near-naifs, who widely adopt weak commitments and subsequently default. Other related theoretical contributions are DellaVigna and Malmendier (2004), Heidhues and Kőszegi (2010) and Eliaz and Spiegler (2006), who address supply-side responses to partial sophistication.

In the realm of commitment savings, the literature has largely focused on positive average effects. I suggest that these effects may be very heterogeneous, including the possibility of a majority being hurt by the product. While not a focus of previous work, my results are in line with previous findings: Ashraf, Karlan and Yin (2006) find that a withdrawal-restriction product increased savings by 81 percent on average, but 50 percent of the clients made no further deposits after the opening balance. Out of 62 clients who selected an amount goal, only six reached this goal within a year. Giné, Karlan and Zinman (2010) offered smokers in the Philippines a commitment savings contract for smoking cessation, in which participants forfeit their savings if they fail a nicotine test after 6 months. Offering the contract increased the likelihood of smoking cessation by 3 percentage points. However, 66 percent of smokers who took up the product failed the nicotine test, forfeiting an average of 277 pesos in savings. In summary, a closer look at the heterogeneity behind average treatment effects reveals that adverse effects of commitment products may be widespread.

Second, the paper provides the first analysis of a commitment savings product with fixed regular installments in a randomized setting. The product design mimics the fixed installment structure found in loan repayment contracts. It is motivated by empirical evidence suggesting that microloans and informal loans are often taken out for consumption purposes, or for recurring business expenditures - rather than as a one-off investment (Ananth, Karlan and Mullainathan (2007), Rutherford (2000)). With loans that are not directly required for income generation, the question arises why individuals are willing to pay substantial loan interest charges rather than choosing to save. Especially for those who borrow in frequent cycles, the long-term difference between expensive loan cycles and equivalent savings cycles reduces to (i) one initial loan disbursement and (ii) a binding fixed-installment structure that is rarely available in savings products.[12] The idea that time-inconsistent agents benefit from commitment to

[12]ROSCAs are inflexible to an individual's needs, and not widely available in the study region. Deposit collectors usually do not commit individuals to fixed installments, and social sanctions can be avoided using small deposits or convenient excuses.

regular fixed installments has been discussed by Fischer and Ghatak (2010), Bauer, Chytilova and Morduch (2012), and John (2017).[13] If a significant share of the demand for loans is a demand for commitment to fixed installments, then we should see that the introduction of a fixed-installment microsavings product will result in (i) substantial increases in saving and (ii) a reduction in the demand for loans. I find strong support for an increase in savings. While less precisely estimated, offering installment savings also facilitates the largest reductions in loan demand over the study period (at the 10th, 20th and 30th percentile). Furthermore, the paper provides the first direct comparison of an installment-savings commitment with a pure withdrawal-restriction commitment. The WR account used in this study is identical to Ashraf, Karlan and Yin (2006)'s SEED, and their effect on savings replicates within a 4 percent margin.[14] The estimated effect of the installment-savings product on bank savings is almost three times higher, consistent with the theoretical work of Amador, Werning and Angeletos (2006): When individuals value both commitment and flexibility, the optimal contract involves a minimum (per-period) savings requirement.

Third, the paper proposes a proxy measure of sophistication that is practicable in a field setting in a developing country. It consists of an interaction of two established measures – one of observed time-inconsistency, and one of self-perceived temptation. Previous literature has often assumed a one-to-one mapping from the take-up of a commitment product to the presence of fully sophisticated time-inconsistency. Such a mapping does not allow for the possibility that individuals may take up commitment products by mistake. Recent advances with prediction tasks (Augenblick and Rabin (2015), Acland and Levy (2015)) are promising, but challenging to implement in a field context.

The paper proceeds as follows. Section 2 outlines a model of commitment under partial sophistication. Section 3 describes the experimental design. Section 4 outlines the empirical strategy. Section 5 presents empirical results. Section 6 discusses alternative explanations. Section 7 concludes and discusses policy implications.

2. Theory: Commitment under Partial Sophistication

The following section presents a simple model of commitment under partial sophistication. It sheds light on (i) why sophisticated hyperbolic discounters can benefit from commitment, (ii) why commitment reduces welfare if it is too weak to be effective, (iii) why partially sophisticated hyperbolic discounters are likely to select into such weak commitment contracts, and (iv) why those with high perceived degrees of time-inconsistency may avoid commitment.

For expositional ease, I use a two-period linear-utility model of consumption and savings. A linear model is sufficient to highlight the basic mechanism, and provides a reasonable approximation for small stakes. All main results hold in a model with concave utility. Concave utility and three or more periods are needed to illustrate the consumption smoothing benefits of an installment-savings product, at the expense of tractability. John (2017) studies these benefits for the case of full sophistication.

2.1. The Model

Consider an agent with a normalized per-period income of $y_t = 1$. The agent chooses whether to save for a nondivisible good which costs the lump-sum $1 < p \leq 2$ and yields a benefit $b > p$. She cannot borrow. The agent lives for 2 periods. In addition, there is a planning period 0, when she may choose to adopt commitment (Section

[13] See also Afzal et al. (forthcoming), who give participants a direct savings vs. loan choice in a lab-in-the-field setting.

[14] Ashraf, Karlan and Yin (2006)'s SEED increases savings by 411 pesos (ITT) after 12 months, while the WR account in this study increases savings by 148 pesos after 4.5 months. Monthly equivalents are 34.25 and 32.88 pesos, respectively.

2.3). Throughout, assume the interest rate is $R=1$ and $\delta=1$ for simplicity. Define s_1 as the amount of savings that she sends from period 1 to period 2, so that $c_1=y_1-s_1\geq 0$. Lifetime utility as evaluated in each period $t\in\{0,1,2\}$ is given by the discounted stream of future consumption:

$$U_t=c_t+\beta\sum_{k=t+1}^{2}E(c_k).$$

For $\beta<1$, the agent is *present-biased*: She exhibits a lower rate of discount over current trade-offs (t vs. $t+1$) than over future trade-offs ($t+s$ vs. $t+s+1$, $s>0$). Up to this point, the savings model resembles the autarky savings framework in Basu (2014) (where autarky refers to the absence of banking). I now generalize the model to allow for partial sophistication and stochastic income, thus creating a need for flexibility. Following O'Donoghue and Rabin (1999), the agent's degree of sophistication about her present bias is captured in the parameter $\tilde{\beta}\in[\beta,1]$, which she believes she will use in all future periods. In particular, the agent believes in period t that her utility function in period $t+s$ will be

$$U_{t+s}=c_{t+s}+\tilde{\beta}\sum_{k=t+s+1}^{2}E(c_k).$$

For a fully sophisticated agent, $\tilde{\beta}=\beta$. A fully naive agent believes she will behave time-consistently in the future, captured in $\tilde{\beta}=1$. A need for flexibility is introduced through stochastic income shocks: With a per-period probability of λ, the agent loses her income in that period, such that $y_t=0$. This shock has a variety of interpretations: It can be interpreted directly as a loss of income, e.g., from redundancy, bad business, or illness of an income-earning household member. With a minor modification, it can be interpreted as a reduced-form taste shock.[15] The implication of a shock is that the agent's lifetime income is reduced to 1 (or zero in case of multiple shocks), which means the nondivisible good can no longer be purchased. When a shock hits, any plans to save are abandoned, and any existing savings are consumed. This results in a third interpretation: More generally, the shock λ corresponds to the probability that, for any time-consistent reason, the agent no longer finds it optimal to save for the good.[16]

Following O'Donoghue and Rabin (1999), an agent's welfare is understood to be the lifetime utility of the period 0 agent from an ex-ante perspective: $W=E[c_1+c_2]$. The advantage of this convention is that no particular period is favoured.

2.2. No-Commitment Equilibrium

The model is easily solved using backward induction. In period 2, the agent will buy the nondivisible whenever she can afford it, i.e., whenever $y_2+s_1\geq p$. Given $p>1$, this requires $y_2=1$ (there is no shock) and $s_1\geq p-1$. Additional savings $s_1>p-1$ are simply consumed, as are insufficient savings $s_1<p-1$. The consumption profile is $c_2=y_2+s_1+(b-p)\cdot\mathbb{1}(y_2+s_1\geq p)$.

In period 1, the agent realizes the good will be bought if and only if she sends $s_1\geq p-1$, and absent shocks. She responds by either sending $s_1=p-1$, or sending zero: In a linear model with $\beta<1$, it is never optimal to shift excess consumption to the future.

[15] Suppose the sudden illness of a family member changes preferences such that utility stays unchanged if a hospital visit (at cost 1) is consumed and paid for, and drops to $u(c)=-\infty$ without a hospital visit.

[16] Time-consistent explanations to abandon savings plans include state-dependent preferences.

Proposition 1. *In the No-Commitment Equilibrium, absent shock realizations, the nondivisible good is bought by sufficiently time-consistent agents, i.e., those with a time-consistency parameter β above a threshold $\hat{\beta}$. The threshold $\hat{\beta}$ increases in the shock frequency λ and the price p, and decreases in the benefit b.*[17](All proofs are in Appendix II.)

The threshold $\hat{\beta} \equiv \frac{p-1}{\lambda(p-1)+(1-\lambda)(b-1)}$ has an intuitive interpretation: It is the ratio of the cost of saving today, $p-1$, to its expected benefit tomorrow, $\lambda(p-1)+(1-\lambda)(b-1)$. Because $b > p$, a time-consistent agent will always save (i.e., $\hat{\beta} < 1$). Note that the no-commitment equilibrium is invariant to the degree of sophistication (see footnote 20).

2.3. Introducing Commitment

For agents with high levels of time-inconsistency, the no-commitment equilibrium is inefficient: Saving for the nondivisible good is welfare-improving regardless of β, since the period 0 planner values consumption in periods 1 and 2 equally, and the benefit of the good exceeds its cost. However, agents with $\beta < \hat{\beta}$ cannot save on their own, creating potential gains from commitment. Suppose the individual is now given the possibility to commit to save, enforced via a penalty for non-compliance which the individual chooses herself ex-ante. While not the only form of commitment (see the discussion in Section 7), self-imposed conditional penalties have many applications: People tell their friends (or worse, their enemies) about a plan to lose weight, and then suffer an embarrassment cost when they are seen eating fast food. They join ROSCAs, knowing there will be social sanctions when they fail to contribute. Monetary penalties appear in most commitment contracts with front-loaded fees (see footnote 5).

In the current setting, individuals commit to a self-chosen default penalty, imposed if they fail to follow a regular installment savings plan. Applied to a simple two-period model, the period 0 agent can choose a penalty D that is imposed on her in period 1 if she fails to meet the savings requirement, $s_1 \geq p-1$. Given D, and absent shock realizations, the period 1 agent now prefers to save iff

$$1-(p-1)+\beta[\lambda(p-1)+(1-\lambda)b] \geq 1-D+\beta(1-\lambda). \tag{1}$$

Equivalently, she is willing to save if the penalty is higher than a minimum effective threshold, denoted D_{min}, which bridges the gap between the current costs and the future benefits of saving. I.e., the agent saves for D such that

$$D \geq D_{min}(\beta) \equiv \underbrace{(p-1)}_{cost\ today} - \beta\underbrace{[\lambda(p-1)+(1-\lambda)(b-1)]}_{benefit\ tomorrow}. \tag{2}$$

Proposition 2. *The minimum penalty that is effective in enforcing the savings plan, denoted D_{min}, strictly decreases in the time-consistency parameter β. Further, D_{min} strictly increases in the shock frequency rate λ.*

Note that $D_{min} \leq p-1 \leq 1$ for all β. A note on enforceability: The model assumes that the penalty D is fully enforceable, even if the agent loses her income to a shock, and is consequently unable to save. In a two-period model, full enforceability is required to prevent the agent from adopting infinitely large penalties, reasoning that she will never incur them on the equilibrium path. Empirical arguments support that commitment penalties can often be enforced irrespective of current financial resources: First, D may be a non-monetary cost, such as embarrassment

[17] Apart from the added shocks λ, the no-commitment equilibrium result is shared with Basu (2014). The models diverge from here.

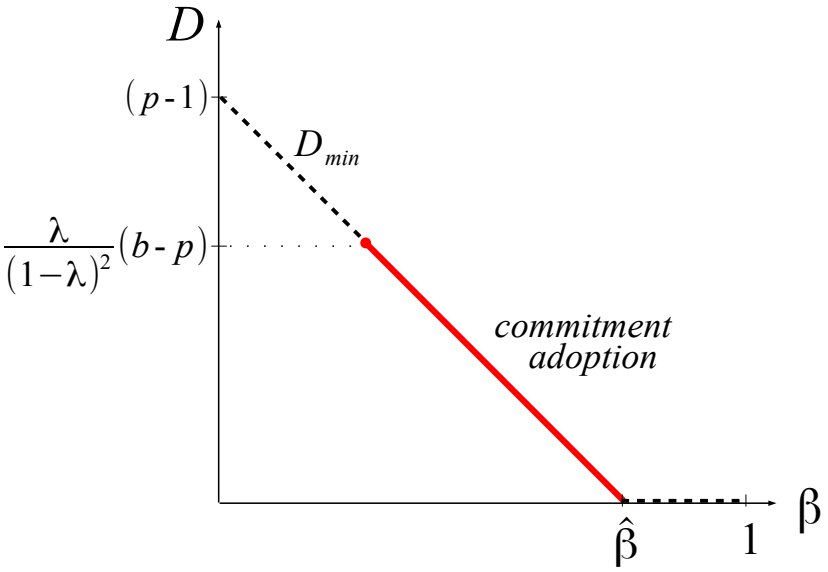

FIGURE 1. COMMITMENT TAKE-UP DECISION (FULL SOPHISTICATION)

Figure 1 illustrates the relationship between time-inconsistency β and commitment adoption D assuming full sophistication, $\tilde{\beta} = \beta$. The dotted line is the minimum default penalty that is effective in enforcing the savings plan, D_{min}. The solid line indicates the range of β for which D_{min} is adopted. Generalizing to partial sophistication ($\beta \leq \tilde{\beta} \leq 1$), β and D_{min} are replaced by $\tilde{\beta}$ and $\tilde{D}_{min}$: Perceived values rather than actual values determine the choice of commitment. Thus, for those unable to save in autarky ($\beta \leq \tilde{\beta} < \hat{\beta}$), commitment is attractive for high $\tilde{\beta}$, and prohibitively expensive for low $\tilde{\beta}$.

towards friends. Second, D may represent a discounted stream of future losses, such as those from losing access to microfinance. Third, in settings with multiple periods, penalties can be enforced using the stock of past savings. What matters for the theoretical results is that the agent incurs the penalty with positive probability on the equilibrium path, due to future uncertainty which cannot be contracted upon.[18] This creates a cost of commitment, and a demand for flexibility.[19]

2.4. Commitment Choice in Equilibrium

The commitment adoption decision has two steps: The period 0 planner first identifies which commitment contract (which penalty) will enable her to save while maximising flexibility. In a second step, she decides whether to adopt this contract. Commitment is costly: If a shock hits in period 1, the agent loses her ability to save. Default becomes unavoidable, and the agent incurs a loss of D. This risk is referred to as 'rational default', as it occurs independently of time-inconsistency. The equilibrium behaviour with full sophistication is summarized in Proposition 3. Figure 1 illustrates the relationship between time-inconsistency and commitment adoption.

[18] Inability to contract on the arrival of shocks may be caused by unobservability or moral hazard.

[19] The commitment versus flexibility trade-off discussed here differs from that in Amador, Werning and Angeletos (2006). In the latter, agents can enforce the desired savings behaviour perfectly, i.e., the implied penalty is infinite, but never incurred. The demand for flexibility comes from taste shocks, rather than from the risk of being financially unable to save.

Proposition 3. *Equilibrium with Full Sophistication: (a) Conditional on adopting commitment, individuals will adopt the minimum effective penalty, D_{min}. (b) Individuals who are sufficiently time-consistent to save in autarky (those with $\beta \geq \hat{\beta}$, see Proposition 1) never adopt commitment. (c) Individuals who cannot save in autarky (those with $\beta < \hat{\beta}$) adopt commitment if i) β is sufficiently high, and ii) the shock frequency rate λ is sufficiently low. The adoption decision is summarized in the condition $\lambda D_{min} \leq (1-\lambda)^2(b-p)$, where λD_{min} represents the expected cost of commitment due to rational default, and $(1-\lambda)^2(b-p)$ captures the expected benefit of a successful savings plan. (d) With full sophistication, offering commitment weakly increases welfare (it strictly increases the expected welfare of adopters).*

A key intuition is that period 1's incentive constraint only depends on whether $D \geq D_{min}(\beta)$, thus choosing the minimum penalty always dominates choosing larger penalties. Choosing $D < D_{min}$ is strictly dominated by choosing no penalty at all, since period 1's incentive constraint is violated, and default occurs with certainty. As a result, the period 0 agent chooses either $D = D_{min}$ or $D = 0$. Further note that the ex-ante benefit of commitment, $(1-\lambda)^2(b-p)$, does not depend on the time-consistency parameter β. In contrast, β determines the cost of commitment, λD_{min}. Perhaps counter-intuitively, commitment is most attractive to those with the *lowest* degree of time-inconsistency (conditional on $\beta < \hat{\beta}$), as the penalty required to enforce the savings plan is small, and poses little risk in the presence of shocks. In consequence, agents adopt commitment for sufficiently high β.

2.5. The Effect of Partial Sophistication

A partially sophisticated agent believes that her future selves will discount the future at rate $\tilde{\beta} > \beta$ rather than at β. This corresponds to the classic pattern of procrastination, where an agent believes she will be patient enough to complete an unpleasant task (dieting, exercising, doing housework) tomorrow, but not today. The current setting examines the simplest form of partial sophistication, where beliefs are deterministic and incorrect. The result is that the period 0 planner underestimates the size of the penalty that will be required to induce her period 1 self to save, resulting in default and welfare losses. The results are robust to more stochastic types of partial sophistication (see Section 2.6).

Proposition 4. *Equilibrium with Partial Sophistication: (a) Conditional on adopting commitment, partially sophisticated individuals will adopt penalties strictly below the required effective minimum, $\tilde{D}_{min} < D_{min}$. As a result, adopters' incentive constraints in period 1 are systematically violated, triggering contract default. (b) Individuals who believe themselves to be sufficiently time-consistent to save in autarky (those with $\tilde{\beta} \geq \hat{\beta}$) never adopt commitment. (c) For those who realize they cannot save in autarky ($\tilde{\beta} < \hat{\beta}$), sophistication negatively predicts commitment adoption. For a given β, commitment is most attractive to those with the largest amount of naiveté, $\tilde{\beta} - \beta$. (d) With partial sophistication, offering commitment weakly decreases welfare. It strictly decreases the expected welfare of adopters.*

All arguments are analogous to the case of full sophistication, except that the period 0 agent believes the period 1 agent will apply $\tilde{\beta} > \beta$ in making intertemporal choices.[20] Holding true β fixed,[21] a higher degree of naiveté

[20] In a two-period model, the degree of sophistication only affects period 0 decisions: Period 2 makes no decisions about the future, and thus period 1's belief $\tilde{\beta}$ is irrelevant. In a multi-period model, additional coordination problems arise, as the agent may consume savings made by past selves rather than to accumulate them over time.

[21] Is is plausible to hold β fixed while $\tilde{\beta}$ increases? While this study is agnostic about the functional relationship between β and $\tilde{\beta}$, measures of both are uncorrelated in the current dataset (see Appendix Figure A7).

$\tilde{\beta} - \beta$ implies that a lower penalty is regarded as effective, which decreases the perceived cost of commitment. Conditional on β, as well as on the agent's perceived inability to save without commitment ($\tilde{\beta} < \hat{\beta}$), adoption monotonically increases with naiveté.

Given its relevance for the experiment, it is worth considering the impact of a lower bound on the penalty, $\underline{D}$. The expected cost of commitment becomes $\lambda max\{\underline{D}, \tilde{D}_{min}\}$, weighted against the expected benefit $(1-\lambda)^2(b-p)$ (analogue to Proposition 3).

Corollary. *Commitment adoption decreases with the introduction of a lower bound on the penalty, $\underline{D}$. Defaults also decrease in $\underline{D}$. Welfare effects of $\underline{D}$ are negative for full sophisticates. For partial sophisticates, introducing $\underline{D}$ increases welfare if it either deters them from adopting commitment, or imposes an incentive-compatible penalty ($\tilde{D}_{min} < D_{min} \leq \underline{D}$). For insufficient bounds $\tilde{D}_{min} < \underline{D} < D_{min}$, it reduces welfare by increasing the cost of default.*

2.6. Discussion and Extensions

The model studies a simple form of partial sophistication. The assumption of a deterministic and incorrect $\tilde{\beta}$ was introduced by O'Donoghue and Rabin (1999), and suffices to highlight the main mechanism of undercommitment. In reality, individuals may have more complex belief distributions about their future preferences, or preferences themselves may be stochastic. Appendix I.1 discusses the case of stochastic sophistication, and argues that the findings observed in the data are hard to reconcile with rational expectations about β: Comparing stochastic with full sophistication, commitment becomes less attractive due to the downside risk that a given penalty will not be effective. This leads to low commitment take-up, high conditional penalties, and low default rates. To reconcile stochastic beliefs about time-inconsistency with the observed low penalties and high default rates, one needs to allow for belief distributions which are systematically biased towards naiveté.

Is it plausible that individuals persistently hold incorrect beliefs about their time preferences, despite being able to observe their own past behaviour? Appendix I.2 discusses why Bayesian learning may fail (see also Ali (2011)). For instance, learning may be specific to context: An individual may realize from past observation whether she is able to save for the nondivisible good by herself (i.e., the inequality $\beta \geq \hat{\beta}$ is observed). However, she may be unfamiliar with her savings behaviour under commitment. Other impediments to Bayesian learning include self-serving beliefs and neuroscientific explanations (such as stress-induced time-inconsistency). Appendix I.3 discusses pessimistic beliefs ($\tilde{\beta} < \beta$) and overcommitment.

3. Experimental Design

3.1. Study Setting and Sample

I designed and implemented the installment-savings commitment product in cooperation with 1st Valley Bank, based in Mindanao, Philippines. 1st Valley Bank is a small rural bank that offers microcredit, agricultural insurance, salary loans, and pension products. The bank agreed to offer both the installment-savings product and the withdrawal-restriction product in two of their branches: Gingoog and Mambajao. Gingoog is a city of 112,000 people in northern Mindanao, and Mambajao is a municipality of 36,000 people on Camiguin Island. For these two branches, the IS and the WR account constituted new product additions.

The sample was obtained through door-to-door visits in all low and middle income areas in proximity to the bank branches. In each household, the survey team identified the person in charge of managing the household budget (usually the mother of the family). The baseline survey was completed with all such individuals who (i) had some form of identification, (ii) claimed to have a large upcoming expenditure (such as school fees, house repairs, or business expansions) and (iii) agreed to receive a visit from a financial advisor (to talk about how to manage household expenses). These screening rules were intentionally minimal: 81 percent of initial respondents were included in the sample.[22] After the baseline survey, individuals were randomly assigned to three groups: 50 percent of individuals were assigned to an 'Installment Savings' (IS) group, and 25 percent each were assigned to a 'Withdrawal Restriction' (WR) and a control (C) group. Participants were not aware that they are taking part in an experiment.[23]

Approximately one week after the baseline survey, individuals received a visit from a bank marketer. Marketers engaged individuals in a conversation about how to manage large lump-sum expenses, and talked about the benefits of saving. Focusing on one particular expenditure, individuals were encouraged to make a formal 'Personal Savings Plan', which contained a purpose, a goal amount, a goal date, and a fixed equal installment plan with due dates (see Appendix Figure A1). The median savings goal was 2400 pesos (close to the median household's weekly income of 2125 pesos), with a median weekly installment of 150 pesos. Common savings goals were school tuition fees, house repairs, and Christmas gifts (see Appendix Table A1). The duration of savings plans was limited to 3–6 months (median: 137 days). In addition, everyone was offered a non-commitment savings account (henceforth called 'ordinary savings account'), containing a free 100 pesos opening balance as a 'welcome gift', along with an encouragement to use this account to save for the expenditure.[24] Table 2 provides an overview of savings plan and account choices.

At the end of the visit, individuals in group IS were asked whether they wanted to commit to the fixed-installment structure outlined in their Personal Savings Plan by taking up the IS product, and the product features were explained. In contrast, individuals in group WR were offered to restrict withdrawals of their savings until they reached either the goal amount or the goal date specified in their Personal Savings Plan, implemented through the WR product. Clients were permitted to revise their savings goals upon accepting a commitment product. Up to the point of offering the commitment products, the marketing script was identical across groups IS, WR, and C. In group IS, 114 clients (out of 423 offered) accepted the IS product.[25] In group WR, 92 (out of 219 offered) accepted the WR product.

Individuals were left to themselves during the savings period, without help from deposit collectors or reminders. After all goal dates had been reached, a comprehensive endline survey asked about savings, outstanding loans, expenditures, changes in income, and various types of shocks experienced. In addition, existing IS clients were offered to 'pre-order' IS for a second round, and told that the bank would continue the product conditional on sufficient demand. While the Pre-Order did not involve a financial commitment, it required the completion of substantial paperwork and a new savings plan (to deter cheap talk).

[22] 5 percent of households were excluded because the household decisionmaker was not available, 6 percent due to lack of ID, 1 percent had no upcoming expenditure, and 6 percent did not agree to a visit by a financial advisor.

[23] The experiment thus constitutes a natural field experiment in the terminology of Harrison and List (2004). This increases generalizability and precludes a "Treatment Specific Selection Bias", as discussed in Al-Ubaydli and List (2015).

[24] The 100 pesos constituted the minimum maintaining balance. During the period of observation, no client closed their account (a 50 peso closing fee applied to ordinary, but not to IS or WR accounts). 18 clients had previously existing 1st Valley Bank savings accounts. For these clients, 100 pesos was deposited in their existing account instead.

[25] One member of the control group was mistakenly offered the IS product, thus a total of 115 IS accounts were opened.

The study suffers from attrition at several stages: Of the 913 individuals who participated in the baseline survey, 852 could be re-located for the marketing visit. Of those reached for marketing, 788 individuals accepted the free ordinary savings account, and 748 agreed to make a savings plan. Several outcome variables (most notably, bank savings by the goal date) are only defined for those who made a savings plan. The endline survey reached 732 respondents. For outcomes which require both endline survey data and a savings plan (such as having purchased the stated savings goal), the sample reduces to their intersection of 615 observations. As shown in Table A3, the orthogonality of attrition to treatment assignment is not rejected at any stage. However, individuals in group WR were reached slightly more often than those in groups IS or C. Table A4 verifies that covariates are balanced across treatment arms at different stages of the sample (Table 1 does this for the full sample).

3.2. Commitment Savings Account Features

The installment-savings product committed clients to a fixed installment plan with weekly (84 percent) or bi-weekly (16 percent) due dates. An account was considered in default from the day the client fell three installments behind. In case of default, the account was closed, an 'Early Termination Fee' was charged, and any remaining savings were returned to the client. A termination fee that is directly linked to the installment structure distinguishes the IS product from withdrawal-restriction or ordinary accounts, and represents its key commitment feature. The amount of the fee was chosen by the client upon signing the IS contract, and donated to charity in case of default.[26] A variety of flexibility features allowed clients to adapt to changing circumstances: First, clients could fall up to two installments behind at any given time, making it theoretically possible to miss every other installment, and pay a double installment in alternate weeks. To encourage timely depositing, a small 10 peso ($0.25) admin fee had to be paid upon making up a missed past installment, but this fee did not accumulate over time. Deposits towards future weeks could be made at any time, as long as they were in increments of the weekly installment. This was a practical requirement, as the client's progress was monitored by making ticks on a collection card for each successful week (see Figure A1). The possibility of making future deposits early effectively provided a form of insurance against uneven income streams. Withdrawals during the savings period were only possible by allowing default to occur. 26

Enforceability of the termination fee was facilitated through the account opening balance: To complete the opening of an IS account, clients had to deposit an opening balance equal to their first weekly installment, but at least 150 pesos (250 pesos) for savings goals below (above) 2500 pesos. Using the same threshold, clients could choose a termination fee as high as they liked, but no lower than a minimum of 150 pesos (250 pesos). Consequently, the minimum termination fee (chosen by 80 percent of clients) could always be enforced. Termination fees above the minimum could be enforced only if the client continued to save, or if their opening balance exceeded the minimum. By construction, all IS accounts were either successfully completed or in default by the goal date, and any remaining savings were transferred to clients' ordinary savings accounts.

The withdrawal-restriction account was simpler in structure: Clients chose to restrict withdrawals before a specified goal was reached. Out of 92 WR clients, 39 chose a goal amount, and 53 chose a goal date. The goal amount can be interpreted as the stronger restriction, since additional deposits need to be made in order to receive

[26] Due to widespread scepticism towards banks, giving the penalty revenues to the bank would have undermined the interpretation of the product as a self-commitment device. Clients could choose between three national Philippine charities, which were intentionally generic and remote. Attitudes towards charities predict neither demand for the IS product nor default (see Table A5), and are controlled for in all regressions in Section 5.3.

TABLE 1. SUMMARY STATISTICS BY TREATMENT ASSIGNMENT

	Installment Savings Treatment	Withdrawal Restr. Treatment	Control	F-stat P-value
Age (yrs)*	43.834	43.449	44.250	0.798
	(0.603)	(0.821)	(0.841)	
Female*	0.941	0.943	0.943	0.991
	(0.011)	(0.015)	(0.015)	
Married	0.862	0.873	0.851	0.795
	(0.016)	(0.022)	(0.024)	
Weekly HH income (pesos)	2890.89	2485.78	3194.43	0.048
	(124.26)	(165.13)	(272.45)	
No. of appliances owned	2.276	2.180	2.250	0.802
	(0.084)	(0.110)	(0.124)	
No. of HH members	5.072	5.180	5.430	0.108
	(0.091)	(0.140)	(0.140)	
Education (yrs)	10.556	10.392	10.564	0.840
	(0.166)	(0.242)	(0.251)	
Received real rewards*	0.503	0.522	0.526	0.837
	(0.023)	(0.033)	(0.033)	
Present Bias*	0.172	0.161	0.156	0.839
	(0.018)	(0.025)	(0.025)	
Perceived	2.384	2.185	2.471	0.225
Temptation (range 0-10)	(0.089)	(0.112)	(0.121)	
Impatience	0.322	0.404	0.333	0.096
	(0.022)	(0.033)	(0.031)	
Faces strong financial	0.393	0.388	0.386	0.987
claims from others*	(0.023)	(0.032)	(0.032)	
Risk aversion (range 0-6)	4.225	4.636	4.132	0.010
	(0.093)	(0.122)	(0.129)	
Cognitive ability	2.937	2.886	2.934	0.887
(range 0-5)	(0.059)	(0.089)	(0.096)	
Financial literacy	1.856	1.838	1.851	0.977
(range 0-5)	(0.047)	(0.068)	(0.069)	
Existing savings account	0.468	0.465	0.425	0.518
	(0.023)	(0.033)	(0.033)	
Donated to charity	0.396	0.386	0.452	0.284
in the last 12 months	(0.023)	(0.032)	(0.033)	
Global Signif. Test (P-value)	0.96	0.18	0.58	
Observations	457	228	228	913

Note: A starred variable indicates that the randomisation was stratified on this variable. Individuals were classified as facing *strong financial claims from others* if they reported strictly above-median values for the financial requests they would face from relatives, friends and neighbours in a hypothetical scenario where they kept cash at home. *Existing savings account* is an indicator for holding a bank account with any bank at the time of marketing. *Real rewards, impatience,* and *present bias* refer to the time-preference elicitation: *Real rewards* is an indicator for receiving monetary incentives. Individuals are defined as *present biased* if the reward needed to make them wait for one month is larger in the present than in the future. They are classified as *impatient* if they always chose the earlier reward in all time preference questions. *Perceived temptation* is the difference between individuals' 'tempted' and their 'ideal' allocation of 10 restaurant nights across 2 years, described in Section 3.3. Risk aversion represents the individual's choice from a set of six lotteries with increasing expected value and variance, where the 'no risk' option yields a score of 6. *Cognitive ability* is the number of correct answers (0-5) from a culture-free intelligence test. *Financial literacy* is the number of correct answers (0-5) to basic numeracy questions.

TABLE 2. SAVINGS PLAN AND ACCOUNT STATISTICS

	Installment Savings Treatment	IS adopters only	Withdrawal Restr. Treatment	WR adopters only	Control	F-stat P-value: Equality IS, WR and C
Accepts ordinary savings account with 100 pesos	0.92	1	0.92	1	0.93	0.95
Accepts commitment account	0.27	1	0.42	1		0.00
Savings Plan Characteristics						
Median Goal Amount (pesos)	2400	2400	2825	2465	3000	0.00
Median Installment Size (pesos)	150	150	200	150	170	0.31
Median Time until Goal Date (days)	137	138	135	133	137	0.79
Commitment Account Characteristics						
Median chosen penalty (IS)		150				
Chooses minimum penalty (IS)		0.80				
Chooses date-based goal (WR)				0.58		
Chooses amount-based goal (WR)				0.42		
Outcomes - Ordinary Accounts (means)						
# Deposits into ordinary account	0.60	1.50	0.25	0.42	0.22	0.01
Makes at least 1 deposit after opening	0.21	0.50	0.07	0.10	0.08	0.00
Savings by goal date (pesos)	75.94	196.75	61.10	95.97	27.16	0.32
Outcomes - Commitment Accounts (means)						
# Deposits into commitment account	1.82	6.75	0.71	1.68		0.00
	successful	*11.98*	*date-based*	*1.68*		
	default	*2.52*	*amount-based*	*1.69*		
Makes at least 1 deposit after opening	0.19	0.69	0.09	0.21		0.00
	successful	*1*	*date-based*	*0.25*		
	default	*0.44*	*amount-based*	*0.15*		
Savings by goal date (pesos), net of penalties	379.85	1233.67	114.30	233.56		0.00
	successful	*2636.25*	*date-based*	*286.09*		
	default (before penalty)	*300.99*	*amount-based*	*162.18*		
	default (after penalty)	*120.51*				
Observations for account data	423	114	219	92	210	
for savings plan data	367	114	188	92	191	

Note: Respondents were allowed to modify their initial savings plan after accepting a commitment product. Deposits refer to physical bank transactions, not to IS installments. Several IS installments could be made in a single deposit. The opening balance for IS and WR counts as the first deposit, while the opening balance for the ordinary accounts was paid by the research team. The sample for account outcomes are those respondents reached in marketing. The sample for savings plan outcomes are those willing to make a savings plan. In the IS group, *'successful'* denotes the subsample who completed their IS contract successfully, while '*default*' is the subsample who defaulted. Similarly, *date-based* and *amount-based* refer to the respective subsample of WR clients. Two outliers are excluded, see footnote 37.

savings back. There was no time limit for reaching the goal amount. However, as is common for Philippine banks, significant dormancy fees were applied after two years of inactivity. The minimum opening balance for the WR account was 100 pesos.[27] For both IS and WR, opening balances were collected one week *after* contract signing. The practical motivation was to give individuals time to prepare for the expense. The theoretical motivation was to free the decisionmaker from temptation in the contract-signing period – a sophisticated hyperbolic discounter should choose a welfare-maximising contract when asked in period 0, but not necessarily when asked in period 1.[28] Finally, both products shared the same emergency provisions: In case of a medical emergency or death in the family, a relocation to an area not served by the bank, or a natural disaster, clients could close their account and access their savings without any penalties. Within the six months of observation, no client exercised this option.

3.3. Eliciting Time Preferences and Sophistication

In the context of a comprehensive baseline survey, I measured time-inconsistent preferences using standard multiple price lists (MPLs):[29] Individuals were asked to choose between a fixed monetary reward (200 pesos) in one period and various mostly larger rewards (180 to 300 pesos) in a later period. A randomly chosen half of the sample received real rewards, for the others the questions were hypothetical (see Appendix V for details of the time preference elicitation). After six questions using a near time frame (now versus one month), the same six questions were asked for a future time frame (one month versus two months). The outcome of interest was the size of the later reward necessary to make the individual switch from preferring the smaller earlier reward. I identify as 'present-biased' those who put a higher premium on waiting for one month in the present than in the future. Individuals who exhibit more patience now than in the future are 'future-biased'. An individual who always prefers the earlier reward in all questions and time frames is classified as 'impatient'. I find 16.6 percent of individuals to be present-biased, 18.9 percent future-biased, and 35.7 percent impatient.[30] Offering real or hypothetical rewards does not significantly affect these proportions in the cross-section. However, incentivized measures do substantially better at predicting later behaviour (IS take-up and default).[31]

In addition to a measure of preference reversals, the analysis requires a measure of sophistication which is not in itself derived from commitment demand. To the author's knowledge, the only study which identifies $\tilde{\beta}$ on an individual level using experimental data is Augenblick and Rabin (2015).[32] Their prediction task requires a sophisticated laboratory setting, which is not practicable in a field experiment with over 900 participants. Instead,

[27] Different minimum opening balances for IS and WR may drive the observed difference in take-up.

[28] See e.g. Benartzi and Thaler (2004). Financially, the late collection just delayed when individuals entered the commitment contract. However, signing the contract was associated with substantial paperwork, as well as a non-financial commitment to the marketers, who personally collected the opening balance. Out of 159 (116) individuals who initially signed the IS (WR) contract, 45 (24) failed to deposit an opening balance.

[29] See e.g. Ashraf, Karlan and Yin (2006). In a developing country field setting, MPLs were more practicable than the convex budget method recently introduced by Andreoni and Sprenger (2012). While MPLs fail to identify point values for β and δ, they do identify time-inconsistency, i.e., $\beta < 1$. Furthermore, the use of monetary payments in measuring consumption preferences has been questioned by Augenblick, Niederle and Sprenger (2015). While their criticism is valid here, severe liquidity constraints are widespread in the low-income study population, creating a tight link between cash inflows and consumption.

[30] In their MPL benchmark, Andreoni and Sprenger (2012) find 16.7 percent present-biased and 10.7 percent future-biased. Kaur, Kremer and Mullainathan (2015) find 17 percent present-biased. Using hypothetical questions, Ashraf, Karlan and Yin (2006) find 27.5 percent present-biased and 19.8 percent future-biased. Since future-biased and time-consistent preferences generate the same predictions for commitment demand, they are not distinguished in the regression analysis.

[31] See Tables A13 and A14 for a comparison of real and hypothetical incentives, and Appendix IV for a discussion.

[32] Related structural approaches include Acland and Levy (2015) and Skiba and Tobacman (2008).

I develop a survey-based proxy measure, adapted from the self-control measure proposed by Ameriks et al. (2007) (henceforth ACLT). ACLT infer sophistication from predictions about future temptations and behaviour. Simple hypothetical survey questions elicit individuals' ideal, tempted, and expected allocation of a fixed resource over time. While designed to identify the parameters of the Gul and Pesendorfer (2001) model, the resulting measure reflects an individual's *perceived* (rather than actual) self-control problems.

The setup is as follows: Respondents were presented with a hypothetical scenario of winning ten certificates for "dream restaurant nights". Each certificate entitled the holder and a companion to an evening at any local restaurant of their choice, including an unlimited budget for food and drink, and all gratuities. The certificates were valid for two years starting immediately, and expired thereafter. Piloting revealed that the low-income setting avoided several confounds which might be present in richer countries: First, participants were used to eating in restaurants only on special occasions, preventing substitution of certificates into everyday consumption. Second, opportunity cost of time was generally low. Third, the restaurants that participants perceived as desirable tended to be family-friendly fast-food chains (e.g. *'Jollibee'*), which required neither advance bookings nor a babysitter. Restaurant vouchers were thus understood as a carefree temptation good. Using the wording of ACLT, I then asked for

1. the *ideal* allocation of the ten certificates to year 1, as opposed to year 2, and

2. the allocation individuals would be *tempted* to consume in year 1.

These questions provide the measure *Perceived Temptation* (from $2.-1.$, $tempted-ideal$).[33] In a $\beta\delta$-model (where costly self-control does not exist), it captures the difference between the ex-ante optimal allocation, and the allocation the agent expects to play in a subgame perfect equilibrium. This difference is a function of perceived time-inconsistency ($\tilde{\beta}$ in Section 2). However, directly using $tempted-ideal$ to proxy $\tilde{\beta}$ is problematic in the presence of 'self-control types' (Toussaert (forthcoming)), who behave time-consistently, but face a cost for resisting temptation. To capture sophistication about present-biased preferences (and filter out time-consistent self-control types), I interact $tempted-ideal$ with an indicator for behaving time-inconsistently in MPLs. The resulting interaction $(tempted-ideal)*presentbias$ proxies $\tilde{\beta}$ for agents with $\beta<1$.[34]

A caveat is that sophistication is modeled as $\tilde{\beta}-\beta$, while $(tempted-ideal)*presentbias$ captures $\tilde{\beta}$ – sophistication for a *given* level of true time-inconsistency, β. When used in regressions, the interpretation of coefficients is facilitated by the fact that measures of *perceived* and *observed* time-inconsistency are empirically orthogonal: Surprisingly, $tempted-ideal$ does not predict present bias in MPLs ($p=0.9$, Figure A7 shows the bivariate distribution). The data suggests that it is plausible to increase $\tilde{\beta}$ while holding β constant. Several factors may explain the lack of a correlation: First, individuals may use costly self-control to appear time-consistent.[35] Second,

[33] A third question asked for the allocation individuals *expected* to consume in year 1, taking into account both the ideal and the temptation. This yields a second measure, $expected-ideal$ (from $3.-1.$). Questions 1, 2 and 3 correspond to questions (a), (c) and (d) in ACLT, respectively. Note that, in a $\beta\delta-$model, the answers to 2. and 3. should be the same. In practice, 82 percent claimed positive differences between *tempted* and *ideal*, while only 22 percent admit differences between *expected* and *ideal.* Suggestive evidence indicates possible image concerns towards the interviewer: Admitting that one is subject to temptation may be less embarassing than admitting that one gives in to these temptations. The main text proceeds by focusing on the tempted allocation. Full details on all self-control questions and their distribution are reported in Appendix V.2. Table A12 reports robustness checks of the main results with respect to the sophistication measure.

[34] Note this interaction assumes domain-generality of β across money and consumption. Temptation values are censored at zero (negative values occurred in 4 of 910 cases).

[35] In a lab experiment, Toussaert (forthcoming) finds between 23 and 36 percent 'self-control types', who are willing to pay for commitment even though they expect to resist temptation, in order to reduce the need for costly self-control.

all variables may be measured with error. Especially the MPL present bias measure is likely understated due to the high fraction (35.7 percent) of always-early choices, in addition to some recent concerns about monetary discounting (Augenblick, Niederle and Sprenger (2015)). Relatedly, $tempted - ideal$ is measured over the consumption domain, while *presentbias* is measured over money. If β is highly domain-specific, a lack of correlation may result. Third, individuals may be pessimistic about their degree of time-inconsistency (discussed in Appendix I.3).

In addition to measuring time-inconsistency, the baseline survey obtained measures of financial claims from others, risk aversion, cognitive ability, financial literacy, intra-household bargaining power, distance to the bank branch (via GPS coordinates), attitudes towards charitable giving, and frequency of income or expenditure shocks, as well as an indicator for having an existing bank account. These measures are discussed in Appendix V.

Table 1 presents summary statistics. Randomisation into treatment groups occurred shortly after the baseline survey. Means were statistically different across treatment groups for income, impatience and risk aversion. Income and impatience have no predictive power in any of the later regressions. In particular, wealthier individuals are no more likely to take up a commitment product than poorer individuals. Risk aversion does have predictive power for the take-up of the withdrawal-restriction account. Robustness checks are reported in Appendix IV.

4. Predictions and Empirical Strategy

The theoretical framework in Section 2 predicts that partially sophisticated hyperbolic discounters are likely to adopt commitment, but then fail to save and default. Full sophisticates can successfully save with commitment, but may be less likely to adopt it, as low stakes are correctly perceived as ineffective, and high stakes may be prohibitive in the face of uncertainty. Going beyond the basic model, time-consistent agents who commit for other reasons, e.g. to protect savings from external claims, are likely to save more with commitment. This generates the following primary outcomes of interest for the empirical analysis, along with predictions from the theory:

a) Average treatment effects of the two commitment products on savings levels: Effects will be ambiguous, depending on the composition of partial and full sophisticates, as well as time-consistents. While not explicitly modelled, installment savings are likely to be more effective than withdrawal-restriction commitments if agents are constrained by self-control problems, since they commit clients to make future deposits. In contrast, withdrawal restrictions will be more effective if 'other-control problems' are important, since they allow clients to safeguard savings (IS permits withdrawals at the cost of default).

b) Heterogeneity in follow-through on installment savings contracts: The presence of partial sophisticates predicts substantial levels of default. With simple degenerate beliefs $\tilde{\beta} > \beta$, default is deterministic – though this prediction softens with stochastic sophistication (Appendix I.1). Full sophisticates and time-consistents will default only in case of shocks.

c) Predictors of commitment demand and follow-through: Demand for commitment will be negatively predicted by sophistication about present bias, but not by present bias on its own.[36] Default on commitment (failure

[36]The relationship between commitment demand and sophistication is monotone and negative if present-biased agents cannot save without commitment, and are able to observe this fact ($\tilde{\beta} < \hat{\beta}$, see part (c) of Proposition 4 and Figure 1). The sample population frequently borrows at very high interest rates to finance lump-sum consumption (this partially motivated the IS design, see introduction), which supports this assumption.

to complete the IS contract) is more likely for present-biased agents with low levels of sophistication ('partial sophisticates'), and less likely for present-biased agents with high levels of sophistication ('full sophisticates').

Section 5.1 addresses part a). I study the average treatment effect of the two commitment products by running OLS regressions on

$$Y_i = \alpha_0 + \alpha_{IS} \cdot IS_i + \alpha_{WR} \cdot WR_i + \varepsilon_i \tag{3}$$

where the primary outcome of interest Y_i are participants' savings at the partner bank, IS_i is an indicator variable for assignment to the 'Installment Savings' group, and WR_i is an indicator variable for assignment to the 'Withdrawal Restriction' group. Secondary outcomes include whether or not the respondent purchased their stated savings goal, and whether they borrowed to do so. A survey-measure of other savings checks for evidence of substitution across savings sources. The coefficients α_{IS} and α_{WR} estimate the Intent-to-Treat effect (ITT) – the causal effect of having been offered the corresponding commitment product.

Section 5.2 addresses b), using a descriptive analysis of the occurence and timing of defaults.

Section 5.3 addresses c). I run probit regressions of the binary choice equation

$$\begin{aligned} Choice_i \quad = \quad & \Phi(\gamma_0 + \gamma_1 \cdot presentbias_i + \gamma_2 \cdot (tempted - ideal)_i * presentbias_i + \gamma_3 \cdot (tempted - ideal)_i \\ & + \gamma_k \cdot X_i^k + \gamma_m \cdot m_i + \varepsilon_i), \end{aligned} \tag{4}$$

where $Choice_i$ is the decision to take up the installment-savings account (within group IS), to take up the withdrawal-restriction account (within group WR), to default on the IS account, or to pre-order IS for a second round. The variables $presentbias_i$ and $(tempted - ideal)_i$ are as defined in Section 3.3. Theory predicts that $\gamma_1 = 0$, $\gamma_2 < 0$ for IS demand (see footnote 36 on monotonicity of γ_2), and $\gamma_1 > 0$, $\gamma_2 < 0$ for IS default. It is silent on other coefficients. X_i is a vector of other individual characteristics, notably a subset which proxies *'Other Control'* or safeguarding motives (facing strong financial claims from others, household bargaining power, education, risk aversion), a subset which proxies past and current shock arrival rates, and finally other demographics (age, female, marital status, an asset index, weekly household income, number of household members, impatience in MPLs, cognitive ability, financial literacy, having donated to charity, distance to the bank branch, and having an existing bank account). All binary choice regressions contain a vector of marketer effects $\gamma_m m_i$ to filter noise from differences in marketer ability.

5. Results

5.1. Average Treatment Effects on Savings

This section presents estimated average effects of the two commitment treatments. The primary outcome variable of interest is clients' total savings balance at the partner bank, summed across ordinary savings accounts and any commitment savings products (IS or WR). The savings period is specific to each individual, starting with the date of the baseline survey, and ending with the goal date specified in an individual's savings plan. The reason for focusing on the goal date is that all savings are expected to be spent on the planned expenditure after this date. The cost of this choice is that it diminishes the sample to those 748 individuals who were willing to make a savings plan (attrition is orthogonal to treatment, see Section 3.1 as well as Tables A3 and A4).

TABLE 3. SAVINGS OUTCOMES (OLS)

	(1) Bank Savings	(2) Purchased Savings Goal	(3) Borrowed to Purchase Goal	(4) Other Savings (survey-based)
Installment Savings Treatment	428.633***	0.116**	0.036	426.811
	(65.587)	(0.049)	(0.0234)	(671.844)
	[0.000]	[0.100]	[0.328]	[0.737]
Withdrawal Restr. Treatment	148.243***	0.134**	0.121***	-328.159
	(40.927)	(0.056)	(0.034)	(705.461)
	[0.086]	[0.075]	[0.003]	[0.631]
Constant	27.160***	0.407***	0.047***	63.451
	(9.399)	(0.040)	(0.017)	(531.028)
R^2	0.04	0.011	0.023	0.00
Observations	746	615	615	603

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Square brackets contain additional p-values corrected for Multiple Hypothesis Testing, following List, Shaikh and Xu (2016). *Bank Savings* is the change in a client's total savings balance at the partner bank, summed across ordinary savings accounts and commitment savings products (IS or WR). Two outliers are excluded, see footnote 37. The period of observation starts with the baseline survey and ends with the goal date specified in an individual's personal savings plan (median duration: 137 days). *Purchased Savings Goal* is an indicator for whether the individual reported having purchased or paid for the savings purpose specified in their savings plan at the time of the endline survey. *Other Savings* is the change in self-reported savings at home and at other banks. Survey-based savings data are truncated at 1 percent.

Column (1) in Table 3 estimates that assignment to the Installment Savings treatment group increased average bank balances by 429 pesos (U.S.$10.20) relative to the control group.[37] This estimate is net of default penalties and the 100 peso gift contained in the ordinary savings account. As shown in Table 2, the effect can be decomposed into a 380 pesos (1.82 deposits) increase in commitment account savings, and a 49 pesos (0.38 deposits) increase in ordinary account savings. Individuals assigned to the Withdrawal-Restriction group saved on average 148 pesos more than the control group – of which 114 pesos (0.71 deposits) in their commitment account, and 34 pesos (0.03 deposits) in their ordinary account. When normalized by the savings duration, the WR treatment effect replicates the result of Ashraf, Karlan and Yin (2006): Their SEED account increased savings by 34.25 pesos per month (411 pesos over 12 months), while the WR account increased savings by 32.88 pesos per month (148 pesos over 4.5 months). The terms and conditions of SEED and the WR account were identical.[38] The small but significant

[37] The ITT estimate excludes two outliers, both in the IS group, whose savings increased by 15 and 18 standard deviations, respectively. Including these outliers changes the ITT estimate to 585 pesos.

[38] The study locations are 70km (2 hours by local bus) apart. The study populations differed: Ashraf, Karlan and Yin (2006) studied a sample of previous clients of the partner bank, while this paper studies a general low-income population with little previous bank exposure.

savings of 27 pesos in the control group may reflect the effect of marketing, specifically the universal provision of savings plans and ordinary savings accounts.[39]

The ex-post probability that a research finding is true depends not only on the level of statistical significance, but also on the prior beliefs, the power of the study, and any replications (Maniadis, Tufano and List (2014)). Appendix A provides such 'Post-Study Probabilities' for a range of priors for the main experimental treatment effect on savings. Consistent with the discussion in Maniadis, Tufano and List (2014), the PSPs strongly increase in the fact that the WR treatment is a replication of Ashraf, Karlan and Yin (2006).

Treatment-on-the-Treated effects (TOT) can be obtained by instrumenting take-up of IS and WR with assignment to the corresponding treatment group. The identifying assumption is that the mere offer of commitment had no effect on savings (other than via encouraging people to use the products). The TOT estimates suggest that taking up the IS product increased savings by 1392 pesos, while taking up the WR product increased savings by 303 pesos. The increased gap in the TOT effects relative to the ITT effects is a result of the higher take-up rate for WR.

Six months after the baseline survey, when all goal dates had been reached, an endline survey asked whether individuals had purchased the good (respectively, paid for the expenditure) they had been saving for. Out of the 615 individuals who had a) made a savings plan and were b) reached by the endline survey (see Table A3), 307 reported to have bought the desired good. When asked how they paid for this expenditure, slightly below 20 percent of these declared to have used loans from formal or informal sources. Columns (2) and (3) in Table 3 present linear probability estimations of the effect of treatment on the likelihood of purchasing the good, and on borrowing for the purchase.[40]

Table 3 confirms that both the IS and the WR treatment increased an individual's chances of purchasing their savings goal. However, unlike the IS group, the WR group was significantly more likely (12.1 percentage points) to borrow in order to obtain the good. One possible mechanism is that signing either commitment contract (IS or WR) psychologically committed clients to purchase their savings goal, but only the Installment Savings product helped them to do so using their own funds.

An important concern is whether savings increases observed at the partner bank constituted new savings, or whether a simple substitution from other sources of savings (at home, or at other institutions) took place. Column (4) of Table 3 reports the change in an individual's total savings balance outside of the partner bank, as measured in the baseline and endline survey: Individuals were asked about their savings at home, money lent out or safekept by others, informal savings, and savings at other institutions. An incentive of 30 pesos was paid for showing an existing bank passbook.[41] The endline survey asked about savings kept around the goal date, as opposed to the survey date. Unfortunately, the survey data is noisy, and coefficients are estimated with substantial imprecision.[42] The available evidence does not suggest a substitution between increased savings at the partner bank, and reduced savings at home or at other institutions. All coefficients are insignificant, and the sign on being offered IS is positive.

Appendix III provides a number of supplementary analyses, including treatment effects on the cumulative distribution of bank savings, total savings, outstanding loans, and expenditures (Figure A4 and Table A11). Table

[39] The effect of marketing as such cannot be identified. However, the non-negativity constraint on bank savings and the 27 pesos control mean suggest that the effect of marketing was likely small.

[40] Probit estimations yield very similar results in terms of effect sizes and significance.

[41] At baseline, 46 percent of the sample reported to have an existing savings or checking account with another institution. This number is partly driven by a formal requirement to open a savings account when obtaining microloans. More than one quarter of bank account holders reported not to have used their account in the last 12 months, and dormant accounts were common.

[42] The savings data has been truncated at 1 percent, reducing the sample from 615 to 603 observations.

A9 presents quantile treatment effects. Table A8 examines treatment effect heterogeneity across a number of covariates, and finds that such heterogeneity is most pronounced for existing savings account holders – consistent with trust and basic familiarity with the banking system.

5.2. Heterogeneity: Descriptive Results

The IS results were strongly bi-modal: 51 IS clients (45 percent) successfully completed their savings contract, with goal dates between December 2012 and April 2013. They completed all scheduled installments with a median of 12 transactions, and reached savings goals between 950 and 7150 pesos (U.S.$170).[43] By design, accounts were closed after completion of the savings plan, and clients could withdraw their savings in order to pay for the planned lump-sum expenditure. Many of these clients pro-actively enquired at the bank to roll over their account into a new IS contract. While this was not an immediate possibility, the repeat marketing stage included the option to 'pre-order' the product for a second round. The pre-order contract was not financially binding, but included substantial official paperwork. Two thirds of the successful clients took up this offer (see Table A2), devised a new savings plan, and chose a new termination fee. The bank has since offered new IS contracts to those enquiring about them.

The situation looked very different for the remaining 63 IS clients (55 percent) who defaulted on their savings contract. After falling three deposits behind, their accounts were closed, and the initially agreed termination fee charged.[44] What happened? Two possibilities emerge: (i) Clients had chosen an IS contract which was optimal for them in *expectation*, and then rationally defaulted upon observing a shock (in other words, a 'bad luck' scenario). Or (ii), clients chose the contract by mistake.

If the 'bad luck' explanation is true, the timing of defaults should depend on the shock distribution: Assume that, as modeled in Section 2, savers get hit by large shocks with a per-period probability of λ. If shocks are i.i.d., a fraction λ of the surviving population should default each period. In other words, the default hazard rate should be constant over time. In sharp contrast, Figure 2 illustrates that clients had a tendency to default either right from the start, or not at all: Out of 63 defaults, 35 clients stopped depositing immediately after the opening balance, 8 clients deposited one more installment, and only 10 clients defaulted after depositing more than five installments. Figure 2 shows the default hazard rate by installments, where the number of defaults at installment t is given by those who discontinued depositing after making t installments. The active population at t includes those who have neither defaulted nor successfully completed their contract at this point. Far from being constant, the default hazard rate spikes after the first installment, and trails off afterwards. The null hypothesis of a constant hazard rate is rejected at $p<0.01$ despite using only 24 installment-observations.[45]

Approximating installments with weeks (84 percent of clients chose weekly installments), Figure 2 further illustrates the predicted default timing given a hazard rate of 0.028 per week. This estimate is obtained from the endline survey: The sample population was questioned about the occurrence of 17 types of common emergencies (sickness, loss of job, bad business, flood damage) including a flexible 'other' category. 45% reported at least one emergency within 6 months, with an average of 0.72 emergencies. A simple proxy of i.i.d. shock arrival with

[43] Savings contracts had a median of 16 scheduled installments, with a range from 8 to 25.

[44] Subject to the enforceability constraint described in Section 3.2. Appendix Figure A3 shows the chosen and charged termination fees.

[45] Using installments like a time axis is a simplification: Individuals were allowed to fall two installments behind at any time. Consequently, time lags occur between the last completed installment and the official time of account closure. Furthermore, individuals could deposit multiple installments in a single transaction.

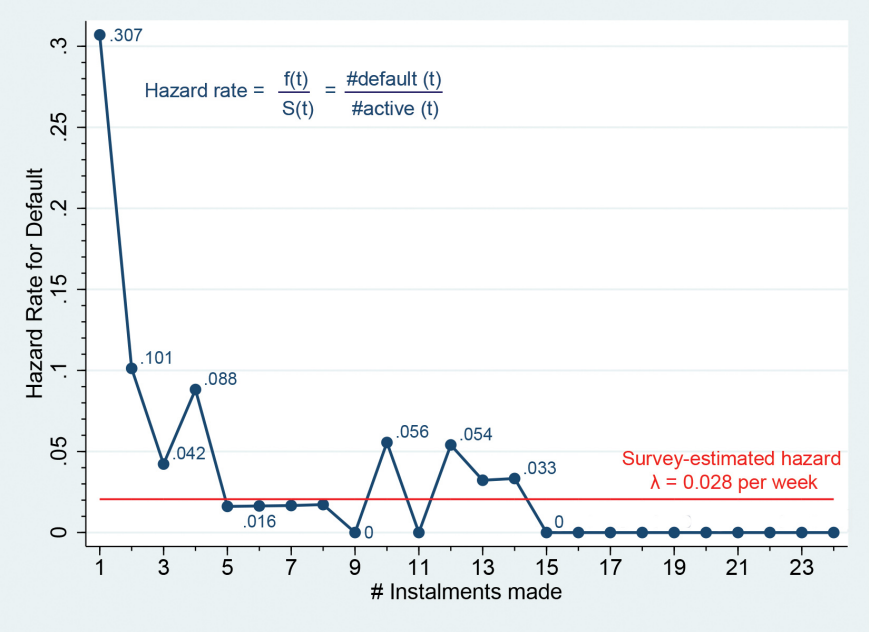

FIGURE 2. HAZARD RATE FOR IS DEFAULT BY NUMBER OF INSTALLMENTS MADE

emergencies yields a hazard rate of 0.028 per week. This hazard rate is neither consistent with the overall frequency of defaults (observed 55 percent versus predicted 29 percent based on a 12 week contract), nor with the steep timing of defaults. I.i.d. shocks are a highly simplistic assumption – more realistically, the shock arrival rate is heterogeneous among individuals, shocks may be correlated, and individuals may be naive about shock arrival. These alternative assumptions are discussed in Sections 6.1 and 6.2. In short, while empirically plausible, they would not be sufficient to generate the observed data.

The second possibility requires a deviation from rational expectations: Individuals could have chosen their contract by mistake. Mistakes (defined as choices that are not optimal under rational expectations) can happen if individuals have incorrect beliefs about their future preferences or their income distribution, including the probability of shocks to either of the two. Section 2 outlines why a time-inconsistent agent with incorrect beliefs about the degree of her time-inconsistency is likely to select into a commitment contract that is too "weak" to overcome her self-control issues, leading to default. Section 6 discusses other types of forecast errors. Looking at the data, it is notable that 80 percent of individuals chose the minimum permissible termination fee for their savings goal (P150 for goals below P2500 (56 percent) and P250 for goals above P2500 (24 percent)) – though the goal choice itself likely captures a margin of commitment strength. The observed combination of minimum penalties and high default rates raises the question whether individuals underestimated the amount of commitment it would take to make them save. This is consistent with the observed tendency to default soon after account opening, as individuals start behaving according to their true degree of time-inconsistency upon entering the depositing phase. Could rational expectations about stochastic future time-inconsistency explain the data? If individuals had correct beliefs on average, they would realize which penalty will be effective for them on average. Building on the discussions in Section 2.6 and Appendix I.1, uncertainty in β makes commitment less attractive. Theory predicts low commitment take-up, high conditional penalties, and low default rates – in sharp contrast to the patterns observed.

For the WR accounts, both benefits and risks were less pronounced: Out of 92 accounts, only five reached the specified goal amount (3 out of 53 date-based accounts, and 2 out of 39 amount-based). Median savings were 100 pesos, equivalent to the minimum opening balance. 79 percent of WR clients (85 percent of amount-based accounts) made no further deposits after the opening balance. Similarly to IS clients, amount-based WR accounts effectively lose their opening balance if they do not continue to deposit. A difference between the two commitment products is that the penalty for discontinuing to save on an amount-based WR account increases with every deposit, while the IS default penalty is fixed. Out of 582 clients who exclusively had an ordinary savings account, one reached their specified goal amount. Summary statistics on account usage can be found in Table 2.

5.3. Heterogeneity: Regressions

In an attempt to resolve the puzzles presented in the previous section, this section analyses empirical predictors of the take-up, default, and pre-order decisions. *A5 A6*

Predicting Commitment Demand & Default: Time Preferences Column (1) of Table 4 presents a probit regression of the IS take-up decision on the time preference measures discussed in Section 3.3. It contrasts these with baseline variables which may capture an 'Other Control' (or safeguarding) motive – the leading alternative explanation for why people demand commitment.

IS take-up is predicted by the proposed measure of sophisticated hyperbolic discounting. Present bias on its own is not a predictor of take-up, consistent with the intuition that *awareness* of time-inconsistency ($\tilde{\beta}$), rather than *actual* time-inconsistency (β), determines demand for commitment. Perhaps more surprisingly, the association of commitment take-up and sophisticated hyperbolic discounting is significant and *negative*. Recall from Section 3.3 that sophistication is measured as the interaction of present bias (observed using MPLs) and self-reported temptation. In other words, those who exhibit hyperbolic preference reversals, but at the same time report *low* levels of temptation, are more likely to take up the product. In contrast, those who report being strongly tempted tend to stay away. The negative link between sophistication and commitment adoption is consistent with the theoretical framework: Commitment is attractive for partially sophisticated agents, who anticipate that a low default penalty will be sufficient to make them save. In contrast, agents who perceive themselves as strongly tempted have two options: Either they choose a sufficiently large penalty, or they stay away from commitment. Non-adoption may be optimal if the required effective penalty is prohibitively high. An additional channel is that penalties were only enforceable up to the current savings level. An agent who anticipates that her first installment is not a sufficient stake to get her to make the second one, should also stay away. Before proceeding to the rest of Table 4, it is worth looking at IS contract choices as an internal margin of commitment demand. Columns (1)-(2) of Table 5 indeed confirm a positive association between sophistication and IS penalty choice, conditional on adoption. Present bias on its own is negatively linked to penalty choice. Aggregating the coefficients, present bias with low levels of sophistication predicts low penalties, while present bias with high levels of sophistication predicts high penalties (note present bias is binary, whereas temptation is in the interval [0,10] with a median of 2). The variation in penalties is partly driven by the variation in savings goals: Choosing a larger savings goal (with a higher minimum penalty) in itself constitutes a stronger commitment choice. However, neither the savings goal choice nor the savings duration are significantly linked to the time preference measures. Due to the selection issues involved in conditioning on take-up, the estimates should be seen as suggestive evidence only.

TABLE 4. PREDICTING COMMITMENT DEMAND & DEFAULT (PROBIT)

Commitment type	Installment Savings				Withdrawal Restr.
Dep. Var.	(1) Take-Up	(2) Default	(3) Default	(4) Pre-Order	(5) Take-Up
	(IS-Sample)	(IS-Sample)	(takeup-Sample)	(takeup-Sample)	(WR-Sample)
Self-Control Motive					
Present Bias	0.0827	0.1119*	0.4837*	-0.4860**	0.1046
	(0.0864)	(0.0654)	(0.2550)	(0.2353)	(0.1301)
Soph. Present Bias	-0.0631**	-0.0453**	-0.1718	0.2375**	-0.0530
(Pres.Bias*Temptation)	(0.0292)	(0.0230)	(0.1299)	(0.1170)	(0.0532)
Perceived Temptation (0-10)	-0.0046	-0.0202*	-0.0655***	0.0109	0.0006
	(0.0125)	(0.0105)	(0.0244)	(0.0301)	(0.0209)
Other-Control Motive					
Faces strong financial	-0.0038	-0.0113	0.0095	0.0097	0.1185*
claims from others	(0.0414)	(0.0330)	(0.0848)	(0.0905)	(0.0646)
HH Bargaining Power (0-5)	0.0053	-0.0116	-0.0778***	0.0792***	0.0456***
	(0.0113)	(0.0090)	(0.0236)	(0.0257)	(0.0167)
Education (yrs)	-0.0094	-0.0030	0.0016	-0.0133	0.0315***
	(0.0064)	(0.0053)	(0.0121)	(0.0142)	(0.0093)
Risk Aversion (0-6)	-0.0059	-0.0181**	-0.0673***	0.0199	0.0500***
	(0.0105)	(0.0084)	(0.0192)	(0.0240)	(0.0168)
Shocks					
#Emergencies last yr	-0.0161	0.0005	0.0558	-0.0855	-0.0215
	(0.0277)	(0.0213)	(0.0601)	(0.0677)	(0.0493)
#Emergencies since baseline		-0.0033	0.1156*	-0.0414	
		(0.0182)	(0.0687)	(0.0658)	
Controls	YES	YES	YES	YES	YES
Marketer FE	YES	YES	YES	YES	YES
Mean Dep. Variable	0.2687	0.1468	0.5463	0.4630	0.4115
Observations	402	402	108	108	209

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Entries in the table represent marginal coefficients. Variables are as described in Table 1. Additional controls include age, gender, marital status, weekly household income, number of appliances owned, number of household members, impatience, cognitive ability, financial literacy, distance to bank (via GPS), having an existing savings account, having donated to charity in the last 12 months. The coefficients for all control variables are reported in Tables A5 and A6. The sample is restricted to clients who could be located for the marketing visit (see Table A3). For robustness checks, see Table A12.

TABLE 5. PREDICTING IS CONTRACT CHOICES

	IS Penalty (pesos)		Goal Amount (pesos)		Savings Plan Duration (days)	
	(1)	(2)	(3)	(4)	(5)	(6)
Present Bias	-75.56**	-69.70**	-1,093	-1,106	-3.263	-2.527
	(34.35)	(31.92)	(672.5)	(765.0)	(17.49)	(18.27)
Soph. Present Bias	30.69**	25.69**	290.7	282.0	0.892	0.224
(Pres.Bias*Temptation)	(13.34)	(12.78)	(263.4)	(296.0)	(10.36)	(10.81)
Perceived Temptation (0-10)	-7.324	-3.443	-241.4	-232.9	-3.019	-2.451
	(9.490)	(7.383)	(156.5)	(182.7)	(2.080)	(2.116)
Demographic Controls	NO	YES	NO	YES	NO	YES
R^2	0.019	0.084	0.042	0.046	0.023	0.042
Observations	109	109	109	109	109	109

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Demographic controls include age, weekly income, and education level. A histogram of the chosen IS penalties is in Figure A3. Savings Plan summary statistics are in Table 2.

Moving back to Table 4, and proceeding with the IS default decision, column (2) predicts IS default among those assigned to the Installment Savings group. Individuals who do not commit cannot default. Thus, the coefficients are best understood as predictors of who took up the commitment product 'by mistake', proxied by take-up and subsequent default. This interpretation notably abstracts from the possibility of rational default following a shock. The results provide further support to the partial sophistication hypothesis: Present-biased individuals are significantly more likely to take up the IS product and then default. This effect is particularly strong for agents who report low levels of temptation, representing naive and partially sophisticated hyperbolics. In contrast, more sophisticated hyperbolics are *less* likely to default: Aggregating the coefficients for present bias (0.11*), sophistication (-0.045**) and temptation (-0.02*) yields a lower likelihood of default for all present-biased agents with temptation values higher than the median of 2.[46]

Columns (3) and (4) of Table 4 restrict the analysis to clients who took up the IS product, and should be interpreted with care: The regressions condition on an endogenous variable, and are likely subject to sample selection bias. In predicting default occurrence, the marginal coefficient on present bias has quadrupled, and kept its significance. The link between present bias and default seems to be stronger than the link between present bias and take-up, consistent with the intuition that *awareness* of time-inconsistency drives commitment adoption, while actual time-inconsistency determines the success of the contract. The temptation measure now has strong predictive power on its own, even when not interacted with present bias. Individuals who report feeling tempted but who do not exhibit hyperbolic preference reversals could be (i) time-inconsistent, but incorrectly classified by the MPLs, for instance because their switching point is beyond the measured range (35.7 percent of respondents chose the earlier reward in all MPL questions). Or (ii), they could have Gul-Pesendorfer preferences: They act time-consistently in MPLs, demand commitment, report high levels of temptation, and don't default on the IS contract (see also Section 3.3).

Moving on to the pre-order (repeat take-up) decision, the coefficients on present bias (-0.49**) and sophistication (0.24**) are large and significant. The aggregate coefficient for a present-biased individual with the median value of perceived temptation is approximately zero. This has a convenient interpretation: Relatively naive hyperbolic discounters (those with below-median reported temptation) are unlikely to take up the IS product again. The same group is most likely to have defaulted on their previous contract. The result is encouraging, suggesting that individuals who 'burnt their fingers' learned about their true preferences. The reverse holds true for present-biased individuals with above-median reported temptation (sophisticated hyperbolics): They were more likely to pre-order IS for a second round, potentially following a positive experience with their first contract.

Predicting Commitment Demand & Default: Other Covariates Table 4 contrasts predictors relating to time preferences ('Self Control Motive') with predictors that are plausibly related to an 'Other Control' (or safeguarding) motive, as well as proxies for shock hazard rates. Both are important alternative mechanisms: Other-control problems are the main theoretical alternative as to why people demand commitment (Anderson and Baland (2002)), and may also explain default in case of household conflicts. Shock arrival rates may moderate commitment demand, and cause defaults (see Section 2).

[46] Among defaulting clients identified as present-biased, 85 percent had below-median values of temptation, i.e. were relatively naive. Only 22 percent of defaulters are identified as present-biased (relative to 17 percent in the whole sample). However, the fact that below-median temptation correlates with default even for those *not* identified as present-biased suggest that the present bias measure may be understated -- see discussion below and in Section 3.3.

Take-up of the IS product is predicted by few factors outside of time preferences: Neither the variables arguably grouped under other-control (facing strong financial claims from others, household bargaining power, education, risk aversion) nor a proxy for the baseline shock arrival rate (*#Emergencies last yr*) predict IS adoption. None of the demographic controls (listed in Table 4's notes) predict adoption, with the exception of a positive relationship with cognitive ability (discussed in Section 6.6), and having an existing bank account at any local bank. Appendix Tables A5 and A6 report the full list of coefficients.

Exploring predictors of IS default outside of time preferences, the most obvious candidate – the occurrence of shocks during the savings period – finds some support in the take-up sample (*#Emergencies since baseline*, column (3)).[47] The positive correlation of defaults with shocks, in combination with the fact that 45 percent of clients successfully completed their IS contract, suggests that a significant portion of clients likely did choose a contract which was optimal for them in expectation. The theoretical prediction that shock realisation should be irrelevant to the pre-order decision (as it does not affect contract optimality in expectation) is supported by the data (see column (4) of Table 4).

Further, at least within the take-up sample, IS default is strongly related to household bargaining power. Individuals may have learned soon after opening their account that it causes household conflicts to put aside a portion of the household budget every week. Clients with low bargaining power are likely to have yielded to these disagreements, and defaulted on their contracts. This can be interpreted as costly experimentation with a new savings technology, and is discussed further in Section 6. The large positive association of household bargaining power with the pre-order decision further supports a learning explanation: Once individuals had learned about the difficulties of regularly diverting a share of the household budget, only those with sufficient autonomy chose to take up the product again.

Risk aversion is negatively associated with IS default, and may mitigate various default risks, including household conflicts and shocks: Risk averse agents may be better managers of household finances, and set aside 'buffers' which can be used to make the IS deposits. Within the set of additional controls, only financial literacy (negative) and cognitive ability (positive) predict IS default (Table A6). Notably for the discussion on alternative mechanisms (Section 6), neither distance to the bank branch nor charity attitudes (footnote 26) significantly relate to IS defaults.

The discussion thus far has focused entirely on the IS product, at the expense of the WR account. If the installment-savings and the withdrawal-restriction commitment product were perceived as close substitutes, then the factors predicting IS take-up should also predict WR take-up. Column (5) of Table 4 presents a probit regression of the WR take-up decision within group WR. In sharp contrast to IS adoption, WR adoption is predicted by high education (in years of schooling), high risk aversion (choosing a safe lottery in Figure A8), high household bargaining power (measured using questions on who decides what in a household), and strong claims from others on own liquid assets. None of the additional control variables predict WR adoption (Table A5). Considering a 94 percent female sample population, this combination of factors is reminiscent of Anderson and Baland (2002), who argue that Kenyan women use commitment devices to protect their savings from intra-household conflicts.[48] Compared to IS, WR was indeed better suited as a safeguarding device: Its withdrawal restriction prevented other

[47] Section 5.2 describes this measure. Clients who defaulted may have had a stronger incentive to report shocks, in order to preserve their self-image or reputation. This would bias the shock coefficient upwards. However, the endline survey was framed as coming from a research organisation, with no direct link to the bank or the IS account. Note that attrition in the endline survey was compensated by imputing the median shock value for those who did not participate.

[48] The estimated linear relationship of commitment take-up with household bargaining power is unable to capture Anderson and Baland (2002)'s inverted U-shape. However, both household bargaining power and female education may be associated with an increased autonomy of the woman in planning to build up savings of her own.

household members from accessing savings, while preserving the woman's flexibility in when to make deposits. In contrast, IS allowed withdrawals at any time, albeit at the cost of defaulting on the account. A reservation must be made with respect to statistical power: Group WR is half the size of group IS, reducing the precision of estimates.

5.4. Welfare

In the theoretical framework outlined in Section 2, welfare implications are clear-cut: Given a simple form of partial sophistication, and a commitment contract where penalty choice is continuous and success is binary, agents are strictly worse off from adopting commitment (but may do so anyway). Even if this mechanism is the correct one, welfare inference is much less obvious empirically: Individuals may benefit from increased savings even if they default later, just like they may smoke less in Giné, Karlan and Zinman (2010) or drink less in Schilbach (2015) despite failing to reach the binary threshold for successful commitment.

A few careful inferences can be made. First, it is informative to look at the ratio of IS deposits to incurred penalties: Out of 63 defaulting clients, 31 clients lost their entire IS balance to the penalty, leaving them with no commitment savings and the penalty loss.[49] For these clients, the IS contract likely reduced welfare. For an additional 9 clients, the incurred penalty exceeded half of their IS balance, equivalent to a 100 percent premium on savings. To compare, the cost of informal loans in the area is between 10 and 20 percent per month. Figure A2 shows the full distribution of the ratio. Second, a case can be made that offering commitment still constituted a Kaldor-Hicks improvement: The 51 successful clients (the "winners" from offering commitment) should be willing to compensate the 63 defaulting clients ("losers") if the gain from reaching their savings goal (captured as $b-p$ in the model from Section 2) exceeds the loss to the defaulters. Splitting the penalties incurred by the 63 defaulters (P11,370) across the 51 successful clients, this is the case if $b-p$ exceeds P223, or 8.2 percent of the average (winners) savings goal. Third, and relatedly, offering IS increased the fraction of people who purchased their savings goal without borrowing by 22 percent (8 pp, see Table 3).

All these considerations are ex-post. A cautious estimate of the frequency of 'ex-ante mistakes' is provided by the pre-order results: 55 percent of all clients (71 percent of defaulting clients and 35 percent of successful clients) chose not to order IS again (see Table A2). An important caveat is that this ignores time effects, i.e., not needing IS again after having saved successfully. Section 7 discusses policy implications for offering commitment, and points to a possible trade-off between achieving large benefits on *average*, and substantial heterogeneity with welfare risks.

6. Alternative Explanations

So far, the paper has focused on time-inconsistency and partial sophistication in attempting to explain why individuals may adopt and subsequently default on commitment contracts. Section 5.2 further argued that the timing of defaults is hard to reconcile with shocks to income or expenditures, at least provided that individuals had correct beliefs about the probability of such shocks. However, numerous other factors could have triggered a wave of defaults shortly after adopting commitment. Aggregate shocks may have affected the study region, or individuals may have been too optimistic about their future income. Another range of explanations can be summarized as

[49] While 35 clients defaulted after the opening balance, only 28 had chosen penalties equal to or greater than the opening balance (see Section 3.2 for design details). 3 clients made additional deposits, but lost them to higher penalties.

'costly experimentation': Individuals were learning what it involves to save each week. How hard would it be to extract 150 pesos from their household every week? How much time will it take to walk to the bank, and how difficult will it be to remember to make installments on time? Finally, how much discipline would it take to curb consumption? While aspects of these questions may be product-specific, they all relate to more general preference and household parameters, as well as prior experiences. How prone is the individual to experience conflict within the household? How far away do they live from the bank, and how constrained is their attention? How much did they struggle to curb consumption, for instance when paying for previous loan installments?

This section will address some main concerns relating to costly product experimentation, specifically, household conflict and transaction costs. It further discusses the evidence for aggregate shocks, income optimism, and persuasion by the bank marketers. As a rule, it is not difficult to think of reasons why individuals may have defaulted once they had entered commitment contracts. However, few of these reasons parsimoniously explain why individuals would demand commitment in the first place. Further, unless individuals were fully naive about potential challenges of saving, measures associated with these challenges should predict selection into (or out of) the contracts. For instance, those with very high shock arrival rates are unlikely to be among the adopters. Table 6 reviews the consistency of the study's findings with a number of alternative explanations.

6.1. Aggregate Shocks

An aggregate shock around the time of account opening may explain the large wave of defaults. The Philippines is a well-known area for earthquakes and tropical storms, and had recently been hit by tropical storm Washi in December 2011, causing 1,268 casualties.[50] The risk of such shocks was thus well-known at the time of marketing in September 2012, possibly affecting take-up rates. Indeed, tropical storm Bopha hit the Mindanao region between December 2 and December 9, 2012. Fortunately, storm Bopha did not cause flash flooding, and the main effect on the study location was a six-day power outage. While this may have affected large businesses, power outages of several hours each day were common in the study area even before the storm, and provisions against power outages were widespread. Because of its limited effect on the area, storm Bopha was not locally classified as a natural disaster (which would have invoked both IS's and WR's emergency provisions). In the endline survey, 20.5 percent of the sample (20.4 percent of defaulting IS clients) reported some damage to their house or crops. Conditional on non-zero damage, the median damage value was 1400 pesos (U.S. $33). Only 3 out of 732 respondents reported to have lost income due to the power outages.

While some negative effects of the storm cannot be ruled out, the timing of the storm does not match the timing of the defaults: The IS accounts were opened between 20 September and 28 October. Out of 63 defaults, 35 made no further deposit after their opening balance, resulting in contract default upon the third missed deposit, usually three weeks later. By December 2, 41 IS accounts were officially in default. No bunching of defaults occured at the time of the storm.

[50] Statistics from the Philippine National Disaster Risk Reduction and Management Council (NDRRMC).

TABLE 6. FINDINGS AND ALTERNATIVE EXPLANATIONS

				Costly Experimentation				
	Aggregate Shocks	**Hetero-geneous shock rates**	**Income Optimism**	**Limited At-tention/ Trans-action Costs**	**Household Conflicts**	**Time-Inconsistency + Partial Sophistication**	**Time-Inconsistency + Full Sophistication**	**Marketer Persuasion**
27% Commitment Take-Up (IS)					Yes, if used to safeguard	Yes	Yes	Yes, but limited by 2-step account opening process
High Average Treatment Effect on Bank Savings					No, predicts higher savings under W-account	Needs full sophistication for part of the population	Yes	
55% Default, incl. 31% immediately after opening	Yes, given shock during account opening time	Ambiguous: High λs discourage take-up	Yes, upon realizing the true income	Ambiguous: High t or scarce attention discourage take-up	Yes, if conflict unexpected	Yes	No, 'rational defaults' only	Yes
Associated Measure Predicts Take-Up?	No evidence of an aggregate shock during account opening.	No, using ex-ante measured λ	Yes, contrary to prediction	No, using distance to bank	No, using HH bargaining power	Yes, sophistication negatively predicts take-up, conditional on β		No, cognitive ability *positively* predicts take-up
Associated Measure Predicts Default?		Yes, using ex-post measured λ. But λs quantitatively too small.	No, using average (predicted–real) income	No, using distance to bank	Yes, using HH bargaining power, but only within take-up sample	Yes, sophistication negatively predicts default, conditional on β. Observed time-inconsistency positively predicts default.		No, cognitive ability *positively* predicts default

Cells in this table summarize whether a given finding would be expected under a given explanation. The design of this table is inspired by DellaVigna and Malmendier (2006).

6.2. Heterogeneous Shock Arrival

Section 5.2 argues that observed default timing is difficult to reconcile with i.i.d. shocks, but individuals may have had heterogeneous shock arrival rates, or the shocks may have been correlated. Specifically, heterogeneous shock frequency rates λ_i would generate a declining rate of defaults, as those with the highest λ_i drop out first. However, a default rate of 31 percent after the first installment requires a substantial fraction of the population to have λs well in excess of 0.31 shocks per week. The maximum observed number of emergencies for an individual was 0.19 per week (0.12 among adopters). Even if this survey measure is flawed, and true λs are much higher, theory predicts that individuals with high λs are likely to stay away from commitment (see Section 2, Proposition 3). The data do not confirm this prediction: An ex-ante measure of individuals' shock frequency from the baseline survey is negatively, but not significantly, related to commitment adoption (Table 4). It is possible that individuals were naive about their latent proneness to shocks. An example of biased beliefs is that one's λ_i corresponds to the average shock frequency $\bar{\lambda}$ in the population. While this would explain a bulk of defaults soon after opening, it does require $\bar{\lambda} = 0.31$, far beyond the elicited frequencies.

6.3. Income Optimism

Following Browning and Tobacman (2007), the link between time-inconsistency in MPL questions and IS default incidence could have been caused by overoptimistic beliefs about future income: If individuals expect their future income to be higher than their current income, they may select the smaller, sooner reward in the 'now vs. 1 month' frame, but choose the larger, later reward in the '1 month vs. 2 months' frame (see Section 3.3). As a result, they would be falsely classified as present-biased.[51] Income optimism could also explain default incidence if it caused individuals to commit to overly ambitious savings plans.

I measure income optimism at the group level by comparing predicted and realized incomes: During the baseline survey, individuals were asked to predict their weekly household income during each of the next six months. Six months later, during the endline survey, individuals were asked to state their realized weekly income for the same time period. At the individual level, a gap between predicted and realized income is not sufficient to indicate systematically biased beliefs, but may simply reflect a bad draw from the income distribution. However, given correct beliefs, the law of large numbers implies that individuals should correctly predict their income *on average*. If the utilized MPL questions capture income optimism rather than time-inconsistency, then individuals classified as 'present-biased' should have higher income prediction gaps. Furthermore, if defaults were caused by systematic mispredictions of future income, then defaulting clients should have higher prediction gaps than successful clients.

The data suggest that moderate income optimism is common across the sample (Table A7 reports full details). However, the average prediction gap is not higher for individuals classified as present-biased. Furthermore, individuals who defaulted on IS did not mispredict their income significantly more than successful IS clients. Surprisingly, the average prediction gap is significantly higher for IS adopters than for non-adopters. This is counterintuitive, as income optimism does not predict a demand for commitment.

[51] A similar pattern could emerge in the presence of liquidity constraints: The MPL questions may capture a desire to borrow from future income, rather than present bias. Implications for commitment demand, default, and combinations with sophisticated hyperbolic discounting are similar to those discussed below for income optimism. Since I do not have an adequate measure of liquidity constraints, I focus on income optimism below.

An explanation that would reconcile most of the evidence is a combination of sophisticated time-inconsistency and income optimism: In the model from Section 2, suppose that the agent believes, counterfactually, that her future per-period income is $y' > 1$ instead of $y = 1$. She believes savings goals up to $p \leq 2y'$ to be achievable. Being sophisticated about her time-inconsistency, she leaves her period 2 self to pay y', and commits her period 1 self to save $p - y'$. Default occurs for $p - y' > 1$, which is never affordable given the true $y = 1$. Note that a model with only fully naive and fully sophisticated agents would predict a positive association between sophistication measures and commitment adoption, and no association between sophistication and default (since all adopters are sophisticated).

Summing up, income optimism *alone* does not explain why individuals demand commitment, nor correlates with observed measures of present bias. A combination of fully sophisticated time-inconsistency and income optimism predicts both a demand for commitment and subsequent default. However, this combination stands in contrast with the observed negative association of sophistication measures (which are robust to income optimism) with take-up and default. The latter could be explained by combining income optimism with partially sophisticated time-inconsistency.

6.4. Limited Attention and Transaction Costs

Clients may have simply forgotten to make their weekly deposits, or may have been deterred from going to the bank by high transaction costs. Indeed, the IS account came with non-negligible transaction and attention costs, and was clearly marketed as such: Clients were presented with an explicit savings plan including due dates for each week (Figure A1), and given the instruction to physically deposit their installments at the bank. Most respondents received their income in cash, and bank transfers were uncommon. The evidence indicates that transaction costs were well-understood: Being 'too busy to go to the bank' was a common reported reason among those who chose not to adopt the IS product.

It is possible that clients were aware of the costs involved, but naive present-biased about their willingness to incur them. This explanation is easily incorporated in the model: A transaction or attention cost c for making the deposit $p - 1$ is equivalent to a higher price p, which makes the nondivisible good less attractive.[52] As long as c is anticipated, all conclusions of the model go through.

Alternatively, naiveté about attention or transaction costs itself may have induced defaults. Suppose that clients fail to anticipate the transaction cost c, or the fact that their attention constraint is binding (e.g. Banerjee and Mullainathan (2008)). By itself, this can explain default even with time-consistent preferences, as an agent may decide to give up on the savings contract if $c > D$ (she may still decide to save at home by herself). However, a time-consistent agent is unlikely to tie herself up in a costly commitment in the first place.

Analogue to income optimism, a combination of sophisticated time-inconsistency and naiveté about transaction costs or attention could explain both commitment adoption and default: Restricting the model in Section 2 to full sophistication, the penalty D_{min} is chosen to make period 1's incentive constraint (equation 1) hold with equality. A small unforeseen cost c thus suffices to violate incentive compatibility, and trigger contract default. If people are truly naive about transaction costs, a proxy for transaction costs should predict default. Empirically, 'distance to the bank branch' does not predict default (measured using GPS coordinates, see Table A5). As with income optimism, explanations based on full sophistication (counterfactually) predict a positive association between sophistication measures and commitment adoption, and no association between sophistication and default.

[52] Strictly, c differs from higher p in the case of a shock in period 2: Savings $p - 1$ can still be consumed, while c is lost.

6.5. Household Conflicts

As discussed in Section 5.3, a desire to safeguard savings from other household members is the leading alternative hypothesis in explaining a demand for commitment savings devices. The installment-savings account did not explicitly restrict withdrawals, but it imposed a cost on them: Individuals could choose to terminate their account and withdraw their savings (less the default penalty) at any time. Arguably, the withdrawal-restriction account was more suitable as a pure safeguarding device. Consistent with this conjecture, measures of 'other-control problems' predict WR take-up but not IS take-up (Table 4). Furthermore, if safeguarding is the driving motive for commitment, the treatment effect on bank savings is likely to be larger under WR than under IS, due to the account's stronger withdrawal restrictions. The opposite holds true in the data (Table 3).

Independent of safeguarding motives, household conflicts may be responsible for the observed defaults: Other household members may not agree with the weekly extraction of household resources. Indeed, Section 5.3 shows that defaults are higher among those with low household bargaining power (defined in Appendix V). As previously argued for other default risks, the anticipation of household conflict should determine selection into commitment. It may also predict a higher optimal default penalty, which can be used to justify the need to honour the commitment to one's family. Neither relationship finds empirical support: There is no correlation of household bargaining power with IS take-up (Table 4), and a near-zero correlation with the penalty choice ($\rho = -0.056$). While household conflicts may have caused some of the defaults, the data suggests that these conflicts were unexpected.

6.6. Marketer Persuasion

The bank marketers received a fixed daily wage (roughly three times the local average) in addition to a small commission. This raises the possibility of persuasion by the marketers - which may explain both commitment take-up and subsequent default. While persuasion cannot be ruled out, three factors worked to contain it: First, marketers were employed and trained exclusively by the research team.[53] A script detailed every aspect of the conversation with a client, and auditors ensured that this script was followed. Strong emphasis was placed on clear explanations of the product features, and clients were encouraged to make sensible, conservative savings plans which were suitable for their income and usual expenditures. Second, opening an IS or WR account was a two-step process: After signing the contract, individuals had a one-week 'cool-down' period before the marketer returned to collect the opening balance. Those who signed the contract but failed to deposit the opening balance (see footnote 28) are considered non-adopters, thus filtering those most likely to have been momentarily swayed. Third, IS take-up is positively predicted by cognitive ability (measured using Raven's matrices, see Table A5 and Figure A9). This is reassuring: The IS product is more complex than traditional savings accounts (though not more complex than a loan contract). A plausible interpretation for a positive relationship with cognitive ability is that those who struggled to understand the rules of an unfamiliar product were more likely to stay away. If clients were manipulated and misled into a product they do not understand, the opposite would be expected.

[53] With permission from the bank, our marketers wore official bank uniforms and IDs. To ensure that participants did not know this was an experiment, the marketers did not reveal any connection to the research organization.

7. Discussion

Commitment devices are receiving substantial attention both in the academic literature and in the public eye, and are generally portrayed as a promising way to overcome intrapersonal conflict. In the context of an installment-savings commitment product in the Philippines, I present evidence that people may fail at choosing commitment contracts which are suitable for their preferences. I argue theoretically that an individual's ability to correctly choose a welfare-improving commitment contract depends on her degree of sophistication about her own time-inconsistency. I observe that a majority of individuals who take up an installment-savings commitment product choose very low stakes for this commitment, and then default on it. Both take-up and default decisions are systematically linked to low measures of sophistication, suggesting that individuals may underestimate the amount of commitment required to make them save. Other types of forecast errors (about attention constraints, household conflicts, or future income) may also have played a role, with similar policy implications. Choosing unsuitable commitment devices may lead to welfare losses – in the current setting, many clients incurred monetary penalties without building up any savings. This mechanism potentially extends to rich country applications such as gym contracts (as suggested by DellaVigna and Malmendier (2006)), diet clubs, and long-term pension savings plans.

From a policy perspective, there is no simple solution. One could offer commitment contracts exclusively with high penalties to ensure incentive-compatibility. However, this would deter commitment adoption – both by individuals who need higher penalties but fail to realize it, and by individuals who genuinely require only small penalties. Alternatively, one could legally mandate commitment savings, as prominently done in state pension schemes in Germany and Australia. In most contexts, this is unlikely to be optimal, as it requires the social planner to have a large amount of information on individual preferences.

The arguments in this paper focus on commitment contracts which impose penalties on binary consumption decisions. Other forms of commitment exist, with potentially different welfare implications. First, it may be possible to condition penalties on continuous events. For instance, choosing a social network full of health-conscious friends may expose an agent to increasing amounts of shame (or unsolicited health advice) the more weight she gains, or the more cigarettes she smokes. When penalties depend on continuous events, partial sophisticates would still choose commitments with insufficient rates of punishment. However, the marginal nature of the involved trade-offs imply that she would move closer to the optimal behaviour, and likely be better off than without commitment.

Second, some commitments work entirely without penalties, by directly eliminating undesirable options from the agent's choice menu. An example is the date-based withdrawal-restriction account featured in this study: Early withdrawals were simply disallowed. All savings were returned on the goal date, regardless of how much the agent had saved.[54] A different example is the decision to go shopping with a limited amount of cash, and no credit cards. At first sight, such commitments appear like a safe choice for partial sophisticates. Two qualifications are necessary: First, it can be hard to actually *eliminate* a bad option from the choice set, as opposed to just increasing its cost. Heidhues and Kőszegi (2009) discuss how individuals find costly ways to overcome the barriers they have set up for themselves. For instance, choosing not to buy alcohol during the weekly grocery shop may result in a last-minute trip to the store in the middle of one's favorite TV show, following a craving for beer. Second, the choices that *can* be elim-

[54] In contrast, amount-based WR accounts imposed a penalty on failing to save by freezing savings indefinitely.

inated may not be sufficient to enforce desirable behaviour. In this study, the WR account restricted withdrawals, but it imposed no pressure to make any savings deposits after opening the account – indeed, 79 percent of clients did not.

The comparison between the installment-savings product and the withdrawal-restriction product may point to a more general trade-off between hard and soft commitments: Offering stronger commitments with more pressure may provide greater benefits on *average* – as observed by a threefold treatment effect of the installment-savings product on average bank savings. However, stronger commitments may imply an increased risk of adverse welfare effects. Softer commitments may be limited in both their risks and their efficacy. Further research is needed to identify ways of overcoming time-inconsistency, which are both effective and at the same time 'safe' for partially sophisticated time-inconsistent agents.

References

Acland, Dan, and Matthew R. Levy. 2015. "Naiveté, Projection Bias, and Habit Formation in Gym Attendance." *Management Science*, 61(1): 146–160.

Afzal, Uzma, Giovanna d'Adda, Marcel Fafchamps, Simon Quinn, and Farah Said. forthcoming. "Two Sides of the Same Rupee? Comparing Demand for Microcredit and Microsaving in a Framed Field Experiment in Rural Pakistan." *Economic Journal*.

Ali, S Nageeb. 2011. "Learning Self-Control." *Quarterly Journal of Economics*, 126(2): 857–893.

Al-Ubaydli, Omar, and John A List. 2015. "On the Generalizability of Experimental Results in Economics." In *Handbook of Experimental Economic Methodology*. , ed. Guillaume R Fréchette and Andrew Schotter. Oxford University Press.

Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. "Commitment vs. Flexibility." *Econometrica*, 74(2): 365–396.

Ameriks, John, Andrew Caplin, John Leahy, and Tom Tyler. 2007. "Measuring Self-Control Problems." *American Economic Review*, 97(3): 966–972.

Ananth, Bindu, Dean Karlan, and Sendhil Mullainathan. 2007. "Microentrepreneurs and Their Money: Three Anomalies." Financial Access Initiative Working Paper.

Anderson, Siwan, and Jean-Marie Baland. 2002. "The Economics of Roscas and Intrahousehold Resource Allocation." *Quarterly Journal of Economics*, 117(3): 963–995.

Andreoni, James, and Charles Sprenger. 2012. "Estimating Time Preferences from Convex Budgets." *American Economic Review*, 102(7): 3333–3356.

Ariely, Dan, and Klaus Wertenbroch. 2002. "Procrastination, Deadlines, and Performance: Self-control by Precommitment." *Psychological Science*, 13(3): 219–224.

Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *Quarterly Journal of Economics*, 121(2): 635–672.

Augenblick, Ned, and Matthew Rabin. 2015. "An Experiment on Time Preference and Misprediction in Unpleasant Tasks." Working Paper.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger. 2015. "Working Over Time: Dynamic Inconsistency in Real Effort Tasks." *The Quarterly Journal of Economics*, 130(3): 1067–1115.

Banerjee, Abhijit V., and Sendhil Mullainathan. 2008. "Limited Attention and Income Distribution." *American Economic Review*, 98(2): 489–493.

Basu, Karna. 2014. "Commitment Savings in Informal Banking Markets." *Journal of Development Economics*, 107: 97–111.

Bauer, Michal, Julie Chytilova, and Jonathan Morduch. 2012. "Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural India." *American Economic Review*, 102(2): 1118–1139.

Benartzi, Shlomo, and Richard H. Thaler. 2004. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy*, 112(1): 164–187.

Besley, Timothy. 1995. "Savings, Credit and Insurance." In *Handbook of Development Economics*. Vol. 3A, , ed. Jere Behrman and T.N. Srinivasan. Elsevier.

Browning, Martin, and Jeremy Tobacman. 2007. "Discounting and Optimism Equivalences." The Wharton School, University of Pennsylvania Working Paper.

Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang. 2016. "Facilitating Savings for Agriculture: Field Experimental Evidence from Malawi." *Economic Development and Cultural Change*, 64(2): 187–220.

Bryan, Gharad, Dean Karlan, and Scott Nelson. 2010. "Commitment Devices." *Annual Review of Economics*, 2(1): 671–698.

DellaVigna, Stefano, and Ulrike Malmendier. 2004. "Contract Design and Self-Control: Theory and Evidence." *The Quarterly Journal of Economics*, 119(2): 353–402.

DellaVigna, Stefano, and Ulrike Malmendier. 2006. "Paying Not To Go To The Gym." *The American Economic Review*, 96(3): 694–719.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *The American Economic Review*, 101(6): 2350–90.

Dupas, Pascaline, and Jonathan Robinson. 2013. "Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics*, 5(1): 163–192.

Eliaz, Kfir, and Ran Spiegler. 2006. "Contracting with Diversely Naive Agents." *The Review of Economic Studies*, 73(3): 689–714.

Fischer, Greg, and Maitreesh Ghatak. 2010. "Repayment Frequency in Microfinance Contracts with Present-Biased Borrowers." London School of Economics EOPP Discussion Papers 21.

Fudenberg, Drew, and David K Levine. 2006. "A Dual-Self Model of Impulse Control." *The American Economic Review*, 96(5): 1449–1476.

Giné, Xavier, Dean Karlan, and Jonathan Zinman. 2010. "Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation." *American Economic Journal: Applied Economics*, 2(4): 213–235.

Gugerty, Mary Kay. 2007. "You Can't Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya." *Economic Development and Cultural Change*, 55(2): 251–282.

Gul, Faruk, and Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica*, 69(6): 1403–1435.

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

Heidhues, Paul, and Botond Kőszegi. 2009. "Futile Attempts at Self-Control." *Journal of the European Economic Association*, 7(2-3): 423–434.

Heidhues, Paul, and Botond Kőszegi. 2010. "Exploiting Naivete About Self-Control in the Credit Market." *American Economic Review*, 100(5): 2279–2303.

John, Anett. 2017. "Can Fixed Regular Deposits Overcome Savings Constraints?" *Revue Économique*, 68(5): 909–924.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. 2015. "Self-Control at Work: Evidence from a Field Experiment." *Journal of Political Economy*, 123(6).

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 112(2): 443–478.

List, John A, Azeem M Shaikh, and Yang Xu. 2016. "Multiple Hypothesis Testing in Experimental Economics." National Bureau of Economic Research Working Paper.

Maniadis, Zacharias, Fabio Tufano, and John A List. 2014. "One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects." *American Economic Review*, 104(1): 277–290.

O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review*, 89(1): 103–124.

Rutherford, Stuart. 2000. *The Poor and Their Money.* Oxford:Oxford University Press.

Schilbach, Frank. 2015. "Alcohol and Self-Control: A Field Experiment in India." Working Paper.

Skiba, Paige Marta, and Jeremy Tobacman. 2008. "Payday Loans, Uncertainty and Discounting: Explaining Patterns of Borrowing, Repayment, and Default." Vanderbilt Law and Economics Research Paper 08-33.

Toussaert, Séverine. forthcoming. "Eliciting Temptation and Self-Control Through Menu Choices: A Lab Experiment." *Econometrica*.