REACTIVE INCENTIVES: HARNESSING THE IMPACT OF SUNK OPPORTUNITY COSTS*

Ian Ayres[*] & Giuseppe Dari-Mattiacci[**]

ABSTRACT

Sunk opportunity costs can causally affect subsequent behavior. Turning down initial temptations might make it easier for people to stay committed in the longer term to a personal goal. We discuss several competing explanations for this enhanced commitment effect – including that resisting temptations might (1)"self-signal" information about individual's own resolve or (2) create prospective risk of cognitive dissonance. Individuals who are subject to what we call "reactive incentives" have superficial incentives to accept a temptation and deviate from some pre-existing goal. But the process of resisting the temptation may cause these individuals to be less likely to subsequently deviate from that goal. We identify two broad categories of reactive incentives: reactive carrots and reactive sticks.

This article tests the causal impact of reactive incentives in a field experiment at a University of Amsterdam gym where new subscribers were randomly assigned to one of several groups. "Temptation" group members were made one-time monetary offers of varying size to quit the gym. None of the temptation group subjects accepted the offered compensation to quit the gym. Subjects who were offered a moderate reactive carrot (a cash payment equaling 110% of their initial subscription price to quit) were statistically more likely to visit the gym, to resubscribe, and to earn higher post-treatment grades. Even though the foregone financial opportunities are sunk, they nonetheless affect our subjects' future behavior. Consistent with the reactive stick hypothesis, we also observe that new subscribers who were exposed to unpleasant weather conditions during their first week of membership were more likely to subsequently visit the gym than new subscribers who did not have to initially confront the reactive stick of bad weather.

---

[*] We are deeply indebted to Maurice Maas at USC and Wouter ter Haar at the University of Amsterdam for their continual support during the experiment. Without them this study would have not been possible. We are thankful to the USC director Theo van Uden for allowing us to run the experiment. We would also like to thank Hellen Volmerink for her assistance in locating data, and Ryann Bubb, Daniel Chen, Jacob Golden, Julian Jamison, Enrico Perotti, Joël van der Weele, Avani Metha Soot, Tess Wilkinson-Ryan and participants in seminars at Tel Aviv University and Leiden University, and the 2017 annual meeting of the European Association for Law & Economics at the University of Liverpool for helpful suggestions. Anthony Cozart, Samuel Dong, Tyler Hill, Pranav Bhandarkar, and Zachary Shelley provided excellent research assistance.

[*] Yale Law School.

[**] Columbia Law School.

This article explores a range of settings where policy-makers might deploy reactive incentives to beneficially enhance self-control– including smoking cessation programs, payday lending, college completion and retirement savings.

INTRODUCTION 3

I. THEORY 5
A. Active versus Reactive Incentives 5
B. Theories of Reactive Incentives 7
1. Gym Screening, Gym Signaling & Third-Party Signaling 7
2. Reactance 8
3. Willpower and Habit Training 9
4. Self-Signaling 10
5. Taste-Based Theories 11
6. Appropriate Contexts 13

II. THE EXPERIMENT 16
A. Experimental Design 16
B. First Wave Results 19
1. Blocking, Balance and Descriptive Statistics 19
2. Treatment Effects 20
C. Second Wave Results 23
D. Using Inclement Weather to Test for Reactive Stick Effects 25

III. POLICY IMPLEMENTATION 26
A. Money-Back Guarantees 27
B. Smoking, Dieting, and Gambling 27
C. Borrowing 30
1. A Possible Critique of Prepayment Penalty Prohibition 30
2. Avoiding Delayed Repayments 31
3. Deterring Equity Stripping 32
D. Saving 33
E. College Completion 35

CONCLUSION 36

INTRODUCTION

One of the most famous exemplars of heroic persuasion in theater is the St. Crispin's Day speech in Shakespeare's *Henry V*. Generations of students learn the lines ("We happy few. We band of brothers.") and analyze how Henry encourages his soldiers to envision a future when they will proudly look back on this day of honor. But less emphasized is how the King's exhortation begins with a perverse financial incentive. When he overhears a lieutenant bemoaning the fact that the English troops are outnumbered 5 to 1, Henry counters:

> Do not wish one more [solider]!
> Rather proclaim it, Westmoreland, through my host
> That he which hath no stomach to this fight,
> Let him depart. His passport shall be made,
> And crowns for convoy put into his purse.

Instead of threatening deserters with execution, Henry offers his soldiers safe passage and money to boot if only they will leave. Instead of using the traditional carrot of a reward if you stay, Henry dangles what might be called a reactive carrot, a financial incentive ("crowns for convoy") to quit.

Henry explains that he would not want to die in the company of any man that would accept such an offer. This explanation for the offer, on its face, sounds like a form of screening.[1] Henry only wants to fight with a certain type of soldier and wants to use the offer to screen out any bad types. But this screening explanation would only work if some soldiers took him up on the offer. No one does.

A more satisfying explanation is that the offer helped steel his soldiers' resolve. Each soldier by resisting Henry's temptation learned something about themselves. They learned that they were not the kind of person who would quit and run for a few crowns. By offering money to quit before battle, Henry's ploy may reduce the chance that soldiers would quit during battle. Resisting the temptation might have played a role in creating Henry's band of brothers.[2]

---

[1] See *infra* at 7 for a discussion of screening theories.

[2] Shakespeare dramatizes a similar strategy in Act I of *Richard III* when Richard, after helping to kill Prince Edward of Lancaster, audaciously courts his widow, Lady Anne. Richard offers to let her kill him with his dagger if she will not succumb to his entreaties ("Take up the sword again or take me up."). Lady Anne, in resisting the temptation to avenge her husband's murder becomes, by degrees, more amenable to Richard's blandishments and by the end of the encounter agrees to wear Richard's ring. Afterward, Richard remarks on the unusual nature of his persuasion:

> Was ever woman in this humour woo'd?
> Was ever woman in this humour won?

The Internet shoe behemoth, Zappos, may not have realized it, but for years it emulated the St. Crispin's Day incentive. After initial training, new employees were offered money to quit. New employees were ushered into a room and made a one-time offer to end their employment in exchange for $2,000. Most employers are trying to reduce turnover, but Zappos offers employees to walk away from their jobs more quickly. Like Henry, the leader of Zappos defended the policy as making sure that the company employs only people who really want to work there.[3] But the deeper explanation might again be the power of resisted temptation. Almost everyone turns down the initial offer to quit,[4] and once you turn down the money, you'd feel like a fool to quit a month later for nothing. So perversely, offering employees/soldiers money to quit, may keep them on the job longer.

It is a scary offer for an employer to make. If many of Henry's soldiers had accepted the offer, Henry would have died at Agincourt. If Zappos employees take the bribe, Zappos would lose a bundle. It is therefore not surprising that more than a decade ago when Zappos first started tempting new recruits to quit, they only offered $100. Only after learning that the vast majority resisted the $100 offer did Zappos start raising the size of the temptation.

Since Amazon's purchase of Zappos in 2009, the Pay to Quit program has been expanded. The offer is now made annually to all Amazon employees and the amount ranges up to $5000 (accompanied by a memo from CEO Jeff Bezos with the subject line "Please Don't Take This Offer").[5] If the fictional Henry had known what Zappos knows now, he might have offered his happy few even more money.

This article explores the power of resisted temptation. To our knowledge, we are the first to test whether sunk opportunity costs can causally affect subsequent behavior.[6] Other studies have tested whether foregone opportunities can impact initial decision-making.[7] But we offer evidence that foregone opportunities can change the choices that people make in the future, and provide theoretical explanations for such behavior. Turning down initial temptations might make it easier for people to stay committed to a goal in the longer term. We discuss several competing

---

[3] Rachel Emma Silverman, *At Zappos, Some Employees Find Offer to Leave Too Good to Refuse*, WALL STREET JOURNAL (May 7, 2015), *available at* https://www.wsj.com/articles/at-zappos-some-employees-find-offer-to-leave-too-good-to-refuse-1431047917.

[4] McGregor, Jena, *Why More Companies Should Pay Employees To Quit*, WASHINGTON POST (April 14, 2014), *available at* https://www.washingtonpost.com/news/on-leadership/wp/2014/04/14/why-more-companies-should-pay-employees-to-quit/.

[5] Ruth Umoh, *Why Amazon Pays Employees $5,000 To Quit*, CNBC (May 21, 2018), *available at* https://www.cnbc.com/2018/05/21/why-amazon-pays-employees-5000-to-quit.html

[6] Opportunity costs have been defined as "the evaluation placed on the most highly valued of the rejected alternatives or opportunities". James M. Buchanan, *Opportunity Cost*, *in* THE NEW PALGRAVE: THE WORLD OF ECONOMICS 520 (John Eatwell, et al., ed. 1991). See also Billy Joel, *Only the Good Die Young* ("the price that you pay for things that you might have done").

[7] For example, wrong choice aversion can cause people to violate Arrow's "independence of irrelevant alternatives" assumption. Adding a similar but clearly inferior choice into a menu of alternatives can drive decision makers to choose the dominating alternative. Paramesh Ray, *Independence of Irrelevant Alternatives*, 41 ECONOMETRICA 987 (1973); Ian Ayres, *Voluntary Taxation and Beyond - the Promise of Social-Contracting Voting Mechanisms*, 19 AM. LAW & ECON. REVIEW 1 (2016).

explanations for this enhanced commitment effect – including that resisting temptations might (1) "self-signal" information about individual's own resolve or (2) create prospective risk of cognitive dissonance. People who are subject to what we call "reactive incentives" have superficial incentives to accept a temptation and deviate from some pre-existing goal. But the process of resisting the temptation may cause the people to be less likely to subsequently deviate from the goal. We identify two broad categories of reactive incentives: reactive carrots and reactive sticks. People who resist reactive carrots and sticks and succeed at not deviating from a chosen path may find it easier to stick to their goals in the longer run.

This article rigorously tests the causal impact of reactive incentives in a field experiment at a University of Amsterdam gym. We conducted the experiment in two waves – first, from September 5th, 2015 to September 15th, 2016, and then, on a second set of subjects from September 15th, 2016 to July 15th, 2017. In each waive, new users were randomly assigned to one of three groups – two of which were offered varying one-time temptations to quit the gym. None of the tempted users accepted the offered temptation. We find evidence that users that received a bonus temptation of 110% of their initial subscription were more likely to go to the gym and had higher post treatment grades than a control group of subjects.

The remainder of this paper is divided into three parts. Part I sets out our theory of reactive carrots and reactive sticks, and relates this theory to prior literature. Part II describes our reactive carrot experiment design and its results, as well as our observational estimates of bad weather as a reactive stick. Finally, Part III discusses a host of other arenas where reactive incentives might beneficially change behavior – including smoking cessation, retirement savings and payday loans.

I. THEORY

A. Active versus Reactive Incentives

The potential power of resisted temptation comes from subjects not just resisting an initial incentive, but from reacting to it by changing some aspect of their future behavior. Importantly, resisted temptations can at times induce subjects to act directly opposite to what is being superficially incentivized. This article's core experiment concerns what we call a "reactive carrot" incentive, which tests whether resisting a contingent reward can induce this opposite behavior. The membership refund offer superficially incentivizes going to the gym less often. We show, however, that this contingent reward can have a "reactive" effect inducing subjects to go to the gym more often. Reactive carrots are used with an expectation that subjects will not only resist the temptation (and act as if they never received the offer), but that the subjects will quit later than they would have if the reactive carrot had not been dangled.

The power of resisted temptation also applies to disincentives. Defying the superficial incentive of a contingent punishment might produce an analogous kind of reactive effect. Like reactive carrots, reactive sticks that provide a superficial incentive for one behavior might, if resisted, cause a subject to do more of the opposite behavior. For example, consider a new gym user who confronts unexpectedly unpleasant weather in the first week of gym membership, which makes it more difficult to travel to the gym. On its face, the unpleasant weather is a short-term deterrent to using the gym – and hence a stick-like disincentive for gym use. However, if the user

resists this disincentive and nonetheless goes to the gym, it is possible that the user will be more likely to continue to go to the gym in the future (compared to the counterfactual where they had not been initially exposed to the bad weather). As with the reactive carrot refund offers, the reactive stick weather disincentives provides the users with an opportunity to learn something about themselves and establish a self-reputation that might make it easier for them to continue gym use in the future. Having gone to the gym in inclement conditions may make it easier to go in the future. Incurring the sunk cost of traveling to the gym during unpleasant weather might improve impacted subjects' willingness to go to the gym in the future. Part III provides suggestive evidence of just this reactive stick effect. We find that new subscribers who experience unexpected inclement weather during the first week of their subscription are more likely than other subscribers to continue using the gym.

Military boot camps and hazing rituals of fraternities may also have this reactive-stick quality. Unpleasant initiation requirements that artificially increase the difficulty of joining might reduce the likelihood of quitting. Reactive sticks exemplify the power of sunk opportunity costs. A new fraternity member might feel silly if he went through all the hazing (declining the opportunity to quit during the initiation period) only to quit a month later.

The possibility of reactive incentive effects analytically doubles the number of incentive tools for policymakers creating a four-fold typology of contingent rewards and deterrents. As shown in Table 1, there are not only traditional "active" carrots and sticks where the incentive's goal is to induce subject compliance with the contingent-reward or punishment, but also two "reactive" incentives where the incentives' goal is to induce subject defiance of the contingency:

Table 1: 4-Fold Typology of Active and Reactive Incentives

Type of incentive		Active	Reactive
Incentive Goal		Compliance	Defiance
Type of Contingency	Rewards	Carrot	Reactive Carrot
	Punishments	Stick	Reactive Stick

The goal of the reactive stick is to have subject persist in doing some designer-desired underlying behavior notwithstanding a contingent punishment, while the goal of reactive carrot is to have the subject reject the contingent reward to persist doing the designer-desired behavior. Both reactive carrots and reactive sticks thus aim to have subjects defy the contingent incentive taking what because of the incentive is the harder path.

The diagonal incentives in the table have similarities with regard to imposition of the incentive: reactive-carrots are like active sticks in that, if they are effective, the incentive does not need to be imposed. In our gym study, subjects resisted the reactive-carrot by turning down the financial incentives to quit, so the proffered refund did not need to be paid. The equilibrium non-imposition quality of effective reactive carrots is analogous to the non-imposition quality of

effective traditional active sticks – for example, when effective threats of criminal punishment deter and hence obviate the need for incarceration. In contrast, the incentives found in the alternative diagonal boxes requires imposition of the incentive to be effective. Reactive sticks are like active carrots in that if they work, the incentive needs to be imposed.[8] The equilibrium imposition effects of these incentives can have important effects on the costs of the incentives. Reactive carrots might be cheaper than active carrots if the reactive temptations are resisted. Zappos rarely had to pay new employees to quit because the take up of its reactive carrots was, as the company's leaders expected, very low. In the gym experiments that follow, we never had to refund money offered as a reactive carrot.

B. Theories of Reactive Incentives

The underlying mechanism for active incentives is straightforward: economic actors are motivated to change their behavior so that they receive contingent rewards and avoid contingent punishments. But the possible underlying mechanisms that might produce the boomerang effects of reactive incentives are diverse and subtly distinguished. In this section, we discuss several possible drivers of reactive responses and assess the likelihood that they would pertain to our experimental context of offers to quit a gym. We begin by describing explanations that are less likely to obtain in this context and end with two theories we believe are more likely explanations for our results.

1. Gym Screening, Gym Signaling & Third-Party Signaling

Offers to quit might have had a screening effect. If gym members who had a lower resolve to exercise accepted the offer and members with a higher resolve rejected the offer, then we might observe an equilibrium in which members who reject the offer go to the gym more on average than the control group (which includes both high and low resolve types).[9] A gym screening theory cannot explain our result, however, because we observed no screening of types in our data. All of our subjects rejected the offers to quit.

The gym's offer to pay money for new users to quit might have a signaling effect. Under this theory, the gym's offer would signal the quality of service to imperfectly informed users. Sellers who have better products and services are more likely to offer money-back guarantees.[10] Under this theory, users who received the offer are more likely to go the gym because they believe

---

[8] Giuseppe Dari-Mattiacci and Gerrit De Geest, *Carrots, Sticks, and the Multiplication Effect*, 26 JOURNAL OF LAW, ECONOMICS, AND ORGANIZATION 365 (2010). A systematic analysis of the advantages and disadvantages of carrots and sticks is provided in Gerrit De Geest and Giuseppe Dari-Mattiacci, *The Rise of Carrots and the Decline of Sticks*, 80 UNIVERSITY OF CHICAGO LAW REVIEW 341 (2013).

[9] Screening models and their operation are described in some detail in Eric Rasmusen and Basil Blackwell, GAMES AND INFORMATION (1994).

[10] The incentive effects of Signaling models are discussed in Roland Bénabou & Jean Tirole, *Intrinsic and Extrinsic Motivation*, 70 REVIEW OF ECON. STUDIES 489 (2003) and Andrei Bremzen, Elena Khokhlova, Anton Suvorov & Jeroen van de Ven, *Bad News: An Experimental Study on the Informational Effects Of Rewards*, 97 REVIEW OF ECONOMICS AND STATISTICS 55 (2015).

the gym has a higher quality than the control-group users who did not receive the offer. We think this gym signaling story is unlikely to explain our results because it is unlikely that gym users had imperfect information about the quality of the gym. In the experiment, we observe post-trial users (who previously attended the gym through a short promotional subscription) had stronger treatment effects than non-trial users, directly opposing the gym signaling prediction. These post-trial users should have had better information they gym's qualities than the non-trial users, and so under a signaling theory we would expect to have smaller treatment effects.

Users rejecting the offer might have also signaled information to third parties. There might be reputational benefits of being seen by others as a type who is resolved to go to the gym. A user might signal this resolve by rejecting the offer. However, third-party signaling is also unlikely to explain our results. As described below, the treated subjects rejected the offer at a private kiosk in a one-off transaction that was unlikely to be viewed by others. Any third-party signaling that occurred by rejecting the offer to quit would not by itself explain why those rejecting users would then go on to use the gym more than the control-group members.

2. Reactance

A boomerang effect of reacting to an incentive by doing less instead of more of the superficially incentivized behavior is consistent with psychological reactance theory.[11] Psychological reactance is a motivation state that is "hypothesized to occur when a freedom is eliminated or threatened with elimination."[12] The theory contends that when a freedom is perceived to be unjustly threatened, the individual will be motivated to reestablish the freedom in a variety of ways including "by increasing liking for the threatened choice."[13] Thus, prohibiting star-cross'd lovers from having a relationship may increase their desire for one.[14]

While our study's offer to quit the gym nominally expand subjects' freedom, subjects may nonetheless feel that the offer constitutes illegitimate manipulation or interference with their life choices. Our study's reactive carrots might have been seen as unreasonable challenges not to go to the gym, which subjects react to by going more.[15]

---

[11] See Stephen A. Rains, The Nature of Psychological Reactance Revisited: A Meta-Analytic Review, 39 HUMAN COMMUNICATION RESEARCH 47 (2013) and Debra Jones Ringold, Boomerang Effects in Response to Public Health Interventions: Some Unintended Consequences in the Alcoholic Beverage Market, 25 JOURNAL OF CONSUMER POLICY 27 (2002).

[12] Sharon S. Brehm & Jack W. Brehm, Psychological Reactance: A Theory Of Freedom And Control 37 (1981).

[13] James Price Dillard & Lijiang Shen, *On the Nature of Reactance and its Role in Persuasive Health Communication*, 72 COMMUNICATION MONOGRAPHS 144 (2005).

[14] Shakespeare, William, ROMEO AND JULIET, *in* THE NORTON SHAKESPEARE (Stephen Greenblatt, et al., ed. 1997). *see also* The Fantasticks, in which fathers concocted a feud for no other reason than to induce a romance between their two children. Harvey Schmidt & Tom Jones, THE FANTASTICKS (1960).

[15] Ruth Engs & David J. Hanson, *Reactance Theory: A Test with Collegiate Drinking*, 64 PSYCHOL. REP. 1083, 1085 (1989) (raising the legal drinking age in the United States from nineteen to twenty-one resulted in increased drinking rates among under-aged college students).

The threat to freedoms and the ensuing reactance-motivated state has been shown to be mediated by both affective, physiological, and cognitive markers.[16] Affectively, unjustly threatening freedoms has been shown to produce a range of negative emotions – including anger, irritation and resentment.[17] Physiologically, threatening subjects' freedoms have been shown to produce increased heart rates and other physical measures of an arousal state.[18] Cognitively, threatening subjects' freedom has been shown to produce "negative cognitions in the form of counterarguments."[19] While we do not have experimental evidence of cognitive counterarguments or physiological response, our results are not supportive of reactance as the causal mechanism underlying our findings of increased gym usage for certain treatment groups because we did not find affective evidence of anger. We asked tempted subjects immediately after the offer to describe their level of satisfaction on a 9-point scale. The treated subjects reported high levels of satisfaction (averaging over 7 on the scale) that were indistinguishable from the untempted control group. If subjects perceived the offer as an illegitimate intrusion on their freedom, we expect that they would have indicated less satisfaction with the gym as a result.

3. Willpower and Habit Training

An influential approach in psychology, proposed by Roy Baumeister, analogizes self-control to a muscle that can be trained.[20] In this framework, resisting a temptation is akin to exercising the self-control muscle and can improve future resistance. The temptation we offer is hence to be viewed as a training session that improved our subjects' future resistance to the urge to substitute gym activity with something more pleasant. However, we find long-lived effects for some treatments and it is difficult to believe that taking a few seconds to resist a one-off temptation would result in enough training to shape future behavior.[21]

This lack of sustained repetition is also why we reject habit formation as a causal mechanism for increased gym usage. Studies have shown that using 2-month commitment contracts can effectively foster gym use and sustain improved gym attendance beyond the short-

---

[16] Christina Steindl, Eva Jonas, Sandra Sittenthaler, Eva Traut-Mattausch & Jeff Greenberg, *Understanding Psychological Reactance: New Developments and Findings*, 223 ZEITSCHRIFT FÜR PSYCHOLOGIE 205 (2015).
[17] Id.
[18] Id.
[19] Id.
[20] Roy F. Baumeister, Kathleen D. Vohs & Dianne M. Tice, *The Strength Model of Self-Control*, 16 CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE 351 (2007).
[21] A second interesting implication of the muscle metaphor is the fact that exercising one's willpower, while having a long-term positive effect on resolve due to training, also has a short-term negative effect due to ego depletion or fatigue. Kathleen D. Vohs, et al., *Making choices impairs subsequent self-control: A limited-resource account of decision making, self-regulation, and active initiative*, 94 JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY 883 (2008). But *see also* Miguel A. Vadillo, Natalie Gold & Magda Osman, *The Bitter Truth About Sugar and Willpower: The Limited Evidential Value of the Glucose Model of Ego Depletion*, 1207 PSYCHOLOGICAL SCIENCE 27 (2016) for evidence to the contrary. This might imply that the gym session that immediately follows the administration of the temptation should be shorter for the treatment group. Unfortunately, our data do not record exit from the gym.

term.[22] But in contrast to a treatment that encourages a 2-month commitment to foster habit formation, our experiment uses a one-off 15-second intervention which would implausibly be sufficient by itself to foster a habit of gym use. However, habit formation may have played a secondary role in causing increased long-term usage. If our initial intervention made people go more to the gym in the initial period (because of one of the effects described below), then this higher initial attendance may have helped the subjects to build a habit of going to the gym.

4. Self-Signaling

Our results are more consistent with a form of self-signaling where the process of resisting temptation teaches our subjects something about their own resolve that helps them establish a "self-reputation" and makes it easier for them to go to the gym in the future.[23] Ronit Bodner and Drazen Prelec have defined self-signaling as "an action chosen partly to secure good news about one's traits or abilities, even when the action has no causal impact on these traits and abilities."[24] They even describe what we would call a reactive stick as an instance of self-signalling: "For example, a person who takes the daily jog in spite of the rain may see that as a gratifying signal of willpower, dedication, or future well being."

Roland Bénabou and Jean Tirole have similarly modelled "willpower activities" like resisting temptation. They hypothesize that resisting temptation is "hard" information that even people with imperfect recall about their state of mind—for example, being committed to going to the gym—can use to build an enduring image of themselves.[25] The unusual nature of our offers to quit likely made them more salient (even though they were one off) and hence more easily recalled as the users made later decision about whether to revisit the gym. Benabou and Tirole explain the power of recall thusly:

> "If I eat this tempting dessert, there goes my whole diet. If I cannot turn down this drink, I might as well admit that I am still a hopeless alcoholic." The fear of *creating precedents* and losing faith in oneself then creates an incentive that helps counter the bias toward instant gratification.[26]

In this framework, people have imperfect recall about their resolve to go to the gym but their resisting temptation is a memorable precedent that signals to themselves of their willpower.

Self-signaling has three complementary effects. First, individuals' motivations to resist the temptation we offer may be strengthened by the fear that failure to go to the gym will create a

---

[22] Heather Royer, Mark Stehr & Justin Sydnor, Incentives, *Commitments, and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company*, 7 AMERICAN ECONOMIC JOURNAL: APPLIED ECONOMICS 51 (2015).
[23] Ronit Bodner & Prelec Drazen, *Self-signaling and diagnostic utility in everyday decision making*, 1 PSYCHOLOGY OF ECONOMIC DECISIONS 105 (2003).
[24] Id. at 1.
[25] Roland Bénabou & Jean Tirole. Willpower and Personal Rules, 112 JOURNAL OF POLITICAL ECONOMY 848 (2004).
[26] Roland Bénabou & Jean Tirole, *Willpower and Personal Rules*, 112 JOURNAL OF POLITICAL ECONOMY 848 (2004).

precedent for the future. Bénabou and Tirole have explained: an initial decision can be seen "as a possible precedent for future ones, so that giving in today raises the probability that he will do the same in the future. By thus tying together sequences of choices, he raises the stakes on each one and better aligns his short-term incentives with his long-run interests."[27] This effect suggests that take-up rates of the offer to quit should be low.

Second, individuals who have resisted the temptation to quit are more likely to go to the gym because resistance signals willpower, which in turn makes future exercise more likely. Third and relatedly, individuals who have resisted temptation may be motivated to reduce cognitive dissonance.[28] Following a decision, dissonance can be reduced by changing your cognition to "remov[e] negative aspects of the chosen alternative or positive aspects of the rejected alternative."[29] Treated subjects may find it easier to appreciate the benefits of gym use (and to discount the benefits of non-use). Behaving inconsistently imposes a cost of ex post rationalization that the individual would gladly avoid. Resisting the temptation thus might create the additional motivation for gym attendance: the urge to avoid cognitive dissonance costs.

A self-signaling mechanism is consistent with our finding of stronger treatment effects for post-trial users than for non-trial users. We were disinclined to embrace a gym signaling explanation for our results because post-trial users likely were better informed about the qualities of the gym (because of their trial visits) than non-trial users. But post-trial users may be more imperfectly informed about their own resolve than non-trial users. By offering new members the option of beginning with a brief trial membership, the gym may have caused users with better information about their resolve to subscribe immediately on a non-trial basis. Hence, consistent with our results, there might have been more opportunity for self-signaling of resolve to produce treatment effects on trial users. On the other hand, individuals who know that they have a self-control problem may have disproportionately used non-trial subscriptions as a commitment device, using pre-payment to overcome this relative lack of resolve. On net then, while the gym's menu of offers may have produced type separation that impacts the effectiveness of self-signaling, it is difficult with our data to definitively predict whether individuals with more imperfect information about their self-control would initially forego trial subscriptions or not.

5. Taste-Based Theories

[27] Id. *see also*, George Ainslie, Picoeconomics: The Interaction of Successive Motivational States 143 (1992).
[28] Eddie Harmon-Jones & Judson Mills, *An Introduction to Cognitive Dissonance Theory and an Overview of Current Perspectives on the Theory*, in COGNITIVE DISSONANCE: PROGRESS ON A PIVOTAL THEORY IN SOCIAL PSYCHOLOGY (1999).
[29] Id. See also Leon Festinger, A THEORY OF COGNITIVE DISSONANCE (1957).

This self-signaling mechanism might be complemented by tastes that users have to avoid making prior actions appear erroneous. These tastes have been understood to underlie sunk cost behavioral effects.[30] For example, in describing a taste for consistency, Erik Eyster has noted:

> Rather than simply regret the past, people may do things in the present to make their past choices look better. For example, someone who has paid a lot of money to join a health club may exercise when injured to justify the money spent.[31]

Tastes for avoiding regret might also underlay the motivational impact of sunk opportunity cost.[32] Users who turned down money to quit the gym may experience even more regret because of the sunk opportunity than control-group users who merely sunk their prepaid gym subscription. A user with a taste for avoiding mistakes might be motivated to go to the gym rather than have to confront the possibility that they made a mistake in turning down the prior offer.

In contrast to the self-signaling idea in which resisted temptation teaches the users something about themselves, the sunk-opportunity cost ideas turn on a user's taste to avoid experiencing this psychic loss. Richard Thaler, in tying sunk-cost effects explicitly to Kahneman and Tversky's prospect theory of attitudes toward loss aversion, presented the classic example:

> A family pays $40 for tickets to a basketball game to be played 60 miles from their home. On the day of the game there is a snowstorm. They decide to go anyway, but note in passing that had the tickets been given to them, they would have stayed home.[33]

---

[30] *See*, e.g., Alexander Morell, et al., *Sticky Rebates: Rollback Rebates Induce Non-Rational Loyalty in Consumers – Experimental Evidence*, 23 PREPRINTS OF THE MAX PLANCK INSTITUTE FOR RESEARCH ON COLLECTIVE GOODS (2009); Robert Pindyck, *Sunk Costs and Real Options in Antitrust*, NBER Working Paper Series (2005), *available at* https://www.nber.org/papers/w11430.

[31] Erik Eyster, RATIONALIZING THE PAST: A TASTE FOR CONSISTENCY, Nuffield College Mimeograph 1 (2002).

[32] Hal R. Arkes & Catherine Blumer, *The Psychology of Sunk Cost*, 35 ORGANIZATIONAL BEHVAIOR AND HUMAN DECISION PROCESSES 124 (1985). The psychological literature has also identified "entrapment" as a specific form of the sunk cost phenomena where subjects "incur small, continuous losses as they seek or wait for an eventual goal." Id. at 137. For example, people who have a longer time for a bus may be less inclined to hail a cab thereby "nullifying all the time you have spent waiting." Id.

[33] Richard Thaler, *Toward a Positive Theory of Consumer Choice*, 1 JOURNAL OF ECONOMIC BEHAVIOR & ORGANIZATION 47 (1980). Thaler also presents evidence an example of how sunk costs can, consistent with prospect theory, cause subjects to prefer a riskier allocation:

> For example, one set of subjects preferred (0) to (-800, 0.2; 200, 0.8), while a different set preferred (- 1,000, 0.2) to (-200). This suggests that the 200 subtracted from the first problem to obtain the second is not viewed as sunk by the subjects.

Id. *at* 47. We posit that sunk opportunity cost frame, for analogous reasons might induce subjects to opt for riskier choices. (It would be possible to test this hypothesis by with an experiment in which some subjects are randomly offered a payment not to participate in the next phase of the experiment. After this temptation is issued to the subset of subjects, all subjects would be asked to choose between (0) and (-500, 0.5; 500, 0.5). We expect that those subjects that were offered (and subsequently rejected) the temptation would choose the latter option more frequently).

A preference to avoid losing $40 might also be actuated if the family had earlier turned down a $40 offer from a scalper to relinquish the tickets. A loss frame might thus be created not only by the prior revealed willingness-to-pay preference but also by a prior revealed willingness-to-accept preference. Again, post-trial users may have experienced larger treatment effects than non-trial users simply due to the type-separating effect of the gym's joint offers of trial and non-trial subscriptions. Our subjects who began with a trial subscription may have chosen the smaller commitment not only because they had lower resolve but also because they knew they would experience more regret if they did not follow through on using the gym. The larger treatment effects that we observe for post-trial users might be driven by their having stronger taste for avoiding loss from wasted sunk costs.[34] We will return to the idea of type-separation below when we explore how treatment effects varied for users who had just opted for longer, more expensive subscriptions at the time of the treatment itself.[35]

6. Appropriate Contexts

While this section has sought to assess a variety of possible mechanisms for reactive effects that might explain our experimental results, we now step back to postulate more broadly on where reactive incentives are most likely to be effective. While the addition of these two new forms of motivation seem superficially to double the instruments in policymakers' toolkits, the beneficial boomerang effects of reactive carrots and sticks are only likely to obtain in certain contexts. Specifically, the foregoing self-signaling theory suggests that reactive effects are more likely to be found when agents (who are subject to the incentive) have self-control problems and are imperfectly informed about the extent of those problems. People with imperfect self-control are often modelled with hyperbolic utility functions and described as having present-biased preferences:

> When costs are immediate people with present-biased preferences tend to *procrastinate*—wait [to do something] when they should do it [now]—while when rewards are immediate they tend to *preproperate*—do it now when they should wait.[36]

Present-biased preferences can cause people to inefficiently put off unpleasant tasks (e.g., doing taxes at the last minute), and to inefficiently accelerate pleasant tasks (e.g., eating the last piece of cake before dinner).

---

[34] While type separation in response to the gym menu of offers might have caused users with larger self-control problems to forego trial subscriptions as a kind of commitment device, we do not see a reason why users with larger loss aversion to inconsistent dynamic behavior would sink larger expenses on a non-trial membership.

[35] See *infra* Figure 4.

[36] Ted O'Donoghue & Matthew Rabin, *Doing It Now or Later*, 89 AMERICAN ECONOMIC REVIEW 103 (1999).

Reactive behavioral effects are more likely in contexts where agents might otherwise succumb to a time-inconsistent taste for immediate gratification.[37] Defying a reactive carrot or stick in the first instance may make it easier for them to exercise will power later on. However, self-signaling turns on agents being imperfectly informed. Accordingly, reactive behaviors are more likely when agents are less informed about a particular dimension of their self-control. We conjecture that reactive effects are more likely to be found from tempting relatively new gym users rather than from treating long-term subscribers. On this dimension, Zappos' initial offer to new trainees before they had even started to work (or Henry V's offer to his troops before they went into battle) is likely to produce more reactive effects than Amazon's annual offer to long-term employees.

While the alternative taste-based theories for generating reactive behavioral effects do not require time-inconsistent preferences of the type that underlay hyperbolic discounting, the beneficial effects stemming from a taste for behavioral consistency across time only arise in contexts where the agent initially resists the reactive incentive. Accordingly, reactive incentives are likely to help agents who naively think they will not have future self-control problems but who in fact would have such problems in the absence of the reactive effect. Thus, for example, the person who has recently finished a smoking cessation program and is flush with confidence that they will never smoke again (even though over 50% of those that quit smoking relapse within a year), may resist a one-off offer to smoke a cigarette for $100.[38] This possibility inverts a standard result of behavioral economics – which suggests that it will be easier induce sophisticated agents (who are aware that they have self-control problems) than naïve agents (who are not aware) to participate in mechanisms to promote better self-control. Sophisticated agents are more likely to take advantage of Christmas club saving devices or opportunities to enter into commitment contracts to reduce the chance of procrastination or preproperation.[39] Sophisticated agents, however, are more likely to accept an initial temptation – in part because they are aware that they have self-control problems and would rather be paid for engaging in the activity that they are likely to undertake and in part because accepting the offer will reduce ex post regret for acting inconsistently. In contrast, naïve agents who do not believe they will experience future self-control problems are more likely to resist and thereby increase their subsequent taste-based costs of acting inconsistently.[40]

---

[37] *See* Ted O'Donoghue & Matthew Rabin, *Incentives and Self-Control*, 42 ECONOMETRIC SOCIETY MONOGRAPHS 215 (2006); David Laibson, *Golden Eggs and Hyperbolic Discounting*, 112 QUARTERLY JOURNAL OF ECONOMICS 443 (1997); Ted O'Donoghue & Matthew Rabin, *Optimal Sin Taxes*, 90 JOURNAL OF PUBLIC ECONOMICS 1825 (2006).

[38] Olaya Garcia-Rodriguez, et al., Probability and Predictors of Relapse to Smoking: Results of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), 132 DRUG AND ALCOHOL DEPENDENCE 479 (2013).

[39] Ayres is a co-founder of www.stickK.com, a for-profit commitment store, where sophisticated users have put more than $45 million at risk to help stick to their goals.

[40] *See also* Michael Abramowitz & Ian Ayres, *Commitment Bonds*, 100 GEO. L.J. 605 (2012). (arguing that naïve agents are also particularly susceptible to traditional dynamic carrot/stick combinations in which agent is offered immediate cash if she will take on commitment contract to pay a greater amount if she later fails to follow through on her goal).

Sunk opportunity costs are only likely to have effects if the foregone opportunities are sufficiently salient. Researchers have shown that in some contexts subjects exhibit "opportunity cost neglect" where decision-makers do not consciously consider what would be the next best alternative to particular decisions.[41] A consumer, for example, might not consider how else they she might have spent the purchase price. In such contacts, experiments have shown that their initial decision to purchase can be reduced by even a "placebic reminder that money preserved by forgoing some purchase will be available for other purchases."[42] Policymakers can enhance the salience of an opportunity cost by foregrounding the option that is being resisted. As Adam Eric Greenberg and Stephen A. Spiller have noted:

> When people neglect opportunity costs, they fail to recognize that the choice is from among multiple options. When opportunity costs are made explicit, whether-or-not decisions transform into which-one decisions.[43]

Accordingly, the specifics of how the menu of choices is presented is likely to be particularly important to the effectiveness of reactive incentives.

Finally, the importance of inducing initial resistance may be influenced by other behavioral factors. For example, the original Zappos offers were made in group settings to multiple trainees where it might be more difficult to step forward and defect from the group's goal.[44] Similarly, the reactive stick hazing rituals in military and fraternity settings are traditionally conducted in group settings where solidarity in behavior is explicitly encouraged.[45] In contrast, our gym experiment offers and the Amazon annual offers are communicated individually where peer pressure effects to conform are likely to be attenuated. Being in a group makes it easier to for individuals to resist a temptation, but it might also produce weaker treatment effects as agents may distinguish later solipsistic decisions. On this and many other dimensions, this article merely opens avenues for future research into both the causes and contexts where reactive behaviors are most likely to obtain.

---

[41] Shane Frederick, Nathan Novemsky, Jing Wang, Ravi Dhar & Stephen Nowlis, *Opportunity Cost Neglect*, 36 JOURNAL OF CONSUMER RESEARCH 553 (2009).

[42] Id. at 553. Our 100% money-back treatment could also be seen as a placebo, under some conditions. Under Dutch law new members who purchased their subscription online would have been allowed to cancel their subscription within 14 days of the purchase (see https://www.acm.nl/nl/bedenktijd) so that, effectively, we simply reminded these subjects of their rights. Most of our subjects, however, purchased their subscription directly at the gym.

[43] Adam Eric Greenberg & Stephen A. Spiller, *Opportunity Cost Neglect Attenuates the Effect of Choices on Preferences*, 27 PSYCHOLOGICAL SCIENCE 104 (2016).

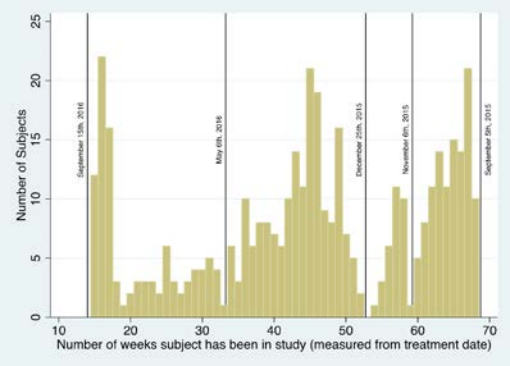
[44] *See* J. M. Levine & R.L. Moreland, *Small groups*, *in* THE HANDBOOK OF SOCIAL PSYCHOLOGY 415-469 (D. T. Gilbert, S. T. Fiske, & G. Lindzey ed. 1998) for a discussion of group functioning.

[45] Gentry R. McCreary & Joshua W. Schutts, Why Hazing? Measuring the Motivational Mechanisms of Newcomer Induction in College Fraternities, 19 JOURNAL OF COGNITION AND CULTURE 343 (2019); Sanford M. Dornsbusch, The Military Academy as an Assimilating Institution, 33 SOCIAL FORCES 316 (1955).

II. THE EXPERIMENT

A. Experimental Design

The study was conducted at the Universitair Sportcentrum Universum, or University Sports Center (USC) "The Universe", a gym affiliated with the University of Amsterdam and located in Amsterdam Science Park.[46] Subjects of the study were new subscribers to the USC whose subscription contract included a provision consent to participate in the study.[47] The USC offers an array of sports activities, we focused on fitness.[48] The testing took place in two waves with two distinct sets of temptations – with the first wave of treatments occurring from September 5, 2015 to September 14, 2016 and the second wave occurring from September 15, 2016 to July 15, 2017. Figure 1 displays the flow of qualifying subjects by week for Wave I participants. Appendix Figure 1 presents the analogous distribution for Wave II participants.



Our subjects were either "non-trial users" of the gym who had just subscribed to one of nine types of subscriptions (ranging from one month to one year in duration), or were "post-trial users" of the gym who had just subscribed to one of these same nine subscription types after finishing a lower-cost, one-month trial subscription. Both new and post-trial users had the option to begin with an initial trial membership, but non-trial users forewent the trial opportunity and subscribed directly on a non-trial basis. Post-trial users accordingly tended to have more information about the gym after their trial visits, but non-trial users, as discussed above, may have displayed a greater commitment to exercise by choosing to contract directly on a non-promotional basis.

---

[46] See footnote 50 for a description of the network of sports facilities affiliated with the University of Amsterdam.

[47] We obtained preliminary approval for the experiment from the Human Subjects Committee at Yale University (May 28, 2015, HSC protocol number 1504015638) and from the Ethics Committee of the Faculty of Economics and Business (EBEC) at the University of Amsterdam (June 5, 2015, EBEC reference number EC 2015-0010) and registered the experimental setup with the American Economic Association Randomized Control Trials Registry (May, 28, 2015, RCT ID AEARCTR-0000650).

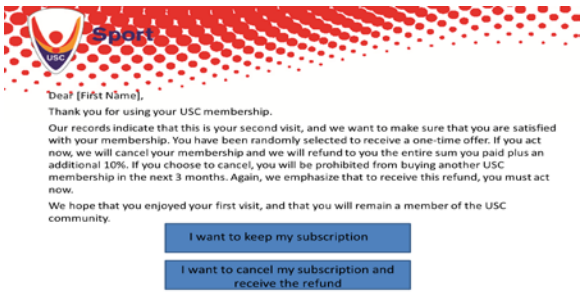
[48] This choice was motivated by both the number of registered users and the ability to control access. The fitness room has a dedicated thumb-print activated gate in addition to the general entrance thumb-print activated gate. Therefore, we could verify that our subjects entered the gym and used the fitness room.

The rationale for choosing subjects who were relatively new gym members was to test individuals who might be more imperfectly informed about their level of self-control with regard to gym use. To qualify for the study, subjects had to visit the gym a second time within one week after subscribing. The second visit within a week requirement was also designed to limit the number of people who accepted our offer to quit the gym by tempting subjects at the beginning of their membership while they were still excited about their new commitment and when they were ready to enter the gym and work out.

As a qualifying subject attempted to enter the gym (on her second visit within a week after subscription) by placing her thumb on an electronic thumb-print reader at the main entry gate, a screen on the electronic gate would direct the subject to a computer kiosk located approximately 10 meters (33 feet) away from the main entry gate. At the computer kiosk, after placing again her thumb on an electronic thumb-print reader, the subject was then shown a screen calling for them to choose Dutch or English as their preferred language.[49]

During the first wave of treatments, subjects were randomly assigned to one of three groups: two treatment groups or a control group. Randomization was implemented with blocking by subscription type. The subjects in the three different groups were then presented with one of three different screens.

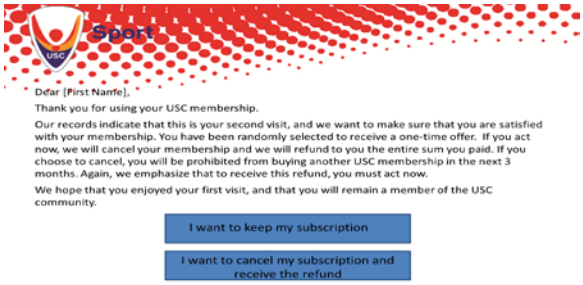
"110% Temptation" subjects were told they had been randomly selected to receive a one-time offer to cancel their membership in return for a refund of "the entire sum you paid plus an additional 10%." This treatment screen emphasized that to receive the refund they must act now and also explained that if a subscriber chose a refund that she would be prohibited from buying another membership in the next three months.[50] Figures 2a through 2e below depict the treatment screens.



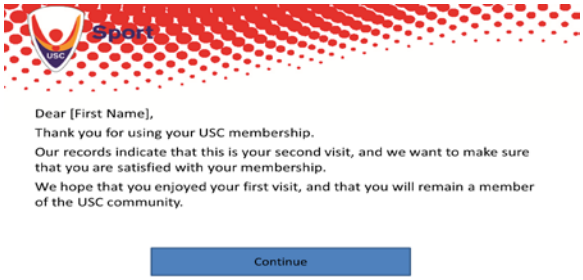
[49] Screenshots of the language preference screen and all Dutch screens can be found in the online appendix. Note that at no point was there any human interaction with subjects in our experiment. If a subject had accepted the temptation to quit, the administration of the refund would have made it necessary to be in contact with a USC employee, but, as emphasized, none of our subjects accepted the temptation.

[50] The USC belongs to a network of 7 sports centers affiliated with the University of Amsterdam. Each of these centers offers a different package of sports activities, so that the packages are partially overlapping. Our experiment focused on fitness, which is offered at 5 different locations, A subscription purchased at any of the centers in the network gives the purchaser the right to access all of the affiliated facilities. In reality, subjects typically purchase a subscription at the facility that they intend to use, which is usually the one that is closest to either where they live or where they take classes. However, in order to avoid strategic behavior, we made sure that accepting the offer to quit resulted in a prohibition to buy a subscription at any of the affiliated centers.

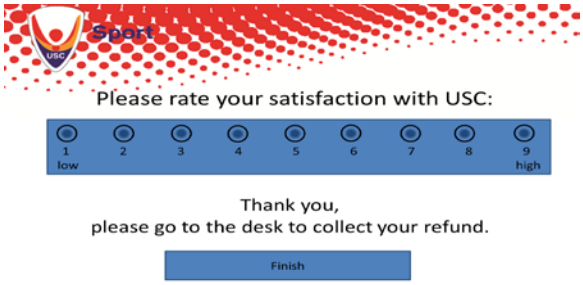
"100% Temptation" subjects received a similar one-time offer, but the refund was limited to "the entire sum you paid."



Control group subjects were presented with identical greeting and ending sentiments.[51] The control group treatment screen, however, simply acknowledging their second visit and expressed that the gym wanted "to make sure that you are satisfied with your membership."



Subjects in either of the temptation groups that clicked on "I want to cancel my subscription and receive my refund" were then shown a screen soliciting their satisfaction with the gym (on an integer scale from 1 through 9) and directing them to the reception desk to collect their refund was presented.



[51] Web Appendix Figures 2a-2i present screenshots of the various screens in English and Dutch.

Subjects in the control group and either of the temptation groups who turned down the temptation offer and all of the control group subjects were shown a screen soliciting their satisfaction.



The computer kiosk and electronic gates were programmed so that qualifying subjects had to complete the kiosk tasks (including indicating a level of satisfaction) before receiving a refund or being able to enter the main gate.

The second wave testing procedure was analogous with subjects being randomly assigned to a control and two temptation treatment groups, except in the second waive the temptation amounts were 125% and 150% of the original subscription price (instead of the 100% and 110% temptation offers of the first wave).

B. First Wave Results

1. Blocking, Balance and Descriptive Statistics

Appendix Figure 3 shows that blocking by the nine subscription types was successfully implemented for the 416 subjects initially qualifying for Wave I testing.[52] We exclude in this section's analysis subjects for whom we had fewer than 16 weeks of gym usage data, reducing the number of subjects analyzed 390.[53]

Appendix Figures 5 provides balance tests of our pre-treatment controls.[54] These tests suggest that we successfully randomized treatment across pre-treatment variables and that the

[52] Appendix Figure 4 shows analogous success blocking the 297 subjects initially qualifying for Wave II testing.
[53] The 16-week limitation analogously reduced the size of the Wave II data analyzed in the Appendix to 223.
[54] Appendix Figure 6 show analogous balance tests for Wave II.

treatment and control groups were well balanced on more than 50 pre-treatment variables.[55] From Appendix Figure 7, we find that subjects in the first wave were primarily (74%) students at the University of Amsterdam (UvA), although the subject pool also included a few university employees (1%) and students at Vrije University (VU – 8%) and Hogeschool van Amsterdam (HvA – 12%), the Amsterdam University of Applied Sciences. 24% of subjects were female and 23% were born in the Netherlands.

2. Treatment Effects

A test of resisted temptation would not be possible if subjects failed to resist. It would still be possible to conduct "intent to treat" analysis on temptation groups if a subset of treated users accepted the offer to quit the gym (for example, by attributing 0 post-treatment visits to all subjects with cancelled memberships).[56] But as intended, each and every tempted subject in both waves of the experiment resisted the temptation, declining our offer to quit.

Figure 3 provides descriptive statistics on outcome characteristics for the 390 subjects from Wave I. Subjects in each group reported mean satisfaction with the gym of at least 7.3 on a nine-point scale (with the 100% temptation group reporting the highest average of 7.47). The table shows that most of gym-related behaviors, like mean gym usage and resubscribing, are positively correlated with the size of the group's temptation. The 110% temptation group has higher outcome means than those of the 100% temptation group, which in turn has higher means than those of the control group. The 110% temptation group visited the gym on average 2.56 times per week post-treatment, while the 100% temptation and control groups visited on average 2.31 and 2.27 times per week, respectively. 41% of the 110% and 100% temptation groups purchased a new subscription post-treatment, while only 37% of the control group did. However, 100% temptation group members spent more on these new subscriptions (53.43 euros on average), than either the 110% temptation (50.05 euros) and control (41.65 euros) group members.[57]

[55] For Wave I, none of the 120 tested differences (60 each for the 100% and 110% temptation groups) are different from the control group at the 1% significance level, 1.7% (2) of the controls are significantly different at the 5% significance level, and 5.0% (6) of the controls are significantly different at the 10% significance level. For Wave II, 1.79% (2) of the 112 controls (56 each for the 125% and 150% temptation groups) are different from the control group at the 1% significance level, 6.25% (7) of the controls are significantly different at the 5% significance level, and 8.0% (9) of the controls are significantly different at the 10% significance level. Appendix Figures 7 and 8 provide detailed summary statistics for pre-treatment variables from subjects in both Wave I and Wave II, respectively. Appendix Figure 9 presents detailed pre-treatment summary statistics for subjects from both waves combined.

[56] *See* Joshua D. Angrist & Jörn-Steffen Pischke, MOSTLY HARMLESS ECONOMETRICS: AN EMPIRICIST'S COMPANION (2009) for discussion of "intent to treat" analysis.

[57] Appendix Figure 10 presents analogous statistics for Wave II.

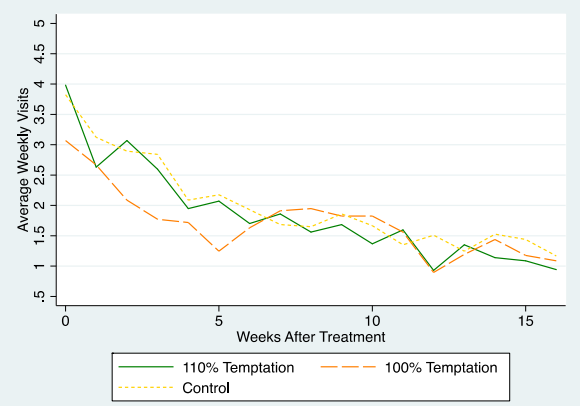
Descriptive Statistics of Core Outcomes: Wave I

Outcome Variable	110% Temptation	100% Temptation	Control	Total
N	131	128	131	390
Satisfaction (1-9 Scale)	7.36	7.47	7.36	7.40
Total post-treatment subscription duration (weeks)*	44.00	43.06	42.88	43.31
Average weekly visits, first 16 weeks post-treatment	2.56	2.31	2.27	2.38
Total visits, first 16 weeks post-treatment	40.77	36.80	36.23	37.94
Subject purchased a subscription post-treatment	0.41	0.41	0.37	0.40
Total Euros spent post-treatment	50.05	53.43	41.65	48.34
Analysis of Subjects with Post-Treatment Grades				
N	59	54	56	169
Average exam grade post-treatment	6.72	6.46	5.96	6.39

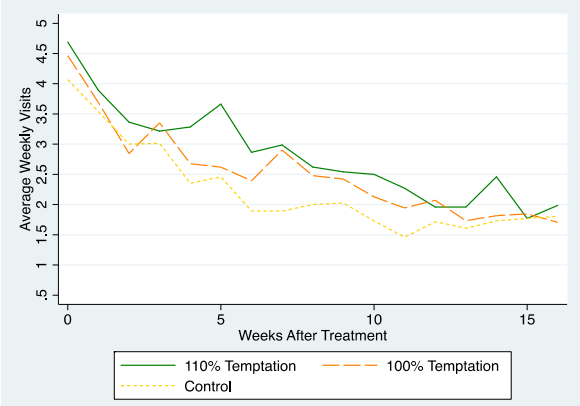
*We exclude subjects present for less than 16 post-treatment weeks

Figure 3 also shows a similar pattern with regard to post-treatment grades for the UvA students for whom we have post-treatment grade information. 110% temptation group members earned (on a ten-point scale) an average grade of 6.72, while 100% temptation and control group members earned on average grades of 6.46 and 5.96, respectively.

Figure 4 depicts the average gym usage of different subject groups over time. Panel A shows the gym usage of post-trial users and Panel B shows the gym usage of non-trial users. In each panel, the average number of visits per week drops over time for all groups and the steepest decline typically occurs in the first two weeks. As highlighted in panel A, the post-trial subjects in the first wave who were treated with the 110% bonus consistently visited the gym more often than post-trial subjects who were assigned to the control group.[58]



Panel A. Visits vs. Time: Wave I, Non-Trial Users



Panel B. Visits vs. Time: Wave I, Post-Trial Users

Figure 5 reports our core tests of treatment effects where we regress outcome variables on treatments and include fixed effects for the 52 weeks of the year and for the initial 9 subscription types on which we blocked. The specifications generally find stronger treatment effects for the

[58] Appendix Figure 11 presents additional panels for all users in Wave I and users in Wave II.

110% treatment than for the 100% treatment, and stronger effects for the post-trial users than for the non-trial users.

Using Seemingly Unrelated Regressions with robust standard errors for our three non-graded outcomes for the full set of 390 subjects, we find marginally significant (p = 6.2%) evidence that 110% temptation group members visited the gym 0.4 more times per week, on average, than control group members treatment. Specification (10) similarly shows that 110% temptation group members have (17.14 euros) higher post-treatment spending than control group members, and that this treatment effect is marginally significant (p = 7.9%).

SUR and OLS Analysis for Wave I: Unpooled Treatments with Week of Year + Blocking Dummies

VARIABLES	(1) Post-Treatment Subscription Duration	(2) Post-Treatment Subscription Duration	(3) Post-Treatment Subscription Duration	(4) Average Post-Treatment Weekly Visits	(5) Average Post-Treatment Weekly Visits	(6) Average Post-Treatment Weekly Visits	(7) Post-Treatment Sale	(8) Post-Treatment Sale	(9) Post-Treatment Sale	(10) Total Spent Post-Treatment	(11) Total Spent Post-Treatment	(12) Total Spent Post-Treatment	(13) Average of Post-Treatment Grades	(14) Average of Post-Treatment Grades	(15) Average of Post-Treatment Grades
110% Temptation	3.149	3.584	5.337*	0.405*	0.201	0.666**	0.0824	0.0203	0.162**	17.14*	22.81*	30.59**	0.838**	0.312	1.482**
	(0.233)	(0.388)	(0.0908)	(0.0623)	(0.458)	(0.0240)	(0.151)	(0.810)	(0.0285)	(0.0791)	(0.0872)	(0.0206)	(0.0197)	(0.539)	(0.0210)
100% Temptation	1.412	3.346	0.0792	0.106	0.0856	0.418	0.0621	0.0595	0.103	18.57*	35.80***	18.04	0.488	-0.251	1.317**
	(0.589)	(0.418)	(0.980)	(0.624)	(0.750)	(0.154)	(0.275)	(0.480)	(0.163)	(0.0548)	(0.00702)	(0.169)	(0.197)	(0.679)	(0.0216)
Constant	7.137	49.82*	5.157	1.053	-0.0761	0.352	0.373	0.595	0.269	-20.78	63.50	-59.29	5.851***	3.498**	1.753
	(0.687)	(0.0693)	(0.759)	(0.470)	(0.966)	(0.823)	(0.334)	(0.288)	(0.495)	(0.751)	(0.472)	(0.399)	(0.000244)	(0.0280)	(0.164)
Observations	390	171	219	390	171	219	390	171	219	390	171	219	169	77	92
R-squared	0.419	0.481	0.568	0.244	0.490	0.329	0.179	0.338	0.308	0.190	0.368	0.307	0.242	0.315	0.454
Non-Trial Users	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y	N
Post-Trial Users	Y	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y	N	Y

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

The first 12 columns come from SUR specifications, while the last 3 come from OLS specifications with robust standard errors. P-values are in parentheses.
The outcomes are number of weeks between treatment and the latest subscription expiry date (Columns 1-3); average weekly visits in the first 16 post-treatment weeks (Columns 4-6); whether there was a sale during the post-treatment duration (Columns 7-9); total Euros spent during the entire post-treatment duration (Columns 10-12); and average post-treatment grades across the entire post-treatment duration (Columns 13-15). Columns 1-12 are restricted to those subjects present for at least 16 post-treatment weeks.

Specification (13) provides evidence that assignment to the 110% temptation group has a positive and statistically significant impact of 0.858 points on average post-treatment grade–which represents an increase of 0.66 standard errors in the average pre-treatment GPA of control group members.[59] This treatment effect on grades is estimated as statistically significant (p = 2.0%), but caution is appropriate in interpreting the causal effect of our one-time, 15 second intervention on examination grades coming often months later. This is especially true because the intervening treatment effects on gym usage are only marginally statistically significant.

The estimated treatment effects for the 110% treatment seem to be largely driven by the larger treatment effects for post-trial users suggested by the panels in Figure 4. Post-trial users who received the 110% treatment visited the gym 0.7 more times per week, were 16% more likely to resubscribe, spent 30.6 euros more post treatment and received grades that were 1.5 points higher, relative to the control group – and all these treatment effects were statistically significant (p < 5%).[60] As argued above,[61] the difference in treatment effects for post-trial and non-trial users might have been the result of post-trial users being more imperfectly informed about their self-control or having stronger loss-aversion tastes than non-trial users who forewent the initial opportunity to subscribe on a less expensive trial basis.

Besides testing whether post-trial users have distinct treatment effects than non-trial users, it is also possible to test whether users who at the time of randomized blocking opted for longer or

[59] Appendix Figure 12 presents estimates [in Specifications (16), (17), and (18)] of the impact of treatment on post-treatment grades with additional controls for average pre-treatment grades, which are estimated to be positive and statistically significant, but do not impact the size sign or significance of the treatment effects reported in Figure 5.
[60] The post-treatment subscription duration estimated in specification (3) is marginally significant (p < 10%).
[61] *See* supra at p.11.

shorter subscriptions had different treatment types. The gym's menu of non-trial subscription durations ranging from as short as one month and as long as one year might have caused different types of users—users with different levels of self-control or different levels of distaste for inconsistent decisions—to opt for different duration contracts. However, as mentioned before, users with larger self-control issues might have either opted for shorter subscriptions because of their willpower concern or might have opted for a longer subscription as a commitment tactic. Figure 6 undertakes this analysis estimating treatment effects for three levels of subscription prices: at the sample mean, and at one standard deviation above and below the sample mean.[62]

Figure 6 shows again that treatment effects are more significant for the 110% temptation than for the 100% temptation – but this specification shows larger estimated treatment effects on each outcome for subjects who selected more expensive/longer subscriptions. For example, users with subscription prices one standard deviations *below* the mean who were treated with the 110% temptation were not estimated to go to the gym statistically more often than the control, but users with subscription prices one standard deviation *above* the mean who were treated with the 110% temptation were estimated to go the gym .65 times more per week than similar control-group users ($p < 5\%$).[63]

Test of Combined Treatment Effect Significance at Various Price Levels, Wave I

Statistic	Subscription Duration		Average Post-Weekly Visits		Sale		Total Spent	
	Point Estimate Value	P-Value	Point Estimate Value	P-Value	Point Estimate Value	P-Value	Point Estimate Value	P-Value
Combined 110% Treatment Effect: All Users								
Price = μ - σ (€ 42.68)	1.17	0.75	0.17	0.59	0.05	0.51	11.18	0.42
Price = μ (€ 102.01)	3.18	0.23	0.41*	0.06	0.08	0.15	17.30*	0.08
Price = μ + σ (€ 161.34)	5.18	0.17	0.65**	0.04	0.11	0.17	23.42*	0.09
Combined 100% Treatment Effect: All Users								
Price = μ - σ (€ 42.68)	2.13	0.56	0.01	0.98	0.04	0.60	31.33**	0.02
Price = μ (€ 102.01)	1.40	0.59	0.10	0.63	0.06	0.28	18.53*	0.05
Price = μ + σ (€ 161.34)	0.67	0.86	0.20	0.50	0.08	0.31	5.72	0.68

μ(Price) ≈ 102.01

σ(Price) ≈ 59.33

The point estimates come from 1) running a SUR analysis with the same specification of AF19; 2) multiplying the treatment-price interaction dummy by the specified price scalar; and 3) conducting a post-estimation test of that estimate's equality to 0. The outcomes are number of weeks between treatment and the latest subscription expiry date (Column 1); average weekly visits in the first 16 post-treatment weeks (Column 2); whether there was a sale during the post-treatment duration (Column 3); and total Euros spent during the entire post-treatment duration (Column 4). The analysis is restricted to those subjects present for at least 16 post-treatment weeks.

C. Second Wave Results

Overall, the results of the first wave provide some causal evidence of reactive effects, with regard to the larger 110% temptation and particularly for users who subscribed after an initial trial subscription. The second wave of our experiment tested whether increased temptations offering

[62] Figure 6 is derived from Appendix Figure 22 which interacted the treatment temptations with a continuous subscription price variable. We then used post estimation tests to evaluate the significance of treatment effects at specific subscription price levels (€102.01, €42.68, and €161.34 representing respectively the median subscription level, one standard deviation below the median, and one standard deviation above the median).

[63] It is also possible to test whether subjects from different demographic subject groups exhibit different treatment effects. Appendix Figure 24 tests for heterogeneous effects with regard to sex, age and national origin for the Wave I 110% temptation group. Notably, we find there that treatment effects are higher for women than men, who were estimated to have the larger and most significant effects on post-treatment subscription duration (10.1 weeks higher than control group), probability of new membership sale (26.9% over control group), and spending on post-treatment memberships (50 euros more than control). Analogous tests for the three other treatment groups are reported in Appendix Figures 25-27.

payments of 125% and 150% of the subscription price overall would produce stronger reactive effects.

An important difference between the first and second wave subjects is that 56% of the first wave subjects were post-trial gym users, while in the second wave this percentage fell below 15%. In the data collection period for the second wave, the gym changed its marketing options so that gym users were not all given the option of beginning with a trial membership. This sharp reduction in the number of post-trial users limited our ability to test for post-trial treatment effects in the second wave data.

An important caveat to our results is that while treatment effects were significant for first-wave post-trial users, our Wave II data produced much less causal evidence of reactive treatment effects. In analogous Wave II regressions that we report in the Appendix,[64] most of the estimated treatment effects were positive, but statistically insignificant. Our Wave II analysis suggests larger treatment effects from the 125% temptation than from the 150% temptation – e.g., showing at least marginally significant ($p < .1$) treatment effects for the 125% temptation group with regard to the likelihood of a post-treatment subscription sale and the duration of the post-treatment subscription.

Figure 7 shows that the treatment effects for the 125% temptation were particularly pronounced for lower-priced, shorter subscriptions.

Test of Combined Treatment Effect Significance at Various Price Levels, Wave II

Statistic	Subscription Duration		Average Post-Weekly Visits		Sale		Total Spent	
	Point Estimate Value	P-Value	Point Estimate Value	P-Value	Point Estimate Value	P-Value	Point Estimate Value	P-Value
Combined 150% Treatment Effect: All Users								
Price = μ - σ (€ 68.43)	0.36	0.86	0.40	0.31	0.13	0.15	-11.89	0.31
Price = μ (€ 115.65)	0.60	0.69	0.40	0.15	0.13**	0.05	1.03	0.90
Price = μ + σ (€ 162.87)	0.83	0.70	0.40	0.32	0.13	0.18	13.96	0.24
Combined 125% Treatment Effect: All Users								
Price = μ - σ (€ 68.43)	7.26***	0.00	0.92**	0.01	0.21**	0.02	7.12	0.53
Price = μ (€ 115.65)	3.43**	0.02	0.38	0.17	0.10	0.11	9.89	0.22
Price = μ + σ (€ 162.87)	-0.39	0.85	-0.16	0.68	-0.00	0.97	12.66	0.28

μ(Price) ≈ 115.65
σ(Price) ≈ 47.22

The point estimates come from 1) running a SUR analysis with the same specification of AF20; 2) multiplying the treatment-price interaction dummy by the specified price scalar; and 3) conducting a post-estimation test of that point estimate's equality to 0. The outcomes are number of weeks between treatment and the latest subscription expiry date (Column 1); average weekly visits in the first 16 post-treatment weeks (Column 2); whether there was a sale during the post-treatment duration (Column 3); and total Euros spent during the entire post-treatment duration (Column 4). The analysis is restricted to those subjects present for at least 16 post-treatment weeks.

We see in this figure that the estimated treatment effects for subscriptions with smaller (€68.43) prices were positive and significant ($p < .05$) for the average post weekly visits, the likelihood of a post-treatment subscription sale and the duration of a post-treatment subscription sale. Our data, however, do not allow us to assess whether the general reduction is estimated Wave II significance was caused by the relative paucity of post-trial users or the change to larger temptation percentages or some other factors.[65]

Stepping back and assessing the core treatment tests across the two waves of data (as well as dozens of robustness regressions reported in our web appendix),[66] we find stronger treatment

[64] See Appendix Figure 13.

[65] The Web Appendix includes Wave II robustness checks analogous to presented for the Wave I results.

[66] Appendix Figures 15 through 23 present results from additional specifications variously pooling the waves into a single specification and adding more detailed time fixed effects (as well as Cox proportion hazard regressions)to test the robustness of our results. The magnitude and significance of the results in these specifications do vary, but broadly support the conclusion that 110% treatment causes an increase in gym visitation and future membership purchases.

effects for the 110% and the 125% temptations than for the larger (150%) or smaller (100%) temptations. We also find conflicting results regarding two different dimensions of user self-selection. We find fairly robust evidence in the first wave that users who forewent the option of beginning with a less expensive trial subscription had smaller treatment effects than users who were treated post-trial. In the second wave, where initial trial contracts were not consistently available, we observe stronger treatment effects for subjects who choose shorter, less-expensive durations as their first subscription. Thus, both first and second wave results are suggestive that subjects who initially self-selected into shorter subscriptions were more impacted by reactive incentives – possibly because these subjects had more self-control problems or greater levels of loss aversion which led them.

In the first Wave, where subjects were confronted with two selection opportunities, subjects who opted for a trial subscription followed by a longer subscription had larger treatment effects than subjects who opted for a trial subscription followed by a shorter subscription duration. At a minimum, our results show the importance and crosscutting separating effects of subscription menus on reactive incentive effects.

D. Using Inclement Weather to Test for Reactive Stick Effects

Finally, we test whether new users who were "treated" with poor weather during the first week of their gym membership were more likely to subsequently visit the gym. As discussed above, inclement weather can be viewed as a stick-like disincentive to visit the gym. Subjects who persist in pursuing their fitness goal in spite of this disincentive may exhibit reactive behaviors of going to the gym more often in the future. Unlike the foregoing randomized experiment, here we report the results of a large, but poorly controlled observational study that can at best provide suggestive evidence of reactive-stick effects.

We use weather data from the KNMI (Koninklijk Nederlands Meteorologisch Instituut, or Royal Dutch Meteorological Institute) weather station located near Schiphol Airport (at about 15 km or 9 miles southwest of the USC) for the months of October through February during our experimental window. The KNMI releases a description of the anticipated weather and an associated color code depending on the severity of the weather. Two severe weather codes, "Geel" (yellow) and "Oranje" (orange) appeared in the data on 2.1 and 0.2 percent of the days, respectively.[67] As our reactive stick measure, we created variables indicating the number of days in a new users first subscription week that there was a yellow or orange severe weather warning. In order to create consistency in the type of inclement weather that we expect new users to face (e.g. snow or rain, rather than a heat wave), we limit our data to cold weather warnings that affect new users who had their first visit from October through February.

Figure 8 reports results from our tests for the impact of poor weather during the first week of gym membership on the number of weekly post-treatment gym visits.[68] For example, specification (2) which includes week-of-year fixed effect to control for the impact on usage of beginning a membership at different times of the year, finds that an increase in the number of "Yellow" weather days is associated with a small (.02 visits) statistically significant ($p < 1\%$)

[67] Appendix Figure 29 reports the frequency of each weather category.
[68] These results exclude individuals that participated in our study.

increase in the average weekly post-treatment gym visits. An increase in the number of "Orange" weather days does not have any effect on future gym visits. A possible explanation could be that subjects were able to resist the mild "Yellow" reactive stick, while they succumbed to the more severe "Orange" reactive stick.[69]

Figure 12. Weather Regressions: Average Number of Weekly Visits in First 16 Post-Treatment Weeks
Oct-Feb

VARIABLES	(1) Average Weekly Visits in First 16 Post-Treatment Weeks	(2) Average Weekly Visits in First 16 Post-Treatment Weeks	(3) Average Weekly Visits in First 16 Post-Treatment Weeks	(4) Average Weekly Visits in First 16 Post-Treatment Weeks
# Yellow Days, First Gym Week	0.0279***	0.0202**		
	(0.00969)	(0.0103)		
# Orange Days, First Gym Week	-0.0351	-0.0621		
	(0.0411)	(0.0443)		
# Yellow or Orange Days, First Gym Week			0.0190**	0.00885
			(0.00779)	(0.00833)
Semester	-0.476***	-0.662***	-0.477***	-0.661***
	(0.0387)	(0.0454)	(0.0387)	(0.0455)
Exam Period	-0.0389*	0.0336	-0.0441**	0.0295
	(0.0207)	(0.0235)	(0.0203)	(0.0234)
Week before Exams	-0.102***	-0.0761***	-0.100***	-0.0756***
	(0.0179)	(0.0211)	(0.0178)	(0.0211)
Resit Period	1.428***	1.203***	1.425***	1.200***
	(0.240)	(0.236)	(0.240)	(0.236)
Constant	1.165***	1.463***	1.166***	1.466***
	(0.0384)	(0.0646)	(0.0384)	(0.0645)
Observations	22,162	22,162	22,162	22,162
R-squared	0.015	0.026	0.015	0.026
Treatment Week-of-Year Indicators	We	Yes	No	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: these analyses exclude those who participated in the Anticarrots experiment.

The finding from this large but poorly controlled regression that inclement weather is associated with increased usage is consistent with the reactive stick idea that being exposed to initial deterrents may lead to longer term goal adherence.

III. POLICY IMPLEMENTATION

The previous two Parts of this article have, respectively, made out a theoretical and empirical case for the possibility that carrots and sticks might produce reactive behaviors. This part turns to a description of various contexts in which reactive incentives might be beneficially introduced. As suggested in the theory section, we are particularly interested in contexts where individuals (i) have self-control problems that might cause them to succumb to time-inconsistent preferences to procrastinate or preproperate and (ii) are relatively inexperienced with the behavior and so might be imperfectly informed about their self-control.

[69] Differently from the reactive carrot treatment, we do not observe directly whether a subject succumbs to the temptation. For instance, we are not able to tell whether a subject stayed home because of the bad weather conditions or because she had already made other plans for that day. Similarly, we do not know whether bad weather induced a subject not to go to the gym or simply to go a day later. We can thus only hypothesize that the "Orange" reactive stick was powerful enough to induce a positive and possibly large portion of the tempted subjects to succumb.

We will show here that the possibility of reactive incentives might lead to circumstances where defaults are preferred to mandatory rules – because defaults provide individuals with the opportunity to resist contracting around. Default rules thus allow private actors to be impacted by sunk opportunity costs in ways that mandatory rules do not. Moreover, affirmative-choice rules which may make the individual's choice to forego a temptation particularly salient and thus might produce stronger reactive effects than defaults that presume the non-temptation result if the individual fails to act.[70] Reactive incentive theory thus adds insights into how to structure choice– including when to use and how to structure mandatory rules, defaults, menus, and altering rules.[71]

C. Money-Back Guarantees

To begin, it is a short step from the gym offers described above to the time-limited money-back guarantees that are offered by sellers on a wide variety of goods and services. These guarantees might beneficially signal seller quality, but the results of the gym experiment suggest that the satisfaction guarantee offers might, if resisted, also induce more intensive buyer use and even increased likelihood of follow-on purchases. If an online retailer that does not typically offer free returns reaches out to customers the day that an item arrives and provides a time-limited offer of free shipping on returns, this offer may decrease the likelihood that customers will seek to return the product afterwards when the shipping is not free. The experiment results showing that offering to refund gym subscriptions did not reduce customer satisfaction is strong evidence that sellers can make buyers offers that they hope the buyers will resist without angering the buyers or making them feel manipulated. We will return to the possible reactive benefits of money-back guarantees when we discuss tuition rebates below.

D. Smoking, Dieting, and Gambling

In this section, we use the example of state-sponsored smoking cessation programs to explore how reactive incentives might help people struggling with substance abuse disorders and other addictive behaviors. Imagine, for example, a standard smoking cessation program that consists of eight weekly group sessions where participants are called upon to gradually reduce their

---

[70] Affirmative-choice rules are a special kind of penalty default that presents an unwanted result unless an individual makes an affirmative choice from a menu of at least two options. *See* Ian Ayres, REGULATING OPT OUT: AN ECONOMIC THEORY OF ALTERING RULES, 121 Yale Law Journal 2032 (2012). For example, with regard to retirement savings, an affirmative-choice rule might require a new hire to make an affirmative savings choice before being eligible to receive their first paycheck. The penalty default is no employment absent the affirmative choice. The gym experiment used a kind of affirmative choice rule, because tempted subjects could not enter the gym unless they made an affirmative choice at the computer kiosk to accept or reject the temptation offer.

One might go further than affirmative choice requirements, and use a temptation default to induce reactive behaviors. For example, one might set default to some temptation that will obtain unless the individual affirmatively opts out (and indeed, one might choosing impeding altering rules to enhance the resisted temptation signal). But, given the iron law of default inertia, temptation defaults are unlikely to be sufficiently resisted to produce net-beneficial reactive results.

[71] Ian Ayres, REGULATING OPT OUT: AN ECONOMIC THEORY OF ALTERING RULES, 121 Yale Law Journal 2032 (2012); and Eyal Zamir (feat. Ayres), MANDATORY RULES (working paper, 2019).

nicotine consumption.[72] In such programs, the last session is in part a celebration of the participants who have succeeded in eliminating tobacco products for the last week or two of this process.[73] The graduates are flushed with pride over their accomplishment. But the sobering truth is that a substantial proportion of participants – over 50% within one year – will still relapse and return to using tobacco.[74]

Dean Karlan, Kenneth Volp and a host of other behavioral economists have conducted experiments testing whether financial incentives – both contingent rewards and punishments can enhance the success rates. Giné, et al. shows that contingent financial punishments, in the form of a savings account that is forfeited if smoking continues, can produce long-term smoking cessation.[75] More recently Halpern, et al. tested the impact of offering thousands of CVS/Caremark employees a $800 reward for smoking cessation sustained through 6 months relative to a program which combined the carrot of $650 reward with the stick of $150 deposit that would be lost if the participant failed to kick the habit.[76] These studies show, to varying degrees, the traditional impacts of carrots and sticks in reducing tobacco consumption.

But the results of the gym experiment suggest that a reactive carrot might also aid in smoking cessation. Imagine at the final group session of a smoking cessation program that the organizers offered successful participants $100 to step forward and smoke a cigarette in front of the peers with whom they completed the program. As with all reactive carrot incentives, there is the fear that the individuals will succumb to the temptation and accept the offer. But one can easily imagine a situation where participants who are flush with pride over recently quitting would be disinclined to light up. Resisting temptation might be especially easy in this group setting where stepping forward would be a sign of disloyalty.

Individuals who resisted this temptation might exhibit reactive behavioral effects. An ex-smoker might experience more displeasure if she caves in and smokes for nothing when she previously turned down an opportunity to smoke for $100. And resisting the temptation, a la Roland Bénabou and Jean Tirole,[77] may be a vivid example of self-control that teaches the individual something about herself that could help keep her from relapsing.

One could imagine a smoking cessation experiment that randomly assigned some cessation participants to a treatment group that would be offered the $100 reactive carrots and other

---

[72]John L. Bottorff, Assessing the feasibility, acceptability and potential effectiveness of an integrated approach to smoking cessation for new and expectant fathers: The Dads in Gear study protocol, 54 Contemporary Clinical Trials (2017).

[73] Id.

[74] Olaya Garcia-Rodriguez, et al., PROBABILITY AND PREDICTORS OF RELAPSE TO SMOKING: RESULTS OF THE NATIONAL EPIDEMIOLOGIC SURVEY ON ALCOHOL AND RELATED CONDITIONS (NESARC), 132 Drug and Alcohol Dependence 479 (2013).

[75] Xavier Giné, Dean Karlan, & Jonathan Zinman, PUT YOUR MONEY WHERE YOUR BUTT IS: A COMMITMENT CONTRACT FOR SMOKING CESSATION, 2 American Economic Journal: Applied Economics 213 (2010).

[76] Scott D. Halpern, et al., RANDOMIZED TRIAL OF FOUR FINANCIAL-INCENTIVE PROGRAMS FOR SMOKING CESSATION, 372 New England Journal of Medicine 2108 (2015).

[77] See our discussion supra at 10.

participants to a control group that would not be so tempted.[78] We imagine that it would be difficult to secure human subject approval for a study that offered to pay some participants to inhale carcinogens. Prohibiting learning of this kind would be wrong. This article has provided both theoretical and empirical support for the possibility that offering a reactive carrot could increase the success of a smoking cessation program. Such programs as currently designed are often unsuccessful, so there is room for improvement. The possibility of beneficial impact from one-time offers to smoke a single cigarette to our minds would justify the risks of experimentation. If the experiment succeeds in showing beneficial reactive behaviors, its results could be scaled to help a large class of individuals. If the experiment fails, its costs would be limited to a small set of subjects, many of whom would have relapsed in any case. Randomized experiments are inappropriate when the relative outcome of different treatments are already known. That's why we don't experimentally test whether parachutes are beneficial.[79] But we don't know whether or not reactive temptations would work to enhance smoking cessation. The real possibility that an offer to smoke could save lives is a classic rationale for experimental test to learn what works.[80]

While this section has focused on smoking cessation, an analogous reactive offer might be applied to cessation programs for other kinds of addictions. One might, for example, offer to subsidize a bet of a recovering problem gambler or offer a recovering alcoholic $100 to take a drink in the hope that the offer will be rejected and steel the offeree against later temptations. The law currently impedes the use of certain reactive carrots. Rehab clinics probably cannot legally offer to pay recovering addicts to use heroin or other controlled substances, but a public or private entity might legally take a group of recovering addict on a field trip to a "shooting gallery" to see if they can resist the temptation to relapse when the opportunity is within easy reach.[81] Such a reactive "treatment" shows that it might be difficult to reliably characterize the baseline from which to distinguish reactive carrots from reactive sticks. To some, a field trip to a shooting gallery is akin to a reactive carrot where the individual is called upon to resist the opportunity to relapse easily, but addicts may experience the difficulty of resisting the siren calls of lotus sellers as a reactive stick. Regardless of the categorization, resisting the temptation to relapse may make staying clean easier in the future.

---

[78] Random assignment at the individual participant level would not allow group offers, but one might individually offer $100 or randomize at the group level with some groups receiving offers and other groups not.

[79] Gordon CS Smith & Jill P Pell, Parachute Use To Prevent Death And Major Trauma Related To Gravitational Challenge: Systematic Review Of Randomised Controlled Trials, 327 BMJ 1459 (2003).

[80] Reasonable readers might draw different boundaries on the potential application of reactive incentives. Instead of smoking cessation, imagine a suicide-prevention program design to help individuals who had recently attempted to take their own life. The risk of additional attempts by people in this group is substantial with one study finding that 1 in 25 make a fatal reattempt within 5 years. Robert Carroll, Chris Metcalfe & David Gunnell, HOSPITAL PRESENTING SELF-HARM AND RISK OF FATAL AND NON-FATAL REPETITION: SYSTEMATIC REVIEW AND META-ANALYSIS, 9 PLOS One (2014). Could we contemplate offering such individuals at the end of successful treatment a one-off financial reward to pay a loved one if they would attempt to kill themselves then and there? We think not (and not because such offers are likely criminal. COMMONWEALTH V. CARTER, No. 15YO0001NE (Mass. Juv. Ct. June 16, 2017). But to our minds the reasons why not cannot lie easily in a priori arguments or from the fact that an accepted offer would cause someone's death, because resisted offers might conversely cause some people not to die. As in the classic trolley problem, we would take affirmative action that would kill some if the action would save more lives.

[81] German Lopez, WHY SOME US CITIES ARE OPENING SAFE SPACES FOR INJECTING HEROIN, CNBC (January 25, 2018), https://www.cnbc.com/2018/01/25/why-some-us-cities-are-opening-safe-spaces-for-injecting-heroin.html.

Finally, one can also imagine reactive carrots applied to dieting. Diet control is distinct from cessation efforts because of the need for ongoing sustenance. Accordingly, tempting a dieter to eat something, even something sweet or outside the diet plan, might not send as vivid a signal of self-control, as the dieter will subsequently need to eat something. Reactive incentives might, nonetheless, be offered on other aspects of weight loss. For example, individuals enrolled in a weight-watchers-like program might be offered money-back or more-than-your-money-back akin to the gym temptations. After a few weeks of program enrolment, subscribers might be offered to refunds of their membership fee or more to end their participation. Resisting such initial temptations might induce more adherence to the program later on.

E. Borrowing

Consumers at times exhibit time-inconsistent preferences in retail credit settings. Just as the choice to save involves delayed consumption, the choice to pay down a loan also forces borrowers to consume present consumption. In this section, we suggest three ways that government policy might deploy reactive incentives to promote more responsible borrowing behavior.

1. A Possible Critique of Prepayment Penalty Prohibition

In 2013, in the aftermath of the mortgage crisis, the Consumer Financial Protection Bureau, pursuant to Dodd-Frank, prohibited prepayment penalties on most residential mortgages.[82] This prohibition of penalties is a mandatory rule that aims at making it easier for borrowers to pay off their debt.[83] But our foregoing results suggest that policy makers might have been able to produce more prepayments if the law allowed lenders to offer loans with a pre-payment penalty that borrowers were able to resist. An alternative policy would have prohibited lenders from solely offering loans with pre-payment penalties, but allowing lenders to include mortgages with prepayment penalties as part of a menu presented to borrowers that also included offers of loans without prepayment penalties. Drafters of Dodd-Frank Act unsuccessfully fought for an analogous kind of menu mandate. They initially proposed that all mortgage lenders be required to offer a "plain vanilla" mortgage with standardized terms (including a 30-year term and a fixed APR) as one of their mortgage products.[84] The goal was to spur competition by making it easier for consumers to compare "plain vanilla" offered by different lenders.[85] Unlike this proposed requirement of offering "plain vanilla" mortgages, our idea of allowing lenders to offer mortgages

---

[82] Prepayment penalties are allowed within the first three years for loans that have an APR that cannot increase, counts as a "qualified mortgage," and have an APR below the Average Prime Offer Rate. WHEN ARE PREPAYMENT PENALTIES ALLOWED IN NEW MORTGAGES?, NOLO (2019), https://www.nolo.com/legal-encyclopedia/when-are-prepayment-penalties-allowed-new-mortgages.html.

[83] Mandatory rules are rules that the parties may not contract around and as general matter tend to be justified by paternalism or externality concerns. Zamir & Ayres, *supra* note 58. Both justifications are likely present as prohibiting prepayment penalties may both protect borrowers from unreasonable fees and, by reducing housing leverage, reduce systemic risk.

[84] Associated Press, CONGRESS WARY OF "PLAIN VANILLA" BANK PROPOSAL (Sept. 22, 2009), http://www.nbcnews.com/id/32968985/ns/business-us_business/t/congress-wary-plain-vanilla-bank-proposal/#.XcHPCzNKhPY.

[85] Id.

with prepayment penalties (if and only if they also offer analogous mortgages without prepayment penalties) is driven by the possibility that resisting the mortgages with prepayment penalties might reactively change a borrower's subsequent payment behavior.

This menu mandate might produce more prepayments than under the current regime that merely prohibits prepayment penalties.[86] For example, imagine that under a mandated menu regime a borrower is offered two identical mortgages at the same APR – except one mortgage offers the borrower a $1000 if she will accept a prepayment penalty. Borrowers who turned down the temptation of $1000 may be more likely to subsequently exercise their prepayment option (than borrowers under the current regime who are not presented with the alternative mortgage). Borrowers who turned down cash to preserve the option of prepayment might feel foolish if they don't follow through and exercise that option.

One might worry that lenders will somehow manipulate the menu offers so that few borrowers will borrow without prepayment penalties. But the prepayment-penalty loans might only be allowed in the menu so long as they were routinely resisted – say, by 90% of borrowers. Placing equilibrium caps on the proportion of loans that could be made with prepayment penalties is reminiscent of how some literal restaurant menus are regulated. Some restaurant liquor licenses (which also mandate that menus offer both food and drink) are only valid so long as a cap on the liquor proportion of revenue is not exceeded.[87]

2. Avoiding Delayed Repayments

In addition to promoting prepayments, policymakers may also want to deter borrowers from delaying their repayment of indebtedness. Self-control problems can lead to the problem of borrowers rolling over payday loans or revolving high-priced credit card balances.[88] Again, it might be possible to tempt borrowers at the beginning of their indebtedness in ways that might reduce the likelihood that they will subsequently delay repayment. Regulating the menu of offered loans might again be used to produce reactive behavioral effects. For example, a law might require car-loan lenders that traditionally offer 5 years loans to also offer 10-year loans with tempting lower monthly payments. A $20,000 loan with a 6% A.P.R. would cost $386 per month with 5-year term, but only $222 per month with a 10-year term. Borrowers who resisted the lower monthly payment temptation initially might be more likely sustain full monthly payments than borrowers who had not turned initially turned down the initial offers. Both types of borrowers would sustain penalties for making partial payments, but the borrower who had resisted the initial temptation might feel an additional cognitive loss of having turned down the right to pay a lower amount without penalty.

Alternatively, the law might require what some mortgage lenders already do and offer a menu where some mortgages include up front points in exchange for lower interest rates. For example, the law might encourage lenders to offer a menu of mortgages with and without points. The mortgages with points (and lower interest rates) would be the temptation offers. Resisting

---

[86] This example illustrates why the regulation of contractual menu can non-trivially impact social welfare. See Ian Ayres, Menus Matter, 73 UNIVERSITY OF CHICAGO LAW REVIEW 3 (2006).

[87] For example, California's Type 47 restaurant liquor license requires that permit holders act as a "bona fide eating place." LICENSE TYPES, California Department of Alcoholic Beverage Control (2019), https://www.abc.ca.gov/licensing/license-types/.

[88] Carol C. Bertaut, Michael Haliassos & Michael Reiter, CREDIT CARD DEBT PUZZLES AND DEBT REVOLVERS FOR SELF CONTROL, 13 Review of Finance 657 (2009).

these mortgages only makes economic sense if the borrower plans to repay relatively quickly. And indeed, enlightened regulation might try to make sure that borrowers understand the break-even time period – that the no-point mortgage is less expensive only if the borrowers succeeds in repaying the loan before a set date.

If borrowers resist the temptation of paying points for lower interest, they might end up repaying the loan faster (than borrowers who had not had turned down the opportunity for a lower interest rate). The resisting borrowers might repay faster not only because their loan has a higher (no-points) interest rate, but because slower payment would make their decision not to buy a lower interest loan an error. As always, reactive carrots can fail if too many agents succumb to the temptation. Enlightened regulation, however, could again only allow the points for lower interest mortgage option if relatively few borrowers opted for it.

3. Deterring Equity Stripping

Another barrier to retirement saving has been the rise of cash-out refinancing. People who build up equity in their home over time as they pay off their original mortgage are often tempted to strip the equity by refinancing so as to obtain cash that can be used on pre-retirement consumption.[89] Before cash-out refinancing became so prevalent, home equity was an important source of savings.[90] People who paid off their mortgages and owned their houses outright had a substantial asset that could be sold when they were ready to downsize. Now homeowners who pay off a bit of their mortgage are barraged with offers to borrow against this new found equity. Mian and Sufi have estimated that cash-out refinancing contributed to the mortgage crisis of 2009 because a much larger proportion of homes were leveraged.[91]

Policymakers who want to help citizens save for retirement (and reduce systemic risk) might well try to put this cash-out genie back in the bottle and recreate the dynamic where more people succeed over the course of their working life in paying off their home mortgage. One way to do this might be to combine a traditional (active) carrot with a reactive temptation. Tax law might be changed so that borrowers can only receive mortgage interest deductions if they agree to treat home their mortgage equity savings like their 401(k) savings. Like 401(k) savings, equity savings that were cashed out before retirement would be subject to a 10% penalty unless the borrower qualified for the hardship exception.

Reactive incentives could be added to this traditional tax incentive to reduce the likelihood that borrowers would (notwithstanding the tax penalty) still choose to cash out equity for pre-retirement consumption. Borrowers might be offered a one-time option to cash out their equity without any penalty early on during their mortgage repayment (when they wouldn't have very much equity built up). Borrowers who resisted the temptation to cash out penalty free might find it easier exercise self-control and stay the course when they become subject to the penalty. Again,

---

[89] Atif R. Mian & Amir Sufi, HOUSE PRICES, HOME EQUITY-BASED BORROWING, AND THE U.S. HOUSEHOLD LEVERAGE CRISIS, NBER Working Paper (2009), https://www.nber.org/papers/w15283.
[90] Id.
[91] Id.

it is possible that borrowers would feel especially foolish paying the penalty when they didn't take advantage of the earlier opportunity to strip some of the equity without any penalty.

Stepping back, it is important to emphasize that we have only raised a series of possible policy implication. Prudent policy-setting would suggest experimentation to assure that the desired reactive behavior effect occurs in each situation.[92] Moreover, reasonable people might disagree with some of our policy objectives. Nudging people toward more timely payment might end up hurting individuals who would have been better off pursuing a more delayed or more flexible repayment pattern. Nonetheless, this section has brought into view a variety of possible policy interventions that could not be seen without the lens of reactive incentives.

F. Saving

Saving money for retirement is another context where self-control problems may reduce well-being. Putting aside money for retirement is after all a form of delayed consumption. Time-inconsistent, present-biased behaviors can undermine the ability of individuals to carry through on plans to save.[93] The behavioral economist David Laibson refers to these present-biased behaviors as "golden egg" problems because some individuals, like the couple in Aesops's fable, have difficulty metaphorically resisting the urge to killing the goose that lays the golden eggs.[94] Federal and state policies use a variety of traditional incentives to induce people to save for retirement. Indeed, economists have recently estimated that the United States "spends more than $100 billion per year on tax subsidies for retirement savings accounts such as 401(k)s and IRAs."[95] Federal law also deploys contingent punishments in the form a 10% ERISA penalty for any early withdrawals from 401(k) plans that do not meet hardship requirements.[96]

Here we discuss two different ways that reactive incentives might be used to enhance retirement savings – first by enhancing the amounts that individuals contribute to their 401(k) accounts, and second, by reducing the amount of dissaving or "leakage" that takes place before individuals retire.

With regard to enhancing contributions, there is a substantial literature showing that retirement plans with non-participation defaults tend to produce lower participation rates than plans where employees by default are enrolled and have contributions deducted from their take-home

---

[92] Michael Abramowicz, Ian Ayres & Yair Listokin, RANDOMIZING LAW, 159 University of Pennsylvania Law Review 929 (2011).

[93] Lin Zhang, Saving and retirement behavior under quasi-hyperbolic discounting, 109 Journal of Economics 57 (2013).

[94] David Laibson, GOLDEN EGGS AND HYPERBOLIC DISCOUNTING, 112 The Quarterly Journal of Economics 443 (1997).

[95] Raj Chetty, et al., Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark, 129 The Quarterly Journal of Economics 1141 (2014); Raj Chetty et al., Active vs. passive decisions and crowd-out in retirement savings accounts (2014), http://www.rajchetty.com/chettyfiles/ret_savings.htm.

[96] RETIREMENT TOPICS – EXCEPTIONS TO TAX ON EARLY DISTRIBUTIONS, IRS (2019), https://www.irs.gov/retirement-plans/plan-participant-employee/retirement-topics-tax-on-early-distributions.

pay.[97] Shlomo Benartzi and Nobel Laureate Richard Thaler have championed the use of a "save more tomorrow" approach that encourages employees to agree in advance to increase the percentage of future salary contributions as they receive raises so that their take-home pay never decreases.[98] The Pension Protection Act of 2006 encourages companies to adopt a version of this approach as their plan default: automatically enroll employees in their 401(k) plans at no less than a three percent savings rate and automatically escalate the rate by at least one percent per year for at least three years.[99]

Automatic enrollment with save-more-tomorrow contribution rates can expect to produce more contributions than a no-participation default. But with save-more-tomorrow, employees are still free to opt out at any point in time. Resisting the temptation to opt out initially might make it easier for them to opt out later on. But a problem with the save-more-tomorrow default is that it doesn't make the act of resisting temptation particularly salient. An affirmative choice rule that forces new employees, as a condition of employment, to choose between a save-more-tomorrow option and the short-sighted option of not participating (which would increase the employee's immediate take-home pay) might provide a more powerful self-signal to help increase employee's subsequent self-control. Indeed, for plans that include matching employer contributions, it might even perversely help employees if the employer made a one-time offer of cash if the employees eschewed participating in the retirement plan. Employees who make an affirmative choice to resist the temptation of immediate cash may feel badly if they later cave in and opt out of participating for nothing.

Reactive incentives might also be deployed to dampen the tendency of employees to dissave from their 401(k) accounts. "Leakage" of 401(k) savings has been identified as a substantial impediment to accumulating retirement nest eggs. Participants too often withdraw their tax deferred savings for non-hardship reasons and are hit with a 10% withdrawal penalty. Alicia Munnell has estimated that about 1.5% of assets leak from 401(k) accounts per year and that over the course of a working life this can reduce accumulated account assets by 25%.[100]

A reactive incentive might be deployed to deter this withdrawal leakage. For instance, after a few months of plan participation, plans might provide a one-time offer to withdraw as much from the account as the individual would like with a bonus from the plan to cover the early withdrawal penalty. Participants who resist this offer to withdraw costlessly may find it more difficult to withdraw money later on when they are subject to the withdrawal penalty. In contrast to the "save more tomorrow" idea discussed above,[101] reactive theory suggests that a "save less today" temptation, if resisted, might also play a role enhancing retirement savings.

---

[97] Brigitte C. Madrian & Dennis F. Shea. THE POWER OF SUGGESTION: INERTIA IN 401 (K) PARTICIPATION AND SAVINGS BEHAVIOR, 116 The Quarterly journal of economics 1149 (2001). (also finding that defaults have a significant impact on contribution rates).

[98] Richard H. Thaler & Shlomo Benartzi, SAVE MORE TOMORROW™: USING BEHAVIORAL ECONOMICS TO INCREASE EMPLOYEE SAVING, 112 Journal of political Economy S164 (2004).

[99] Pension Protection Act of 2006, Pub. L. 109–280, § 902(a), 120 Stat. 780 (2006).

[100] Alicia H. Munnell, 401(k)/IRA Holdings in 2013: An Update from the SCF, 14-15 CTR. FOR RETIREMENT RES. AT B.C. 4–5 (2014). See also Ian Ayres & Jacob Hacker, Social Security Plus, 26 Elder Law Journal 261 (2019);

[101] *See* Thaler and Benartzi, *supra* note 76, at 30.

G. College Completion

People can also save by investing in human capital. Enrolling in higher education can increase human capital, but often entails deferred consumption as individuals delay entering the job market (and take on debt). From a behavioral perspective, it shouldn't be surprising that pursuing higher education can pose problems of self-control. A substantial number of people who matriculate to college do not end up graduating. Only 60% of students graduate college within six years of matriculating in their first full-time program.[102] The problem of non-completion is at times particularly pronounced for minority students. While 64% of white students graduate within the six years of enrollment, only 40% of black students do.[103]

Colleges and other institutions of higher education might explore using reactive incentives to reduce these drop-out rates. Making the first weeks of school artificially more difficult than subsequent weeks might produce reactive stick effects. Surviving these harder initial weeks when new students are still flush with enthusiasm might increase their ability to persist in subsequent periods when initial enthusiasm fades. The existing practice at many universities of more severely grading the initial writing assignments may produce reactive benefits for at-risk students.[104]

Alternatively, we might consider using tuition-back guarantees as reactive carrots to reduce subsequent attrition. Akhil Amar and Ian Ayres have argued that law schools should offer tuition refunds to students shortly after they receive their first-year grades.[105] Their proposal is driven in large part by the consumer-protection objectives of alleviating the cost to students who mistakenly matriculated and providing schools with stronger incentives to admit qualified students and serve their needs. Their tuition-back offer is a traditional ("active") carrot intended to induce some acceptance, but time-limited offers of tuition refunds, made shortly after students arrive on campus might produce reactive behavioral effects. A student who turned down the refund offer might be less likely to subsequently quit without the refund. By offering during the first several weeks of school, students may still be excited enough that they are unlikely to quit.

Higher education might be a context where enlightened policy might include some selection when deciding which students receive the offer. There may be certain types of students for whom the acceptance rate of an initial offer is so high that the offer on net increases the dropout rate. For example, it might be that the intended reactive effect is strongest for students with an intermediate likelihood of subsequent attrition. Other groups might be made the offer in order to accelerate their dropping out – possibly because, a la Amar and Ayres, the program of study was ill advised—or because accelerating their early attrition will reduce the contagion effect on remaining students.

---

[102] FAST FACTS: GRADUATION RATES, National Center for Education Statistics (2019), https://nces.ed.gov/fastfacts/display.asp?id=40.

[103] Table 326.10: Graduation rate from first institution attended for first-time, full-time bachelor's degree- seeking students at 4-year postsecondary institutions, by race/ethnicity, time to completion, sex, control of institution, and acceptance rate: Selected cohort entry years, 1996 through 2011, National Center for Education Statistics (2019), https://nces.ed.gov/programs/digest/d18/tables/dt18_326.10.asp.

[104] Wayne A. Grove & Tim Wasserman, The Life-Cycle Pattern of Collegiate GPA: Longitudinal Cohort Analysis and Grade Inflation, 35 JOURNAL OF ECONOMIC EDUCATION 162 (2004).

[105] Akhil Reed Amar & Ian Ayres, PAYING STUDENTS TO QUIT LAW SCHOOL, Slate (November 18, 2011), https://slate.com/news-and-politics/2011/11/law-schools-should-pay-students-to-quit.html.

Our goal in this Part has been merely to sketch possibilities. Our focus has been reactive carrots but we imagine that there would also be opportunities to beneficially deploy reactive sticks. The Part has also made more explicit how reactive thinking enriches and complicates questions of default choice and menu regulation. Our goal here is to begin a conversation and spur further empiricism.

CONCLUSION

This article has tried to make both a theoretical and empirical case that sunk opportunity costs can impact the future behavior of people who decide to forego present temptations. Behavioralists have long understood the powerful impact that sunk costs can have on future behavior. For example, the governments of France and Britain continued to invest in the Concorde, long after it was clear that the project was not financially viable because they had "[t]oo much invested to quit."[106] But we show here that sunk *opportunity* costs can also make it easier for people to stay the course and conform their future behavior to their past behavior. These behavioral effects create the possibility for two new incentive categories. People who resist the temptation to accept a "reactive carrot" experience a sunk opportunity cost in foregoing the incentive. People who resist a "reactive stick" and persist in a short-term behavior in spite of a contingent punishment experience a sunk opportunity cost by foregoing the opportunity to avoid the punishment. Both reactive carrots and reactive sticks thus aim to have subjects defy the contingent incentive and take what, because of the incentive, is the harder path. Each of these sunk opportunity cost incentives can impact future behavior because initial resistance can both help people "self-signal" information about their own resolve and increase the costs of subsequently acting inconsistently.

Reactive incentives open up new possibilities for policymakers. Where self-control problems impede individuals from pursuing their own self-interest, reactive incentives provide a new kind of choice architecture to enhance personal resolve. The last section suggested ways that reactive carrots might help people to kick a bad habit or put their financial house in better order.

Our discussion of policy interventions also shows how the possibility of reactive effects can enrich behavioral theories of foundational components of legal choice. Reactive effects have implications for the choice between mandatory and default rules, as well as when to deploy affirmative choice rules. In particular, our analysis suggests that policy makers would do well to pay more attention to the salience of foregone opportunities. This article's initial exploration suggests that salience of foregone opportunities might be seen as a policy choice that can be impacted by implementing regulation. Reactively-informed policymaking is likely to lead to better theories of contractual menus and to novel interventions – such as our proposal that the law at time allow sellers to add tempting items to their menu of offers if and only if most buyers resist the temptation.

We began this article with Shakespeare's tale of Henry V financially incentivizing his soldiers to quit and suggested that this stratagem might have caused them to fight harder at the

[106] Allan I. Teger, TOO MUCH INVESTED TO QUIT (1980).

battle of Agincourt. Appreciating that foregone opportunities can affect future behavior can guide us to new readings of literature. Consider for example, Robert Frost's poem, *The Road Not Taken,*[107] the title of which focuses our attention on the foregone opportunity. A traditional reading of the poem is that choosing the road "less travelled by" has impacted the speaker's life because different paths inevitably lead to different experiences. But the power of sunk opportunity costs suggest an additional possible reading. It may be that the speaker's subsequent life was impacted because she forewent the other path. Through the lens of this article, we can see that the speaker's affirmative choice of the less-travelled road might have led her to make different life choices than she would have made if there had been only the one less-travelled road in those yellow woods.[108]

Frost's protagonist imagines the future impact of the road not taken "[s]omewhere ages and ages hence," but the future impact a foregone opportunity is explicitly depicted in William Styron's novel, *Sophie's Choice*. As with Frost's poem, the title itself alludes to the centrality of opportunity cost to the narrative. To choose is necessarily to pay some cost of foregone opportunity. In Styron's novel, we learn that the defining moment in Sophie's past was when she was forced to choose which one of her two children would be immediately sent to be killed in a crematorium and which would be sent to the children's camp with the possibility of survival. The narrative suggests that Sophie's decision to forego the opportunity to protect her daughter, Eva, and instead to send her son, Jan, to the children's camp had devastating consequences in Sophie's later life, including her decision to ultimately take her own life much later after she had emigrated to the United States. It is hard to conjure a more powerful exemplar of the impact of foregone opportunity. If the sadistic officer had directly taken Sophie's daughter to her death, without forcing Sophie to decide, Sophie might have been less devastated and made different subsequent decisions for the rest of her life.[109] Sophie's choice taught her something about herself – something that Sophie couldn't live with.

We end with this tragic tale as a cautionary note. This article has laid out the idea that carrots and sticks can at times induce reactive behavioral effects. People may resist the short-term incentive end up to do more of some behavior that was initially disincentivized. Offering to pay someone to smoke a cigarette now might reduce the chance that they smoke in the future. But these boomeranging reactive effects, as our experiments suggest, will not always obtain. Moreover, Sophie's Choice shows that sunk opportunity costs might induce a wide range of subsequent behavioral effects that are not exclusively beneficial. Reactive behavioral incentive

---

[107] Robert Frost, THE ROAD NOT TAKEN, *in* The Road Not Taken: A Selection of Robert Frost's Poems (Robert Frost & Louis Untermeyer, ed. 1991).

[108] Indeed, this reactive reading of the poem provides a new interpretation of the poem's last lines. The speaker famously concludes:

> Two roads diverged in a wood, and I—
> I took the one less traveled by,
> And that has made all the difference.

It would be a standard interpretation to read "that" as referring to "travelling on the less traveled road." But it is possible that the speaker believes that *the making of a choice*, rather than the path itself, has made the difference. This possibility is strengthened by the speaker earlier telling us that the chosen road might not really have been less traveled (the previous passersby had "worn them really about the same" and "both that morning equally lay in leaves"). Resisting the temptation to follow the crowd on what the speaker somewhat arbitrarily labels as the more travelled path may have allowed the speaker – in the terms of our theory – to self-signal something that she carried with her into the future.

[109] Jennifer Gerarda Brown, The Sophie's Choice Paradox and the Discontinuous Self: Two Comments on Wertheimer, 74 Denver University Law Review 1255 (1996).

hold promise as new policy tool, but the variety of potential behaviors that sunk opportunity costs might induce counsels toward proceeding experimentally and with some degree of care.