Theoretical Economics 15 (2020), 279–317 1555-7561/20200279

# Willpower and compromise effect

Yusufcan Masatlioglu
Department of Economics, University of Maryland

Daisuke Nakajima
Department of Economics, Otaru University of Commerce

Emre Ozdenoren
Department of Economics, London Business School

This paper provides a behavioral foundation for modeling willpower as a limited cognitive resource that bridges the standard utility maximization and Strotz models. Using the agent's ex ante preferences and ex post choices, we derive a representation that captures key behavioral traits of willpower-constrained decision making. We use the model to study the pricing problem of a profit-maximizing monopolist who faces consumers with limited willpower. We show that the optimal contract often consists of three alternatives and that the consumer's choices reflect a form of the "compromise effect," which is induced endogenously.

Keywords. Willpower, self-control, temptation, compromise effect, exploitative Contracts.

JEL classification. D86, D91.

## 1. Introduction

Standard theories of decision making assume that people choose what they prefer and prefer what they choose. However, introspection suggests that implementation of choice may not be automatic, and there is often a wedge between preferences and actual choices. Recently, psychologists and economists have emphasized the lack of self-control in decision making as an important reason for this wedge.[1] When people face temptation, they make choices that are in conflict with their commitment preferences.

Yusufcan Masatlioglu: yusufcan@umd.edu
Daisuke Nakajima: nakajima@res.otaru-uc.ac.jp
Emre Ozdenoren: eozdenoren@london.edu

We would like to thank Kfir Eliaz, Collin Raymond, Larry Samuelson, Joel Sobel, Tomasz Strzalecki, and Neslihan Uler. We are grateful to the three anonymous referees for the insightful comments and suggestions, which greatly improved the paper. Parts of this research were conducted while Masatlioglu was visiting the Institute for Advanced Studies at Hebrew University. Nakajima appreciates the financial support provided by Japan Society for the Promotion of Science (JSPS KAKENHI Grant Number 26780113).

[1] Models of self-control problems include quasi-hyperbolic time discounting (e.g., Laibson (1997), O'Donoghue and Rabin (1999)), temptation costs (e.g., Gul and Pesendorfer (2001, 2004)), and conflicts between selves or systems (e.g., Shefrin and Thaler (1988), Bernheim and Rangel (2004), Fudenberg and Levine (2006)).

© 2020 The Authors. Licensed under the Creative Commons Attribution-NonCommercial License 4.0. Available at http://econtheory.org. https://doi.org/10.3982/TE3086

People do not always succumb to temptation and are sometimes able to overcome temptations by using cognitive resources. This ability is often called willpower. There is a growing experimental psychology literature demonstrating that willpower is a limited resource, and it is more than a mere metaphor (e.g., Baumeister and Vohs (2003), Faber and Vohs (2004), Muraven et al. (2006)). Motivated by these experiments, economists have used limited willpower to explain patterns of consumption over time (Ozdenoren et al. (2012), Fudenberg and Levine (2012)).

Our goal is to characterize a simple and tractable model of limited willpower that is suitable to study a wide range of economic problems. The characterization uses a novel data set given by the agent's ex ante preferences ($\succsim$) and ex post choices ($c$).[2] Temptation and self-control have been studied using the preference over menus framework pioneered by Kreps (1979) and Gul and Pesendorfer (2001), where the agent's second-period choices are inferred from his preferences over menus. In contrast, in our framework the modeler *directly* observes both components of the data, namely, ex ante preferences and ex post choices. The menu preferences framework provides a powerful tool to elicit a very rich set of behaviors at the ex ante stage but also has limitations, as highlighted by Spiegler (2013). For example, it assumes that the agent can predict the degree of his future temptations and his future choices.[3] Although it is useful to study the sophisticated benchmark, many interesting applications with policy-related implications emerge from the assumption of naivety.[4] By relying on directly observed rather than predicted behavior, our model is tailored to study these applications. This is illustrated in Section 5, where we apply the model to monopolistic contracting where consumers have limited willpower but are unaware of their willpower problems.

In Section 2, we propose two models. We refer to the more general of the two as the limited willpower model. We refer to the second and simpler model as the constant willpower model, which is based on three ingredients. The first, commitment utility $u$, represents the agent's commitment preferences. The other two ingredients are temptation values $v$ and the willpower stock $w$, which jointly determine how actual choices depart from what commitment utility would dictate. The key to determining the actual choice is the willpower constraint. This constraint is determined by the most tempting available alternative and the willpower stock. The agent is able to consider an alternative $x$ in $A$ if he can overcome the temptation, i.e., $\max_{y\in A} v(y) - v(x) \leq w$. Otherwise, he does not have enough willpower to choose this alternative. He then picks the alternative that maximizes his commitment utility from the set of alternatives that satisfies the willpower constraint. Formally, the ex post choice from a set $A$ is the outcome of the

[2]Ahn and Sarver (2013) also use two kinds of behavioral data. As opposed to ours, their data include both the menu preferences and the *random* ex post choices from menus.

[3]Ahn et al. (2019) combine menu preferences with data on ex post choice to detect the agent's degree of naivety.

[4]Evidence suggests that consumers who have self-control problems are often, at least partially, naive about this fact. For example, there are "hot–cold empathy gaps" where individuals are not able to recognize the intensity of temptation, or other visceral urges, at an ex ante state. Loewenstein and Schkade (1999) review several studies that find that people tend to underestimate the influence on their behavior of being in a hot state, such as hunger, drug craving, curiosity, and sexual arousal, among others.

following maximization problem:

$$\max_{x\in A} u(x) \quad \text{subject to} \quad \max_{y\in A} v(y) - v(x) \leq w.$$

The more general limited willpower model is similar to the constant willpower model except that it allows the willpower stock to depend on the chosen alternative.

The constant willpower model bridges the standard utility maximization and Strotz models. When the willpower stock is very large, the willpower-constrained agent behaves like a standard agent who chooses the most preferred alternative (according to $u$). When the willpower stock is lower, the constraint starts to bind and a wedge between preferences and choices appears—the agent can only choose alternatives that are close enough, in terms of temptation, to the most tempting alternative. At the other extreme, when the willpower stock is very low, the agent behaves like a Strotzian agent who always succumbs to temptation.[5] Note that the agent's choice will satisfy WARP in the two extreme cases for different reasons. While in the former, choices reflect only the ex ante preference, in the latter temptation ranking alone determines the choices. In comparison, in the limited willpower case, the agent might not have enough willpower to choose the least tempting alternative but might have enough willpower to choose the moderately tempting alternative. Such choices reflect a compromise between the ex ante preference and the temptation ranking and violate WARP. Similar examples of WARP violations are featured in Fudenberg and Levine (2006), Dekel et al. (2009), Noor and Takeoka (2010), Chandrasekher (2010).[6] Lipman and Pesendorfer (2013) note that such behavior can provide a new perspective on the compromise effect.

The costly self-control models in Gul and Pesendorfer (2001), Noor and Takeoka (2010, 2015), Grant et al. (2017) are closely related to the constant willpower model. In the costly self-control models, the agent experiences a utility cost that can be a convex function of the difference between the temptation values of the most tempting alternative and the chosen alternative from a menu.[7] One can express a willpower constraint as the limiting case of a convex cost function that is zero if the temptation difference is less than the willpower stock and infinite otherwise. Hence, just as the Strotz model can be viewed as a limiting case of costly self-control with linear costs (Gul and Pesendorfer (2001)), the constant willpower model is a limiting case of costly self-control with convex costs.[8] Similar to the Strotz model, constant willpower is an important case methodologically and for applications, and its behavioral characterization uses a different domain, primitives, and axioms. In Section 4, we compare the linear and convex self-control models in terms of their implied second-stage choice.

[5] Strotz (1955) proposed a model of changing tastes with consistent planning. In a framework where the agent has preferences over finite decision trees, Gul and Pesendorfer (2005) provide an axiomatization of the Strotz model of consistent planning.

[6] In a menu preference framework, Dillenberger and Sadowski (2012) note that when agents anticipate experiencing guilt or shame when they deviate from a social norm, their choices can also violate WARP.

[7] The idea that self-control costs can be convex has been introduced by Fudenberg and Levine (2006).

[8] In the remainder of the paper, we refer to the costly self-control models with linear and convex costs as linear and convex self-control models, respectively.

In Section 5, we solve the pricing problem of a profit-maximizing monopolist who faces consumers with constant willpower. Monopolistic contracting is a key application of self-control models and was first studied by DellaVigna and Malmendier (2004). Our treatment is similar to Eliaz and Spiegler (2006) who study unconstrained contracting with dynamically inconsistent consumers. We consider a two-period model of contracting between a monopolist and a consumer. In the first period, the monopolist offers the consumer a contract that the consumer can accept or reject. If the consumer accepts the contract, in the second period, he chooses an offer from the contract and pays its price to the monopolist. We assume that both parties are committed to the contract once accepted.

This framework fits into many real world situations, such as signing up for a phone plan, gym membership, or a credit card, making a hotel reservation, or renting a car, among others. In all of these examples, consumers often agree to a contract that specifies a basic level of consumption that can be "upgraded" at the time of consumption. It has been noted that these contracts can be exploitative. The literature has focused on contracts that offer two alternatives, which we call "indulging contracts." At the time of signing an indulging contract, the consumer believes that he will consume a basic and cheap alternative but ends up choosing the more expensive alternative at the time of consumption. However, in practice, it is common to see contracts that offer more than one upgrade to the consumer. For example, when renting a car, consumers often choose a basic car when signing the contract, but they are offered multiple levels of upgrades when picking up the car. Our analysis shows that such contracts can be optimal for naive consumers who have limited willpower or convex self-control costs.

Specifically, we show that the optimal contract consists of three alternatives and that the consumer's choices reflect a form of the "compromise effect," which is induced endogenously by the contract offered by the monopolist. Interestingly, the optimal contract includes a tempting alternative that neither the consumer nor the firm believes would be chosen from an ex ante perspective and is indeed not chosen ex post. When the consumer has limited willpower, we show that indulging contracts can never be optimal. We also show that for low levels of willpower, the monopolist exploits the consumer as if he has no willpower. For high enough willpower, the monopolist sells the efficient alternative at an exploitative price. When the consumer's willpower is very high, there is no exploitation and the monopolist uses a commitment contract. Profits are lower and the consumer is better off if he has more willpower. In Section 5.3, we compare these results with those obtained under other related models of self-control.

## 2. Model

The agent's ex ante preferences $\succsim$ are over a finite set of alternatives $X$. These preferences can be interpreted as the agent's commitment preferences. The agent's ex post choices are captured by a choice correspondence $c$ that assigns a nonempty subset of $A$ to each $A \in \mathcal{X}$ where $\mathcal{X}$ is the set of all nonempty subsets of $X$.

| | $w=1$ | $w=3$ | $w=5$ |
|---|---|---|---|
| $c(e,m,l)$ | $l$ | $m$ | $e$ |
| $c(e,m)$ | $m$ | $e$ | $e$ |

TABLE 1. Choices for different levels of the willpower stock.

We say that $(\succsim, c)$ has a *limited willpower representation* if there exists $(u, v, w)$ where $u : X \to R$ represents preference $\succsim$ and $c$ is given by

$$c(A) = \arg\max_{x \in A} u(x) \quad \text{subject to } \max_{y \in A} v(y) - v(x) \leq w(x)$$

where $v : X \to R$ captures the temptation values and $w : X \to \mathbb{R}_+$ is the willpower function. If $w$ is a constant function, we call it a *constant willpower representation*.

In the standard model in which there is no willpower problem, a decision maker chooses the alternative that maximizes the commitment utility, $u$, from any menu. An agent who has constant willpower also maximizes $u$ but faces a constraint. The willpower requirement of alternative $x$ is given by the difference between the temptation value of the most tempting alternative on the menu, $\max_{y \in A} v(y)$, and the temptation value of $x$. The agent can choose $x$ only if its willpower requirement is less than the willpower stock, $w$. Otherwise, he does not have enough willpower to choose this alternative. Notice that the willpower requirement is menu dependent. This is because willpower depletion not only depends on how tempting the chosen alternative is but also on the most tempting alternative on the menu.

As a simple example, suppose that a consumer is offered three alternatives at rental car pick-up: an economy car ($e$), a mid-size sedan ($m$), and a luxury car ($l$). Suppose ex ante that $e \succ m \succ l$, and $v(e) = 0$, $v(m) = 2$, $v(l) = 4$. Table 1 shows the agent's choices from two sets, $\{e, m, l\}$ and $\{e, m\}$, for varying levels of willpower stock.[9] When the willpower stock is high, $w = 5$, the agent chooses according to his ex ante preferences. When the willpower stock is low, the agent also behaves like a standard preference maximizer, except that he chooses the most tempting alternative. When the willpower stock is intermediate, $w = 3$, then the model has interesting implications: decisions can be driven by a compromise between the ex ante preference and temptation. To see this, suppose that all three alternatives are available. The agent is not able to choose $e$ since $v(l) - v(e) = 4 > 3 = w$. In this case, he chooses the compromise alternative $m$ since $v(t) - v(b) = 2 < 3$. However, when only $e$ and $m$ are available, there is no need to compromise (since $v(m) - v(e) = 2 < 3$), and the agent chooses $e$.

## 3. BEHAVIORAL CHARACTERIZATION

In this section, we introduce the axioms and provide two representation theorems. Our first axiom is standard.

[9]We will abuse the notation and write $c(x, y, \dots)$ instead of $c(\{x, y, \dots\})$. Similarly, we omit braces and write $A \cup x$ instead of $A \cup \{x\}$.

A 1. *$\succ$ is a complete, transitive, and asymmetric binary relation.*

The second axiom, Independence from (Unchoosable) Preferred Alternative (IPA), says that better options that are not chosen can be removed without affecting the actual choice.

A 2. *(IPA) If $x \succ y$ and $y \in c(A \cup x)$ then $c(A) = c(A \cup x)$.*

This axiom can be viewed as a relaxation of WARP. Recall that WARP allows any unchosen alternative to be dropped without affecting actual choices. In contrast, IPA allows only preferred unchosen alternatives to be dropped without affecting actual choices. IPA is based on the intuitive notion that when a tempting alternative is also the most preferred available alternative, it should be chosen. Hence, any unchosen alternative that is strictly preferred to the chosen one must have a relatively low temptation value. IPA states that dropping such alternatives should not affect the actual choice. Let us revisit the example in Section 2 with three alternatives: $e$, $m$, and $l$ with $e \succ m \succ l$. Suppose that a mid-size sedan is chosen when all three options are available, i.e., $m = c(e, m, l)$. This means that the most preferred alternative ($e$) is not chosen, and hence, it is not the most tempting alternative and is irrelevant in the sense that dropping it from the menu should not affect the choice behavior of the agent. That is, we must have $c(e, m, l) = c(m, l)$. On the other hand, it is possible that removing $l$, the least preferred alternative, might influence the choice. If $m$ is not as tempting as $l$, the agent can choose the best alternative $e$ when $l$ is removed, i.e., $e = c(e, m) \neq c(e, m, l)$. Hence, WARP is not satisfied in the presence of limited willpower.

Axiom 1 and Axiom 2 together imply that the choice is unique, $|c(A)| = 1$ for all $A$. To see this, assume that $x$ and $y$ are chosen from $A$. By Axiom 1, assume $x \succ y$ without loss of generality. Then Axiom 2 implies that $x \notin c(A)$.

The next axiom is Choice Betweenness (CB): the choice from the union of two sets is "between" the choices made separately from each set with respect to preference.

A 3. *(Choice Betweenness) If $c(A) \succsim c(B)$, then $c(A) \succsim c(A \cup B) \succsim c(B)$.*

To understand this axiom, take the union of two choice sets $A \cup B$, and w.l.o.g., suppose that $A$ contains one of the chosen alternatives from $A \cup B$. Consider two (not necessarily mutually exclusive) cases. First, suppose that $A$ contains the most tempting item in $A \cup B$. In this case, the agent should not be able to choose a strictly better alternative from $A$ (since he needs to overcome the same temptation from $A \cup B$ as from $A$) but should still be able to choose the alternative originally chosen from $A \cup B$, i.e., $c(A) \sim c(A \cup B)$.[10] Note that in this case the axiom is automatically satisfied since $c(A \cup B)$ must be in between $c(A)$ and $c(B)$ in terms of preference. As a second case, suppose that $B$ contains the most tempting item in $A \cup B$. In this case, the agent should

[10]Implicit in these arguments is that only the most tempting alternatives matter in influencing the agent's choices. Clearly, this is also the case in the representation since only the alternative with the highest $v$ value matters in determining which alternatives are choosable from a choice set.

be able to choose at least as preferred an alternative from $A$ as he can from $A \cup B$ since he needs to overcome a weaker temptation from $A$. Moreover, the alternative chosen from $B$ cannot be strictly preferred since the most tempting alternative is contained in $B$. Thus, the axiom should also be satisfied in this case.

In fact, $c(A \cup B)$ can be strictly between $c(A)$ and $c(B)$. Continuing with our earlier example, let $A = \{e, m\}$ and $B = \{l\}$. Recall that both $e$ and $m$ are strictly better than $l$, so $c(A) \succ c(B)$. The choice from all three options, $m$, is strictly better than $l$, the worst alternative, so $c(A \cup B) \succ c(B)$. Moreover, from the set $A$, $e$ is chosen; thus, $c(A) \succ c(A \cup B) \succ c(B)$.

Next, we present our first representation theorem.

THEOREM 1. $(\succ, c)$ *satisfies A1–A3 if and only if it admits a limited willpower representation.*

Theorem 1 provides a characterization of limited willpower with three simple and intuitive behavioral postulates in a novel domain. Although the axioms are simple, the proof of the theorem uses a new approach in this literature that relies on the concept of interval orders, so we provide a sketch of the proof next. Suppose that $P$ is a binary relation. Let $\Gamma_P(A)$ be the set of undominated alternatives according to $P$ in $A$. The key result we use is that a binary relation $P$ is an interval order (i.e., it is irreflexive and $xPb$ or $aPy$ holds whenever $xPy$ and $aPb$) if and only if there exist functions $v$ and $w$ such that

$$\Gamma_P(A) = \left\{x \in A : \max_{y \in A} v(y) - v(x) \leq w(x)\right\}.$$

In our framework, a natural binary relation to consider is $\rhd'$. We say that $x \rhd' y$ if $y \succ x = c(x, y)$. Whenever this is the case, we say that $x$ "blocks" $y$. We interpret this situation as $x$ being much more tempting than $y$ so that $x$ is chosen even though $y$ is better with respect to the ex ante preference. This relation is not an interval order because it cannot identify all the blocking pairs. Specifically, there are situations where $x$ is more tempting than $y$ (and hence should block it), but $x$ is better than $y$ according to the ex ante preference. To identify these pairs, we introduce a second binary relation $\rhd''$. We say that $x \rhd'' y$ if $x \succ y$ and there exist $a$ and $b$ such that $a \rhd' y$, $x \rhd' b$, and $a \not\rhd' b$. To see the intuition, assume that $a \rhd' y$ and $a \not\rhd' b$. Then $a$ is more tempting than $y$ but not tempting enough to block $b$. Since $x$ blocks $b$, it must block $y$. Since $x$ is better than $y$, our original relation cannot identify this. Our new relation $\rhd''$ identifies all these cases. Then the bulk of the proof shows that $\rhd'' \cup \rhd'$ is an interval order, which characterizes the temptation ranking and willpower stock. The utility ranking comes from the ex ante preference.

Our next goal is to characterize the constant willpower model. To do this, we need one more assumption. Consider four alternatives $x, y, z, t \in X$. Suppose that $y \succ c(y, z)$, i.e., the agent prefers $y$ to $z$ but is unable to choose it. Intuitively, this means that $z$ is more tempting than $y$. If, in addition, $c(t, z) = t$, then $t$ must also be more tempting than $y$; otherwise, the agent would not be able to choose $t$. If $x \succ c(x, y)$, then the agent prefers $x$ but cannot choose it against $y$ because $y$ is too tempting. Since $t$ is even more

tempting than $y$, the agent should not be able to choose $x$ against $t$ either. This intuitive conclusion would hold for the constant willpower model, but it is not implied by IPA and CB. This is our next axiom, Consistency.

A 4. *(Consistency) Let $y \succ c(y,z)$ and $c(t,z) = t$. If $x \succ c(x,y)$, then $c(x,t) = t$.*

Now, we are ready to state the main representation theorem.

THEOREM 2. *$(\succ, c)$ satisfies A1–A4 if and only if it admits a constant willpower representation.*

The proof of Theorem 2 relies on the concept of semiorders ($P$ is a semiorder if it is an interval order and $xPt$ or $tPz$ for any $t$ whenever $xPyPz$). The key result we use is that a binary relation $P$ is a semiorder if and only if there exists a function $v$ and a scalar $w$ such that

$$\Gamma_P(A) = \left\{x \in A : \max_{y \in A} v(y) - v(x) \leq w\right\}.$$

We use the Consistency axiom to show that one can construct a semiorder by properly modifying the interval order we created in the proof of Theorem 1.

We now discuss how to identify the utility and temptation ranking in our model. The utility ranking is ordinally unique. Although the temptation ranking is not, we now discuss the extent to which it can be identified. As we discussed above, in our model, whenever $y \succ x = c(x,y)$ (i.e., $x \rhd' y$), we can conclude that $x$ is more tempting than $y$ by at least the amount $w(y)$, i.e., $v(x) - v(y) \geq w(y) > 0$. Hence, in these cases, it is safe to claim that $x$ is more tempting than $y$ ($v(x) > v(y)$). This revelation is true for both of our models.

There are in fact more nontrivial revelations about the temptation ranking coming from $\rhd''$. To illustrate this, consider the following data: $y = c(y,t) \succ x = c(x,y) \succ z \succ t = c(z,t)$. This data immediately reveal that $x$ is more tempting than $y$ and $t$ is more tempting than $z$. In addition, in the general model, $x$ must be more tempting than $t$. To see this, assume that $v(t) \geq v(x)$. Then we have

$$v(t) \geq v(x) > v(y) + w(y),$$

which implies that $t = c(y,t)$, which is a contradiction. Therefore, we can conclude that

$$v(x) > v(t) > v(z) \quad \text{and} \quad v(x) > v(y)$$

This example also illustrates the limits of what the choice data are able to reveal about the temptation ranking. Here, we cannot reveal the temptation ranking of $y$ compared to $t$ or $z$.

In the constant willpower model, we can further identify the temptation ranking. For example, $y = c(y,t) \succ x \succ z = c(x,z) \succ t = c(z,t)$ immediately reveals that $t$ is more tempting than $z$, which is more tempting than $x$. However, in the general model, these data cannot reveal the temptation ranking of $y$ relative to all other alternatives. Either

$y$ is the most tempting alternative, or $y$ is the least tempting alternative but with high $w$. On the other hand, in the constant willpower model, it is revealed that $y$ is more tempting that both $z$ and $x$. To see this, assume that $v(z) \geq v(y)$. Then we have

$$v(t) - w > v(z) \geq v(y),$$

which implies that we must have $t = c(y, t)$, which is a contradiction. Therefore, we can conclude that

$$v(y), v(t) > v(z) > v(x).$$

Almost all rankings are identified except the ranking between $y$ and $t$. While there are 24 different possible temptation rankings in this example, only two of them are consistent with our data.

## 4. Related literature

In this section, we compare our model with the models of Gul and Pesendorfer (2001), Noor and Takeoka (2010, 2015), Grant et al. (2017). At the methodological level, these models operate under different domains (lotteries) and primitives (preferences over menus). We consider a generic set of alternatives, and our primitives consist of ex ante preferences ($\succsim$) and ex post choices ($c$). Avoiding any a priori structure on the set of alternatives and focusing on ex post choices enhances the potential testability of our model.

Since there is no axiomatic characterization of self-control driven models in our domain and primitives, a direct comparison is difficult. To facilitate the comparison, we consider ex post (i.e., second-stage) choices for models based on menu preferences. Consider the following general formulation:

$$c(A) = \operatorname*{argmax}_{x \in A} \left\{ u(x) - \psi\left(\max_{y \in A} v(y)\right) \phi\left(\max_{y \in A} v(y) - v(x)\right) \right\} \tag{1}$$

where $u$ represents the ex ante preference $\succ$, $v$ is the temptation value, $\phi$ measures the cost of temptation, which is an increasing function of temptation frustration, $\max_{y \in A} v(y) - v(x)$ and $\psi$ is the level of temptation, which is an increasing function of the highest level of temptation.

To express the limited willpower model in terms of equation (1), we set the cost and level functions as

$$\phi_W(a) = \begin{cases} 0 & a \leq w \\ \infty & a > w \end{cases} \quad \text{and} \quad \psi_W \equiv \text{constant}$$

Gul and Pesendorfer (2001) provide a characterization for a linear self-control model where $\phi_L(a) = a$ and $\psi_L \equiv 1$. Noor and Takeoka (2010) provide two models of convex self-control. In their first model, the cost function is more general than the one in equation (1). In this formulation, the cost of choosing the product in the choice set is a function of both the product and the most tempting alternative. They impose other

assumptions such that this function can be interpreted as a cost function. Their second model is in line with equation (1) where the cost function is strictly increasing and convex with $\phi_C(0) = 0$ and $\psi_C \equiv 1$. Noor and Takeoka (2015) also consider a similar model in which the cost function is menu dependent where $\phi_M(a) = a$ and $\psi_M$ is arbitrary.

Finally, Grant et al. (2017) consider a hybrid model of both costly self-control and limited willpower. In this model, the cost function is a combination of the cost functions of linear costly self-control and limited willpower.

$$\phi_H(a) = \begin{cases} a & a \leq w \\ \infty & a > w \end{cases} \quad \text{and} \quad \psi_W \equiv 1$$

Next, we compare our model with the implied choices generated by the models of Gul and Pesendorfer (2001) and Noor and Takeoka (2010). We impose that $u$ is a one-to-one function so that Axiom 1 is automatically satisfied by all the models.

In the linear self-control model of Gul and Pesendorfer (2001), the agent's ex post choices are given by

$$c(A) = \operatorname*{argmax}_{x \in A} \{u(x) + v(x)\}$$

that satisfy WARP. Moreover, Axiom 2 is violated unless $u + v$ is a one-to-one function. To see this, consider $A = \{x, y\}$ such that $u(x) > u(y)$ and $u(x) + v(x) = u(y) + v(y)$, implying that $x \succ y$ and $c(A) = A$. Hence, $c(A)$ cannot be equal to $c(y)$, violating Axiom 2. Once we assume that $u + v$ is a one-to-one function, it is routine to show that the linear self-control model satisfies Axioms 3 and 4, and hence, it is a special case of the constant willpower model. We can also establish this fact directly from the representations. Take a pair of real functions $(u, v)$ defined on $X$, where both $u$ and $u + v$ are one-to-one. Then the constant willpower model $(u, u + v, w = 0)$ yields the same second-stage choices as the linear self-control model $(u, v)$.

We next consider the convex self-control model of Noor and Takeoka (2010). In this model, ex post choices are given by

$$c(A) = \operatorname*{argmax}_{x \in A} \Big\{ u(x) - \phi\Big(\max_{y \in A} v(y) - v(x)\Big) \Big\}$$

Like the linear model, this model violates Axiom 2 unless $u(x) - \phi(\max_{y \in A} v(y) - v(x))$ has a unique maximizer. In the discussion below, we impose this assumption.

We first illustrate with a simple example that the convex self-control model violates Axiom 3. Suppose that $\phi(a) = a^2$, and consider Table 2.

| | $x$ | $y$ | $z$ | $t$ |
|---|---|---|---|---|
| Utility values ($u$) | 5.5 | 2.5 | 1 | 0 |
| Temptation value ($v$) | 0 | 1 | 2 | 2.5 |

Table 2. Utility and temptation values of $\{x, y, z, t\} \subseteq X$.

It is straightforward to check that $c(x, z) = x$, $c(y, t) = y$ but $c(x, y, z, t) = z$, violating Axiom 3. Note that, in this example, the most tempting alternative is $t$. When alternatives $y$ and $z$ are compared with $t$, they are both ex post ranked above $t$, but their ex ante and ex post rankings do not agree: $y$ is ranked above $z$ according to $u$ but ex post $z$ is ranked above $y$ when $t$ is in the choice set. This type of reversal does not occur in the willpower model because the DM always ranks feasible alternatives, which are ex post ranked above the most tempting alternative, according to the ex ante ranking $u$. However, such reversals can occur in the costly self-control model when the DM experiences a utility cost due to temptation. Note also that violations of Axiom 3 by the convex self-control model require at least four distinct alternatives.

The next proposition generalizes this example to show that such examples can be generated for any convex self-control model with a nonlinear cost function.

Proposition 1. *Fix a convex self-control model $(u, v, \phi)$ where $u : X \to [0, \bar{u}]$, $v : X \to [0, \bar{v}]$, $\phi : [0, \bar{v}] \to \mathbb{R}$ and $\bar{u} \geq \phi(\bar{v})$. Assume $\phi$ is not linear and the set of alternatives $X$ is rich in the sense that for every $a \in [0, \bar{u}]$ and $b \in [0, \bar{v}]$ there exists an alternative $x \in X$ such that $u(x) = a$ and $v(x) = b$. Then $(u, v, \phi)$ violates Axiom 3.*

De Clippel and Rozen (2018) note an "overfitting problem" that arises in modeling boundedly rational decision making and show that with a limited data set, observed choices may be incorrectly attributed to a particular model. In Appendix B, we show that for a *fixed finite* set of alternatives, choices generated by the constant willpower model can be attributed to the convex self-control model.[11] Proposition 1 shows that this attribution is due to a similar overfitting problem that arises when the set of alternatives is not rich enough.

## 5. Monopoly pricing

In this section, we apply our representation to the pricing problem of a profit-maximizing monopolist who faces consumers with limited willpower. Our results extend the existing literature on contracting with consumers with self-control problems (DellaVigna and Malmendier (2004), Eliaz and Spiegler (2006), Heidhues and Kőszegi (2010)). The previous literature has highlighted that the monopolist can exploit naive consumers by offering them indulging contracts with two alternatives. One of the alternatives serves as a bait that the consumer believes he will choose at the time of signing the contract. At the time of consumption, the consumer ends up choosing the other more indulging alternative.

Our analysis extends these results in two ways. First, for limited willpower preferences, we show that the optimal contract may require three alternatives, which we call compromising. The consumer believes he will choose the bait, but at the time of consumption is unable to do so due to the existence of a tempting alternative. Instead, he chooses the third, compromise alternative. Hence, consumers' choices reflect a form

[11] Axiomatic characterization of the finite version of the costly self-control model is still an open question.

of the "compromise effect," which is induced endogenously by the contract offered by the monopolist. To our knowledge, exploitation of the compromise effect by the monopolist has not been studied previously. Second, we show that for limited willpower preferences, indulging contracts are never optimal and are dominated by compromising contracts. Third, we show that when the consumer has sufficient willpower, exploitation and efficiency can go hand in hand. The monopolist uses a compromising contract to sell the product that would be sold under a commitment contract at an exploitative price. Finally, this application shows how willpower stock can be used as a natural comparative static for the consumer's level of self-control. We provide a comparison with related models of self-control in Section 5.3, where we further discuss this point.

Following the previous literature, we focus on naive consumers who do not necessarily recognize the extent of their self-control problem. In our model, this means that consumers believe that they have more willpower than they actually do. We assume that the monopolist knows that consumers have limited willpower and devises the contract that best exploits the consumers' naivety about their willpower limitation.

Let us denote the finite set of alternatives available to the monopolist by $X$. A contract $C$ is a menu of offers, where each offer is an alternative with an associated price, i.e., $C = \{(s, p(s)) : s \in S \subset X\}$. We assume that a contract cannot offer the same alternative with two different prices. That is, if $(x, p)$ and $(x, p')$ are both in $C$, then $p$ must be equal to $p'$. We consider a two-period model of contracting between a monopolist and a consumer. In the first period, the monopolist offers the consumer a contract $C$. The consumer can accept or reject the contract. If the consumer accepts the contract, in the second period, he chooses an offer from the contract and pays its price to the monopolist. If the consumer rejects the contract, then he receives his outside option normalized to zero. We assume that both parties are committed to the contract once accepted.

We denote the cost to the monopolist of providing alternative $s$ by $c(s)$, its utility to the consumer by $u(s)$, and its temptation value by $v(s)$. We assume that the consumer's utility and temptation values are both quasilinear in prices. Broadly speaking, the idea that temptation would decrease in price seems reasonable in many situations. When the price of a good increases, the consumer must forego other potentially tempting goods. Moreover, when the price is sufficiently high, the good might become unaffordable. Quasilinearity of temptation values in prices is clearly a partial equilibrium way of capturing the impact of prices on temptation and a restrictive assumption. However, it provides tractability and is implicitly invoked in the literature on changing tastes, where it is usually assumed that both the present and future utilities are quasilinear in prices.

The monopolist's profit from selling alternative $s$ at price $p(s)$ is $p(s) - c(s)$. The production cost is incurred only for the service that the consumer chooses from the menu. Following Eliaz and Spiegler (2006) and Spiegler (2011), we assume that the consumer is naive in the sense that he believes he has no self-control problem, i.e., he believes that from a contract $C$, he will choose the offer $(s, p(s))$ that maximizes $u(s) - p(s)$. In reality, the consumer might be tempted by the other offers in the contract $C$, and his second-period choices are governed by the limited willpower model. This means that from $C$,

the consumer chooses the offer $(s, p(s))$ that maximizes $u(s) - p(s)$ subject to

$$\max_{(s',p(s'))\in C} \big(v(s') - p(s')\big) - \big(v(s) - p(s)\big) \leq w$$

where $w$ is the consumer's willpower stock. We assume that the monopolist knows that the consumer has limited willpower and can perfectly predict the consumer's second-period choices.[12]

To simplify the analysis, we assume that $u - c$ and $v - c$ have unique maximizers $x^u$ and $x^v$ in $A$. In other words, $x^u$ and $x^v$ are the most efficient alternatives with respect to $u$ and $v$. To make the problem interesting, we assume that $x^u \neq x^v$. We define the difference between the temptation value and the utility value of an alternative $s$ as its *excess temptation* and denote it by $e(s) \equiv v(s) - u(s)$. We further assume that $e$ has a unique maximizer, $z^*$, and a unique minimizer, $y^*$, in $A$. Then it is easy to see that

$$\begin{aligned} e(x^v) &= v(x^v) - u(x^v) = v(x^v) - c(x^v) - \big[u(x^v) - c(x^v)\big] \\ &> v(x^u) - c(x^u) - \big[u(x^v) - c(x^v)\big] \\ &> v(x^u) - c(x^u) - \big[u(x^u) - c(x^u)\big] = v(x^u) - u(x^u) = e(x^u). \end{aligned}$$

Therefore, we have

$$e(z^*) \geq e(x^v) > e(x^u) \geq e(y^*).$$

Figure 1 illustrates the model graphically, which we use below to provide intuition for our results. In the figure, the vertical axes measure price and the horizontal axes measure the temptation value $v$. On the horizontal axes, we indicated the excess temptation

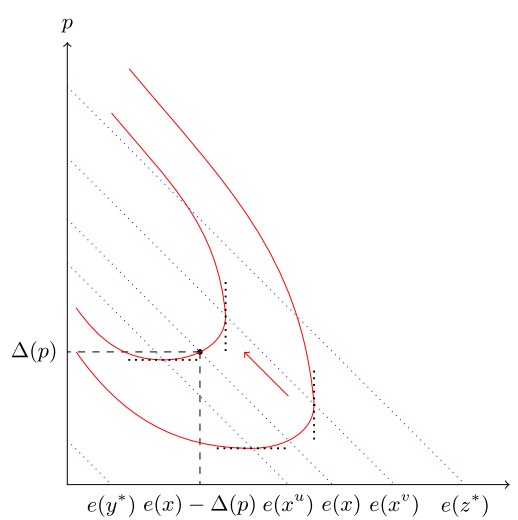



FIGURE 1. Iso-$e$ and iso-profit lines.

[12]More precisely, we solve for the optimal contract for the monopolist given its beliefs about the consumer's behavior. To do this, we do not need to know whether the monopolist (or the consumer) holds correct beliefs about the consumer's second-period behavior.

values $e(z^*)$, $e(x^v)$, $e(x^u)$, and $e(y^*)$. Any other alternative, $x$, corresponds to a point between these values on the horizontal axes in terms of its excess temptation $e(x)$. We call the dotted lines with slope $-1$, iso-$e$ lines where $e$ refers to excess temptation. Consider an alternative $x$ with excess temptation $e(x)$ and price $p = u(x) + \Delta(p)$. In the figure, the point $(\Delta(p), e(x) - \Delta(p))$ corresponds to this alternative. By moving down the iso-$e$ line, we can read its excess temptation value $e(x)$. In other words, excess temptation is given by the sum of the two coordinates. Its $y$-coordinate, $\Delta(p)$, captures the difference between its price $p = u(x) + \Delta(p)$ and its utility value $u(x)$. Its $x$-coordinate, $e(x) - \Delta(p)$, captures its temptation value at price $p$.

In the figure, we also illustrate two iso-profit lines for the firm. As we described above, any point $(b, a)$ corresponds to an alternative $e^{-1}(a + b)$ with price $u(e^{-1}(a + b)) + a$.[13] Hence, the corresponding profit is given by $u(e^{-1}(a + b)) + a - c(e^{-1}(a + b))$. The red iso-profit line traces all points with constant profit. Profits increase as the iso-profit lines move northwest. The iso-profit lines have two important properties. First, on the iso-$e$ line for $x^u$, the iso-profit lines have zero slope. Second, on the iso-$e$ line for $x^v$, the iso-profit lines have an infinite slope. These properties are illustrated in Figure 1. To see the intuition for the first property, suppose the monopolist maximizes profit subject to the consumer achieving a certain utility level. If the agent is a $u$ maximizer, then indifference curves are horizontal lines. At the optimum, iso-profit line and the indifference curve must be tangent. At the point of tangency, $u - c$ must be maximized, which occurs at $x = x^u$. In the figure, this corresponds to the iso-e line for $x^u$, implying that the iso-profit lines have zero slope on the iso-e line for $x^u$. If the agent is a $v$ maximizer, then indifference curves are vertical lines, and the second property follows from similar steps.

### *Optimal contract with sophisticated consumers*

As a benchmark case, consider a consumer who perfectly understands what he chooses once he accepts the contract. In this case, the monopolist's maximization problem is

$$\max_{s \in A,\, p(s) \geq 0} p(s) - c(s)$$

subject to

$$u(s) - p(s) \geq 0.$$

Clearly, due to the participation constraint, the firm sets the price of $s$ equal to $u(s)$. In other words, the monopolist extracts the entire surplus (in terms of $u$.) Then the optimal contract offers only the efficient alternative $x^u = \arg\max u(s) - c(s)$ at price $u(x^u)$.

### *Optimal contract with naive consumers*

Next, we solve for the optimal contract with a naive consumer. We perform this in three steps. First, we show that, without loss of generality, we can restrict attention to contracts that offer at most three alternatives. That is, for any contract that sells $x$ at price

[13] We assume that the excess temptation function is invertible. This is without loss of generality because we can restrict attention to alternatives with the largest excess temptation.

$p(x)$, there is another contract that sells $x$ at the same price that contains at most three of the alternatives from the original contract. Second, for each alternative $x$, we find a contract that sells $x$ at the highest possible price. Third, we identify the profit-maximizing alternative and the associated optimal contract that sells this alternative.

To establish the first step, let $C = \{(s, p(s)) : s \in S \subset X\}$ be an arbitrary contract that sells $x$ at price $p(x)$. This means that (i) $(x, p(x))$ is offered in $C$, (ii) the consumer accepts the contract, i.e., there exists an offer $(s^*, p(s^*))$ in $C$ such that $u(s^*) - p(s^*) \geq 0$, and (iii) $(x, p(x))$ is the best offer for the consumer given his willpower constraint. Now, identify two offers from the contract $C$ such that

$$y = \underset{(s,p(s))\in C}{\arg\max}\ u(s) - p(s) \quad \text{and} \quad z = \underset{(s,p(s))\in C}{\arg\max}\ v(s) - p(s)$$

We illustrate that the contract $C' = \{(x, p(x)), (y, p(y)), (z, p(z))\}$ sells $x$ at price $p(x)$. First, this contract is a subset of $C$ and $(x, p(x))$ is offered. The consumer accepts this contract since $u(y) - p(y) \geq u(s^*) - p(s^*) \geq 0$. Finally, because the most tempting offer is the same, the constraint he faces is unaffected. Therefore, $(x, p(x))$ is still the best offer for the consumer.

When $x$, $y$, and $z$ are distinct, the consumer believes that he will choose $(y, p(y))$ but cannot because he does not have enough willpower to choose $(y, p(y))$ when $(z, p(z))$ is available. However, rather than completely indulging in $(z, p(z))$, the consumer chooses the second-best $(x, p(x))$ in terms of $U$. We refer to a contract with these features as a *compromising contract*. If $x$ is equal to $y$, then the contract could be simpler. Indeed, the contract $\{(x, p(x))\}$ sells $x$.[14] We refer to a contract that includes a single offer as a *commitment contract*. If $x$ is equal to $z$, then the contract reduces to $\{(x, p(x)), (y, p(y))\}$. In this case, the consumer believes that he will choose $(y, p(y))$, which provides the highest utility. However, he actually chooses $(x, p(x))$ because the consumer does not have enough willpower to choose $(y, p(y))$ when $(x, p(x))$ is available. We refer to this type of contract as an *indulging contract*.

We now investigate the revenue-maximizing contract that sells $x$ by focusing on the three types of contracts we identified.

The Commitment Contract: The monopolist's problem is to maximize $p(x)$ subject to $u(x) \geq p(x)$. Then the monopolist sets the price $p(x)$ equal to $u(x)$. Hence, the highest revenue from selling $x$ using a commitment contract is $u(x)$.

The Indulging Contract: The monopolist's problem is to choose $p(x)$, $y$, and $p(y)$ to maximize $p(x)$ subject to

$$u(y) - p(y) \geq 0 \tag{2}$$

$$v(x) - p(x) \geq v(y) - p(y) + w \tag{3}$$

Constraint (2) guarantees that the naive consumer is willing to accept the contract. Constraint (3) implies that the consumer does not have enough willpower to resist the temptation to choose $(x, p(x))$. Clearly, both constraints are binding, so the maximization

[14]Note that the contract $\{(x, p(x))\}$ does *not* always sell $x$ because it may not be acceptable to the consumer at time 0. However, when $x$ is equal to $y$, provided that the original contract is acceptable, $\{(x, p(x))\}$ is also acceptable.

problem is equivalent to choosing $p(x)$, $y$, and $p(y)$ to maximize $v(x) - e(y) - w$. The monopolist sets $y = y^*$ to minimize $e(y) + w$, which implies $p(y^*) = u(y^*)$ by constraint (2). Then constraint (3) implies that the optimal price for selling $x$ with an indulging contract is

$$p^{\text{ind}}(x) = v(x) - e\big(y^*\big) - w.$$

The Compromising Contract: The monopolist's problem is now to choose $p(x)$, $y$, $p(y)$, $z$, and $p(z)$ to maximize $p(x)$ subject to

$$u(y) - p(y) \geq 0 \tag{4}$$

$$v(z) - p(z) \geq v(y) - p(y) + w \tag{5}$$

$$v(x) - p(x) + w \geq v(z) - p(z) \tag{6}$$

$$u(x) - p(x) \geq u(z) - p(z) \tag{7}$$

Constraints (4) and (5) guarantee that the consumer signs the contract believing that he will choose $y$ but does not actually do so. Constraint (6) implies that the consumer has enough willpower to choose $x$ over $z$, and constraint (7) means that choosing $x$ over $z$ is also desirable. The first two constraints (4) and (5) are clearly binding, which implies

$$p(y) = u(y) \quad \text{and} \quad p(z) = v(z) - e(y) - w$$

The remaining two constraints become

$$p(x) \leq v(x) - e(y) \quad \text{and} \quad p(x) \leq u(x) - e(y) + e(z) - w$$

Both constraints can be relaxed by choosing $y = y^*$, and the second constraint can be relaxed by choosing $z = z^*$.[15] Thus, the optimal price for selling $x$ with a compromising contract is

$$p^{\text{comp}}(x) = \min\big\{v(x) - e\big(y^*\big), u(x) - e\big(y^*\big) + e\big(z^*\big) - w\big\}. \tag{8}$$

Figure 2 illustrates the revenue-maximizing commitment, indulging, and compromising contracts for selling $x$ for relatively low willpower stock $w$ such that $e(x) - e(y^*) > w$. The commitment contract sells the alternative $(x, u(x))$ ($x$ at price $u(x)$). This is represented by point $A$ in the figure. In the indulging contract, the firm chooses the price of $x$ in a way that the consumer does not have the willpower to choose $(y^*, u(y^*))$ over $(x, p^{\text{ind}})$. Hence, the monopolist increases the price of $x$, moving up its iso-$e$ line until its temptation value reaches the vertical dashed line on the right, which corresponds to constraint (3). Hence, the indulging contract contains $y^*$ at price $u(y^*)$ and $x$ at price $p^{\text{ind}}$, which is represented by point $B$. Finally, in the compromising contract, the firm uses a third alternative, $(z^*, p(z^*))$, to eliminate alternative $(y^*, p(y^*))$. Hence, the monopolist increases the price of $z^*$, moving up its iso-$e$ line until its temptation value reaches the vertical dashed line on the right. The consumer can choose all the points

[15]When the second constraint is not binding it may not be necessary to set $z = z^*$ in the contract. In this case, the monopolist can choose any $z$ that satisfies $e(z) \geq e(x) + w$.

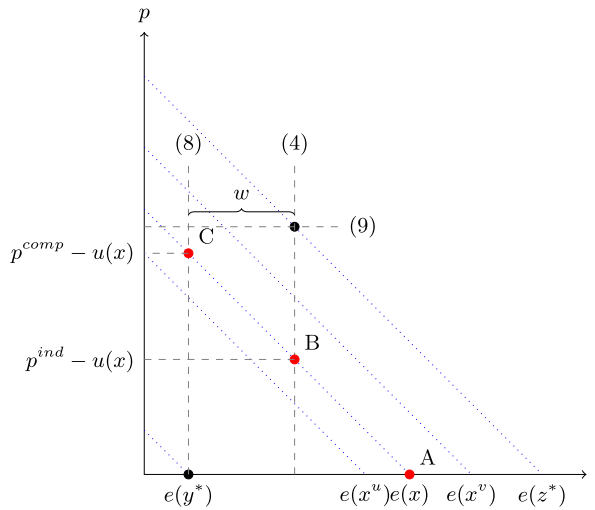



FIGURE 2. Revenue-maximizing commitment ($A$), indulging ($B$) and compromising ($C$) contracts for selling $x$.

on the iso-$e$ line of $x$ to the right of the dashed vertical line on the left, which corresponds to constraint (6). Moreover, only the alternatives below the horizontal dashed line can be chosen because the consumer prefers $(z^*, p(z^*))$ to any alternative above this line, which corresponds to constraint (7) (notice that this constraint does not bind for this particular $x$). Hence, the compromising contract contains $y^*$ at price $u(y^*)$, $z^*$ at price $p(z^*)$, and $x$ at price $p^{\text{comp}}$, which is represented by point $C$.

Figure 2 makes it clear that either constraint (6) or constraint (7) binds. If $e(z^*) - e(x) \geq w$, then constraint (6) is binding and the revenue is $v(x) - e(y^*)$. If $e(z^*) - e(x) \leq w$, then constraint (7) is binding, and the revenue is $u(x) - e(y^*) + e(z^*) - w$. Hence, the highest revenue from selling $x$ using a compromising contract is

$$\begin{aligned} v(x) - e\big(y^*\big) \quad & \text{if } e\big(z^*\big) - e(x) \geq w \\ u(x) + e\big(z^*\big) - e\big(y^*\big) - w \quad & \text{if } e\big(z^*\big) - e(x) \leq w \end{aligned}$$

From Figure 2, when $e(z^*) - e(x) \geq w$, comparing the price of $x$ under different contracts, we see that the compromising contract generates the highest revenue and that the commitment contract generates the least revenue. When $e(z^*) - e(x) \leq w$, then constraint (3) implies that the price of $x$ must be lower than $u(x)$ so that the consumer does not have the willpower to choose $(y^*, u(y^*))$. Hence, the commitment contract dominates the indulging contract. Nevertheless, as long as $w \leq e(z^*) - e(y^*)$, the compromising contract still generates strictly higher revenue than the commitment contract. Finally, when $w > e(z^*) - e(y^*)$, constraint (7) becomes very stringent and is violated as soon as the price of $x$ goes above $u(x)$. In this case, the commitment contract $\{(x, u(x))\}$ becomes weakly better than the compromising contract.

The next proposition summarizes the discussion above.

Proposition 2. *If $e(z^*) - e(y^*) < w$, the revenue-maximizing contract is the commitment contract $\{(x, u(x))\}$. If $e(z^*) - e(y^*) \geq w$, the revenue-maximizing contract is the compromising contract $\{(x, p^{\text{comp}}), (y^*, u(y^*)), (z^*, v(z^*) - e(y^*) - w)\}$.*

Note that when the compromising contract is optimal, consumer's choices violate WARP because $c((x, p^{\text{comp}}), (y^*, u(y^*)), (z^*, v(z^*) - e(y^*) - w)) = (x, p^{\text{comp}})$ and $c((x, p^{\text{comp}}), (y^*, u(y^*))) = (y^*, u(y^*))$. These WARP violations play a key role in delivering the endogenous compromise effect in Proposition 2. They allow the monopolist to block the bait $y^*$ by a third tempting alternative $z^*$ and increase the price of the "middle" alternative $x$ above the price that the monopolist could charge by using the indulging contract.

### 5.1 *The optimal contract*

Now that we have identified the revenue-maximizing contract for each alternative, the remaining task is to find which alternative the monopolist should sell to maximize its profit. From Proposition 2, we know that if $e(z^*) - e(x) \leq w$, then the monopolist's revenue is either $u(x)$ or $u(x) + e(z^*) - e(y^*) - w$. Hence, from the set $\{x : e(z^*) - e(x) \leq w\}$, it is optimal to sell the maximizer of $u(x) - c(x)$. If $e(z^*) - e(x) \geq w$, the monopolist's revenue is $v(x) - e(y^*)$. Hence, from the set $\{x : e(z^*) - e(x) \geq w\}$, it is optimal to sell the maximizer of $v(x) - c(x)$. Combining the two cases, the optimal contract sells either

$$\operatorname*{arg\,max}_{x:e(z^*)-e(x)\leq w} u(x) - c(x) \quad \text{or} \quad \operatorname*{arg\,max}_{x:e(z^*)-e(x)\geq w} v(x) - c(x),$$

whichever generates the higher profit shown by Proposition 2. These observations lead to the following proposition.

Proposition 3. *(i) If $e(z^*) - e(x^v) \geq w$, then the optimal contract is the best compromising contract selling $x^v$ at $v(x^v) - e(y^*)$. The consumer's welfare is the same as when he had no willpower at all.*

*(ii) If $e(z^*) - e(x^v) < w < e(z^*) - e(x^u)$, then the optimal contract is the best compromising contract, which sells an alternative other than $x^u$ or $x^v$.*

*(iii) If $e(z^*) - e(x^u) \leq w < e(z^*) - e(y^*)$, then the optimal contract is the best compromising contract that includes $y^*$ and $z^*$ but actually sells the efficient service $x^u$ at a price exceeding $u(x^u)$. The consumer is exploited, but the degree of the exploitation declines as his willpower increases.*

*(iv) If $e(z^*) - e(y^*) \leq w$, then the optimal contract is the commitment contract selling the efficient service $x^u$ at price $u(x^u)$. The consumer is not exploited even though he is naive.*

Figure 3 illustrates the optimal contract in each of the four cases of Proposition 3, which characterizes the optimal contract as the willpower stock $w$ increases. The dashed lines in each case correspond to constraints (6) and (7). Figure 3a is when only (6) is

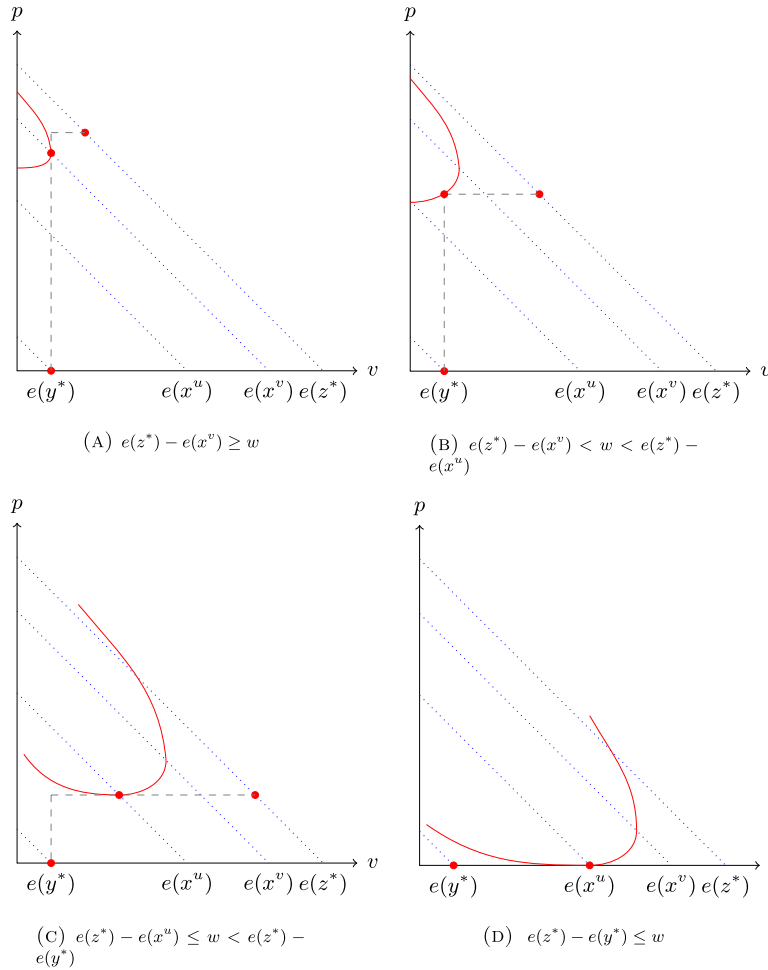



FIGURE 3. Optimal contracts for different levels of willpower stock.

binding and the optimal contract sells $x^v$. Figure 3b is when both constraints are binding and the optimal contract sells a product with excess temptation between $e(x^u)$ and $e(x^v)$. Figure 3c is when only (7) is binding and the optimal contract sells $x^u$ at an exploitative price. Finally, Figure 3d shows that when the willpower stock is high enough, both constraints disappear, and the optimal contract is the commitment contract that sells $x^u$ without exploitation.

Next, we will consider how the optimal contract, the profit, and the naive consumer's welfare (measured with respect to his ex ante preference $u$) varies with the consumer's willpower stock. To do this, we use Figure 4, which contains information about the optimal contract, the monopolist's profit, and the consumer's welfare as $w$ varies. Starting from the red dot in the upper-left corner, the thick red line, which we call the contract curve, traces the product sold and its price as $w$ increases. For any point on the contract curve, by moving down the iso-$e$ line, we can determine the product sold to the consumer by the optimal contract. The y-coordinate of the point gives us the price of this

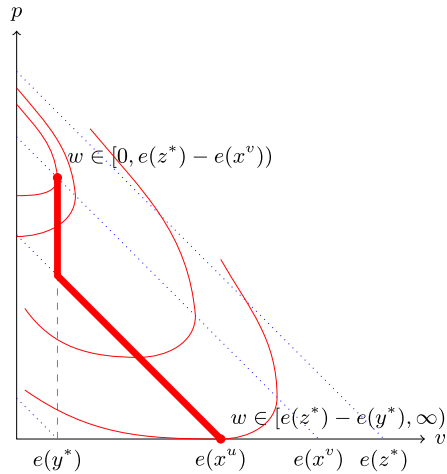



FIGURE 4. Profit for different levels of willpower stock.

product in excess of its utility value. Hence, as we move down the contract curve, the monopolist's profit declines and the consumer's welfare increases.

We highlight an interesting feature of our model. When the willpower stock is below a certain level, the product sold under the optimal contract and its price remain the same, which indicates that a small amount of willpower does not help the consumer at all. This occurs in the upper-left corner of the contract curve in Figure 4, which corresponds to $w \in [0, e(z^*) - e(x^v))$. In this range, since the monopolist sells $x^v$ at the same price, both the monopolist's profit and the consumer's welfare do not change.

As $w$ increases, we enter the range $[e(z^*) - e(x^v), e(z^*) - e(x^u))$, which corresponds to the vertical portion of the contract curve. In this range, the excess temptation of the product sold decreases from $e(x^v)$ to $e(x^u)$, its price and the monopolist's profit drop, and the consumer's welfare increases.

The next range $[e(z^*) - e(x^u), e(z^*) - e(y^*))$ corresponds to the linear, decreasing portion of the contract curve. Here, efficiency and exploitation go hand-in-hand since the optimal contract sells the efficient alternative $x^u$ at an exploitative price exceeding its utility value. In this range, the excess temptation of the product remains constant at $e(x^u)$, its price and the monopolist's profit decline, and the consumer's welfare increases.

Finally, in the lower-right corner of the contract curve $[e(z^*) - e(y^*), \infty)$, the nature of the optimal contract changes since the monopolist sells the efficient alternative using a commitment contract and there is no exploitation. When the consumer's willpower stock exceeds a certain level, naivety does not hurt the consumer at all.

The next proposition summarizes the key points of above discussion.

PROPOSITION 4. *(i) The monopolist's profit is weakly decreasing (strictly decreasing if $w \in (e(z^*) - e(x^v), e(z^*) - e(y^*))$) in the consumer's willpower.*

*(ii) The consumer's welfare is weakly increasing (strictly increasing if $w \in (e(z^*) - e(x^v), e(z^*) - e(y^*))$) in his willpower.*

*(iii) If the consumer's willpower is below a threshold ($w \in [0, e(z^*) - e(x^v)]$), the monopolist can earn the same profit as when the consumer has no willpower at all.*

*(iv) For relatively high willpower, the monopolist sells the efficient product with an exploitative price.*

*(v) When the consumer has high enough willpower, there is no exploitation despite the consumer's naivety.*

### 5.2 *Partial naivety about the willpower stock*

In this section, we briefly consider the case in which the consumer is partially naive about his future behavior. Our approach is analogous to that in the Strotzian case, where a partially naive consumer understands that his preferences change but underestimates the extent of the change (see Chapter 4, in Spiegler (2011)). Similarly, our consumer understands that he needs willpower to resist the offers that are more tempting than the one he initially plans to choose but *overestimates* his willpower stock. We assume that the firm knows the true willpower stock. It turns out that in our model, as in the Strotzian case, the optimal contract with a partially naive consumer is the same as that with a fully naive consumer. The key observation is that partial naivety does not affect how the consumer behaves after signing the contract because his behavior is determined by the true willpower stock. However, it will affect how the consumer evaluates a contract before signing it, which depends on how much willpower he believes he has. The optimal contract for fully naive consumers set prices such that $y^*$ is marginally not choosable over $z^*$. Thus, anyone who overestimates his amount of willpower still believes that he can choose $y^*$. Thus, he is willing to sign the contract, the optimal contract does not change, and all previous results carry through as long as consumer overestimates his willpower.

### 5.3 *Optimal contracting under related models*

In this section, we compare our results with those obtained when consumers have convex self-control preferences. We refer the interested reader to Masatlioglu et al. (2019) for a more complete analysis.

Under convex self-control preferences, compromising contracts are always strictly optimal (i.e., strictly better than both the commitment and the indulging contracts). This result shows that the optimality of indulging contracts with linear self-control preferences is a knife-edge case. Second, with costly self-control, exploitation always sacrifices efficiency. This is because in costly self-control models, in contrast to limited willpower models, consumers never become $u$-maximizers in the second stage.

The general version of the costly convex self-control model does not have an obvious analogue for the willpower parameter that controls the consumer's level of self-control. However, it has the following intuitive special case:

$$\varphi(x) = \begin{cases} lx & \text{if } x \leq \hat{w} \\ k(x - \hat{w}) + l\hat{w} & \text{if } x > \hat{w} \end{cases}$$

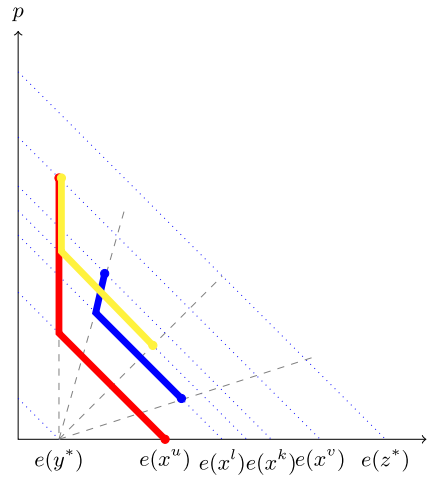



FIGURE 5. Contract curves for different models.

where $k > 1 > l > 0$. In this specification, $\hat{w}$ can be interpreted as the analogue of the willpower stock in the constant willpower model. As one would expect, when $l$ goes to zero and $k$ goes to $\infty$, the solution to the contracting problem approaches the solution given in Proposition 3.

In Figure 5, the thick blue line is the contract curve for the piecewise linear model. For comparison, the figure also shows the contract curve for the constant willpower model (thick red line). The two curves show that optimal contracts share some properties.[16] For example, when $\hat{w}$ is below a certain level, the product sold under the optimal contract and its price remain the same. As $\hat{w}$ increases, the monopolist's profits decrease and the consumer's welfare increases. However, there are key differences. For example, in the piecewise linear model, the indulging contract can be optimal when $\hat{w}$ is large enough (represented by the blue dot at the bottom of the contract curve). In contrast, in the constant willpower model, the indulging contact is never optimal. (Instead, when the consumer has enough willpower, the commitment contract becomes optimal.) We also see that the optimal contract in the piecewise linear model never sells the $u$-efficient or $v$-efficient alternative.

## 6. CONCLUSION

Since, Kreps (1979), researchers have studied a two-period choice model in which an agent selects one of several menus in the planning period under the assumption that the agent will make a choice from each menu in the consumption period. This new and rich data set has allowed researchers to study phenomena such as temptation, guilt, and shame, among others.

In this paper, to derive the limited willpower representation, we use ex ante preferences and ex post choices as our data. Revealing the ex ante preferences over *alternatives*

[16] Figure 5 also displays the contract curve for the Grant et al. (2017) (thick yellow line), which provides a hybrid solution.

is a simpler and more natural task than revealing ex ante preferences over *all menus of alternatives*. More important, our data set allows us to remain agnostic about whether the agent is sophisticated or naive about anticipating his ex post choices. To derive the representation, we introduce a new axiom called Choice Betweenness. We show that this axiom is independent of the Set Betweenness axiom that is commonly invoked in the menu preferences domain.

Although the model is simple and tractable, it is rich enough to generate new insights in applications. We demonstrate this in an application to monopolistic contracting. Finally, we would like to highlight an important avenue for exploration in future work which is the implications of limited willpower in a dynamic setting with multiple tasks. In the current manuscript, we consider a model in which willpower is needed in a single choice task. In fact, people often use willpower in multiple tasks, and using more willpower in one task might mean less willpower is left for another. Moreover, the model is static. In reality, there are dynamic effects in the sense that the amount of willpower used in one period can affect the willpower stock in the next period. Incorporating these considerations into an axiomatic framework could lead to new insights about behavior and a rich set of testable implications.

## Appendix A: Proofs

### *Proofs of Theorems 1 and 2*

Before we provide the proofs of Theorems 1 and 2, we provide a brief sketch. To prove Theorem 1, we first define a binary relation $\rhd'$. We say that $x \rhd' y$ if $y \succ x = c(xy)$. In words, $x$ blocks $y$ if $x$ is worse than $y$ but the agent cannot choose $y$ when $x$ is available. Next, we define a second binary relation $\rhd''$. We say that $x \rhd'' y$ if $x \succsim y$ and there exist $a$ and $b$ such that $a \rhd' y$, $x \rhd' b$, and $a \not\rhd' b$. We say that $x \rhd y$ if $x \rhd' y$ or $x \rhd'' y$. Next, we show that $\rhd$ is an interval order, i.e., it is irreflexive and $x \rhd b$ or $a \rhd y$ holds whenever $x \rhd y$ and $a \rhd b$. The binary relation $\rhd$ is an interval order if and only if there exist functions $v$ and $w$ such that

$$\Gamma_{\rhd}(S) = \{x \in S : \max v(y) - v(x) \leq w\}.$$

Finally, to complete the proof of the first step, we show that $S$ is indifferent to the $\succsim$-best element in $\Gamma_{\rhd}(S)$.

In the proof of Theorem 2, we use consistency to show that we can construct a semiorder $\overline{\rhd}$ (i.e., $\overline{\rhd}$ is an interval order and if $x\overline{\rhd}y\overline{\rhd}z$ then $x\overline{\rhd}t$ or $t\overline{\rhd}z$ for any $t$) by properly modifying $\rhd$ such that $S$ is indifferent to the $\succsim$-best element in $\Gamma_{\overline{\rhd}}(S)$. To complete the proof, we note that the binary relation $\overline{\rhd}$ is a semiorder if and only if there exist a function $v$ and a scalar $w$ such that

$$\Gamma_{\rhd}(S) = \{x \in S : \max v(y) - v(x) \leq w\}.$$

### *Proof of Theorem 1*

We first show that Axioms 1–3 imply an important implication of our model.

Claim 1. *Suppose* $(\succ, c)$ *satisfies Axioms 1–3. Then if* $x \succ c(A \cup x)$, *then* $c(B) = c(B \cup x)$ *for all* $B \supset A$.

Proof. Let $L(n)$ stand for the statement of Claim 1 that is restricted to when $|B - A| \leq n$. Notice that Axiom 2 is $L(0)$. First, we shall show $L(1)$. That is, $x \succ c(A \cup x)$ (so $c(A) = c(A \cup x)$ by Axiom 2) implies $c(A \cup y) = c(A \cup x \cup y)$ for any $y$.

**Case 1:** $y \succ c(A \cup x \cup y)$: By Axiom 2, $c(A \cup x) = c(A \cup x \cup y)$. By the assumption, we have $x \succ c(A \cup x) = c(A \cup x \cup y)$. By applying Axiom 2, we get $c(A \cup y) = c(A \cup x \cup y)$.

**Case 2:** $y \prec c(A \cup x \cup y)$: By Axiom 3, $y \prec c(A \cup x \cup y) \precsim c(A \cup x) \prec x$. By Axiom 1, we get $c(A \cup x \cup y) \prec x$. Then by Axiom 2, we get the desired result, $c(A \cup y) = c(A \cup x \cup y)$.

**Case 3:** $y \sim c(A \cup x \cup y)$: We have three subcases:

- If $y = c(A \cup y)$, then $c(A \cup y) = c(A \cup x \cup y)$.
- If $y \succ c(A \cup y)$, then Axiom 2 implies $c(A \cup y) = c(A)(= c(A \cup x))$. Applying Axiom 3, we get $c(A \cup x \cup y) = c(A \cup y)$, which is a contradiction because $c(A \cup x \cup y) = y \succ c(A \cup y)$.
- If $y \prec c(A \cup y)$, then Axiom 3 implies $(c(A \cup x) =)c(A) \succsim c(A \cup y)$. Applying Axiom 3 again, it must be $c(A \cup x \cup y) \succsim c(A \cup y) \succ y$, which is a contradiction because $c(A \cup x \cup y) = y$.

Now suppose that $L(k)$ is true up when $1 \leq k \leq n-1$. We shall prove $L(n)$. Assume $x \succ c(A \cup x)$ and let $B = A \cup \{y_1, y_2, \dots, y_n\}$ where all of $y_i$'s are distinct and excluded from $A$. Our goal is to show $c(B) = c(B \cup x)$. Without loss of generality, assume $y_1 \succ y_2 \succ \cdots \succ y_n$.

**Case 1:** $y \succ c(A \cup x \cup y)$ for some $y \in \{y_1, y_2, \dots, y_n\}$: Since $(B \setminus y) \cup x \supset A \cup x$ and the difference of their cardinality is $n-1$, we can utilize $L(n-1)$. Then we get $c((B \setminus y) \cup x) = c((B \setminus y) \cup x \cup y)(= c(B \cup x))$. Applying $L(1)$ to $x \succ A \cup x$, we have $(y \succ)c(A \cup x \cup y) = c(A \cup y)$. Applying $L(n-1)$ to this yields $c(B \setminus y) = c((B \setminus y) \cup y) = c(B)$. Notice that $c(B \setminus y) = c((B \setminus y) \cup x)$ because $x \succ c(A \cup x)$ and $L(n-1)$. These three equalities imply $c(B) = c(B \cup x)$.

**Case 2:** $y \prec c(A \cup x \cup y)$ for some $y \in \{y_1, y_2, \dots, y_n\}$: By Axiom 3, we have $c(A \cup x) \succsim c(A \cup x \cup y) \succ y$. Since $x \succ c(A \cup x)$ and Axiom 1, we have $x \succ c(A \cup x \cup y)$. Because $|B \setminus (A \cup y)| = n-1$, by applying $L(n-1)$ we have $c(B) = c(B \cup x)$.

**Case 3:** $y_i = c(A \cup y_i \cup x)$ for all $i = 1, \dots, n$: In this case, we have

$$y_1 = c(A \cup y_1 \cup x) \succ y_2 = c(A \cup y_2 \cup x) \succ \cdots \succ y_n = c(A \cup y_n \cup x)$$

Since $c(A \cup y_i \cup x) = c(A \cup y_i)$ by $L(1)$, the above relations still hold when $x$ is removed:

$$y_1 = c(A \cup y_1) \succ y_2 = c(A \cup y_2) \succ \cdots \succ y_n = c(A \cup y_n)$$

Recursively applying Axiom 3 implies

$$\big(c(A \cup y_1 \cup x) =\big)y_1 \succsim c\big(A \cup \{y_1, y_2, \dots, y_n\}\big)\big(= c(B)\big) \succsim y_n\big(= c(A \cup y_n \cup x)\big)$$

In other words,

$$c(A \cup y_1 \cup x) \succsim c(B) \succsim c(A \cup y_n \cup x)$$

Since $(A \cup y_1 \cup x) \cup B = B \cup x$, Axiom 3 implies $c(B \cup x) \succsim c(B)$. Similarly, since $(A \cup y_n \cup x) \cup B = B \cup x$, Axiom 3 implies $c(B) \succsim c(B \cup x)$. Therefore, by Axiom 1, $c(B) = c(B \cup x)$. □

For any binary relation $R$, let $\Gamma_R(S)$ be the set of $R$-undominated elements in $S$, i.e.,

$$\Gamma_R(S) = \{x \in S : \text{ there exists no } y \in S \text{ such that } yRx\}$$

Instead of constructing $v$ and $w$, we shall construct a binary relation over $X$, denoted by $\rhd$ such that $c(S)$ is the $\succ$-best element in $\Gamma_{\rhd}(S)$.[17] It is known (Fishburn (1970)) that, if (and only if) $\rhd$ is an interval order,[18] there exist functions $v$ and $\varepsilon$ such that

$$\Gamma_{\rhd}(S) = \left\{x \in S : v(y) - v(x) \leq w \ \forall y \in S\right\} = \left\{x \in S : \max_{y \in S} v(y) - v(x) \leq w\right\}$$

so that we can get the desired representation.

Now, for any $x \neq y$, we define $x \rhd y$ when either $x \rhd' y$ or $x \rhd'' y$ where $\rhd'$ and $\rhd''$ are defined as follows:

1. $x \rhd' y$ if $y \succ x = c(xy)$
2. $x \rhd'' y$ if $x \succ y$ and there exist $a$ and $b$ such that $a \rhd' y$, $x \rhd' b$, and $a \not\rhd' b$.

Note that $x \rhd' y$ and $x \rhd'' y$ cannot happen at the same time. In addition, $\rhd'$ and $\rhd''$ are both irreflexive.

We need to show that (i) $\rhd$ is an interval order and (ii) the $\succ$-best element in $\Gamma_{\rhd}(S)$ is equal to $c(S)$.

Claim 2. *$\rhd'$ is asymmetric and transitive.*

Proof. By construction, $x \rhd' y$ and $y \rhd' x$ cannot happen at the same time. Suppose $x \rhd' y$ and $y \rhd' z$, i.e., $z \succ c(yz) = y \succ c(xy) = x$. Then by Claim 1, $c(xyz) = c(xz)$ because $y \succ c(xy)$. By Axiom 3, $(z \succ)c(yz) \succsim c(xyz) \succsim c(xy)$. Hence, we have $z \succ c(xyz) = c(xz)$. Hence, we have $z \succ x = c(xz)$, so $x \rhd' z$. □

Claim 3. *If $x \rhd' y$ and $a \rhd' b$ but neither $x \rhd' b$ or $a \rhd' y$, then it must be $x \rhd'' b$ or $a \rhd'' y$ but not both.*

Proof. First, we shall show that $x \rhd'' b$ and $a \rhd'' y$ cannot happen at the same time. Suppose it does. Then by definition of $\rhd'$ and $\rhd''$, we have $y \succ x \succ b \succ a \succ y$. Axiom 1 is violated.

[17] In our framework, the $\succ$-best element is equal to the $\succsim$-best element.

[18] $\rhd$ is called an interval order if it is irreflexive and $x \rhd b$ or $a \rhd y$ holds whenever $x \rhd y$ and $a \rhd b$.

Now, we shall show that either $x \rhd'' b$ or $a \rhd'' y$ must be defined. Suppose not. Then, along with the definition of $\rhd'$, we have $b \succ x = c(xy)$, and $y \succ a = c(ab)$. Therefore, $c(xyab)$ must be weakly worse than $x$ or $a$ because it must be weakly worse than $c(xy)$ or $c(ab)$ by Axiom 3.

Since neither $(x, b)$ nor $(a, y)$ belongs to $\rhd'$ or $\rhd''$, we have $c(xb) = b \succ x$, and $c(ay) = y \succ a$. By Axiom 3, $c(xyab)$ must be weakly better than $c(xb)$ or $c(ay)$ so it must be weakly better than $y$ or $b$.

Hence, either $x$ or $a$ must be weakly better than either $y$ or $b$. Since we have already seen $b \succ x$ and $y \succ a$, the only possibilities are $a \succsim b$ or $x \succsim y$, neither of which is possible because $a \rhd' b$ and $x \rhd' y$. □

Claim 4. *$\rhd$ is an interval order.*

Proof. We need to show that $\rhd$ is irreflexive. By definition, we cannot have (i) $x \rhd' y$ and $y \rhd' x$, (ii) $x \rhd' y$ and $y \rhd'' x$, or (iii) $x \rhd'' y$ and $y \rhd'' x$. Hence, $\rhd$ is irreflexive.

Next, we show that $x \rhd b$ or $a \rhd y$ holds whenever $x \rhd y$ and $a \rhd b$. We shall prove this case by case.

**Case 1:** $x \rhd' y$ and $a \rhd' b$: If we have $x \rhd' b$ or $a \rhd' y$, then we are done. Assume not, then Claim 3 implies we must have $x \rhd'' b$ or $a \rhd'' y$ (not both). Then $x \rhd b$ or $a \rhd y$.

**Case 2:** $x \rhd' y$ and $a \rhd'' b$: In this case, by definition of $\rhd''$ and Claim 3, there exist $s$ and $t$ such that $a \rhd' t$ and $s \rhd' b$ but not $s \rhd t$. Focus on $x \rhd' y$ and $a \rhd' t$, we must have either $a \rhd y$ (it is done in this case) or $x \rhd t$ (so either $x \rhd' t$ or $x \rhd'' t$). If $x \rhd' t$, then by looking at $x \rhd' t$ and $s \rhd' b$ the claim requires $x \rhd b$ because it is not $s \rhd t$. Thus, we consider the final subcase: $x \rhd'' t$. If so, we have $x \rhd' y$ and $s \rhd' b$ so it must be either $x \rhd b$ (then done) or $s \rhd y$. If $s \rhd y$, then it must be $s \rhd' y$ (i.e., not $s \rhd'' y$) because $y \succ x \succsim t \succ a \succsim b \succ s$. Therefore, we have $s \rhd' y$ and $a \rhd' t$ with not $s \rhd t$. Hence, it must be $a \rhd y$.

**Case 3:** $x \rhd'' y$ and $a \rhd'' b$: By definition of $\rhd''$, there exist $s$ and $t$ such that $x \rhd' t$ and $s \rhd' y$ with not $s \rhd t$. Then by focusing on $x \rhd' t$ and $a \rhd'' b$, we must have either $x \rhd b$ (done) or $a \rhd t$. Suppose the latter. Then we have $s \rhd' y$ and "$a \rhd' t$ or $a \rhd'' t$," so the previous two cases are applicable so we conclude $a \rhd y$ because it is not $s \rhd t$. □

Claim 5. *$c(S)$ is equal to the $\succ$-best element in $\Gamma_{\rhd'}(S)$.*

Proof. First, we prove that $\Gamma_{\rhd'}(S)$ does not include any element that is strictly better than $c(S)$. Suppose $x \in \Gamma_{\rhd'}(S)$. Let $S'$ and $S''$ be the subsets of $S \setminus x$ that consist of elements that are better than $x$ and strictly worse than $x$, respectively. That is,

$$S' := \{y \in S : y \succ x\} \quad \text{and} \quad S'' := \{y \in S : x \succ y\}.$$

Then we have $c(S' \cup x) \succ x$ by definition of $c$ and $x = c(xy) \succ y$ for all $y \in S''$ by the definition of $\rhd'$. Then by applying Axiom 3, we get $c(S'' \cup x) = x$. Thus, $c(S' \cup x) \succ c(S'' \cup x)$ implies $c((S' \cup x) \cup (S'' \cup x)) = c(S) \succsim x$ again by Axiom 3.

Next, we shall show that $c(S) \in \Gamma_{\rhd'}(S)$. Suppose not. Then there exists $y \in \Gamma_{\rhd'}(S)$ such that $y \rhd' c(S)$ by Claim 2 (especially $\rhd'$ is transitive). That is $c(S) \succ c(\{c(S), y\}) = y$. Thus, by Claim 1, we have $c(S \setminus c(S)) = c((S \setminus c(S)) \cup c(S)) = c(S)$, a contradiction.

Combining the first and second results, the $\succ$-best element in $\Gamma_{\rhd'}(S)$ is equal to $c(S)$. □

Claim 6. *$c(S)$ is equal to the $\succ$-best element in $\Gamma_{\rhd}(S)$.*

Proof. Since $\rhd \supseteq \rhd'$ by construction, we have $\Gamma_{\rhd}(S) \subseteq \Gamma_{\rhd'}(S)$. Therefore, by Claim 5, it is enough to show is that the $\succ$-best elements in $\Gamma_{\rhd'}(S)$ (which is $c(S)$) is included in $\Gamma_{\rhd}(S)$. Suppose $c(S) \notin \Gamma_{\rhd}(S)$. Since $\rhd$ is an interval order, it is automatically transitive. Therefore, there exists $y \in \Gamma_{\rhd}(S)$ such that $y \rhd c(S)$ but not $y \rhd' c(S)$. Therefore, it must be $y \rhd'' x$ so $y \succ c(S)$. Since $y \in \Gamma_{\rhd'}(S)$, $y$ cannot be strictly better than $c(S)$ (see the proof of Claim 5). □

**(The Representation $\Rightarrow$ The Axioms)**

Showing that the first axiom is necessary is straightforward. For the second axiom, if $x \succ c(A \cup x)$ then $A$ must have an element $y$ with $v(y) > v(x) + w(x)$, so its superset $B$ also includes $y$ so $\Gamma(A) = \Gamma(B)$, so $c(A) = c(B)$.

The third axiom: Let $x^*$ be the $u$-best element in $\Gamma(A \cup B)$. Then it must be in $\Gamma(A)$ or $\Gamma(B)$ as well so it is not possible that $A \cup B$ is strictly preferred to both $A$ and $B$. Now we show that the union cannot be strictly worse than both. Let $x_A$ and $x_B$ be the $u$-best elements in $A$ and $B$, respectively, and take $v_A$ and $v_B$ be the maximum values of $v$ in $A$ and in $B$, respectively. Then we have

$$v_A \leq u(x_A) + \varepsilon(x_A) \quad \text{and} \quad v_B \leq u(x_B) + \varepsilon(x_B)$$

Therefore, the maximum value of $v$ in $A \cup B$ is the higher one between $v_A$ and $v_B$, either $x_A$ or $x_B$ must be in $\Gamma(A \cup B)$ so $c(A \cup B)$ must be weakly better than either $c(A)$ or $c(B)$.

*Proof of Theorem 2*

We are now done proving the sufficiency of the axioms for the representation in Theorem 1. Next, we show the sufficiency of Axioms 1–4 for the representation in Theorem 2.

Claim 7. *If $x \succ c(xy) \succ c(yz)$, then for all $t$, $c(xyzt)$ is either $c(xt)$ or $c(yz)$.*

Proof. Assume $x \succ c(xy) \succ c(yz)$, then it must be $x \succ y \succ z$. Consider $c(zt)$. If $c(zt) = t$ then by Axiom 4, we get $c(xt) = t$. Since $y \succ c(yz)$, by Claim 1, we have $c(zt) = c(yzt)$. By Axiom 3, we have $c(xt) = c(xyzt) = c(yzt)$. Hence, $c(xyzt) = c(xt)$.

Now assume $c(zt) = z$. Since $x \succ c(xy)$, by Claim 1, we have $(z =)c(yz) = c(xyz)$. By Axiom 3, we have $c(zt) = c(xyzt) = c(xyz)(= c(yz))$. Hence, $c(xyzt) = c(yz)$. □

Again as in the proof of Theorem 1, instead of defining $v(\cdot)$ and $w > 0$, we shall construct a binary relation over $X$, denoted by $\overline{\rhd}$ such that $c(S)$ is equal to the $\succ$-best element in $\Gamma_{\overline{\rhd}}(S)$ (i.e., the set of $\overline{\rhd}$-undominated elements in $S$). It is known (Fishburn

(1970)) that if (and only if) $\overline{\rhd}$ is a semiorder,[19] which is a special type of an interval order, there exist function $v$ and positive number $w$ such that

$$\Gamma_{\overline{\rhd}}(S) = \left\{ x \in S : \max_{y \in S} v(y) - v(x) \leq w \right\}$$

so we get the desired representation.

Next, we define $(i, j)$-representation of an arbitrary binary relation $P$.

Definition 1. Two functions $i : X \to N$ and $j : X \to N$ where $i(x) \geq j(x)$ for all $x \in X$ represents a binary relation $P$ if $xPy$ if and only if $i(x) < j(y)$.

Let $\rhd$ be the interval order that is defined in the proof of Theorem 1. First, we argue that $\rhd$ has an $(i, j)$-representation without any gaps as described in the following claim.

Claim 8. *Any interval order, $P$, has an $(i, j)$-representation (see Figure 6) if there exist two functions $i : X \to N$ and $j : X \to N$ such that:*

*(i) For all $x \in X$, $i(x) \geq j(x)$;*

*(ii) The ranges of $i$ and $j$ have no gap: That is, if there exist $x$ and $y$ such that $i(x) > i(y)$ then for any integer $n$ between $i(x)$ and $i(y)$ there is $z$ with $i(z) = n$. Similarly, for $j(\cdot)$;*

*(iii) $xPy$ if and only if $i(x) < j(y)$.*

Proof. The following proof is based on Mirkin (1979). Given an interval order, $P$, ($xPy$ and $zPw$ imply $xPw$ or $zPy$), we can show that, for all $x$ and $y$ in $X$, $L(x) \subseteq L(y)$ or $L(y) \subseteq L(x)$, and $U(x) \subseteq U(y)$ or $U(y) \subseteq U(x)$, where $L(x)$ and $U(x)$ are lower and

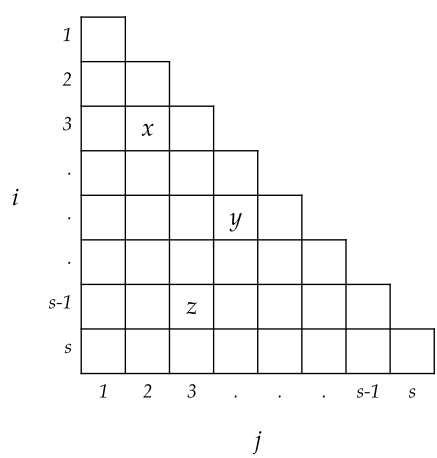



Figure 6. The graph of the $(i, j)$-representation. Condition (ii) implies every row and column (not every cell) includes at least one alternative. Condition (iii) implies $(x, y) \in P$ but $(x, z) \notin P$.

[19] $\overline{\rhd}$ is a semiorder if it is an interval order and if $x \overline{\rhd} y \overline{\rhd} z$ then $x \overline{\rhd} t$ or $t \overline{\rhd} z$ for any $t$.

upper contour sets of $x$ with respect to $P$, respectively. That is, $L(x) = \{y \in X : xPy\}$ and $U(x) = \{y \in X : yPx\}$. Irreflexivity indicates that there is a chain with respect to lower contour sets (this is also true for upper contour sets), i.e., relabel elements of $X$, $|X| = n$ such that $L(x_j) \subseteq L(x_i)$ for all $1 \le i \le j \le n$. Moreover, we can include strict inclusions such as there exists $s \le n$ such that $\varnothing = L(x_s) \subset L(x_{s-1}) \dots L(x_2) \subset L(x_1)$ where $\{x_1, x_2, \dots, x_s\} \subseteq X$. For all $k \le s$, define

$$I_k = \{x \in X : L(x_k) = L(x)\}$$

$I_k$ is not empty for any $k$ since $x_k \in I_k$ by construction. Clearly, the system $\{I_k\}_1^s$ is a partition of the set $X$, i.e., $\cup_{k=1}^s I_k = X$, $I_k \cap I_l = \varnothing$ when $k \ne l$. Define

$$i(x) := k \quad \text{if } L(x) = L(x_k) \text{ for some } x_k \text{ in } X.$$

Now construct another family of nonempty sets $\{J_m\}_1^s$, as follows:

$$J_s = L(x_{s-1}) \setminus L(x_s), \dots, J_2 = L(x_1) \setminus L(x_2), J_1 = X \setminus L(x_1)$$

Clearly, the system $\{J_m\}_1^s$ is another partition of the set $X$. Most importantly, we have $\varnothing = U(y_1) \subset U(y_2) \dots U(y_{s-1}) \subset U(y_s)$ where $y_i \in J_i$ for all $i \le s$. Define

$$j(x) := k \quad \text{if } x \in J_k.$$

To see Condition (i) holds, let $i(x) = i$. That means $x \in I_i$. If there exists no element $z$ such that $zPx$, i.e., $U(x) = \varnothing$, then $j(x) = 1 \le i(x)$. Otherwise, find the largest integer $j$ such that $x \in L(x_j)$. Note that $j$ must be strictly less than $i$. Then by definition, $j(x) = j + 1$, which is less than $i = i(x)$.

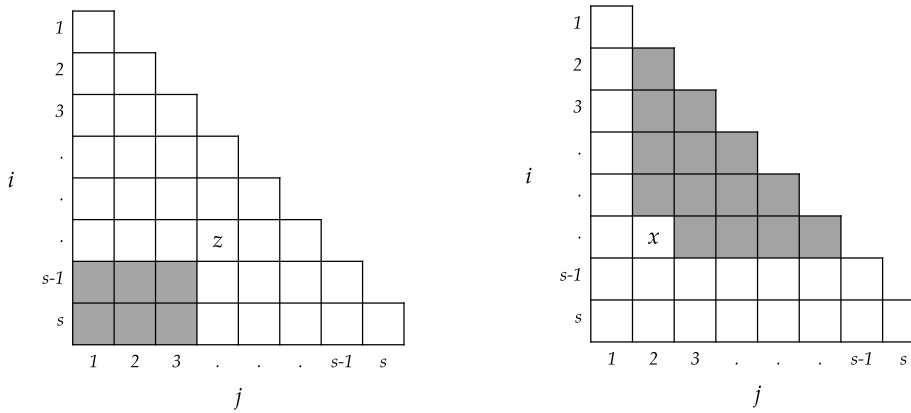
Since both $\{I_k\}_1^s$ and $\{J_k\}_1^s$ are partitions of $X$, there is no gap (Condition (ii)). Finally, we have Condition (iii) since $xPy \Leftrightarrow y \in L(x) \Leftrightarrow j(y) \ge i(x) + 1 > i(x)$. □

Let $\rhd$ be the interval order that is defined in the proof of Theorem 1. By Claim 8, it has an $(i, j)$-representation. We now modify the $(i, j)$-representation of $\rhd$ so that the resulting binary relation is a semiorder, say $\overline{\rhd}$, such that $c(S)$ is equal to the $\succ$-best element in $\Gamma_{\rhd}(S)$. In other words, we construct a semiorder based on the interval order we created without affecting the representation. To do this, we prove several claims that relate to the $(i, j)$-representation with the preference $\succ$.

Claim 9. *If $i(x) = j(y) - 1$, it must be $y \succ x$.*

Proof. Since $i(x) < j(y)$, we know that $x \rhd y$. If $x \rhd' y$, then we are done since in that case $y \succ x = c(xy)$. So suppose that $x \rhd'' y$. Then by definition of $\rhd''$, there exist $\alpha$ and $\beta$ such that $\alpha \rhd' y$ and $x \rhd' \beta$ and $\alpha \not\rhd' \beta$. Moreover, by Claim 3 $\alpha \not\rhd'' \beta$. So $\alpha \not\rhd \beta$. Since $\alpha \rhd y$ and $x \rhd \beta$, $i(\alpha) < j(y)$ and $i(x) < j(\beta)$. Since $\alpha \not\rhd \beta$, $i(\alpha) \ge j(\beta)$. Therefore, it must be $i(x) \le j(y) - 2$, a contradiction. □

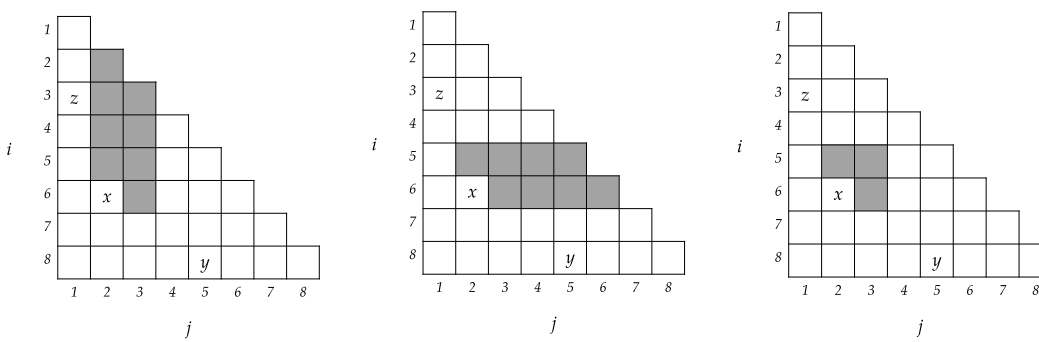
Definition 2. $(i, j)$ is called a prohibited cell if there exists $z$ such that $i(z) < i$ and $j(z) > j$. Otherwise, it is called a safe cell (see Figure 7a).




(A) Prohibited cells because of $z$

(B) Movable cells for $x$

FIGURE 7. Prohibited and movable cells.




(A) Movable cells if $x \succ z, y$

(B) Movable cells if $y, z \succ x$

(C) Movable cells if $y \succ x \succ z$

FIGURE 8. Examples of movable cells with different preferences.

To obtain a semiorder representation, we need to move each alternative that is in a prohibited cell to a safe cell and still the representation holds. The next definition describes a way in which alternatives can be moved.

DEFINITION 3. An alternative $x$ can be moved to the cell $(i, j)$ where $i \geq j$ if $(a)$ $i \leq i(x)$ and $j \geq j(x)$, $(b)$ $x \succ y$ for all $y$ with $i < j(y) \leq i(x)$, $(c)$ $z \succ x$ for all $z$ with $j(x) \leq i(z) < j$.

Definition requires that the alternatives in prohibited cells must move up and right (Condition (a)). As an outcome $x$ is moved a new cell, $(i, j)$, it is possible that there exists $y$ such that $i(x) \geq j(y)$ but $i < j(y)$. Condition (b) requires that in this case $x \succsim y$. Suppose to the contrary that $y \succ x$. Since $i < j(y)$, in the new representation $x \rhd' y$. But in the original representation, we have $x \not\rhd y$. So, the two representations must represent different preferences. Condition (c) can be understood similarly.

To understand this definition, we provide three examples (Figure 8). In Figure 8a, we have $x \succ y, z$. Since we have $z \not\rhd x$ and $x \succ z$, $x$ cannot be moved a cell where $z$

will eliminate $x$ (Condition (c)). That is, $j \leq i(z) = 3$. On the other hand, since $x \succ y$, there is no restriction on movement on $i$. In Figure 8b, we have completely opposite situation $y, z \succ x$. Since we have $x \not\rhd y$ and $y \succ x$, $x$ cannot be moved a cell where $x$ will eliminate $y$ (Condition (b)). That is, $i \geq j(y) = 5$. On the other hand, since $z \succ x$, there is no restriction on movement on $j$. Finally, we provide an example where both Condition (b) and (c) induce restrictions because we have $y \succ x \succ z$.

Claim 10. *Suppose $\beta \succ y \succ \alpha$ and $\alpha \rhd y \rhd \beta$. If there exists $x$ such that $x \not\rhd \beta$ and $\alpha \not\rhd x$, then $x \succ \beta$ or $\alpha \succ x$.*

Proof. Suppose $\beta \succ x \succ \alpha$ and we shall get a contradiction. Then we have $\beta = c(\beta x)$ and $c(x\alpha) = x$ because $x \not\rhd \beta$ and $\alpha \not\rhd x$. By the assumption, we have $\beta \succ c(\beta y) \succ c(\alpha y)$. By Claim 7, $\alpha\beta xy$ must be equal to either $c(\beta x) = \beta$ or $c(\alpha y) = \alpha$. Consider $c(\beta y)$ and $c(x\alpha)$, both of which are strictly worse than $\beta$ and strictly better than $\alpha$. Axiom 3 dictates that $\beta \succ \beta y \succsim c(\alpha\beta xy) \succsim x\alpha \succ \alpha$. Hence, $c(\alpha\beta xy)$ cannot be equal to $\beta$ or $\alpha$, which is a contradiction. □

Given the assumptions of Claim 10, we have $j(x) \leq i(\alpha) < j(y)$ and $i(y) < j(\beta) \leq i(x)$. This means that $(i(x), j(x))$ is a prohibited cell because of $y$. This means that $x$ needs to be moved. Claim 10 illustrate that $x$ can be moved because $x \succ \beta$ or $\alpha \succ x$. The next claim shows that there is a unique way to move $x$. That is, $x$ can be moved to either $(i(y), j(x))$ or $(i(x), j(y))$ but not to both.

Claim 11. *If there exist two alternatives $x$ and $y$ such that $i(x) > i(y)$ and $j(x) < j(y)$, then $x$ can be moved to either $(i(y), j(x))$ or $(i(x), j(y))$ but not to both.*

Proof. There exist two alternatives $\alpha$ and $\beta$ such that $i(\alpha) = j(y) - 1$ and $j(\beta) = i(y) + 1$.[20] By Claims 8 and 9, we have $\beta \succ y \succ \alpha$ and $\alpha \rhd y \rhd \beta$. Since $j(\beta) \leq i(x)$ and $j(x) \leq i(\alpha)$, $x \not\rhd \beta$ and $\alpha \not\rhd x$ by Claim 8. Thus, by Claim 10, we have $x \succ \beta \succ \alpha$ (so $x$ cannot be moved to $(i(x), j(y))$ because of $\alpha$) or $\beta \succ \alpha \succ x$ (so $x$ cannot be moved to $(i(y), j(x))$ because of $\beta$). Therefore, all we need to show is that $x$ can be moved to either of them.

**Case I:** $x \succ \beta$. We show that $x$ can be moved to $(i(y), j(x))$. First, Condition (a) holds trivially: $i(y) \leq i(x)$ and $j(x) \geq j(x)$. For Condition (b), take an element $z$ such that $i(y) < j(z) \leq i(x)$ (so $y \rhd z$ but $x \not\rhd z$). Then it must be either $y \succ z$ (which implies $x \succ z$) or $z \succ y = c(yz)$ in which case we have $z \succ y \succ \alpha$ and $\alpha \rhd y \rhd z$ (with $x \not\rhd z$ and $\alpha \not\rhd x$). By Claim 10, we should have $x \succ z$ or $\alpha \succ x$. Since we are considering the case $x \succ \beta (\succ \alpha)$, it must be $x \succ z$. Condition (c) is trivially satisfied because $j = j(x)$.

**Case II:** $\alpha \succ x$: Condition (a) and (b) will be now trivial while Condition (c) can be proven in the same way how we prove Condition (b) in case I. □

Claim 12. *Let*

$$U_x = \big\{y : i(x) > i(y),\ j(x) < j(y) \text{ and } x \text{ can be moved to } \big(i(y), j(x)\big)\big\} \cup \{x\}$$

$$R_x = \big\{y : i(x) > i(y),\ j(x) < j(y) \text{ and } x \text{ can be moved to } \big(i(x), j(y)\big)\big\} \cup \{x\}$$

[20]This is because since neither $i(y)$ is the smallest nor $i(y)$ is the largest integer within the range of $i$.

*and let*

$$i_x = \min_{y \in U_x} i(y) \quad \textit{and} \quad j_x = \max_{y \in R_x} j(y)$$

*Then (i) $x$ can be moved to $(i_x, j_x)$, and (ii) $(i_x, j_x)$ is a safe cell. That is, there is no $z$ with $i(z) < i_x$ and $j(z) > j_x$.*

Proof. Notice that by the definitions of movability, $i_x \le i(x)$ and $j_x \ge j(x)$.

(i) Clearly, $i_x \le i(x)$ and $j_x \ge j(x)$ as $x \in U_x, R_x$. First, we show that $i_x \ge j_x$. Take an alternative $y \in U_x$ such that $i(y) = i_x$. Since $y \in U_x$, $x$ cannot be moved to $(i(x), j(y))$ by Claim 11. By the definition of movability, $x$ cannot be moved to $(i(x), j)$ if $j \ge j(y)$. Hence, for all $z \in R_x \setminus \{x\}$, $j(z) < j(y)$, which means $j_x = \max_{z \in R_x} j(z) \le j(y)$. Since $i(y) \ge j(y)$, we have $j_x \le j(y) \le i(y) = i_x$.

Since $x$ can be moved to $(i_x, j(x))$, then the second condition of the movability is satisfied. Similarly, we can prove the third requirement as well. Therefore, $x$ can be moved to $(i_x, j_x)$

(ii) If $z \notin U_x, R_x$, then by Claim 11, it must be $(i_x \le) i(x) \le i(z)$ or $j(z) \le j(x) (\le j_x)$. If $z \in U_x$, then $i(z) \ge i_x$. If $z \in R_z$, then $j(z) \le j_x$. □

Now, define $x \bar{\rhd} y$ if and only if $j_y > i_x$.

Claim 13. *$\bar{\rhd}$ is a semiorder.*

Proof. Since $i_x \ge j_x$ by Claim 12 for all $x$, $\bar{\rhd}$ is an interval order.

Next, we shall show that if $(i, j)$ is a safe cell, there is no element $x$ such that $i_x < i$ and $j_x > j$. Suppose there is such $x$. Notice that it must be $i \le i(x)$ or $j \ge j(x)$ because $(i, j)$ is a safe cell, so it must be $i_x < i(x)$ or $j > j(x)$. Suppose $i_x < i(x)$. Then there exists $y$ such that $i(y) = i_x$ and $j(x) < j(y)$ such that $x$ can be moved to $(i(y), j(x))$. By Claim 11, $x$ cannot be moved to $(i(x), j(y))$, so it cannot be moved to $(i(x), j')$ for any $j' \ge j(y)$. Since $(i, j)$ is a safe cell and $i(y) = i_x < i$, it must be $j(y) \le j (< j_x)$. Hence, $x$ cannot be moved to $(i(x), j_x)$, which contradicts the definition of $j_x$ unless $j_x = j(x)$. But if so, $j(y) > j_x > j$ but this contradicts that $(i, j)$ is a safe cell. Analogously, we can show a contradiction if $j > j(x)$.

By Claim 12, all elements have been moved to safe cells, so there is no pair of elements $x$ and $y$ such that $i_x < i_y$ and $j_x > j_y$. Therefore, if $i_x < j_y \le i_y < j_z$ (i.e. $x \bar{\rhd} y \bar{\rhd} z$) then for any $w$, it must be either $j_w > j_y$ or $i_w \le i_y$, which implies $j_w > i_x$ or $i_w < j_z$ (i.e., $x \bar{\rhd} w$ or $w \bar{\rhd} z$). Therefore, $\bar{\rhd}$ is a semiorder. □

Claim 14. *If $x \rhd y$, then $x \bar{\rhd} y$.*

Proof. By definitions of $i'$ and $j'$, $i_x \le i(x)$ and $j_x \ge j(x)$ for all $x$. Therefore, if $x \rhd y$, then $i_x \le i(x) < j(y) \le j_y$ so we have $x \bar{\rhd} y$. □

Claim 15. *If $x \bar{\rhd} y$ but not $x \rhd y$, then $x \succ y$.*

Proof. First, we shall note that both $x$ and $y$ must be in prohibited cells. If neither of them is in, $i_x = i(x)$ and $j_y = j(y)$ so $x \overline{\rhd} y$ and not $x \rhd y$ cannot happen at the same time. If only $x$ is in a prohibited cell, then $i_x < j(y) \leq i(x)$ so $x$ cannot be moved to $(i_x, j_x)$. Similarly, we can prove that it is not possible that only $y$ is in a prohibited cell.

Next, we shall show that $i_x < i(x)$ and $j_x > j(x)$. Since $x$ can be moved to $(i_x, j_x)$ while $y \succ x$, it must be $i_x \geq j(y)$ because $j(y) \leq i(x)$ (i.e. not $x \rhd y$). Combined with $x \overline{\rhd} y$, we get $j_y > j(y)$. Flipping $x$ and $y$, one can prove $i_x < i(x)$.

Therefore, there must exist $z$ and $z'$ with $i(z) \in [j(y), j_y - 1]$ and $j(z') \in [i_x - 1, i(x)]$ (notice that these intervals are nonempty). Furthermore, we can take such $z$ and $z'$ so that $i(z) = j(z') - 1$ because $i_x - 1 < j_y - 1$ and $i(x) > j(y)$. Thus, $z' \succ z$ by Claim 9. Since $x$ is movable to $(i_x, j_x)$, we have $x \succ z'$. Similarly, we have $z \succ y$. Therefore, we conclude $x \succ y$. □

Claim 16. *$c(S)$ is equal to the $\succ$-best element in $\Gamma_{\bar{\rhd}}(S)$.*

Proof. We know $\bar{\rhd}$ is transitive, $\bar{\rhd} \supseteq \rhd$ and $x \succ y$ whenever $x \bar{\rhd} y$ but not $x \rhd y$. It is easy to see that this claim can be proven in the exactly same way as Claim 6. □

**(Representation ⇒ Axioms 1–4)** Showing that the first axiom is necessary is straightforward. Let

$$\Gamma(A) = \left\{ x \in A : \max_{y \in A} v(y) - v(x) \leq \overline{w} \right\}$$

For Axiom 2, if $x \succ c(A \cup x)$ then $A$ must have an element $y$ with $v(y) > v(x) + \overline{w}$, so it is clear that $\Gamma(A) = \Gamma(A \cup x)$ so $c(A) = c(A \cup x)$.

Axiom 3: Let $x^*$ be the $u$-best element in $\Gamma(A \cup B)$. Then it must be in $\Gamma(A)$ or $\Gamma(B)$ so it is not possible that $c(A \cup B)$ is strictly preferred to both $c(A)$ and $c(B)$. Now we show that the union cannot be strictly worse than both. Let $x_A$ and $x_B$ be the $u$-best elements in $\Gamma(A)$ and $\Gamma(B)$, respectively, and take $v_A$ and $v_B$ be the maximum values of $v$ in $A$ and in $B$, respectively. Then we have

$$v_A \leq v(x_A) + \overline{w} \quad \text{and} \quad v_B \leq v(x_B) + \overline{w}$$

Therefore, the maximum value of $v$ in $A \cup B$ is the higher one between $v_A$ and $v_B$, either $x_A$ or $x_B$ must be in $\Gamma(A \cup B)$ so $c(A \cup B)$ must be weakly better than either $c(A)$ or $c(B)$.

Finally, we show that the representation implies Axiom 4. Suppose $x \succ c(xy) \succ c(yz)$, then it must be $x \succ y \succ z$, $v(y) - v(x) > \overline{w}$ and $v(z) - v(y) > \overline{w}$. Therefore, $v(z) - v(x) > 2\overline{w}$.

Since $c(tz) = t$, we must either "$z \succ t$ and $v(t) - v(z) > \overline{w}$" or "$t \succ z$ and $v(z) - v(t) \leq \overline{w}$." In both cases, we have $v(t) - v(x) > \overline{w}$, hence we have $c(xt) = t$.

## *Proof of Proposition 1*

We know that $\phi$ is convex, increasing and not linear, i.e., $\phi'$ is not constant. Since $\phi$ is convex, $\phi(\delta) \leq \phi(\bar{v}) - \phi(\bar{v} - \delta)$ for all $\bar{v} > \delta > 0$. There must exist $\delta > 0$ such that the

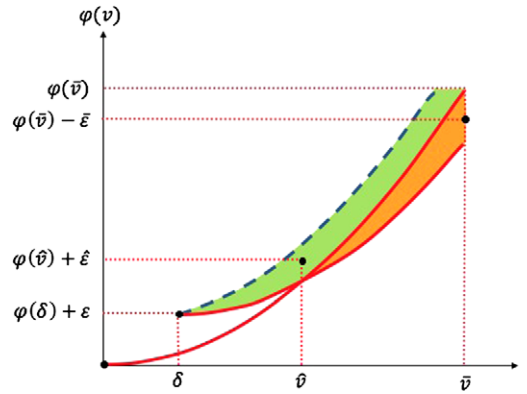



FIGURE 9. Construction of the four alternatives in Proposition 1.

inequality is strict since otherwise

$$\frac{\phi(\delta)}{\delta} = \frac{\phi(\bar{v}) - \phi(\bar{v} - \delta)}{\delta}.$$

Taking $\delta$ to zero, this implies that $\phi'(0) = \phi'(\bar{v})$, and $\phi$ must be linear, a contradiction.

Fix $\delta > 0$ such that the inequality is strict, and given $\delta$ choose $\epsilon > 0$ such that $\phi(\bar{v} - \delta) + \phi(\delta) + \epsilon < \phi(\bar{v})$. Let $\hat{\phi}(a) = \phi(a - \delta) + \phi(\delta) + \epsilon$ for all $a \in [\delta, \bar{v}]$. Note by construction $\hat{\phi}(\delta) > \phi(\delta)$ and $\hat{\phi}(\bar{v}) < \phi(\bar{v})$. Hence, there exists $\hat{v} \in (\delta, \bar{v})$ such that $\hat{\phi}(\hat{v}) = \phi(\hat{v})$. Next, we define four alternatives, $x$, $y$, $z$, and $t$ in terms of their $u$ and $v$ values. The richness of $X$ guarantees that the four alternatives exist. Let $t \equiv (u_t, v_t) = (0, \bar{v})$ and $z \equiv (u_z, v_z) = (\phi(\delta) + \epsilon, \bar{v} - \delta)$. Let $\hat{\epsilon} \in (0, \epsilon)$ and define $y \equiv (u_y, v_y) = (\phi(\hat{v}) + \hat{\epsilon}, \bar{v} - \hat{v})$. Let $\bar{\epsilon} \in (0, \phi(\bar{v}) - \hat{\phi}(\bar{v}))$ and define $x \equiv (u_x, v_x) = (\phi(\bar{v}) + \bar{\epsilon}, 0)$. See Figure 9 for the construction of the four alternatives.

By construction $u_x > u_y > u_z > u_t$. It is easy to see that $c(x, z) = x$, $c(y, t) = y$ and $c(x, y, z, t) = z$ and $(u, v, \phi)$ violates Axiom 3.

### *Proof of Proposition 2*

PROOF. We first show that, for any $x$, the best compromising contract is strictly better than the best indulging contract as long as $w > 0$ and $x \neq z^*$. Indeed, it is easy to see that, if $w = 0$ then the two contracts generate the same revenue and if $x = z^*$, the compromising contract reduces to an indulging contract. Now suppose $w > 0$ and $x \neq z^*$. If the revenue from the compromising contract is $v(x) - e(y^*)$, the result is immediate. Suppose the revenue from the best compromising contract is $u(x) + e(z^*) - e(y^*) - w$. Since, by definition of $z^*$, $u(x) + e(z^*) \geq v(x)$, it exceeds $v(x) - e(y^*) - w$ which is the revenue from the best indulging contract.

Now, let us compare the best commitment contract and the best compromising contract. If $e(z^*) - e(x) > w$, then the best compromising contract yields $v(x) - e(y^*)$, which is weakly greater than $u(x)$ since $e(x) \geq e(y^*)$. If $e(z^*) - e(y^*) > w \geq e(z^*) - e(x)$, the best compromising contract yields $u(x) + e(z^*) - e(y^*) - w$. Since $e(z^*) - e(y^*) - w > 0$, the best compromising contract is strictly better than the best commitment contract.

If $e(z^*) - e(y^*) \leq w$, then $e(z^*) - e(x) \leq e(z^*) - e(y^*) \leq w$. This means the best compromising contract yields $u(x) + e(z^*) - e(y^*) - w$. Since $e(z^*) - e(y^*) - w \leq 0$, the best commitment contract is better. □

*Proof of Proposition 3*

Proof. (i) Assume $e(z^*) - e(x^v) \geq w$. By Proposition 2, the optimal contract must be a compromising contract. Since $e(z^*) - e(x^v) \geq w$, $x^v$ provides the highest revenue among the alternative satisfying $e(z^*) - e(x) \geq w$. Now consider $x$ such that $e(z^*) - e(x) < w$. Then we have $e(z^*) - v(x) + u(x) < w$. This implies that $u(x) + e(z^*) - e(y^*) - w < v(x) - e(y^*)$. By definition, $u(x) + e(z^*) - e(y^*) - w < v(x) - e(y^*) \leq v(x^v) - e(y^*)$ for all $x$ such that $e(z^*) - e(x) < w$. This establishes the fact that $x^v$ provides the highest revenue overall.

(ii) Assume $e(z^*) - e(x^v) < w < e(z^*) - e(x^u)$. By Proposition 2, the optimal contract must be a compromising contract.

(iii) Assume $e(z^*) - e(x^u) \leq w < e(z^*) - e(y^*)$. By Proposition 2, the optimal contract must be a compromising contract. Since $e(z^*) - e(x^u) \leq w$, $x^u$ provides the highest revenue among the alternative satisfying $e(z^*) - e(x) \leq w$. Now consider $x$ such that $e(z^*) - e(x) > w$. Then we have $e(z^*) - v(x) + u(x) > w$. This implies that $u(x) + e(z^*) - e(y^*) - w > v(x) - e(y^*)$. By definition, $u(x^u) + e(z^*) - e(y^*) - w > u(x) + e(z^*) - e(y^*) - w > v(x) - e(y^*)$ for all $x$ such that $e(z^*) - e(x) > w$. This establishes the fact that $x^u$ provides the highest revenue overall. The price, which is $u(x^u) + e(z^*) - e(y^*) - w$, is strictly higher than $u(x^u)$ since $w < e(z^*) - e(y^*)$. Therefore, the efficient service $x^u$ is sold at a price exceeding $u(x^u)$.

(iv) Assume $e(z^*) - e(y^*) \leq w$. By Proposition 2, the optimal contract is the commitment contract selling the efficient service $x^u$ at price $u(x^u)$.

□

*Proof of Proposition 4*

Proof. To see part (i), recall that, by equation (8), the maximum revenue for selling $x$ is $\min[v(x) - e(z^*), u(x) - e(y^*) + e(z^*) - w]$. Hence, the monopolist's revenue for selling any alternative is weakly decreasing in $w$. This implies that the monopolist's optimal profit is weakly decreasing in $w$.

Next, we show part (ii). That is, we show total surplus is weakly increasing in $w$. Suppose $w > w'$. If $w \geq e(z^*) - e(y^*)$, the optimal contract sells $x^u$ at price $u(x^u)$. This maximizes the total surplus $(u - c)$ and gives 0 to the consumer. Clearly, the optimal contract under $w'$ does not generate more total surplus or consumer's ex ante welfare. Thus, we focus on the case where $w < e(z^*) - e(y^*)$. Let

$$E_1 = \{s : v(s) - e(y^*) - c(s) \leq u(s) - e(y^*) + e(z^*) - w - c(s)\}$$

$$E_2 = \{s : u(s) - e(y^*) + e(z^*) - w - c(s) \leq v(s) - e(y^*) - c(s) \leq u(s) - e(y^*) + e(z^*) - w' - c(s)\}$$

$$E_3 = \{s : u(s) - e(y^*) + e(z^*) - w' - c(s) \leq v(s) - e(y^*) - c(s)\}$$

Clearly, these there sets cover the entire alternative set. With this definition, notice that the monopolist's highest profits from selling $s$ is given by

when the willpower is $w$:

$$v(s) - e(y^*) - c(s) \quad \text{when } s \in E_1 \qquad u(s) - e(y^*) + e(z^*) - w - c(s) \quad \text{when } s \in E_2 \cup E_3$$

and when the willpower is $w'$:

$$v(s) - e(y^*) - c(s) \quad \text{when } s \in E_1 \cup E_2 \qquad u(s) - e(y^*) + e(z^*) - w' - c(s) \quad \text{when } s \in E_3$$

Let $x$ and $x'$ be the alternative sold in the optimal contract and $\pi$ and $\pi'$ the profit generated by the optimal contract under $w$ and $w'$, respectively.

**Case 1:** $x, x' \in E_1$

In this case, the both of them maximize $v(s) - e(y^*) - c(s)$. By the assumption (the uniqueness of the optimal alternative), $x = x'$. Thus the total surplus must be equal.

**Case 2:** $x \in E_1$**,** $x' \in E_2$

In this case, $x$ maximizes $v - e(y^*) - c$ in $E_1$ while $x'$ maximizes the same object in $E_1 \cup E_2$. Thus, $\pi \leq \pi'$. By Proposition 3(1), it must be $\pi = \pi'$. Since the monopolist can earn the same profit by selling $x$ under $w'$, it must be $x = x'$ (so $x, x' \in E_1 \cap E_2$) by the assumption.

**Case 3:** $x \in E_1$**,** $x' \in E_3$

Since $x \in E_1$, $\pi = v(x) - e(z^*) - c(x) \leq u(x) - e(y^*) + e(z^*) - w - c(x)$. If $x'$ was sold under $w$, it would generate $u(x') - e(y^*) + e(z^*) - w - c(x')$, which must be smaller than $\pi$. Thus $u(x) - c(x) > u(x') - c(x')$.

**Case 4:** $x \in E_2 \cup E_3$**,** $x' \in E_1$

In this case, $\pi = u(x) - e(y^*) - e(z^*) - w - c(x) \leq v(x) - e(y^*) - c(x)$ and $\pi' = v(x') - e(y^*) - c(x')$. However, the monopolist could have obtained the profit $\pi'$ by selling $x'$ even under $w$ so $\pi \geq \pi'$. By Proposition 3(1), $\pi = \pi'$. By the assumption, $x = x'$.

**Case 5:** $x, x' \in E_2 \cup E_3$

In this case, $\pi = u(x) - e(y^*) + e(z^*) - w - c(x)$. By selling $x'$ under $w$, the monopolist's profit would be $u(x') - e(y^*) + e(z^*) - c(x') - w$, which cannot be greater than $\pi$ so $u(x) - c(x) \geq u(x') - c(x')$.

Finally, to see part (iii), note that consumer's welfare is total surplus minus monopolist's profit. Since total surplus is weakly increasing and profit is weakly decreasing, consumer's welfare is weakly increasing in $w$. □

## Appendix B

Suppose $c$ defined on a *fixed finite* set of alternatives $A$ is generated by the limited willpower $(u, v, w)$. We now demonstrate that $c$ can be generated by the convex self-control model.

For any $x, y \in A$ such that $0 \leq v(y) - v(x) \leq w$, choose $l > 0$ such that $u(x) - l(v(y) - v(x)) > u(y)$ if $u(x) > u(y)$; otherwise, $u(x) - l(v(y) - v(x)) < u(y)$. The second inequality holds for any $l > 0$. If the first inequality holds for $l$, it also holds for any $l'$ such that

$0 < l' < l$. Since we have finitely many inequalities (hence finitely many $l$'s), let $l^* > 0$ be small enough so that all the inequalities hold for $l^*$.

Similarly, for any $x$, $y$ such that $v(y) - v(x) > w$, choose $k > 0$ such that $u(x) - k(v(y) - v(x) - w) + l^* w < u(y)$. If the inequality holds for $k$, it also holds for $k' > k$. Since we have finitely many inequalities (hence finitely many $k$'s), let $k^* > 0$ be large enough so that all the inequalities hold for $k^*$.

Finally, define the cost function as

$$\phi_{(l^*,k^*,w)}(x) = \begin{cases} l^* x & \text{if } x \leq w \\ k^*(x - w) + l^* w & \text{if } x > w. \end{cases}$$

By construction, the convex self-control model $(u, v, \phi_{(l^*,k^*,w)})$ generates $c$.

## References


Ahn, David, Ryota Iijima, Yves Le Yaouanq, and Todd Sarver (2019), "Behavioral characterizations of naivete for time-inconsistent preferences." (Forthcoming in Review of Economic Studies). [280]

Ahn, David S. and Todd Sarver (2013), "Preference for flexibility and random choice." *Econometrica*, 81, 341–361. [280]

Baumeister, Roy and Kathleen Vohs (2003), "Willpower, choice, and self-control." In *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*, 201–216, Russell Sage Foundation Publications. 6. [280]

Bernheim, B. Douglas and Antonio Rangel (2004), "Addiction and cue-triggered decision processes." *American Economic Review*, 94, 1558–1590. [279]

Chandrasekher, Madhav (2010), *A Model of Temptation-Induced Compromise*. Unpublished paper, Department of Economics, W.P. Carey School of Business. [281]

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini (2009), "Temptation-driven preferences." *Review of Economic Studies*, 76, 937–971. [281]

DellaVigna, Stefano and Ulrike Malmendier (2004), "Contract design and self-control: Theory and evidence." *Quarterly Journal of Economics*, 119, 353–402. [282, 289]

Dillenberger, David and Philipp Sadowski (2012), "Ashamed to be selfish." *Theoretical Economics*, 7, 99–124. [281]

Eliaz, Kfir and Ran Spiegler (2006), "Contracting with diversely naive agents." *Review of Economic Studies*, 73, 689–714. [282, 289, 290]

Faber, Ronald J. and Kathleen D. Vohs (2004), "To buy or not to buy?: Self-control and self-regulatory failure in purchase behavior." In *Handbook of Self-Regulation: Research, Theory, and Applications*, 509–524, Guilford Press, New York, New York. [280]

Fishburn, Peter (1970), *Utility Theory for Decision Making*. New York, NY. [303, 305, 306]

Fudenberg, Drew and David K. Levine (2006), "A dual-self model of impulse control." *American Economic Review*, 96, 1449–1476. [279, 281]

Fudenberg, Drew and David K. Levine (2012), "Timing and self-control." *Econometrica*, 80, 1–42. [280]

Grant, Simon, Sung-Lin Hsieh, and Meng-Yu Liang (2017), "Costly self-control and limited willpower." (Forthcoming in Economic Theory). [281, 287, 288, 300]

Gul, Faruk and Wolfgang Pesendorfer (2001), "Temptation and self-control." *Econometrica*, 69, 1403–1435. [279, 280, 281, 287, 288]

Gul, Faruk and Wolfgang Pesendorfer (2004), "Self-control and the theory of consumption." *Econometrica*, 72, 119–158. [279]

Gul, Faruk and Wolfgang Pesendorfer (2005), "The revealed preference theory of changing tastes." *Review of Economic Studies*, 72, 429–448. [281]

Heidhues, Paul and Botond Kőszegi (2010), "Exploiting naïvete about self-control in the credit market." *American Economic Review*, 100, 2279–2303. [289]

Kreps, David M. (1979), "A representation theorem for "preference for flexibility"." *Econometrica*, 47, 565–577. [280, 300]

Laibson, David (1997), "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics*, 112, 443–478. [279]

Lipman, Barton and Wolfgang Pesendorfer (2013), "Temptation." In *Advances in Economics and Econometrics, Tenth World Congress, Volume 1, Economic Theory*. Econometric Society Monographs, 243–288, Cambridge University Press. 8. [281]

Loewenstein, George and David Schkade (1999), "Wouldn't it be nice? Predicting future feelings." chapter 16, 85–105, Russell Sage Foundation, New York, New York. [280]

Masatlioglu, Yusufcan, Daisuke Nakajima, and Emre Ozdenoren (2019), "Contract design with costly convex self-control." Unpublished paper, arXiv:1907.07628. [299]

Mirkin, Boris G. (1979), *Group Choice*. Halsted Press, Washington, DC. Translated by Yelena Oliker. [306]

Muraven, Mark, Dikla Shmueli, and Edward Burkley (2006), "Conserving self-control strength." *Journal of personality and social psychology*, 91, 524–537. [280]

Noor, Jawwad and Norio Takeoka (2010), "Uphill self-control." *Theoretical Economics*, 5, 127–158. [281, 287, 288]

Noor, Jawwad and Norio Takeoka (2015), "Menu-dependent self-control." *Journal of Mathematical Economics*, 61, 1–20. [281, 287, 288]

O'Donoghue, Ted and Matthew Rabin (1999), "Doing it now or later." *American Economic Review*, 89, 103–124. [279]

Ozdenoren, Emre, Stephen Salant, and Dan Silverman (2012), "Willpower and the optimal control of visceral urges." *Journal of the European Economic Association*, 10, 342–368. [280]

Shefrin, Hersh and Richard Thaler (1988), "The behavioral life-cycle hypothesis." *Economic Inquiry*, 26, 609–643. [279]

Spiegler, Ran (2011), *Bounded Rationality and Industrial Organization*. Oxford University Press. [290, 299]

Spiegler, Ran (2013), "Comments on "behavioral" decision theory." In *Advances in Economics and Econometrics,Tenth World Congress, Volume 1, Economic Theory*, 289–304, Cambridge University Press. 9. [280]

Strotz, Robert H. (1955), "Myopia and inconsistency in dynamic utility maximization." *Review of Economic Studies*, 23, 165–180. [281]

---

Co-editor Ran Spiegler handled this manuscript.

Manuscript received 2 December, 2017; final version accepted 22 July, 2019; available online 25 July, 2019.