Compulsive and Addictive Consumption

Subir Bose* Miltiadis Makris†

February 2010

Abstract

We present a new theory of rational addiction based on four premises. First, addictive consumption is compulsive. Second, occurrence of cravings is stochastic and depends on past behavior. Third, addicts try to rationally manage the process of cravings through their consumption. Fourth, our decision maker is not fully aware of the addictive properties of the substance. In our model there is scope for campaigns that inform consumers about the addictive properties of the various substances. Drugs with stronger withdrawal syndromes are associated with lower consumption. Moreover, our theory provides mico-foundations for nonlinear intrinsic habit-forming behavior by starting from a standard model of fully rational decision maker with intertemporally separable preferences. Our model may also give rise to extrinsic habit-forming behavior. In fact, in our model, due to addiction being harmful, we may have *contrarian* behavior due to informational cascades. Our model can feature failed attempts to quit and occasional use as a process of experimentation without relying on 'behavioral' assumptions. Finally, on a technical level, we derive some new results on monotonicity properties of the value function of a stochastic dynamic programming problem.

Keywords: Rational Addiction, Cue-triggered consumption, stochastic intertemporal optimization

JEL: TBC

*Department of Economics, University of Leicester

†*Corresponding Author.* Department of Economics, University of Leicester, University Road, Leicester LE17RH, UK, E-mail: M.Makris@le.ac.uk.

1 Introduction

Often consumption is addictive and (hence) compulsive.[1] Moreover, often consumers are not fully aware of the addictive properties (vis-a-vis themselves) of certain consumption goods. Consider for instance smoking which arguably has both obvious withdrawal syndromes ("I am dying for a smoke") and more subtle ones ("I would love to have a smoke right now", "smoking helps me to relax/concentrate"). However, existing literature ignores this. This paper attempts to fill this gap.

The seminal paper on rational addiction, Becker and Murphy (1988) (BM hereafter), investigates a nonlinear intrinsic habit formation model where past consumption increases the value of future consumption. The main insight is that higher future price of the addictive substance reduces current consumption. In that model, the decision maker is rational and fully aware of the implications of her consumption decisions when maximizes her intertemporal payoff. Thus, government policy should only depend on externalities imposed by addicts. Moreover, the decision maker would never choose to avoid cues or enter rehabilitation. Orphanides and Zervos (1995) (OZ hereafter) investigate the implications of the decision-maker facing subjective uncertainty about future health costs of addictive consumption. In that work, the consumer learns about true costs from past experience. If past consumption of the addictive substance has been costly, then the decision-maker becomes fully aware of the true costs of addictive consumption and the environment becomes identical to that in BM. OZ emphasize the importance of information about health consequences of addictive consumption. Gruber and Köszegi (2001) (GK hereafter) introduce dynamic inconsistency in BM by means of hyperbolic discounting. That work is motivated by evidence of unrealized intentions to quit at some time in the future and the search for self-control devices to help quit.[2] In their context government policy should also depend on the "internalities" imposed by dynamically inconsistent addicts. Moreover, commitment is valuable when it changes future behavior. However, the habit-forming preferences these papers deploy are ad hoc: they lack axiomatic foundations.[3]

[1] We would like to thank, for the very useful comments and discussions, participants in seminars at Exeter and Athens University of Economics and Business. The usual disclaimer applies.

[2] See Gruber and Köszegi (2001) for a discussion of this evidence.

[3] For an axiomatization of the linear intrinsic habit formation model see Rozen (2009).

Gul and Pesendorfer (2007) in turn characterize axiomatically preferences over menus of streams of consumption rather than on streams themselves and investigate their implications for rational addiction. In their context, past consumption affects cost of current self-control and the decision maker can choose consumption and future options that go against temptations. The latter are defined there as consumption bundles and future options that are costly to ignore (i.e. not choose). There as well taxes are harmful without externalities as in BM, while rehab is desirable as in GK, but as a commitment device to reduce temptation. Therefore, prohibitive policies are beneficial. Moreover, the demand for temporary commitment generates rehab cycles. Bernheim and Rangel (2004) (BR hereafter) investigate also addictive consumption in a framework that allows for micro-founded preferences,[4] but under "cue–induced mistakes". Specifically, the decision maker operates in a stochastic environment where she may be 'hit by temptations' that carry very high physiological and psychological costs of being ignored. In this case, she will always consume the addictive substance.[5] However, addicts can engage in activities that reduce the exposure to temptations whenever they are in a "cool" state, i.e. whenever they do not have a cue-triggered impulse. Thus rehab cycles arise due to a form of "consumption-smoothing".

In this paper, we focus on the case of health and monetary costs of addictive consumption being fully known by consumers (in contrast to OZ) and we present a new theory of rational addiction based on four central premises. First, addictive consumption is compulsive in that consumption decisions are influenced by the presence of temptations. At this stage, we should empha-

Rustichini and Siconolfi (2005) also axiomatize dynamically consistent habit formation over consumption streams, but do not offer a particular structure for the utility or form of habit aggregation.

[4]In fact, the model in BR can also accommodate intrinsic habit-forming preferences as in BM. However, most of the insights in BR do not rely on such preferences; the insights they offer can also be derived in a context with no intertemporal preference complementarities (see pp. 1567-1568 in BR).

[5]In fact BR describe their model in terms of a 'cold' and a 'hot' state. In the former, the decision maker matches actions to preferences, while in the latter the decision maker consumes the substance with no reference to her preferences. In this way the latter behavior may diverge from preferences. However, this behavior is also observationally very similar to the one that would arise when the decision maker always matches actions to preferences but in some states the cost of abstaining from the addictive consumption is very high that 'forces' the decision maker to consume the substance. For a related point see footnote 18 and the last paragraph in p. 1563 in BR.

size that one can interpret the models of MB, OZ, GK and BR as focusing on an environment where temptations are always present. This follows from their assumption that for any given future consumption levels, future health consequences of current consumption and strength of future temptations, abstention is always inferior to consumption. In these models, it is the degree of the temptations that depends on past consumption: in MB, OZ and GK this is described by the habit-forming process, while in BR this is described by the presence of a 'cold' and a 'hot' state with the latter being more likely the higher past consumption is. However, in our model (as in GP) temptations may not be present and in the absence of temptations, consumption of the addictive substance is inferior to abstention due to the presence of high health costs. Furthermore, in contrast to all the above papers, temptations here are cue-triggered taste-shocks/cravings as in Laibson (2001) (L hereafter). Second, temptations are endogenous in that their occurrence depends on past behavior (as in BR and GP but crucially in contrast to L[6]). Third, addicts understand their susceptibility to cravings (as in GP, BR and L) and try to rationally manage the process through their consumption even under a temptation (as in GP and L but crucially in contrast to BR). In fact, our decision maker would like to pay for not having a craving.[7] However, once she gets an urge then she will indulge in her craving as long as the (net of health and monetary costs) benefits from doing so outweigh the expected future addiction costs from increasing the likelihood of future cravings. The fourth premise of our theory, and what differentiates substantially our work from the existing literature, is that individuals are not fully aware of the easiness to quit. The reason for this is not some kind of dynamic inconsistency as in GK. We need to emphasize that the central premise of our theory

[6]Note here that in L the probability of occurrence of the various cues is exogenously determined, which is not the case here or in BR. In L, however, it is past cue-conditioned consumption of the substance that affects the degree of cue-triggered impulse to consume the substance (ie. the marginal utility from consumption). In this sense, preferences in L belong to the intrinsic habit-formation paradigm.

[7]This echoes the beneficial effect of removing temptations in GP and L. However, in our model, as in L, eliminating options would not be beneficial for consumers. Note also that we abstain from analyzing external ("lifestyle") activities that reduce exposure to temptations (which are the focus of BR, GK and to some extend in L). We recognize that individuals may manage their addiction through lifestyle activities. We abstract from this possibility to focus on the novel aspect of our theory, which is management of addiction through consumption in the face of uncertainty over the addictive properties of the substance.

is motivated by the fact that often the addictive properties of consuming certain goods are not well understood due to lack of relevant research or conclusive evidence: smoking, so-called 'comfort' eating, kleptomania, exposure to sexual content, shopping etc. are some examples. Hence our theory is consistent with the assumption that our decision maker is a Bayesian expected utility maximizer with standard discounting who, however, faces uncertainty over the likelihood of future temptations. In other words, nothing prevents our decision maker from being a standard *homo economicus* who nevertheless lacks some information about the addictive properties of the substance. Therefore, when she consumes, the decision maker in effect experiments and tries to infer the addictive properties of the substance by using her experience in terms of current and past cravings and past consumption. However, the theory is also consistent with the consensus that seems to be emerging in the neuroscience and psychology literature[8] that the way individuals associate levels of pleasure/pain (i.e. utility) with certain (consumption) activities (and hence the way individuals make decisions) is influenced by two mechanisms: cognitive control and impulses generated by forecasting.[9] The former associates activities and pleasure/pain in an undistorted way by "identifying alternative courses of action or projecting the future consequences of choices" (from BR p. 1563), while the latter may distort these associations by over-relying on recent experiences and repetitive associations in the past. Thus, the latter mechanism may give rise to misperceptions about the actual (current and future) pleasure/pain associated with the (current) consumption of addictive substances. Moreover, forecasts of utility from consumption may be influenced by the presence (or not) of certain environmental cues and complementarities between activities and consumption.[10]

We should also emphasize here that the theory we present may not be applicable to all addictive substances. For instance, we focus on cue-conditioned impulses that do not defeat higher cognitive control (while the existence of cue-conditioned cravings that override cognitive control is the focus of BR). Moreover, there are addictive goods whose addictive properties are well understood and relevant information is publicly available. However, there are also addictive goods for which the latter is not true and cue-triggered mis-

[8] For this see the excellent discussion in pp. 1562-1565 in BR and in pp. 84-86 in L.

[9] The latter is referred to in BR as impulses generated by the "Hedonic Forecasting Mechanism".

[10] See also L for a related discussion of cue-triggered consumption.

takes are not common[11] (smoking, sex, shopping, food - addictions and kleptomania are some examples). Therefore, we view our work as an important complement of the existing literature in understanding (rational) addiction.

The four central premises of our theory differentiate it substantially from the received literature. On the one hand, some of its predictions (briefly discussed below) are new and offer some very interesting insights for addictive consumption. On the other hand, our model of rational addiction gives rise to a decision problem which has been previously unexplored. Specifically, while the value function that characterizes the decision-maker's problem can be shown, using standard techniques, to be well-defined, the study of the decision-maker's choices requires proving certain monotonicity properties for the value function that do not have a counterpart in the papers mentioned above and the received dynamic programming literature.

In our model of rational addiction there is scope for informational policies that take the form of campaigns informing consumers about the addictive properties of the various substances (and not about the health and monetary costs of addictive consumption as, following the received literature, these are assumed to be known by consumers). Moreover, our theory provides microfoundations for nonlinear intrinsic habit-forming behavior by starting from a standard model of fully rational decision maker with intertemporally separable preferences. It does so by virtue of uncertainty over the likelihood of future temptations which is endogenous in that it depends on past behavior and the history of past and current temptations. Interestingly, our model may also give rise to extrinsic habit-forming behavior.[12] The reason is that in our context the past and current consumption of peers may provide valuable information about the addictive properties of the substance and thereby influence the consumption of our decision maker. In fact, in our model, due to addiction being harmful, we may have *contrarian* behavior as a result of informational cascades. Finally, our model can feature failed attempts to quit and occasional use despite the fact that we do not deploy any 'behavioral' assumption. Our model can feature such consumption patterns as a process of information acquisition (not due to dynamic inconsistency or cue-induced mistakes).

[11] On this see also footnote 18 in BR.

[12] In models of extrinsic habit formation, individuals derive utility from their relative position in society, as in the "catching up with the Joneses" effect of Abel (1990). This differentiates these models from the intrinsic formation paradigm where evaluation of own consumption uses as a reference point own past consumption.

Our analysis has also a number of other interesting implications. First, consumption patterns depend on the inherent addictive properties of the substance as well as on the family and social environment of individuals when they make their first consumption decision. The reason for the latter is that the environment will have a big impact on the prior of the decision maker about the addictive properties of the substance, which in turn will influence the first and subsequent consumption decisions. Second, more addictive drugs may be associated with lower consumption among more experienced users and higher consumption among new users (in contrast to BR). This will be the case if the degree of addiction is increasing in past consumption due to the fact that in our model higher (perceived) degree of addiction leads to lower consumption to reduce future cravings. Third, our model predicts that addictive substances with higher self-control costs are associated with lower consumption. Finally, it provides a theoretical foundation behind an existing method to stop smoking with, arguably, many beneficiaries (one of the authors is one).[13] This method in fact has at its centre the fact that smoking has very subtle cravings which give rise to forecasts of benefits that can be mistaken as actual benefits. It is also against the use of quit aids because, and our model agrees to this, they perpetuate addiction (though admittedly at lower health costs). Instead, subsidies for aids could be spent on information campaigns about the addictive properties of smoking that may reduce smoking significantly and thereby both health and future monetary costs.

The organization of the paper is as follows. The next section describes the basic model which is used to build our intuition and derive in Section 3 most of our results. The robustness of the insights of this basic model is the topic of Section 4, where, among others, the relation of our model to the intrinsic habit formation literature is also discussed. Finally, Section 5 concludes

2 Model

In this section we consider the simplest model that can capture our story. We consider some extensions later in Section 4 .

There are two periods $t = 1, 2$. Let δ be the discount factor. A consumer chooses action a_t in each period t where $a_t \in \{0, 1\}$; in what follows, we will

[13] See "The Easy Way to Stop Smoking" by Alan Carr, Penguin Books Ltd, 3rd Revised edition (1999).

use the words "action" and "consumption" interchangeably. $a_t = 1$ and 0 represent consumption of the addictive substance and abstention in period t respectively. We will often use smoking to describe the model, though our framework can describe other addictive goods as well as we have mentioned in the Introduction.

Net (of monetary and non-monetary short and long run costs) utility per period is given by the bounded function $u(a_t, x_t)$. The random variable $x_t \in \{0, 1\}$ is used to capture "urge" or "craving". In any period t, $x_t = 1$ (respectively, $x_t = 0$) represents the state when the urge is present (respectively, absent). The following two inequalities on the function u depict the basic assumption that in the absence of any intertemporal effects the optimal action of the consumer would be to choose $a_t = x$ in state x.

Assumption 1 $u(0, 0) > u(1, 0)$

Action $a_t = 1$ is costly and hence in the absence of any urge, $a_t = 0$ is the best action ceteris paribus. This is captured by assumption (1). This assumption also differentiates our work from BR.

Assumption 2 $u(1, 1) > u(0, 1)$

Even though $a_t = 1$ is costly, when the craving happens, the urge is sufficiently strong so as to make $a_t = 1$ the best action all other things equal. This is captured by assumption (2). Note that in our case $u(1, 1) - u(0, 1)$ is bounded. If it was unbounded then $x = 1$ could be thought of as a cue that defeats cognitive control along the lines of BR.

The above describe the ex-post (period t) preferences given the state x_t. However, we want to model the effect that the consumer is aware that compulsive consumption (we focus here) is bad. This is captured by assumption (3), that shows that (even after taking into account the ex-post preferences) the consumer would prefer not to get the urge.

Assumption 3 $u(0, 0) > u(1, 1)$.

As we see shortly, while assumptions (1) and (2) drive the second period optimal choice, the first period action is influenced by assumption (3) as well.

At this stage note that the above assumptions imply that the welfare costs of the addictive consumption are known and well understood by the decision maker. However, before we continue, we need to point out here that

the above formulation abstracts from any long run cumulative cost-effects of the compulsive consumption. We choose this formulation not because we think that such costs are not important, but in order to emphasize that our results do not rely on any cumulative welfare costs. In particular, note that $u(1,0)-u(0,0)$ represents the welfare cost (due to health and monetary costs) of compulsive consumption, while $u(0,0)-u(0,1)$ represents the welfare cost associated with having a craving and not consuming (ie. the cost of self-control). We discuss this, as also the issue of how our model relates to the habit formation models, in more detail in section 4.

We now describe the beliefs of the consumer about the evolution of the state x_t. The consumer does not know the true stochastic process but attaches a probability distribution (representing his prior) over the set of possible processes. For simplicity, we restrict attention to the situation when this set consists of the following two processes only. The first is an i.i.d process where in both periods, the probability that $x_t = 1$ is equal to p with $1 > p > 0$. The other process, which we refer to as the addictive process, depends on the past and current consumption. More specifically, under the latter process, the probability that $x_2 = 1$ is given by $f(h_1, a_1)$, with f being continuous, where h_1 is a measure of past consumption prior to period 1. For brevity, we will refer to h_1 as history in this and the next sections. Higher h_1 represents a higher level of past (passive) consumption, with $h_1 = 0$ representing no consumption in the past. We assume the following:

Assumption 4 *$f(h,a)$ is increasing in both arguments.*

Note that for the two period model it suffices to assume only that f is increasing in a. However, for the more general model, we also need the monotonicity with respect to h. Denote by f_0 the probability that $x_1 = 1$ which depends positively on consumption up to and including (some artificial) period zero.[14] It is the second process that makes consumption addictive: higher current or past consumption makes higher future consumption more likely through the effect of increasing the likelihood of occurrence of future cravings. The prior that the process is the addictive one is given by μ_0.

We also make the following intuitive assumption on the stochastic processes.

[14] In other words, we can think of an artificial period $t = 0$ and write f_0 as $f(h_0, a_0)$. We write f_0 to simply avoid notational cluttering.

Assumption 5 $f(0,0) = p$

Zero current and past consumption under the addictive process gives the same likelihood for craving in the next period as the i.i.d process. Put differently, for someone who has never consumed the addictive good, the chances of getting an urge next period is the same under the addictive process as it is under the i.i.d process. Note that assumptions (4) and (5) imply that $f \geq p$. Note also that assumption (5) implies that as long as the consumer's subjective belief puts some (initial) probability on the addictive process being the true process, she does expect an increase in the likelihood of future cravings whenever there is (additional) consumption of the addictive good. Note here that expressing the model in terms of a decision maker who knew that the probability of having a craving is $f(h, a, \theta)$ (with f being increasing in h and a for any θ) but did not know her addictive type $\theta \in \{\theta_l, \theta_h\}$, with $f(0,0,\theta_l) = f(0,0,\theta_h) \equiv p$, $f(h,a,\theta_h) > f(h,a,\theta_l)$ for $h > 0$ and/or $a > 0$, and $f(h,a,\theta_h)/f(h,a,\theta_l)$ being increasing in h and a, would give qualitatively similar results.

The sequence of events in any period is as follows. In any period t, the consumer starts with the prior μ_{t-1} and past behavior summarized by h_{t-1} and a_{t-1}.The state x_t is realized and is observed by the consumer, who then uses the realized value of x_t to update her prior to arrive at the posterior $\mu(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})$, where h_0 and a_0 are the measure of past consumption and actual consumption in (the artificial) period zero (and so $f(h_0, a_0) = f_0$). The consumer also chooses action a_t, given her beliefs, in order to maximize her intertemporal welfare. Being a rational consumer, she takes into account the possible implications of her current choice of action on the likelihood of the future occurrences of cravings. An important point to note is that the consumer cannot commit to future actions. In addition, given that between the observation in period t of state x_t and the observation in period $t+1$ of state x_{t+1} there is no new information about the state-generating process, we have for period $t+1$ prior μ_t that $\mu_t = \mu(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})$. That is, period $t+1$ prior is (endogenous and) equal to the posterior of the previous period.

For later use, we define

$$\begin{aligned} D &= u(0,0) - u(1,1) \\ B &= u(1,1) - u(0,1) \end{aligned}$$

As we will see in more detail shortly, D captures, in a way, the (long run)

benefit from not having the urge. B on the other hand represents the short run benefit from indulging to the urge.

3 Optimal consumption

In this section we analyze the optimal choices of the consumer. Here, and for the rest of the paper as well, we make the tie-breaking assumption that if in any period the consumer is indifferent between choosing $a_t = 1$ or $a_t = 0$, she will choose $a_t = 0$. It can easily be checked that no qualitative result is affected if one were to break the tie in the other way.

We start by considering the second period. Since there is no future period to consider, the posterior beliefs about the stochastic process are actually irrelevant, and the optimal consumption in the second period is determined solely by the second period state x_2. Clearly, given assumptions (1) and (2), the consumer chooses $a_2 = 1$ if and only if $x_2 = 1$. Therefore the welfare benefit of not having a craving in the second period is D.

Turning to the first period choices, assumptions (1) and (3) imply that the short run and the long run incentives are not in conflict when $x_1 = 0$ (The action $a_1 = 1$ is costly in terms of current period payoff. Moreover, for any posterior beliefs, it (weakly) worsens the future expected payoff by making cravings (weakly) more likely.[15]) Thus, if $x_1 = 0$, the optimal action is $a_1 = 0$.

The problem is more interesting when $x_1 = 1$. In this case, the posterior beliefs are important. Let for the rest of this section only, with some abuse of notation, μ_1 stand for the posterior in question.[16] Payoff from action a_1 is

$$u(a_1, 1) + \delta u(0, 0) - \delta D[(1 - \mu_1)p + \mu_1 f(h_1, a_1)] \tag{1}$$

The first term is the current period payoff while the second is the discounted payoff in the absence of any future urge. The third term represents the discounted expected cost when the craving occurs next period, which occurs with perceived probability $(1 - \mu_1)p + \mu_1 f(h_1, a_1)$.

[15] And, if $\mu_0(1 - f_0) > 0$, the future expected payoff is in fact strictly lower.

[16] By Bayes rule, the period-1 posterior (belief) that the true process is the addictive one given that $x_1 = 1$ is given by

$$\frac{\mu_0 f_0}{\mu_0 f_0 + (1 - \mu_0)p}$$

As expression (1) shows, when $x_1 = 1$, the current period payoff is maximized by choosing $a_1 = 1$. However, this raises the future cost by increasing the (posterior) likelihood of a craving (since $f(h_1, 1) > f(h_1, 0)$). Therefore when deciding whether or not to smoke the rational consumer balances the current "kick" from satisfying the urge, B, versus the decrease in discounted future utility, δD, from exposing herself to a greater risk of being addicted.

The resolution of this trade-off is then that non-smoking (i.e. $a_1 = 0$) is the optimal response even in the presence of craving if and only if

$$\frac{B}{\delta D} \leq \mu_1 \left[f(h_1, 1) - f(h_1, 0) \right] \tag{2}$$

The left hand side reflects the trade-off between the current "kick" and the future cost of addiction. The right hand side is the perceived increase in the probability of being addicted, i.e., of getting a craving in the future as a result of current smoking. Thus, if the increase in future perceived probability of craving is sufficiently high then no consumption is optimal. We collect the results on optimal consumption in the Proposition below.

Proposition 1 *Optimal Consumption for the two periods is given by*

(a) For $t = 2$, $a_2 = x_2$

(b) For $t = 1$, $a_1 = 0$ when $x_1 = 0$. When $x_1 = 1$, $a_1 = 0$ if and only if inequality (2) holds.

Proof. Parts (a) and the first part of (b) follows from the discussion above. To see the last part of (b), note that when $x_1 = 1$ the payoff in (1) implies that action $a_1 = 0$ is preferred to $a_1 = 1$ when

$$\begin{aligned} & u(0,1) + \delta u(0,0) - \delta D[(1-\mu_1)p + \mu_1 f(h_1, 0)] \\ \geq & \\ & u(1,1) + \delta u(0,0) - \delta D[(1-\mu_1)p + \mu_1 f(h_1, 1)] \end{aligned}$$

which after some straightforward algebraic manipulation, and substituting B for $u(1,1) - u(1,0)$ gives the inequality (2). ■

For the discussion that follows note that D can be rewritten as $D = u(0,0) - u(0,1) - B$, where $u(0,0) - u(1,0)$ represents the withdrawal cost, which is assumed to be independent of health and monetary costs of addictive consumption. Thus, B/D is increasing in B. Inequality (2) confirms certain results of the received literature. First, higher future non-addiction (ie. monetary or health) costs lower current consumption (consider a decrease in B

). Second, more addictive drugs are associated with lower consumption for any given beliefs (consider an increase in $f(h_1, 1) - f(h_1, 0)$ for any given h_1). Third, direct peer effects can reduce self-control (B increases); peer effects are discussed further later. Finally, reducing the occurrence of cravings and exposure to cues (e.g. change habits/environment) is beneficial for any given priors (consider a decrease in the true probability of $x_1 = 1$).

Importantly, Proposition (1), and in particular inequality (2), provide us with some new implications and insights. First, they illustrate the role of (a policy of providing) information (about the addictive properties of the compulsive consumption) in our model. Suppose the true process is in fact the addictive one but that the consumer does not know this and consequently has beliefs such that $\mu_1 < 1$. Then, if for some history h_1 we have $B < [f(h_1, 1) - f(h_1, 0)]\,[\delta D]$ but inequality (2) is not satisfied, the consumer will choose to smoke in period $t = 1$ but will stop (voluntarily) if provided with (credible) information about the true process. It is important to note that similar to the standard rational addiction model, (many) other types of policy interventions - for example, forcing the consumer not to smoke,[17] or raising taxes on the addictive good - are not welfare enhancing policy interventions in our set up where there are no consumption externalities and/or no merit good arguments.

Second, in this model, the prior μ_0 can be thought of arising as a result of the environment in which the decision maker has been raised (and prior to the first instance she has to decide whether to consume the compulsive good or not). As such, it can capture the cultural and family environment and habits as well as peer and own experience. Thus, for a young, and maybe excessively impulsive, youngster who lives in a family and social environment where smoking is the norm, μ_0 may be relatively low. In this case, the posterior μ_1 will be relatively low and hence, all other things equal, the above result implies that the likelihood of smoking is relatively high.

Third, consider a decision-maker with some prior μ_0 and (positive) past consumption such that $[B/\delta D]/[f(h_1, 1) - f(h_1, 0)] \leq \mu_0 < \frac{\mu_0 f_0}{\mu_0 f_0 + (1-\mu_0)p}$. One could think that in this case the decision maker would pledge before the occurrence of state x_1 to quit smoking and indeed abide by it regardless of the realized state. Next period coming though, the decision maker will smoke if $x_2 = 1$. Thus, our model features failed attempts to quit and occasional

[17] For instance, extraneous resistance to cravings (physical confinement, "special medicines" - see BR for discussion of such measures)

use as a process of information acquisition (not due to dynamic inconsistency or cue-induced mistakes). Of course, one might argue at this stage that this is an artifact of the two-period horizon of our basic model. However, as we show in the next section the insights in this section carry over in a model with more than two periods under certain plausible conditions. In fact, we show there that if there is a sufficiently long period with no cravings (and hence no consumption) then the decreasing posterior will ultimately be low, which in turn implies that if the decision maker is eventually 'hit' by a craving she may indeed indulge and get her short run 'kick'.

Fourth, quit aids are beneficial in reducing health costs but reduce self-control as they sustain addiction (consider an increase in B and hence an increase in the likelihood of wanting to smoke).

Fifth, drugs with stronger withdrawal syndromes are associated with lower consumption. This follows directly from recalling that self-control costs are represented by $u(0,0) - u(0,1)$ and $D = u(0,0) - u(0,1) - B$.

Finally, consider addictive substances such that the degree of addiction (for given beliefs μ_1) $f(h_1, 1) - f(h, 0)$ is increasing in past consumption history h_1. In this environment, more addictive drugs (for given beliefs μ_1) are associated with a higher likelihood that consumption will not take place for more experienced users. Loosely speaking, thus, more addictive drugs are associated with lower consumption among more experienced users and higher consumption among new users (in contrast to BR).

Our model features experimentation and learning of the addictive properties of a certain consumption good if μ_0 is different than 0 or 1. Note however that absence of learning does not change the essence of our main message insofar $\mu_0 < 1$. In fact, our case of information campaigns would be stronger if $\mu_0 = 0$ (and knowing that the process is addictive leads to no smoking). We conclude this section by discussing some new implications of our model for peer effects.

3.1 Peer Effects

Our model features two kinds of peer effects:

(1) Simple peer effect. Can be modelled as an increase in $u(1,1)$ which would imply an increase in B (and a decrease in D). This would give the standard peer effect where an agent has a higher tendency to smoke if more people around him also smoke. Alternatively, this can also be modelled as an increase in $f(h_{t-1}, a_{t-1})$ and p in period t for any given history. In other

words, for any given history, the more people smoke, the more likely it is that a particular individual of that group will face an urge to smoke.

(2) There is a less obvious peer effect which arises out of a process of information acquisition as a result of observing others' smoking practices. In particular suppose whether the process is addictive or not is the same for all individuals. The prior μ_0 is now the probability that the typical process is addictive. In such an environment observing others' consumption gives an individual further information about addictiveness of her own process. Given our assumptions on D and B, it is still the case that Lemma will still be true and hence $a_t = 1$ implies that $x_t = 1$. However, if $a_t = 0$ it may still be the case that $x_t = 1$. Hence as in herd behavior models actions will not perfectly reveal the state/signal x. It still remains that an agent does not smoke in the absence of an urge (i.e. $a_t = 0$ whenever $x_t = 0$). As in the herding behavior literature (e.g. Bikhchandani et al. 1992, and Banerjee 1992), an informational cascade can arise for $x_t = 1$. In more detail, an agent who would otherwise smoke may decide not to smoke if many of his peers have been smoking frequently as the latter would imply that the process is very likely to be addictive. Similarly, an agent who would otherwise not smoke may decide to smoke if many of his peers have not been smoking frequently as the latter would imply that the process is very likely not to be addictive.

Note the different implications of the two different forms of peer group effects. In the preference related peer group effect we get the standard conclusion of reinforcing smoking behavior in a group. In the information related peer group effect however, we can get contrarian behavior due to informational cascades. Therefore, our model can be thought of as providing some mico-foundations for extrinsic habit formation (ie. for consumption history of peers to affect own tendency to consume).

4 Extensions

We start by considering a model with $T > 2$ periods (where T can be finite or infinite) and then move to a discussion of cumulative welfare effects.

4.1 Longer Horizon

For any period $1 \leq t \leq T$, and given any history of consumption represented by the vector $(a_0, ..., a_{t-1})$, let the scalar h_t be a measure of past consumption,

which, hereafter, we will simply call the period t history. For example, for some $\rho \in (0, 1/2]$, we could have $h_t = \sum_{i=0}^{t-1} a_i(\rho)^{t-i}$. In this case, h_t belongs to the interval $[0, 1]$ for all t and $h_{t+1} = \rho(h_t + a_t)$. More generally, for any h_t, we postulate that $h_{t+1} = g(h_t, a_t)$ with $g(0, 0) = 0$, and the function g being continuous and increasing in both arguments. The fact that g is increasing in h differentiates also our model from that in GP. We also assume that for any $h \in [0, H]$, where H is a well-defined scalar, and $a \in \{0, 1\}$, $g(h, a) \in [0, H]$. Moreover, $h_1 \in [0, H]$.

Next, we describe the (natural) extensions of the corresponding terms and definitions of the two period model. Let $f(h_t, a_t)$ denote the probability that $x_{t+1} = 1$ under the addictive process. We retain all the assumptions from the 2-period model about function f. As in the two period model, we let μ_{t-1} and $\mu(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})$ denote period-t prior and posterior beliefs that the true process is the addictive one. As usual, the latter is obtained by updating μ_{t-1} (using Bayes rule) upon observing the realized value of x_t for any $t \geq 1$.[18] For reasons similar to those in the two-period model we have $\mu_t = \mu(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})$.

To simplify the discussion hereafter let as focus on the case when $0 < \mu_0 f_0 + (1 - \mu_0)p < 1$ and $f(h, 1) < 1$ for any h (and hence Bayes rule can be applied in any period t - even if $t > 1$ and $\mu_{t-1} = 1$ and $x_t = 0$).

Note then that as long as $1 > \mu_{t-1} > 0$, assumption (5) implies that, except for the case $h_{t-1} = a_{t-1} = 0$, the period-t posterior belief is increasing in x (otherwise is independent of x_t).[19] Furthermore, the period t posterior is increasing in the prior μ_{t-1}. Finally, if $1 > \mu_{t-1} > 0$, then $\mu(1, \mu_{t-1}, h_{t-1}, a_{t-1})$ is increasing, while $\mu(0, \mu_{t-1}, h_{t-1}, a_{t-1})$ is decreasing, in h_{t-1} a_{t-1} (otherwise they are independent of past consumptions).

It is helpful to define here the posterior probability that $x_{t+1} = 1$ given h_t, a_t and prior μ_t. Let this posterior probability be denoted by $\pi(a_t, h_t, \mu_t)$

[18] Formally,

$$\mu(1, \mu_{t-1}, h_{t-1}, a_{t-1}) = \frac{\mu_{t-1} f(h_{t-1}, a_{t-1})}{\mu_{t-1} f(h_{t-1}, a_{t-1}) + (1 - \mu_{t-1})p}$$

and

$$\mu(0, \mu_{t-1}, h_{t-1}, a_{t-1}) = \frac{\mu_{t-1}(1 - f(h_{t-1}, a_{t-1}))}{\mu_{t-1}(1 - f(h_{t-1}, a_{t-1})) + (1 - \mu_{t-1})(1 - p)}.$$

[19] In fact, after a straightforward rearrangement, one can see that $\mu(1, \mu_{t-1}, h_{t-1}, a_{t-1}) - \mu(0, \mu_{t-1}, h_{t-1}, a_{t-1})$ is proportional to $\mu_{t-1}(1 - \mu_{t-1})(f(h_{t-1}, a_{t-1}) - p)$.

where

$$\pi(a_t, h_t, \mu_t) = [1 - \mu_t]p + \mu_t f(h_t, a_t)$$

The optimal action in any period t depends on period-t history h_t, posterior μ_t and observed state x_t. The beliefs about the future occurrences of the cravings depend on these variables and the chosen action a_t. The observed state x_t influences also the current payoff. Let $a(h_t, x_t, \mu_t)$ denote the optimal period-t consumption for any t.

Let $Z(a_t, x_t, h_t, \mu_t)$ denote the period-t expected discounted payoff when in period t the state is x_t, the history is h_t, the posterior is μ_t and the action is a_t. That is,

$$\begin{gathered} Z(a_t, x_t, h_t, \mu_t) \equiv u(a_t, x_t) + \\ +\delta\pi(a_t, h_t, \mu_t) \times \\ \times Z(a(g(h_t, a_t), 1, \mu(1, \mu_t, h_t, a_t)), 1, g(h_t, a_t), \mu(1, \mu_t, h_t, a_t)) + \\ +\delta[1 - \pi(a_t, h_t, \mu_t)] \times \\ \times Z(a(g(h_t, a_t), 0, \mu(0, \mu_t, h_t, a_t)), 0, g(h_t, a_t), \mu(0, \mu_t, h_t, a_t)) \end{gathered}$$

Let finally

$$V(x, h, \mu) \equiv Z(a(h, x, \mu), x, h, \mu)$$

be the equilibrium expected discounted payoff given posteriors μ, history h and craving-state x, where we suppress (unless stated otherwise to avoid confusion) the dependence of it on the horizon T.

For similar reason to that in the two-period model, if T is finite, then $a(h_T, x_T, \mu_T) = x_T$. Moreover, we have $V(x_T, h_T, \mu_T) = u(x_T, x_T)$.

To discuss optimal consumption for any period $t < T$, $a(h_t, x_t, \mu_t)$, we need to understand first the properties of the period-t value function $V(x_t, h_t, \mu_t)$.

Consider the following properties:

(A) $V(x_{t+1}, h_{t+1}, \mu_{t+1})$ is nonincreasing,

(B) $V(0, g(h_t, a_t), \mu(0, \mu_t, h_t, a_t))$ is nonincreasing in h_t and a_t,

(C) $V(0, g(h_t, a_t), \mu(0, \mu_t, h_t, a_t)) - V(1, g(h_t, a_t), \mu(1, \mu_t, h_t, a_t)) \geq D$.

Property (A) states that at optimum the decision maker is (weakly) better off from the point of view of period $t+1$ if there is no craving, history is low and posterior is low. Property (B) states that for given period-$t+1$ prior μ_t, the period-$t+1$ value function conditional on not having a craving is lower if past consumptions are higher. Finally, property (C) states that for given period-$t+1$ prior μ_t and period-t history and action h_t and a_t, the gain in period-$t+1$ value function from not having a craving is at least equal to D.

To appreciate the implications of the above properties note first that they are all trivially true when $t = T-1$ (ie. T is finite) due to $a(h_T, x_T, \mu_T) = x_T$ and thereby $V(x_T, h_T, \mu_T) = u(x_T, x_T)$.

Second, in the Appendix B we can prove the following Lemma:

Lemma 2 *If properties (A) and (B) are true for $t \geq n$, $2 \leq n+1$, then $a(h_n, 0, \mu(0, \mu_{n-1}, h_{n-1}, a_{n-1})) = 0$ and property (C) is true for $t = n-1$.*

That is, if the value function in all periods from $n+1$ (inclusive) and onwards satisfies properties (A) and (B), then property (C) is true for period n and the optimal action in period n when there is no craving is to abstain from consumption. In fact, the latter result, in conjunction with properties (A) and (B) being true for period T, has been used in the simple two-period model for $n = 1$.

Third, one can also show that property (A) is true for $t = T-2$ (ie. T is finite). This follows directly from the above Lemma, that $u(0,0) > u(1,1) > u(0,1)$, the definition of optimum, that $\pi(a_{T-1}, h_{T-1}, \mu_{T-1})$ is nondecreasing and, finally, that $V(x_{T-1}, h_{T-1}, \mu_{T-1}) = u(a(h_{T-1}, x_{T-1}, \mu_{T-1}), x_{T-1}) + \delta u(0,0) - \delta\pi(a(h_{T-1}, x_{T-1}, \mu_{T-1}), h_{T-1}, \mu_{T-1})D$. The reader is referred to Appendix A for the details.

Fourth, property (B) holds for $t = T-2$ if $\pi(0, g(h,a), \mu(0, \mu, h, a))$ is nondecreasing in h and a. To see this, note, due to Lemma 2, that $V(0, g(h_{T-2}, a_{T-2}), \mu(0, \mu_{T-2}, h_{T-2}, a_{T-2})) = u(0,0) + \delta u(0,0)$

$- \delta\pi(0, g(h_{T-2}, a_{T-2}), \mu(0, \mu_{T-2}, h_{T-2}, a_{T-2}))D$. Note however that the monotonicity of $\pi(0, g(h,a), \mu(0, \mu, h, a))$ with respect to h and a depends on the properties of g and f.

Finally, the applicability of properties (A) and (B) for any period $t \leq T-3$ when $T > 3$ ($t < T-3$ when $T \equiv \infty$) is not guaranteed because for such t the value function depends on posteriors and higher past consumptions have two opposite effects on posteriors depending on the craving state: positive if there is an urge and negative if there is no craving.

Bearing the above in mind, let us restrict hereafter attention to environments where

Assumption 6 *f and g are such that (for any $t \geq 1$ and $\prod_{i=2}^{1} y \equiv 1$) if*

$h'' \geq h'$ and $a'' \geq a'$ (with at least inequality strict) then

$$\begin{aligned} & \frac{(1-f(h',a'))}{(1-f(h'',a''))}\prod_{i=2}^{t}(1-f(L^{t-i}g(h',a'),0)) \\ \leq \quad & \frac{f(L^{t-1}g(h'',a''),0)}{f(L^{t-1}g(h',a'),0)}\prod_{i=2}^{t}(1-f(L^{t-i}g(h'',a''),0)). \end{aligned}$$

In the above assumption, L is an operator on functions $g(h,a)$ and $\mu(0,\mu_j,h,a)$ defined by

$$\begin{aligned} Lg(h,a) &= g(g(h,a),0) \text{ and} \\ L\mu(0,\mu_j,h,a) &= \mu(0,\mu(0,\mu_j,h,a),g(h,a),0). \end{aligned}$$

Moreover, L^t denotes the tth application of the L operator after using the convention that $L^0g(h,a) = g(h,a)$ and $L^0\mu(0,\mu_j,h,a) = \mu(0,\mu_j,h,a)$. That is, in general, $L^tg(h,a) = g(L^{t-1}g(h,a),0)$ and $L^t\mu(0,\mu_j,h,a)$

$= \mu(0,L^{t-1}\mu(0,\mu_j,h,a),L^{t-1}g(h,a),0)$. To understand the notation, consider the example with $t=2$. In this case we have: $L^2g(h,a) = Lg(g(h,a),0)$ $= g(g(g(h,a),0),0)$ and $L^2\mu(0,\mu_j,h,a) = L\mu(0,\mu(0,\mu_j,h,a),g(h,a),0) = \mu(0,\mu(0,\mu(0,\mu_j,h,a),g(h,a),0),g(g(h,a),0),0) = \mu\,(0,L\mu(0,\mu_j,h,a),Lg(h,a),0)$. Similarly for any $t>2$. Essentially the tth application of the operator gives the posterior belief and the history in period $j+t+1$ given that in period j the prior was μ_j, history was h and action was a, and that in any period after j up to period $t+j+1$ the craving-state has been $x=0$ and the action has been to abstain from addictive consumption.

We show in Appendix C, that the above assumption implies that for any $1 \leq t \leq T-j-1$

(i) $\pi(0,L^{t-1}g(h,a),L^{t-1}\mu(0,\mu_j,h,a))$ is non-decreasing in h and a and

(ii) $\mu(1,L^{t-1}\mu(0,\mu_j,h,a),L^{t-1}g(h,a),0)$ is non-decreasing in h and a,

which are what really needed for the results below.

To see how strong the above assumption is and how it relates to the existing literature, we also show in Appendix C that the set of g and f functions that are consistent with assumption 6 is not empty. In fact, we also show that functions g and f which are consistent with (BM, GK, L and BR) must be sufficiently responsive to h relative to a.

We turn to discussing optimal consumption when T is finite.

4.1.1 The Case of Finite Horizon

The following Lemma extends in the $T-$period model certain properties of the 2-period model.

Lemma 3 *Assume 6. Then properties (A)-(C) and* $a(h_{t+1}, 0, \mu(0, \mu_t, h_t, a_t)) = 0$ *also hold for any* $0 \leq t \leq T-1$.

Proof. Under assumption 6, in Appendix B we show that if properties (A) and (B) are true for any $t \geq n$, then they are also true for $t = n-1$, for any given $2 \leq n+1 \leq T$. Induction as well as the fact that, as we have mentioned earlier, properties (A)-(C) hold in the last period, and Lemma 2 imply directly the desired result. ■

Thus, under assumption 6, we have, as intuition would suggest given the dependence of occurrences of future cravings on history, that higher consumption now or in the past is (weakly) detrimental for future welfare for any given future state and posteriors. Moreover, given that cravings make consumption and hence future cravings more likely given beliefs, we have that the benefit in terms of overall future welfare from not having an "urge" next period (i.e. period $t+1$) is (weakly) higher than the one-period welfare benefit of not having a craving next period D. Furthermore, for reasons similar to those in the two-period model, choosing $a_t = 1$ when $x_t = 0$ is not optimal.[20] Hence, in what follows, we restrict attention to analyzing the situation when $x_t = 1$.

To do so let us lighten notation first. Even though the period-t posterior depends on $\mu_{t-1}, h_{t-1}, a_{t-1}$, as well as on the realized value of x_t, for expositional simplicity, and unless there is a risk of confusion, we suppress the dependence of period-t posteriors on $\mu_{t-1}, h_{t-1}, a_{t-1}$ and let $\mu_t(x_t) = \mu(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})$ (the dependence in question is captured simply by the time-subscript of $\mu_t(x)$). Similarly, we set, whenever there is no risk of confusion, $a_t(h_t, x_t) \equiv a(h_t, x_t, \mu_t(x_t))$ and $V_t(x_t, h_t) \equiv V(x_t, h_t, \mu_t(x_t))$. Finally, let us set $\pi_t(h_t, a_t) \equiv \pi(a_t, h_t, \mu_t(1))$.

We start by introducing first some definitions. Let $\widehat{D}_{t+1}(h_t)$ be given by

$$\widehat{D}_{t+1}(h_t) = V_{t+1}(0, g(h_t, 1)) - V_{t+1}(1, g(h_t, 1))$$

[20] Put differently, since smoking is costly (monetarily and otherwise) from both short and long run point of view, there is no reason for the consumer to indulge in smoking in the absence of any urge.

Note that the above Lemma implies that $\widehat{D}_{t+1}(h_t) \geq D > 0$. $\widehat{D}_{t+1}(h_t)$ has a similar interpretation to D. In particular, it is the expected gain - evaluated at $t+1$ - from not having an urge in the future given that the consumer smokes in the current period and optimally responds to the presence or absence of an urge in the future. Contrasting $\widehat{D}_{t+1}(h_t)$ with D, note that the kind of current consumption is irrelevant for the definition of the latter. The reason is that in a two-period model, the second-period consumption is determined entirely by the temporal payoff and is independent of the first- period consumption, and we have $\widehat{D}_2(h_1) = D$. With more than two periods, the consumption in any period $t < T - 1$ affects future consumption incentives.

The latter also lies behind the existence of the extra term $\Psi_{t+1}(h_t)$ defined below. Let us first define $\Gamma_{t+1}(h, x)$ as

$$\Gamma_{t+1}(h_t, x) = V_{t+1}(x, g(h_t, 0)) - V_{t+1}(x, g(h_t, 1))$$

Note that the above Lemma implies that $\Gamma_{t+1}(h_t, x) \geq 0$. $\Gamma_{t+1}(h_t, x)$ represents the expected future gain - evaluated at $t + 1$ - from not smoking in the current period if the future state is x. $\Psi_{t+1}(h_t)$ is given by

$$\Psi_{t+1}(h_t) = \pi_t(h_t, 0)\Gamma_{t+1}(h_t, 1) + (1 - \pi_t(h_t, 0))\Gamma_{t+1}(h_{t,}0)$$

$\Psi_{t+1}(h_t) \geq 0$ represents the expected value of $\Gamma_{t+1}(h_t, x)$ conditional on not smoking in period t. Note that in the two-period model, $\Gamma_2(h_1, 1) = \Gamma_2(h_1, 0) = 0$ (and hence $\Psi_2(h_1) = 0$) because, as we have already mentioned, first-period consumption does not affect the optimal consumption in the second period (that is, $a_2(g(h_1, a_1), x_2)$ is independent of a_1) and because there is no future in the second period to be affected by consumption up to period 2 (that is, $V(x_2, h_2, \mu_2)$ is independent of h_2 and μ_2).

Note now that payoff from action a_t in period t (when $x_t = 1$) is then given by

$$\begin{aligned} & u(a_t, 1) + \delta V_{t+1}(0, g(h_t, a_t)) \\ & -\delta \pi_t(h_t, a_t) \left[V_{t+1}(0, g(h_t, a_t)) - V_{t+1}(1, g(h_t, a_t)) \right] \end{aligned} \tag{3}$$

We have:

Proposition 4 *No smoking is the optimal action in period t (given that $x_t = 1$) if and only if the following inequality holds:*

$$B \leq \delta\{\mu_t(1)\left[f(h_t,1)-f(h_t,0)\right]\widehat{D}_{t+1}(h_t)+\Psi_{t+1}(h_t)\} \tag{4}$$

Proof. Using expression (3), we see that when $x_1 = 1$, action $a_1 = 0$ is preferred to $a_1 = 1$ when

$$\begin{aligned}
& u(0,1)+\delta V_{t+1}(0,g(h_t,0)) \\
& -\delta\pi_t(h_t,0)\left[V_{t+1}(0,g(h_t,0))-V_{t+1}(1,g(h_t,0))\right] \\
\geq & \\
& u(1,1)+\delta V_{t+1}(0,g(h_t,1)) \\
& -\delta\pi_t(h_t,1)\left[V_{t+1}(0,g(h_t,1))-V_{t+1}(1,g(h_t,1))\right]
\end{aligned}$$

which after some straightforward algebraic manipulation, that includes adding up at both sides the term $\pi_t(h_t,0)\left[V_{t+1}(0,g(h_t,1))-V_{t+1}(1,g(h_t,1))\right]$, and using the definitions for B, $\widehat{D}_{t+1}(h_t)$ and $\Psi_{t+1}(h_t)$, gives the inequality. ■

Comparing inequality (4) with inequality (2), we can see two differences. First, the term D is replaced by $\widehat{D}_{t+1}(h_t)$. Second there is an extra term $\Psi_{t+1}(h_t)$. The reason for this is reflected in our earlier discussion of the interpretation of these terms.

At this point, we can ask how the presence of a longer horizon affects the likelihood of smoking in any period t. To answer this, recall that for any $t < T-1$, $\Psi_{t+1}(h_t) \geq 0$ and that $\widehat{D}_{t+1}(h_t) \geq D$ (while $\Psi_T(h_{T-1}) = 0$ and $\widehat{D}_T(h_{T-1}) = D$). Therefore, a longer (finite) horizon makes smoking less likely as it (weakly) increases the relative cost of smoking (see the right hand side of inequality (4) and compare it with inequality (2)).

In the next subsection, we show that the above insight carries forward to the infinite horizon case. However, proving this intuitive result turns out to be technically more challenging than one might have conjectured.

Before we do so, we demonstrate here that our multi-period model can feature failed attempts to quit and occasional use as a process of experimentation. That is, in what follow, the objective is to show that our model is capable of generating behavior that shows like recidivism. To start with, note that if all the decision maker is doing is choosing $a_t = x_t$ every period (for example, the optimal policy if the decision maker thought $\mu_t = 0$, i.e. the process is non-addictive for sure), to an outside observer who does not observe x_t it is trivially true that it might look like the decision maker stops consuming for some period and then relapses. We therefore look at the more interesting case where the outside observer can observe the value if x_t (for

example x_t represents situations, may be cues, such that it is "known" that the decision maker becomes more tempted to consume in such situations).

Suppose first that we are in period $T-3$, where T is finite, and $x_{T-3}=1$. Suppose also that in the past the decision maker has consumed the substance and hence $h_{T-3}>0$. Suppose also that h_{T-3} such that we have

$$B=\delta\{\mu_{T-3}(1)\left[f(h_{T-3},1)-f(h_{T-3},0)\right]\widehat{D}_{T-2}(h_{T-3})+\Psi_{T-2}(h_{T-3})\}$$

Hence, according to the above proposition, the decision maker chooses $a_{T-3}=0$. In this case, an outside observer could easily claim that the decision maker is attempting to quit. Suppose now that we look at a path (of realizations of x_t) such that $x_{T-2}=0$ and $x_{T-1}=1$. We know from the above lemma that $a_{t-2}=0$. Since T is the terminal period, according to the above proposition, the decision maker chooses $a_{T-1}=1$ if (and only if)

$$B>\delta\mu_{T-1}(1)\left[f(h_{T-1},1)-f(h_{T-1},0)\right]D$$

Focus on an environment with (weak) depreciation of the stock of addiction: $g(h,0)\leq h$. Assume also that $f(h,1)-f(h,0)$ is (weakly) increasing in h. The former is the maintained assumption in the rational addiction literature. Note also that for the parametric example where

$$\begin{aligned} f &= p+g(h,a) \\ g(h,a) &= \rho(h)+\rho_2 a, \\ \rho_2 &> 0,\ \rho(0)=0,\ \rho(h)\leq h, \rho'(h)\geq 0 \end{aligned}$$

for any history h, the expression $f(h,1)-f(h,0)$ is independent of h and hence trivially weakly increasing in h. As we also discuss in Appendix C this example is consistent with the models in the received literature on rational addiction. Note thus that if $a_{T-3}=a_{T-2}=0$, then $h_{T-1}\leq h_{T-3}$ and $f(h_{T-1},1)-f(h_{T-1},0)\leq f(h_{T-3},1)-f(h_{T-3},0)$.

Since $\Psi_{t+1}(h_t)\geq 0$ and $\widehat{D}_{t+1}(h_t)\geq D$, the (additional) condition for $a_{t-3}=0$ and $a_{t-1}=1$ (i.e. a failed attempt to quit) is therefore that $\mu_{T-1}(1)$ be strictly less than $\mu_{T-3}(1)$. Now, let $f_{T-2}=f(h_{T-3},0)$ and $f_{T-1}=$

$f(g(h_{T-3},0),0)$ and note that

$$\begin{aligned}
& \mu_{T-1}(1) - \mu_{T-3}(1) \\
= \; & \frac{\mu_{T-3}(1)(1-f_{T-2})f_{T-1}}{\mu_{T-3}(1)(1-f_{T-2})f_{T-1} + (1-\mu_{T-3}(1))(1-p)p} - \mu_{T-3}(1) \\
= \; & \mu_{T-3}(1)\left[\frac{(1-f_{T-2})f_{T-1}}{\mu_{T-3}(1)(1-f_{T-2})f_{T-1} + (1-\mu_{T-3}(1))(1-p)p} - 1\right] \\
= \; & \mu_{T-3}(1)(1-\mu_{T-3}(1))\left[\frac{(1-f_{T-2})f_{T-1} - (1-p)p}{\mu_{T-3}(1)(1-f_{T-2})f_{T-1} + (1-\mu_{T-3}(1))(1-p)p}\right]
\end{aligned}$$

the sign of which depends on the sign of $(1-f_{T-2})f_{T-1} - (1-p)p$. Note further than since $f_{T-1} \leq f_{T-2}$,

$$\begin{aligned}
& (1-f_{T-2})f_{T-1} - (1-p)p \\
\leq \; & (1-f_{T-2})f_{T-2} - (1-p)p
\end{aligned}$$

Suppose now that $p = \frac{1}{2}$, which seems to be a natural case for 'random' cravings. Since $f_{T-2} > p$, we have $(1-f_{T-2})f_{T-2} - (1-p)p < 0$ and hence $\mu_{T-1}(1) < \mu_{T-3}(1)$.

4.2 Infinite Horizon

Turning to the case of $T \to \infty$, the first issue that arises is whether the value function $V(x,h,\mu)$ is well-defined. The second issue is whether it satisfies properties (A) and (B) (in which case, by Lemma 2, it will also satisfy property (C) and that the best response to no craving is no zero consumption). If these are still true, our results, and in particular Lemma 3 and the above proposition, will be robust to allowing for infinite horizon.

By using standard dynamic programming techniques one can show that $V(x,h,\mu)$ is indeed well-defined, continuous and bounded. Furthermore, with a somewhat more involved analysis, one can also that the value function satisfies also properties (A) and (B) under assumption 6. The details are in Appendix D.

4.3 Cumulative Welfare Effects

Here we consider environments where the period t payoff given craving state x_t, action a_t and history of past consumption h_t is given by $v(a_t, x_t, h_t)$

with $v(1,1,h_t) > v(0,1,h_t)$, $v(1,0,h_t) < v(0,0,h_t)$, $v(1,1,h_t) > v(0,0,h_t)$ and $v(a_t, x_t, h_t)$ being nondecreasing in h_t. These assumptions maintain the properties of compulsive consumption, while introducing cumulative welfare costs of past consumptions. Let also $D(h) = v(0,0,h) - v(1,1,h)$ and $B(h) = v(1,1,h) - v(0,1,h)$, and note that these replace D and B, respectively (with $D(h_1)$ and $B(h_1)$ being the relevant variables in the two-period model).

Focus on the case of infinite horizon to facilitate comparisons with most of the received literature on rational addiction. By repeating the steps in Appendix B one can easily show that Lemma 2 is still valid here. As we also show in Appendix D the value function is still well-defined. Moreover, most of our results,[21] and in particular the above proposition and Lemma 3 are robust to the introduction of cumulative welfare results.

At this stage we can discuss how our work relates to the habit formation literature. Intrinsic formation models assume, in general, an intertemporal utility where past consumption affects valuation of current and future consumption. Our discussion above emphasizes that our results do not rely on the presence or absence of such intertemporal effects.

Many of the intrinsic habit-formation models seem to be able to explain phenomena that standard intertemporally separable preferences cannot. Some examples are the following: Constantinides (1990) helps understand data indicating that individuals are far more averse to risk than might be expected; Boldrin, Christiano, and Fisher (2001) who combine habit formation and intersectoral inflexibilities in a model of real business cycles to suggest an explanation for why consumption growth is connected strongly to income, but only weakly to interest rates; Uribe (2002) who gives an explanation for the contractions in consumption that are observed before the collapse of exchange rate stabilization programs. The above literature however postulates the habit formation preferences in an ad hoc manner. In fact, until recently there have been no theoretical underpinnings of habit formation preferences.[22] Rozen (2009), who axiomatizes the so-called linear habit formation model used in some of the papers above, Rustichini and Siconolfi (2005), who axiomatize dynamically consistent habit formation over consumption

[21] The only results that would require, in order to still hold, further assumptions on the monotonicity with respect to history of $B(h)/D(h)$ have to do with the relative behavior of groups that differ in terms of their history.

[22] However, there is a large literature on the axiomatization of static reference dependence.

streams, but do not offer a particular structure for the utility or form of habit aggregation, and Gul and Pesendorfer (2007), who also axiomatize a dynamically consistent non-linear habit formation model by considering preferences on menus of streams of consumption rather than on streams themselves, are recent notable exceptions. In relation to this strand of literature, our model generates non-linear habit forming preferences, but by starting from a standard intertemporally separable discounted utility. The reason is that our fully rational decision maker lacks information about the determination of the state of the world (the "urge" to consume in the future), with the perceived mechanism depending on an endogenous (due to the addictive nature of consumption) Markov chain. To see this, revert to our model with no welfare cumulative effects where utility is given by $u(a, x)$. Note then that at time t the utility is given by $u(a_t, x_t) + \sum_{i=1}^{T-t} \delta^i E_t[u(a_{t+i}, X_{t+i}) \mid h_{t-1}, a_{t-1}, a^{t+i-1}, x_t]$, where $a^{t+i-1} = (a_t, a_{t+1}, ..., a_{t+i-1})$ and $E_t[u(a_{t+i}, X_{t+i}) \mid h_t, a^{t+i-1}, x_t]$ denotes the expectations operator with respect to X_{t+i} given the $t-$period prior μ_{t-1}, past history h_{t-1} and consumption a_{t-1} (and hence h_t), observed state x_t and the consumption stream a^{t+i-1} (which will determine the history stream $(h_{t+1}, h_{t+2}, ..., h_{t+i})$ and hence the perceived probability distribution over X_{t+i}, $p + \mu(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})[f(h_{t+i-1}, a_{t+i-1}) - p]$, for all $i = 1, ..., T - t$). Thus, our *expected* utility falls under the rubric of *nonlinear* habit formation models.

5 Conclusions

We have presented a theory of rational addiction that complements the received literature in an important way. In particular, our theory of rational addiction is based on four central premises. First, addictive consumption is compulsive. Second, cue-triggered cravings are endogenous in that their occurrence depends on past behavior. Third, addicts understand their susceptibility to cravings and try to rationally manage the process through their consumption even under a temptation. Fourth, and what differentiates substantially our theory from existing work, is that consumers are not fully aware of the easiness to quit because they lack some information about the addictive properties of the substance.

In our context, there is scope for campaigns that inform consumers about the addictive properties of the various substances. Moreover, our theory provides some mico-foundations for habit-forming behavior by starting from a

standard model of fully rational decision maker with intertemporally separable preferences, but with uncertainty over the likelihood of future temptations which is endogenous in that it depends on past behavior and the history of past and current temptations. Our analysis has also a number of other interesting implications. These include that consumption patterns depend on the inherent addictive properties of the substance as well as on the family and social environment of individuals when they make their first consumption decision. Moreover, failed attempts to quit and occasional use can emerge as a process of information acquisition. Finally, our model predicts that drugs with stronger withdrawal syndromes are associated with lower consumption.

6 References

ABEL, A. (1990), "Asset Prices Under Habit Formation and Catching Up With the Joneses", American Economic Review, 80, 38–42.

BANERJEE, A.V. (1992), "A Simple Model of Herd Behavior", Quarterly Journal of Economics, 107, pp. 797-817.

BECKER, G. S. and MURPHY, K. M. (1988), "A Theory of Rational Addiction", Journal of Political Economy, 96, 675–700.

BERNHEIM, B. D. and RANGEL, A. (2004), "Addiction and Cue-Conditioned Cognitive Processes", American Economic Review, 94, 1558–1590.

BIKHCHANDANI, S. HIRSHLEIFER D. and WELCH, I. (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", Journal of Political Economy, 100, 992-1026.

BOLDRIN, M., Christiano, L. and Fisher J. (2001), "Habit Persistence, Asset Returns, and the Business Cycle," American Economic Review, 91, 149–166.

CONSTANTINIDES, G. (1990), "Habit Formation: A Resolution of the Equity Premium Puzzle", Journal of Political Economy, 98, 519–543.

GRUBER, J. and KOSZEGI, B. (2001), "Is Addiction Rational? Theory and Evidence", Quarterly Journal of Economics, 116, 1261–1303.

GUL, F. and PESENDORFER, W. (2007), "Harmful Addiction", Review of Economic Studies, 74, 147–172.

LAIBSON, D. (2001), "A Cue Theory of Consumption", Quarterly Journal of Economics, 66, 81–120.

ORPHANIDES, A. and ZERVOS, D. (1995), "Rational Addiction with Learning and Regret", Journal of Political Economy, 103, 739-758.

ROZEN, K. (2009), "Foundations of Intrinsic Habit Formation", mimeo.

RUSTICHINI, A., and SICONOLFI, P. (2005), "Dynamic Theory of Preferences: Taste for Variety and Habit Formation", mimeo.

STOCKEY, N.L. and LUCAS, R.E. Jr. with PRESCOTT, E. (1989), Recursive Methods in Economic Dynamics, Harvard University Press.

URIBE, M. (2002), "The Price-Consumption Puzzle of Currency Pegs," Journal of Monetary Economics, 49, 533–569.

7 Appendix A

Here we prove property (A) for $t+1 = 1 = T-1$

Proof. Let $(x''_{T-1}, h''_{T-1}, \mu''_{T-1}) \geq (x'_{T-1}, h'_{T-1}, \mu'_{T-1})$ and $a'_{T-1} = a(h'_{T-1}, x'_{T-1}, \mu'_{T-1})$ and $a''_{T-1} = a(h''_{T-1}, x''_{T-1}, \mu''_{T-1})$. We have by the definition of optimum that when $x''_{T-1} = x'_{T-1} \equiv x_{T-1}$ we have

$$\begin{aligned} & V(x_{T-1}, h'_{T-1}, \mu'_{T-1}) \geq \\ = \; & u(a''_{T-1}, x_{T-1}) + \delta u(0,0) - \delta\pi(a''_{T-1}, h'_{T-1}, \mu'_{T-1})D \geq \\ & u(a''_{T-1}, x_{T-1}) + \delta u(0,0) - \delta\pi(a''_{T-1}, h''_{T-1}, \mu''_{T-1})D \end{aligned}$$

where the last inequality follows from $\pi(a, h'_{T-1}, \mu'_{T-1}) \leq \pi(a, h''_{T-1}, \mu''_{T-1})$. Note now that by definition $V(x_{T-1}, h''_{T-1}, \mu''_{T-1}) = u(a''_{T-1}, x_{T-1}) + \delta u(0,0) - \delta\pi(a''_{T-1}, h''_{T-1}, \mu''_{T-1})D$. Thus, $V(x_{T-1}, h''_{T-1}, \mu''_{T-1}) \leq V(x_{T-1}, h'_{T-1}, \mu'_{T-1})$. Finally, we have when $h''_{T-1} = h'_{T-1} \equiv h_{T-1}$ and $\mu''_{T-1} = \mu'_{T-1} \equiv \mu_{T-1}$ by definition of optimum and Lemma 2 that

$$\begin{aligned} & V(0, h_{T-1}, \mu_{T-1}) = \\ = \; & u(0,0) + \delta u(0,0) - \delta\pi(0, h_{T-1}, \mu_{T-1})D \geq \\ & u(a''_{T-1}, 1) + \delta u(0,0) - \delta\pi(1, h_{T-1}, \mu_{T-1})D \end{aligned}$$

where the last inequality follows from $u(0,0) > u(a,1)$ and $\pi(0, h_{T-1}, \mu_{T-1}) \leq \pi(1, h_{T-1}, \mu_{T-1})$. Thus, $V(1, h_{T-1}, \mu_{T-1}) \leq V(0, h_{T-1}, \mu_{T-1})$. ■

8 Appendix B

Here we prove Lemma 2 and part of Lemma 3

Lemma 2

Proof. We start with proving that $a(h_n, 0, \mu(0, \mu_{n-1}, h_{n-1}, a_{n-1})) = 0$

For any given μ_{n-1}, h_{n-1} and a_{n-1} and corresponding period n history h_n and posteriors $\mu_n(x_n) \equiv \mu(x_n, \mu_{n-1}, h_{n-1}, a_{n-1})$, let $x_n = 0$. Then, if $a_n = 0$, the expected discounted payoff in period n is

$$\begin{aligned} & u(0,0) \\ & +\delta\pi(0, h_n, \mu_n(0))V(1, h', \mu(1, \mu_n(0), h_n, 0)) \\ & +\delta(1-\pi(0, h_n, \mu_n(0)))V(0, h', \mu(0, \mu_n(0), h_n, 0)) \\ \equiv\ & u(0,0) + E_{\pi(0,h_n,\mu_n(0))}V(X, h', \mu(X, \mu_n(0), h_n, 0)) \end{aligned}$$

where $h' = g(h_n, 0)$. On the other hand, if $a_n = 1$, the payoff is

$$\begin{aligned} & u(1,0) \\ & +\delta\pi(1, h_n, \mu_n(0))V(1, h'', \mu(1, \mu_n(0), h_n, 1)) \\ & +\delta(1-\pi(1, h_n, \mu_n(0)))V(0, h'', \mu(0, \mu_n(0), h_n, 1)) \\ \equiv\ & u(1,0) + E_{\pi(1,h_n,\mu_n(0))}V(X, h'', \mu(X, \mu_n(0), h_n, 1)) \end{aligned}$$

where $h'' = g(h_n, 1)$. Note that $\pi(0, h_n, \mu_n(0)) \geq \pi(1, h_n, \mu_n(0))$. Note by the property (A) for $t = n$ and $\mu(1, \mu_n(0), h_n, a) \geq \mu(0, \mu_n(0), h_n, a)$ that $V(1, g(h_n, a), \mu(1, \mu_n(0), h_n, a)) \leq V(0, g(h_n, a), \mu(0, \mu_n(0), h_n, a))$, by the property (A) for $t = n$ and $\mu(1, \mu_n(0), h_n, 1) \geq \mu(1, \mu_n(0), h_n, 0)$ that $V(1, h', \mu(1, \mu_n(0), h_n, 0)) \geq V(1, h'', \mu(1, \mu_n(0), h_n, 1))$ and by the property (B) for $t = n$ that $V(0, h', \mu(0, \mu_n(0), h_n, 0)) \geq V(0, h'', \mu(0, \mu_n(0), h_n, 1))$. Therefore, $E_{\pi(1,h_n,\mu_n(0))}V(X, h'', \mu(X, \mu_n(0), h_n, 1))$

$\leq E_{\pi(0,h_n,\mu_n(0))}V(X, h', \mu(X, \mu_n(0), h_n, 0))$. This alongside $u(0,0) > u(1,0)$ proves that $a_n(h_n, 0, \mu_n(0)) = 0$.

We now prove property (C) for $t = n-1$

Given the previous result, we have, for any given μ_{n-1}, h_{n-1} and a_{n-1} and corresponding period n history h_n and posteriors $\mu_n(x_n)$,

$$\begin{aligned} & V(0, h_n, \mu_n(0)) \\ =\ & u(0,0) \\ & +\delta\pi(0, h_n, \mu_n(0))V(1, h', \mu(1, \mu_n(0), h_n, 0))) \\ & +\delta(1-\pi(0, h_n, \mu_n(0)))V(0, h', \mu(0, \mu_n(0), h_n, 0)) \\ \equiv\ & E_{\pi(0,h_n,\mu_n(0))}V(X, h', \mu(X, \mu_n(0), h_n, 0)) \end{aligned}$$

where $h' = g(h_n, 0)$ and similarly

$$\begin{aligned} & V(1, h_n, \mu_n(1)) \\ =\ & \max_a \{u(a,1) + E_{\pi(a,h_n,\mu_n(1))}V(X, g(h_n, a), \mu(X, \mu_n(1), h_n, a))\} \end{aligned}$$

Note, due to $u(0,0) - u(0,1) > u(0,0) - u(1,1) = D$, that property (C) for $t = n-1$ is proved if $E_{\pi(1,h_n,\mu_n(1))}V(X, h'', \mu(X, \mu_n(1), h_n, 1))$

$\leq E_{\pi(0,h_n,\mu_n(1))}V(X, h', \mu(X, \mu_n(1), h_n, 0))$

$\leq E_{\pi(0,h_n,\mu_n(0))}V(X, h', \mu(X, \mu_n(0), h_n, 0))$, where $h'' = g(h_n, 1)$. These follow directly after noting that (a) $\pi(0, h_n, \mu_n(0)) \leq \pi(0, h_n, \mu_n(1)) \leq \pi(1, h_n, \mu_n(1))$, (b) by the assumption that property (A) holds for $t = n$ and $\mu(1, \mu_n(x_n), h_n, a) \geq \mu(0, \mu_n(x_n), h_n, a)$ we have $V(1, g(h_n, a), \mu(1, \mu_n(x_n), h_n, a)) \leq V(0, g(h_n, a), \mu(0, \mu_n(x_n), h_n, a))$, (c) by the assumption that property (A) holds for $t = n$ and $\mu(1, \mu_n(1), h_n, 1)$ $\geq \mu(1, \mu_n(1), h_n, 0) \geq \mu(1, \mu_n(0), h_n, 0)$ we have $V(1, h', \mu(1, \mu_n(0), h_n, 0)) \geq V(1, h', \mu(1, \mu_n(1), h_n, 0)) \geq V(1, h'', \mu(1, \mu_n(1), h_n, 1))$, (d) by the assumption that property (B) holds for $t = n$ we have $V(0, h', \mu(0, \mu_n(1), h_n, 0)) \geq V(0, h'', \mu(0, \mu_n(1), h_n, 1))$, and (e) by the assumption that property (A) holds for $t = n$ and $\mu(0, \mu_n(0), h_n, 0) \leq \mu(0, \mu_n(1), h_n, 0)$ we have $V(0, h', \mu(0, \mu_n(0), h_n, 0))$ $\geq V(0, h', \mu(0, \mu_n(1), h_n, 0))$. ■

Result needed for Lemma 3

Proof. Assume that properties (A) and (B) hold for $t = n$.

We first prove property (A) for $t = n-1$

Let $(x''_n, h''_n, \mu''_n) \geq (x'_n, h'_n, \mu'_n)$ and $a'_n = a(h'_n, x'_n, \mu'_n)$ and $a''_n = a(h''_n, x''_n, \mu''_n)$. We have by the definition of optimum that when $x''_n = x'_n \equiv x_n$

$$\begin{aligned} & V(x_n, h'_n, \mu'_n) \geq \\ = \; & u(a''_n, x_n) \\ & +\delta\pi(a''_n, h'_n, \mu'_n)V(1, g(h'_n, a''_n), \mu(1, \mu'_n, h'_n, a''_n)) \\ & +\delta(1 - \pi(a''_n, h'_n, \mu'_n))V(0, g(h'_n, a''_n), \mu(0, \mu'_n, h'_n, a''_n)) \\ \equiv \; & u(a''_n, x_n) + E_{\pi(a''_n,h'_n,\mu'_n)}V(X, g(h'_n, a''_n), \mu(X, \mu'_n, h'_n, a''_n)) \end{aligned}$$

Note now that $V(x_n, h''_n, \mu''_n) = u(a''_n, x_n) + E_{\pi(a''_n,h''_n,\mu''_n)}V(X, g(h''_n, a''_n), \mu(X, \mu''_n, h''_n, a''_n))$. Thus, $V(x_n, h''_n, \mu''_n) \leq V(x_n, h'_n, \mu'_n)$ follows directly if

$E_{\pi(a,h''_n,\mu''_n)}V(X, g(h''_n, a), \mu(X, \mu''_n, h''_n, a)) \leq E_{\pi(a,h'_n,\mu'_n)}V(X, g(h'_n, a), \mu(X, \mu'_n, h'_n, a))$. This follows after observing (a) $\pi(a, h'_n, \mu'_n) \leq \pi(a, h''_n, \mu''_n)$, (b) by the assumption that property (A) holds for $t = n$ and $\mu(1, \mu_n, h_n, a) \geq \mu(0, \mu_n, h_n, a)$ we have $V(1, g(h_n, a), \mu(1, \mu_n, h_n, a)) \leq V(0, g(h_n, a), \mu(0, \mu_n, h_n, a))$, (c) by the assumption that property (A) holds for $t = n$ and $\mu(1, \mu'_n, h'_n, a) \leq \mu(1, \mu''_n, h'_n, a) \leq \mu(1, \mu''_n, h''_n, a)$ we have $V(1, g(h'_n, a), \mu(1, \mu'_n, h'_n, a)) \geq$

$V(1, g(h'_n, a), \mu(1, \mu''_n, h'_n, a)) \geq V(1, g(h''_n, a), \mu(1, \mu''_n, h''_n, a))$, and (d) by the assumption that properties (B) and (A) hold for $t = n$ alongside $\mu(0, \mu'_n, h'_n, a) \leq \mu(0, \mu''_n, h'_n, a)$ we have $V(0, g(h'_n, a), \mu(0, \mu'_n, h'_n, a)) \geq V(0, g(h'_n, a), \mu(0, \mu''_n, h'_n, a))$

$\geq V(0, g(h''_n, a), \mu(0, \mu''_n, h''_n, a))$.

To conclude the proof of this part let $h'_n = h''_n \equiv h_n$, $\mu'_n = \mu''_n \equiv \mu_n$ and $x'_n = 0 < 1 = x''_n$ and note by Lemma 2 that

$$\begin{aligned} & V(0, h_n, \mu_n) = \\ = \; & u(0,0) \\ & +\delta\pi(0, h_n, \mu_n)V(1, g(h_n, 0), \mu(1, \mu_n, h_n, 0)) \\ & +\delta(1 - \pi(0, h_n, \mu_n))V(0, g(h_n, 0), \mu(0, \mu_n, h_n, 0)) \\ \equiv \; & u(0,0) + E_{\pi(0,h_n,\mu_n)}V(X, g(h_n, 0), \mu(X, \mu_n, h_n, 0)) \end{aligned}$$

Note now that $V(1, h_n, \mu_n) = u(a''_n, 1) + E_{\pi(a''_n,h_n,\mu_n)}V(X, g(h_n, a''_n), \mu(X, \mu_n, h_n, a''_n))$ and that $u(0,0) > u(1,1) \geq (a''_n, 1)$. Thus, $V(1, h_n, \mu_n) \leq V(0, h_n, \mu_n)$ follows directly if $E_{\pi(a''_n,h_n,\mu_n)}V(X, g(h_n, a''_n), \mu(X, \mu_n, h_n, a''_n)) \leq$

$E_{\pi(0,h_n,\mu_n)}V(X, g(h_n, 0), \mu(X, \mu_n, h_n, 0))$. This follows after observing (a) $\pi(0, h_n, \mu_n) \leq \pi(a''_n, h_n, \mu_n)$, (b) by the assumption that property (A) holds for $t = n$ and $\mu(1, \mu_n, h_n, a) \geq \mu(0, \mu_n, h_n, a)$ we have $V(1, g(h_n, a), \mu(1, \mu_n, h_n, a))$ $\leq V(0, g(h_n, a), \mu(0, \mu_n, h_n, a))$, (c) by the assumption that property (A) holds for $t = n$ and $\mu(1, \mu_n, h_n, a''_n) \geq \mu(1, \mu_n, h_n, 0)$ we have $V(1, g(h_n, 0), \mu(1, \mu_n, h_n, 0))$ $\geq V(1, g(h_n, a''_n), \mu(1, \mu_n, h_n, a''_n))$, and (d) by the assumption that property (B) holds for $t = n$ we have $V(0, g(h_n, 0), \mu(0, \mu_n, h_n, 0))$

$\geq V(0, g(h_n, a''_n), \mu(0, \mu_n, h_n, a''_n))$.

Finally, we prove property (B) for $t = n - 1$

Recalling the definition of the operator L, note that property (B) for $t = n - 1$ would be implied directly by setting $j = 0$ in the following statement:

$V(0, L^j g(h, a), L^j \mu(0, \mu_{n-1}, h, a))$ is nonincreasing in h and a for any $T-n \geq j \geq 0$.

In what follows we prove the above statement under the assumption that properties (A) and (B) hold for all $t \geq n$.

This is done by induction on j.

Clearly the above statement is true for $j = T - n$

due to $V(0, L^{T-n}g(h, a), L^{T-n}\mu(0, \mu_{n-1}, h, a)) = u(0,0)$ after recalling that in the last period the optimal action follows the craving state. Assume that it is also true for some admissible $j = i$. For $j = i - 1$, we then have, after using Lemma 2, that

$$V(0, L^{i-1}g(h, a), L^{i-1}\mu(0, \mu_{n-1}, h, a)) = u(0,0)$$
$$+E_{\pi(0,L^{i-1}g(h,a),L^{i-1}\mu(0,\mu_{n-1},h,a))}V(X, L^i g(h, a), \mu(X, L^{i-1}\mu(0, \mu_{n-1}, h, a), L^{i-1}g(h, a), 0)).$$

Recall that $L^i g(h,a) = g(L^{i-1}g(h,a),0)$ and $L^{i-1}g(h,a)$ is increasing in h and a. Recall also that $\mu(0, L^{i-1}\mu(0,\mu_{n-1},h,a), L^{i-1}g(h,a),0) = L^i\mu(0,\mu_{n-1},h,a)$. Note that $L^{i-1}g(h,a)$ and $L^{i-1}\mu(0,\mu_{n-1},h,a)$ refer to history and posteriors in period $n+i-1$. Clearly then, using in the above expectation the inductive assumption, property (A) for $t = n+i-1$, and recalling that assumption 6 implies that

(i) $\pi(0, L^{i-1}g(h,a), L^{i-1}\mu(0,\mu_{n-1},h,a))$ is nondecreasing in h and a and

(ii) $\mu(1, L^{i-1}\mu(0,\mu_{n-1},h,a), L^{i-1}g(h,a),0)$ is nondecreasing in h and a,

we have that the above expectation is nondecreasing in h and a and thereby the desired result. ■

9 Appendix C

We start with (i). Note that $L^t g(h,a)$ is increasing in h and a. Note also that by assumption $L^{t-1}g(h,a) \geq L^t g(h,a)$. We have for any $1 \leq t \leq T-j$ (and $\prod_{i=2}^{1} y \equiv 1$):

$$\pi(0, L^{t-1}g(h,a), L^{t-1}\mu(0,\mu_j,h,a)) = p + L^{t-1}\mu(0,\mu_j,h,a)[f(L^{t-1}g(h,a),0) - p]$$
$$= p + \frac{L^{t-2}\mu(0,\mu_j,h,a)(1-f(L^{t-2}g(h,a),0))[f(L^{t-1}g(h,a),0)-p]}{L^{t-2}\mu(0,\mu_j,h,a)(1-f(L^{t-2}g(h,a),0)) + (1-L^{t-2}\mu(0,\mu_j,h,a))(1-p)}$$

Thus, for $h' \leq h''$ and/or $a' \leq a''$ we have that

$$\pi(0, L^{t-1}g(h',a'), L^{t-1}\mu(0,\mu_j,h',a')) - \pi(0, L^{t-1}g(h'',a''), L^{t-1}\mu(0,\mu_j,h'',a''))$$

has the sign of

$$
\begin{aligned}
&L^{t-2}\mu(0,\mu_j,h',a')(1-f(L^{t-2}g(h',a'),0))[f(L^{t-1}g(h',a'),0)-p]\times \\
&\{L^{t-2}\mu(0,\mu_j,h'',a'')(1-f(L^{t-2}g(h'',a''),0))+(1-L^{t-2}\mu(0,\mu_j,h'',a''))(1-p)\} \\
&-L^{t-2}\mu(0,\mu_j,h'',a'')(1-f(L^{t-2}g(h'',a''),0))[f(L^{t-1}g(h'',a''),0)-p]\times \\
&\{L^{t-2}\mu(0,\mu_j,h',a')(1-f(L^{t-2}g(h',a'),0))+(1-L^{t-2}\mu(0,\mu_j,h',a'))(1-p)\} \\
&= \\
&\{L^{t-2}\mu(0,\mu_j,h',a')(1-f(L^{t-2}g(h',a'),0))L^{t-2}\mu(0,\mu_j,h'',a'')(1-f(L^{t-2}g(h'',a''),0))\times \\
&\times[f(L^{t-1}g(h',a'),0)-f(L^{t-1}g(h'',a''),0)]\} \\
&+(1-p)\times \\
&\{L^{t-2}\mu(0,\mu_j,h',a')(1-f(L^{t-2}g(h',a'),0))[f(L^{t-1}g(h',a'),0)-p](1-L^{t-2}\mu(0,\mu_j,h'',a'')) \\
&-L^{t-2}\mu(0,\mu_j,h'',a'')(1-f(L^{t-2}g(h'',a''),0))[f(L^{t-1}g(h'',a''),0)-p](1-L^{t-2}\mu(0,\mu_j,h',a'))\}
\end{aligned}
$$

Recalling the monotonicity properties of $L^{t-1}g(h,a)$, we have that the sign of the first term above is non-positive. The sign of the second term above is also non-positive if

$$
\begin{aligned}
&\frac{L^{t-2}\mu(0,\mu_j,h',a')}{(1-L^{t-2}\mu(0,\mu_j,h',a'))}\times \\
&\frac{(1-f(L^{t-2}g(h',a'),0))}{(1-f(L^{t-2}g(h'',a''),0))}\frac{[f(L^{t-1}g(h',a'),0)-p]}{[f(L^{t-1}g(h'',a''),0)-p]} \\
\leq\ &\frac{L^{t-2}\mu(0,\mu_j,h'',a'')}{(1-L^{t-2}\mu(0,\mu_j,h'',a''))}
\end{aligned}
$$

which can be rewritten as

$$
\begin{aligned}
&\frac{L^{t-3}\mu(0,\mu_j,h',a')(1-f(L^{t-3}g(h',a'),0))}{(1-L^{t-3}\mu(0,\mu_j,h',a'))(1-p)}\times \\
&\frac{(1-f(L^{t-2}g(h',a'),0))}{(1-f(L^{t-2}g(h'',a''),0))}\frac{[f(L^{t-1}g(h',a'),0)-p]}{[f(L^{t-1}g(h'',a''),0)-p]} \\
\leq\ &\frac{L^{t-3}\mu(0,\mu_j,h'',a'')(1-f(L^{t-3}g(h'',a''),0))}{(1-L^{t-3}\mu(0,\mu_j,h'',a''))(1-p)}
\end{aligned}
$$

and hence, by iterating backwards, as

$$\begin{aligned}
&\frac{\mu_j(1-f(h',a'))}{(1-\mu_j)(1-p)}\frac{\prod_{i=2}^{t}(1-f(L^{t-i}g(h',a'),0))}{\prod_{i=2}^{t}(1-f(L^{t-i}g(h'',a''),0))}\frac{[f(L^{t-1}g(h',a'),0)-p]}{[f(L^{t-1}g(h'',a''),0)-p]}\\
\leq\ &\frac{\mu_j(1-f(h'',a''))}{(1-\mu_j)(1-p)}.
\end{aligned}$$

This is true of $\mu_j = 0$ or, otherwise, if

$$\begin{aligned}
&\frac{(1-f(h',a'))}{(1-f(h'',a''))}\frac{\prod_{i=2}^{t}(1-f(L^{t-i}g(h',a'),0))}{\prod_{i=2}^{t}(1-f(L^{t-i}g(h'',a''),0))}\\
\leq\ &\frac{[f(L^{t-1}g(h'',a''),0)-p]}{[f(L^{t-1}g(h',a'),0)-p]}.
\end{aligned} \tag{5}$$

We turn to (ii). Note due to $p \leq f(h,a)$, that $L^{t-1}\mu(0,\mu_j,h,a) \geq L^t\mu(0,\mu_j,h,a)$. Note also that $L^t\mu(0,\mu_j,h,a)$ is nonincreasing in h and a. We have, for any $t \leq T-j-1$:

$$\begin{gathered}
\mu(1, L^{t-1}\mu(0,\mu_j,h,a), L^{t-1}g(h,a),0) = \\
\frac{L^{t-1}\mu(0,\mu_j,h,a)f(L^{t-1}g(h,a),0)}{L^{t-1}\mu(0,\mu_j,h,a)f(L^{t-1}g(h,a),0)+(1-L^{t-1}\mu(0,\mu_j,h,a))p}
\end{gathered}$$

Thus,

$$\mu(1, L^{t-1}\mu(0,\mu_j,h',a'), L^{t-1}g(h',a'),0) - \mu(1, L^{t-1}\mu(0,\mu_j,h'',a''), L^{t-1}g(h'',a''),0)$$

has the sign of

$$\begin{aligned}
&L^{t-1}\mu(0,\mu_j,h',a')f(L^{t-1}g(h',a'),0)(1-L^{t-1}\mu(0,\mu_j,h'',a''))p\\
&-L^{t-1}\mu(0,\mu_j,h'',a'')f(L^{t-1}g(h'',a''),0)(1-L^{t-1}\mu(0,\mu_j,h',a'))p
\end{aligned}$$

The sign of this is non-positive if

$$\frac{L^{t-1}\mu(0,\mu_j,h',a')}{(1-L^{t-1}\mu(0,\mu_j,h',a'))}\frac{f(L^{t-1}g(h',a'),0)}{f(L^{t-1}g(h'',a''),0)} \leq \frac{L^{t-1}\mu(0,\mu_j,h'',a'')}{(1-L^{t-1}\mu(0,\mu_j,h'',a''))}$$

or, equivalently, if

$$\begin{aligned}
& \frac{L^{t-2}\mu(0,\mu_j,h',a')}{(1-L^{t-2}\mu(0,\mu_j,h',a'))}\frac{(1-f(L^{t-2}g(h',a'),0))}{(1-f(L^{t-2}g(h'',a''),0))}\frac{f(L^{t-1}g(h',a'),0)}{f(L^{t-1}g(h'',a''),0)} \\
\leq \quad & \frac{L^{t-2}\mu(0,\mu_j,h'',a'')}{(1-L^{t-2}\mu(0,\mu_j,h'',a''))}
\end{aligned}$$

By backward iteration (recall the steps above), the latter is true if $\mu_j = 0$ or, otherwise, if ,

$$\begin{aligned}
& \frac{(1-f(h',a'))}{(1-f(h'',a''))}\frac{\prod_{i=2}^{t}(1-f(L^{t-i}g(h',a'),0))}{\prod_{i=2}^{t}(1-f(L^{t-i}g(h'',a''),0))} \\
\leq \quad & \frac{f(L^{t-1}g(h'',a''),0)}{f(L^{t-1}g(h',a'),0)}.
\end{aligned} \tag{6}$$

Comparing (5) and (6), and noting that $\frac{f(L^{t-1}g(h'',a''),0)}{f(L^{t-1}g(h',a'),0)} \leq \frac{f(L^{t-1}g(h'',a''),0)-p}{f(L^{t-1}g(h',a'),0)-p}$, we thus have that a necessary and sufficient condition for both assumptions (i) and (ii) to be true for any μ_j is (6).

To see how restrictive (6) is and how it relates to the literature, consider the following law of motion: $g(h,a) = \rho(h) + \rho_2 a$ with $\rho_2 > 0$, $\rho(h)$ positive and increasing with $\rho(0) = 0$, $\rho(H) + \rho_2 \leq H$ (recall our requirement that $g(h,a) \leq H$) and $h_1 < h = \rho(h) + \rho_2$ (so that consumption raises history).[23] This encompasses the law of motions (conditional on history h be bounded and $\rho(h)$ being increasing) in BM (where $\rho(h) = \rho_1 h$, $1 \geq \rho_1 \geq 0$ and $\rho_2 = 1$), GK (where $\rho(h) = \rho_1 h$ and $\rho_1 = \rho_2 < 1$), L (where $\rho(h) = \rho_1 h$ and $\rho_1 + \rho_2 = 1$). In addition, if $\rho(\rho_2) = \rho_2$, $\rho(h) = h$ for $h < \rho_2$ and $\rho(h) = \rho_2 + \rho_1(h - \rho_2)$ for $h > \rho_2$ with $0 \leq \rho_1 < 1$, then it also shares in a simple manner the qualitative characteristics of the law of motion in BR; in particular that there is depreciation ($\rho_1 < 1$) and once consumption takes place history never reverts to the 'clean state' $h = 0$ (here the lower history of someone who has ever tried the substance is $\rho_2 > 0$ and in BR it is $\rho_2 = 1$).

[23] Interestingly, (6) is not satisfied if $\rho(h) = 0$ for any h (as in GP) and $f(h'a') < f(h'',a'')$: in this case $f(L^{t-i}g(h,a),0) = p$ for any $i = 1,...,t$. Thus, our assumption that g is increasing in both a and h is crucial for our results when $T > 2$.

Furthermore, consider $f(h,a) = \hat{f}(g(h,a))$ (with $\hat{f}(g)$ increasing, $\hat{f}(0) = p$ and $\hat{f}(g(H,1)) < 1$), which is consistent with BR (for a given "lifestyle activity").[24] A simple special case of this is $f(h,a) = p + g(h,a)$ with $g(H,1) < 1-p$ (recall our requirement that $f(h,a) < 1$).

For such fundamentals, focus on the case of $h_1 \geq \rho_2$ (thus, our decision maker has already consumed once the substance - or in the case of smokers our decision maker has been a passive smoker). We can thus restrict further attention to the case of $g(h,a) = c(a) + \rho_1 h + \rho_2 a$, with $c \equiv 0$ (as in BM, GK and L) or $c(1) = 0$ and $c(0) = (1-\rho_1)\rho_2 > 0$ (as in BR) and $\rho_2 > 0$ and $0 < \rho_1 < 1$.

We then have that $f(h,a) = p + c(a) + \rho_1 h + \rho_2 a$. Moreover, $f(L^j g(h,a), 0) = p + c(0) + \rho_1 L^j g(h,a)$, and $Lg(h,a) = c(0) + \rho_1 g(h,a) = c(0) + c(a)\rho_1 + \rho_1^2 h + \rho_1\rho_2 a$, $L^2 g(h,a) = c(0) + \rho_1 Lg(h,a) = c(0)(1+\rho_1) + c(a)\rho_1^2 + \rho_1^3 h + \rho_1^2\rho_2 a$, and continuing the iteration, $L^j g(h,a) = c(0)\sum_{i=0}^{j-1}\rho_1^i + c(a)\rho_1^j + \rho_1^{j+1}h + \rho_1^j\rho_2 a$.

Thus, $f(L^j g(h,a), 0) = p + c(0)\sum_{i=0}^{j}\rho_1^i + c(a)\rho_1^{j+1} + \rho_1^{j+2}h + \rho_1^{j+1}\rho_2 a$.

Therefore, after using convention $\sum_{\kappa=0}^{-1} y = 0$, $(1-f(h,a))\prod_{i=2}^{t}(1-f(L^{t-i}g(h,a),0)) = \prod_{i=2}^{t+1}(1-p-c(0)\sum_{\kappa=0}^{t-i}\rho_1^\kappa - c(a)\rho_1^{t-i+1} - \rho_1^{t-i+2}h - \rho_1^{t-i+1}\rho_2 a)$ and (6) can be rewritten as

$$\frac{\prod_{i=2}^{t+1}(1-p-c(0)\sum_{\kappa=0}^{t-i}\rho_1^\kappa - c(a')\rho_1^{t-i+1} - \rho_1^{t-i+2}h' - \rho_1^{t-i+1}\rho_2 a')}{\prod_{i=2}^{t+1}(1-p-c(0)\sum_{\kappa=0}^{t-i}\rho_1^\kappa - c(a'')\rho_1^{t-i+1} - \rho_1^{t-i+2}h'' - \rho_1^{t-i+1}\rho_2 a'')}$$

$$\leq \frac{p + c(0)\sum_{i=0}^{t-1}\rho_1^i + c(a'')\rho_1^t + \rho_1^{t+1}h'' + \rho_1^t\rho_2 a''}{p + c(0)\sum_{i=0}^{t-1}\rho_1^i + c(a')\rho_1^t + \rho_1^{t+1}h' + \rho_1^t\rho_2 a'}$$

The above alongside $\rho_1 H + \rho_2 < 1-p$ (to ensure $f(h,a) < 1$) and $\rho_1 H + \rho_2 \leq H$ (to ensure that $g(h,a) \leq H$) place restrictions on $\rho_1 > 0$ and $\rho_2 > 0$.

Clearly for $\rho_1 \downarrow 0$ (and hence $\rho_2 \leq H$ and $\rho_2 < 1-p$) the above is violated

[24] Recall that in BM, GK and GP we have $f(h,a) = 0$, while in L we have $\mu_0 = 0$ and/or $f(h,a) = p$ for any h, a.

if $a' < a''$ and $c \equiv 0$. However, it is satisfied if $c(1) = 1$ and $c(0) = \rho_2(1-\rho_1)$ (note that in this case we have $f(h,a) = p + \rho_2$ as $\rho_1 \downarrow 0$).

For $\rho_1 \uparrow 1$ and hence $\rho_2 \downarrow 0$ and $H < 1-p$ and $c(0) \downarrow 0$, the above is satisfied (recall also that $c(1) = 0$) if

$$\frac{(1-p-h')^t}{(1-p-h'')^t} \leq \frac{p+h''}{p+h'}$$

which is satisfied if $h' = h''$ and $a' < a''$. Moreover, if $a' = a''$ and $h' < h''$, we have that the left hand side of the above inequality is decreasing in t. Thus, the above inequality is satisfied for any $t \geq 1$ if and only if

$$\begin{aligned} \frac{(1-p-h')}{(1-p-h'')} &\leq \frac{p+h''}{p+h'} \Longrightarrow \\ (1-2p-h')h' &\leq (1-2p-h'')h'' \end{aligned}$$

This, in turn is satisfied for any $h', h'' \in [h_1, H]$ such that $h'' > h'$ if and only if $(1-2p-h)h$ is nondecreasing, which is true if and only if

$$1 - 2p - 2H \geq 0$$

Note that the latter implies $1 - p > H$.

Accordingly, by continuity the fundamentals in question satisfy (6) if $1-2p \geq 2H$, ρ_1 is sufficiently high and ρ_2 is sufficiently low (and $\rho_1 H + \rho_2 < 1-p$ and $\rho_1 H + \rho_2 \leq H$).

10 Appendix D: The Infinite Horizon Case

To accommodate the extension where the temporal payoff depends also on past consumption, let us define here the decision-maker's utility to be $v(a_t, x_t, h_t)$, $t = 1, ..., \infty$, with $v(0,0,h) > v(1,0,h)$ and $v(0,0,h) > v(1,1,h) > v(0,1,h)$ and $v(a,x,h)$ being nonincreasing in h for any a and h. Assume that v is bounded and continuous functions on $\{0,1\} \times \{0,1\} \times [0,H]$.

Recall that the law of motion for period t consumption history is

$$h_t = g(h_{t-1}, a_{t-1})$$

with h_0 and a_0 predetermined, $g(0,0) = 0$ and g being continuous and strictly increasing.

Recall also that (for given p and continuous and strictly increasing $1 > f(h,a) \geq p$) Bayesian updating implies the following difference equation for the period t posterior:

$$\mu_t = M(\mu_{t-1}, x_t, h_{t-1}, a_{t-1})$$

with μ_0 predetermined ($f(h_0, a_0) \equiv f_0$) and

$$\begin{aligned} M(\mu_{t-1}, 1, h_{t-1}, a_{t-1}) &= \frac{\mu_{t-1} f(h_{t-1}, a_{t-1})}{(1-\mu_{t-1})p + \mu_{t-1} f(h_{t-1}, a_{t-1})} \text{ and} \\ M(\mu_{t-1}, 0, h_{t-1}, a_{t-1}) &= \frac{\mu_{t-1}(1 - f(h_{t-1}, a_{t-1}))}{(1-\mu_{t-1})(1-p) + \mu_{t-1}(1 - f(h_{t-1}, a_{t-1}))} \end{aligned}$$

Despite the seemingly nonstationary nature of the probability measure over the stochastic state x, one can re-write it in a way that beliefs over the next-period's craving shock can be represented by a stationary and continuous mapping. In more detail, note that the probability that $x_{t+1} = 1$ given past consumptions and craving shocks is equal to

$$\pi(a_t, h_t, \mu_t) \equiv p + \mu_t[f(h_t, a_t) - p]$$

Note that by including the posterior probability in the set of state/predetermined variables the above probability becomes stationary. However this comes at the expense of having an additional law of motion with no clear-cut monotonicity properties: that for posteriors. To see this define the following law of motion (from the decision-maker's point of view) of the period t craving-state

$$x_t = \chi(\mu_{t-1}, h_{t-1}, a_{t-1}, \omega_t)$$

with ω_t being a uniformly distributed random variable in $[0,1]$, $\chi(\mu_{t-1}, h_{t-1}, a_{t-1}, \omega)$ being strictly decreasing in ω and $\chi(\mu_{t-1}, h_{t-1}, a_{t-1}, \omega) = 1$ when $\omega \leq \pi(a_{t-1}, h_{t-1}, \mu_{t-1})$ and zero otherwise. Note that χ is nondecreasing in μ_{t-1}, h_{t-1} and a_{t-1}. We then have that the period t posterior is in effect a stochastic state variable as well with law of motion

$$\mu_t = \hat{M}(\mu_{t-1}, h_{t-1}, a_{t-1}, \omega_t) \equiv M(\mu_{t-1}, \chi(\mu_{t-1}, h_{t-1}, a_{t-1}, \omega_t), h_{t-1}, a_{t-1})$$

Clearly, $\hat{M}$ is nondecreasing in μ_{t-1}. $\hat{M}$ is also nondecreasing in h_{t-1} and a_{t-1} if $\omega_t \leq \pi(a_{t-1}, h_{t-1}, \mu_{t-1})$. However, if $\omega_t > \pi(a_{t-1}, h_{t-1}, \mu_{t-1})$, the (contradicting) monotonicity properties of M and χ with respect to h_{t-1} and

a_{t-1} imply that the corresponding monotonicity of $\hat{M}$ requires more structure into the problem (ie. the functions M and χ). This is what complicates the derivation of the properties of the (shown below to be well-defined) value function, and the additional structure is accomplished by assumption 6.

We are now ready to proceed.

10.1 V is well-defined for $T \equiv \infty$

Our first task is to show that a well-defined value function $V(x, h, \mu)$ exists. Letting then the period$-t$ vector of state/predetermined variables be $s_t = (x_t, h_t, \mu_t)$, the Bellman equation is

$$V(x_t, h_t, \mu_t) =$$
$$\max_{a_t \in \{0,1\}} \left\{ v(a_t, x_t, h_t) + \delta \left\{ \begin{array}{c} \pi(a_t, h_t, \mu_t) V(1, g(h_t, a_t), \frac{\mu_t f(h_t, a_t)}{(1-\mu_t)p + \mu_t f(h_t, a_t)}) \\ + (1 - \pi(a_t, h_t, \mu_t)) V(0, g(h_t, a_t), \frac{\mu_t (1 - f(h_t, a_t))}{(1-\mu_t)(1-p) + \mu_t (1 - f(h_t, a_t))}) \end{array} \right\} \right\}$$

Define the set S of bounded and continuous functions γ of x, h and μ, with $x \in \{0, 1\}$, $h \in [0, H]$ and $\mu \in [0, 1]$. Define then the function of x, h and μ

$$(\hat{L}\gamma)(x, h, \mu) = \max_{a \in \{0,1\}} \{ v(a, x, h) + \delta \left\{ \begin{array}{c} \pi(a, h, \mu) \gamma(1, g(h, a), \frac{\mu f(h, a)}{(1-\mu)p + \mu f(h, a)}) \\ + (1 - \pi(a, h, \mu)) \gamma(0, g(h, a), \frac{\mu (1 - f(h, a))}{(1-\mu)(1-p) + \mu (1 - f(h, a))}) \end{array} \right\} \}$$

Note that if $\gamma \in S$, then $\hat{L}\gamma$ is also bounded and continuous by Berge's theorem of maximum. Thus, the above defines a mapping $\hat{L}$ from the set S into itself. Moreover the set S with the sup norm, $\|\gamma\| = \sup_{x,a,\mu} |\gamma(x, a, \mu)|$ is a complete normed vector space (see Theorem 3.1 in Stockey and Lucas (1989) p. 47).

Define the metric $\lambda(z, y) = \|z - y\|$ and thereby the complete metric space (S, λ). We then know from the Contraction Mapping Theorem (Theorem 3.2 in Stockey and Lucas (1989) p. 50) that if $\hat{L} : S \to S$ is a contraction mapping with modulus β, then (a) $\hat{L}$ has exactly one fixed point, call it V, in S, and (b) for any $V_0 \in S$, $\lambda(\hat{L}^n V_0, V) \leq \beta^n \lambda(V_0, V)$, $n = 0, 1, 2, ...$ To show then that the Bellman equation above is uniquely defined, we only have to show that $\hat{L}$ is a contraction mapping with modulus β (i.e. that for some $\beta \in (0, 1)$, $\lambda(\hat{L}z, \hat{L}y) \leq \beta \lambda(z, y)$ for all $z, y \in S$). For this we make use of Blackwell's sufficient conditions for a contraction (Theorem 3.3 in Stockey

and Lucas (1989) p. 54): Let $Q \subseteq R^l$, and let $B(Q)$ be a space of bounded functions $V : Q \to R$ with the sup norm. Let $L : B(Q) \to B(Q)$ be an operator satisfying (a) monotonicity: $z, y \in B(Q)$ and $z(q) \leq y(q)$, for all $q \in Q$, implies $(Tz)(q) \leq (Ty)(q)$, for all $q \in Q$,(b) discounting: there exists some $\beta \in (0,1)$ such that $(T(f+d))(q) \leq (Tf)(q) + \beta d$, for all $f \in B(X)$, $d \geq 0$, $q \in Q$, were $(f+d)(q)$ is the function defined by $(f+d)(q) = f(q)+d$. Then, L is a contraction mapping with modulus β.

Applying this to our case we have that the monotonicity requirement is trivially satisfied because $(\hat{L}\gamma)(x,h,\mu)$ is the maximized value of the function

$$w(a,x,h,\mu;\gamma) \equiv v(a,x,h)+\delta \left\{ \begin{array}{c} \pi(a,a,\mu)\gamma(1, g(h,a), \frac{\mu f(h,a)}{(1-\mu)p+\mu f(h,a)}) \\ +(1-\pi(a,h,\mu))\,\gamma(0, g(h,a), \frac{\mu(1-f(h,a))}{(1-\mu)(1-p)+\mu(1-f(h,a))}) \end{array} \right\},$$

and if $z(x,h,\mu) \leq y(x,h,\mu)$, then $w(a,x,h,\mu;y)$ is uniformly higher than $w(a,x,h,\mu;z)$. In more detail, after defining $a_f \equiv \max_a w(a,x,h,\mu,f)$, we have that if $z(x,h,\mu) \leq y(x,h,\mu)$ then $(\hat{L}y)(x,h,\mu) \geq w(a_z,x,h,\mu;y) \geq w(a_z,x,h,\mu;z) = (\hat{L}z)(x,h,\mu)$. The discounting requirement is also trivially satisfied as $(\hat{L}(V+d))(x,h,\mu) = (\hat{L}TV)(x,h,\mu)+\delta d$. Therefore, the mapping $\hat{L} : S \to S$ is a contraction mapping with modulus δ. Hence, the Bellman equation and the value function it defines implicitly are well-defined.

10.2 Properties of V

Given Lemma 2, showing Lemma 3 and Proposition 4 amounts to showing that properties (A) and (B) are valid when we move to infinite horizon (with cumulative welfare effects of past consumptions - where Lemma 2 is still true as we also mention in the relevant subsection).

After recalling the definition of the operator L on functions g and μ, let S^* be the subset of S with all functions γ that are weakly decreasing in x, h, μ and satisfy the following property:

$$\gamma(0, L^j g(h,a), L^j \mu(0, \mu_{n-1}, h, a)) \text{ is nonincreasing in } h \text{ and } a \text{ for any } j \geq 0,\ n \geq 1$$

Note that S^* is nonempty (as it includes all constant functions) and all functions in S^* satisfy properties (A) and (B) (that latter follows by setting $j = 0$).

Given that $\hat{L}$ is uniformly contracting on the complete space S (endowed with the sup norm), we have that if $\hat{L} : S^* \to S^*$ and S^* is closed, then the unique value function defined by the Bellman equation lies in S^*.

10.2.1 Proof that $\hat{L} : S^* \to S^*$

Here we show that $\hat{L} : S^* \to S^*$.

To show this, consider $\gamma \in S^*$ and let $(x'', h'', \mu'') \geq (x', h', \mu')$. Note that the (assumed) monotonicity properties of $\pi(a, h, \mu)$, $\mu(x, \mu, h, a)$,

$\pi(0, L^j g(h, a), L^j \mu(0, \mu_{n-1}, h, a))$, $\mu(1, L^j \mu(0, \mu_{n-1}, h, a), L^j g(h, a), 0)$ and $L^j \mu(0, \mu_{n-1}, h, a), L^j g(h, a)$, for all $j \geq 0$ and $n \geq 2$, imply that

$$\begin{aligned} E_{\pi(a,h,\mu')}[\gamma(X_n, g(h, a), \mu(X_n, \mu', h, a)] &\geq E_{\pi(a,h,\mu'')}[\gamma(X_n, g(h, a), \mu(X_n, \mu'', h, a))] \\ E_{\pi(a,h',\mu)}[\gamma(X_n, g(h', a), \mu(X_n, \mu, h', a)] &\geq E_{\pi(a,h'',\mu)}[\gamma(X_n, g(h'', a), \mu(X_n, \mu, h'', a))] \end{aligned}$$

and, after setting $\mu'_n(0) = \mu(0, \mu_{n-1}, h', a')$ and $\mu''_n(0) = \mu(0, \mu_{n-1}, h'', a'')$ and recalling $\mu(0, L^j \mu''_n(0), L^j g(h'', a''), 0)) = L^{j+1} \mu''_n(0)$, that

$$\begin{aligned} & E_{\pi(0,L^j g(h',a'),L^j \mu'_n(0))}[\gamma(X_{n+j+1}, L^{j+1} g(h', a'), \mu(X_{n+j+1}, L^j \mu'_n(0), L^j g(h', a'), 0))] \\ \geq\ & E_{\pi(0,L^j g(h'',a''),L^j \mu''_n(0))}[\gamma(X_{n+j+1}, L^{j+1} g(h'', a''), \mu(X_{n+j+1}, L^j \mu''_n(0), L^j g(h'', a''), 0))] \end{aligned}$$

Note then that the definitions of maximum and $a(.)$, the properties of v, and the above properties imply (after setting $a' = a(x', h', \mu')$ and $a'' = a(x'', h'', \mu'')$) that

$$\begin{aligned} & (\hat{L}\gamma)(x, h, \mu') \\ =\ & v(a', x, h) + \delta \left\{ \begin{array}{c} \pi(a', h, \mu')\gamma(1, g(h, a'), \frac{\mu' f(h,a')}{(1-\mu')p+\mu' f(h,a')}) \\ + (1 - \pi(a', h, \mu'))\, \gamma(0, g(h, a'), \frac{\mu'(1-f(h,a'))}{(1-\mu')(1-p)+\mu'(1-f(h,a'))}) \end{array} \right\}\} \\ \geq\ & v(a'', x, h) + \delta \left\{ \begin{array}{c} \pi(a'', h, \mu')\gamma(1, g(h, a''), \frac{\mu' f(h,a'')}{(1-\mu')p+\mu' f(h,a'')}) \\ + (1 - \pi(a'', h, \mu'))\, \gamma(0, g(h, a''), \frac{\mu'(1-f(h,a''))}{(1-\mu')(1-p)+\mu'(1-f(h,a''))}) \end{array} \right\} \\ \geq\ & v(a'', x, h) + \delta \left\{ \begin{array}{c} \tilde{\pi}(a'', h, \mu'')\gamma(1, g(h, a''), \frac{\mu'' f(h,a'')}{(1-\mu'')p+\mu'' f(h,a'')}) \\ + (1 - \tilde{\pi}(a'', h, \mu''))\, \gamma(0, g(h, a''), \frac{\mu''(1-f(h,a''))}{(1-\mu'')(1-p)+\mu''(1-f(h,a''))}) \end{array} \right\} \\ =\ & (\hat{L}\gamma)(x, h, \mu'') \end{aligned}$$

Similarly,

$$
\begin{aligned}
& (\hat{L}\gamma)(x,h',\mu) \\
= \;& v(a',x,h') + \delta \left\{ \begin{array}{c} \pi(a',h',\mu)\gamma(1,g(h',a'),\frac{\mu f(h',a')}{(1-\mu)p+\mu f(h',a')}) \\ +(1-\pi(a',h',\mu))\,\gamma(0,g(h',a'),\frac{\mu(1-f(h',a'))}{(1-\mu)(1-p)+\mu(1-f(h',a'))}) \end{array} \right\}\} \\
\geq \;& v(a'',x,h') + \delta \left\{ \begin{array}{c} \pi(a'',h',\mu)\gamma(1,g(h',a''),\frac{\mu f(h',a'')}{(1-\mu)p+\mu f(h',a'')}) \\ +(1-\pi(a'',h',\mu'))\,\gamma(0,g(h',a''),\frac{\mu(1-f(h',a''))}{(1-\mu')(1-p)+\mu'(1-f(h',a''))}) \end{array} \right\} \\
\geq \;& v(a'',x,h'') + \delta \left\{ \begin{array}{c} \pi(a'',h'',\mu)\gamma(1,g(h'',a''),\frac{\mu f(h'',a'')}{(1-\mu)p+\mu f(h'',a'')}) \\ +(1-\pi(a'',h'',\mu))\,\gamma(0,g(h'',a''),\frac{\mu(1-f(h'',a''))}{(1-\mu)(1-p)+\mu(1-f(h'',a''))}) \end{array} \right\} \\
= \;& (\hat{L}\gamma)(x,h'',\mu)
\end{aligned}
$$

The above in turn imply from Lemma 2 that $a(0,h,\mu)=0$. Thus, we have (by using again the definition of optimum and the above properties) that:

$$
\begin{aligned}
& (\hat{L}\gamma)(0,L^j g(h',a'),L^j\mu_n'(0)) \\
& = v(0,0,h') \\
+\delta & \left\{ \begin{array}{c} \pi(0,L^j g(h',a'),L^j\mu_n'(0))\gamma(1,L^{j+1}g(h',a'),\mu(1,L^j\mu_n'(0),L^j g(h',a'),0)) \\ +(1-\pi(0,L^j g(h',a'),L^j\mu_n'(0)))\,\gamma(0,L^{j+1}g(h',a'),L^{j+1}\mu'(0)) \end{array} \right\}\} \\
& \geq v(0,0,h'') \\
+\delta & \left\{ \begin{array}{c} \pi(0,L^j g(h'',a''),L^j\mu_n''(0))\gamma(1,L^{j+1}g(h'',a''),\mu(1,L^j\mu_n''(0),L^j g(h'',a''),0)) \\ +(1-\pi(0,L^j g(h'',a''),L^j\mu_n''(0)))\,\gamma(0,L^{j+1}g(h'',a''),L^{j+1}\mu_n''(0)) \end{array} \right\} \\
& = (\hat{L}\gamma)(0,L^j g(h'',a''),L^j\mu_n''(0))
\end{aligned}
$$

Thus, we have that $\hat{L}$ maps S^* into itself. It remains to show that S^* is closed. We show this next.

10.2.2 Proof that S^* is closed

Let $\gamma(x,h,\mu)$ be a function from $\{0,1\}\times[0,H]\times[0,1]$ to R and recall that S denotes the subset of all continuous and bounded γ. The subset of S consisting of all weakly decreasing functions is denoted by S^d. We will endow all the function spaces we will deal with hereafter with the sup norm. We first (partly) show a standard result which is that S^d is a closed subset of S. The reason to have this argument explicitly is that we will need to do various

iterations on S^d and we will be using the same type of arguments there as well.

S^d is a closed subset of S We first want to show that S^d is a closed subset of S.

Consider a sequence of functions $\{\gamma_n\}_n$, where each $\gamma_n \in S^d$, that converges to a limit function $\overline{\gamma}$. We want to show that $\overline{\gamma} \in S^d$. Showing continuity and boundedness of $\bar{\gamma}$ is standard, so the only thing to show is that $\overline{\gamma}$ is weakly decreasing.

So, suppose to the contrary that it is not. More specifically, suppose, without loss of generality[25], that there are two points in the domain (x, h', μ) and (x, h'', μ) such that $h'' > h'$ but that $\overline{\gamma}(x, h', \mu) < \overline{\gamma}(x, h'', \mu)$. To be specific, suppose $\overline{\gamma}(x, h'', \mu) - \overline{\gamma}(x, h', \mu) = \varepsilon > 0$.

Now, since γ_n converges uniformly to $\overline{\gamma}$, there is a value of n, say $\widehat{n}$ such that for all $n > \widehat{n}$,

$$|\gamma_n(x, h'', \mu) - \overline{\gamma}(x, h'', \mu)| < \frac{\varepsilon}{2}$$

and

$$|\gamma_n(x, h', \mu) - \overline{\gamma}(x, h', \mu)| < \frac{\varepsilon}{2}$$

From the above two inequalities we have,

$$\gamma_n(x, h'', \mu) > \overline{\gamma}(x, h'', \mu) - \frac{\varepsilon}{2} \tag{7}$$

and,

$$-\gamma_n(x, h', \mu) > -\overline{\gamma}(x, h', \mu) - \frac{\varepsilon}{2}$$

which can be rewritten as

$$-\gamma_n(x, h', \mu) > -\overline{\gamma}(x, h', \mu) - \frac{\varepsilon}{2} \tag{8}$$

Combining (7) and (8), we get

$$\gamma_n(x, h'', \mu) - \gamma_n(x, h', \mu) > \overline{\gamma}(x, h'', \mu) - \overline{\gamma}(x, h', \mu) - \varepsilon = 0$$

which contradicts the fact that γ_n is a nonincreasing function.

[25] It is more convenient to take the two points in the domain that differ in the h component only. However, as will be clear shortly from the argument, this choice of two points is not essential to the argument.

$S^{1,d}$ **is a closed subset of** S^d We now consider a subset $S^{1,d}$ of S^d that has some further properties.

Specifically, take any function $\gamma \in S^d$. Consider any point in the domain of the form $(0, h, \mu)$. For any admissible (according to assumption 6) g and f, consider the function $\widetilde{\gamma}^1$ which maps h, a, μ to R as:

$$\widetilde{\gamma}^1(h, a, \mu) \equiv \gamma(0, g(h, a), \mu(0, h, \mu, a)) \tag{9}$$

Define then the set of functions, $S^{1,d} \subseteq S^d$ such that $\gamma \in S^{1,d}$ if and only if $\widetilde{\gamma}^1$ is weakly decreasing in h and a.

We need to show that $S^{1,d}$ is closed. Towards that end consider a sequence of functions $\{\gamma_n^1\}_n$, where $\gamma_n^1 \in S^{1,d}$, such that this sequence has a limit function $\overline{\gamma}^1$. We need to show that $\overline{\gamma}^1 \in S^{1,d}$.

Notice however that since $S^{1,d}$ is a subset of S^d and we have shown that S^d is closed, any limit function then must be in S^d so the only way $\overline{\gamma}^1$ can be not in $S^{1,d}$ is because its corresponding $\overline{\widetilde{\gamma}}^1$ is not weakly decreasing in h and a.

Fix a μ and consider $h'' \geq h'$ and $a'' \geq a'$, with at least one inequality strict, such that $\overline{\widetilde{\gamma}}^1(h'', a, \mu) - \overline{\widetilde{\gamma}}^1(h', a, \mu) = \varepsilon$, for some $\varepsilon > 0$.

Let $g' = g(h', a')$ and $g'' = g(h'', a'')$. Similarly, $\mu' = \mu(0, \mu, h', a')$ and $\mu'' = \mu(0, \mu, h'', a'')$. From (9), we have, $\gamma_n^1(0, g', \mu') = \widetilde{\gamma}_n^1(h', a', \mu)$ and $\gamma_n^1(0, g'', \mu'') = \widetilde{\gamma}_n^1(h'', a'', \mu)$. Moreover $\bar{\gamma}_n^1(0, g', \mu') = \overline{\widetilde{\gamma}}_n^1(h, a', \mu)$ and $\bar{\gamma}_n^1(0, g'', \mu'') = \overline{\widetilde{\gamma}}_n^1(h, a'', \mu)$.

Since γ_n^1 converges to $\overline{\gamma}^1$ uniformly, there exists a positive integer $\widehat{n}$, such that for all $n > \widehat{n}$,

$$|\gamma_n^1(0, g'', \mu'') - \overline{\gamma}^1(0, g'', \mu'')| < \frac{\varepsilon}{2}$$

and

$$|\gamma_n^1(0, g', \mu') - \overline{\gamma}^1(0, g', \mu')| < \frac{\varepsilon}{2}$$

From the above two inequalities we have,

$$\gamma_n^1(0, g'', \mu'') > \overline{\gamma}^1(0, g'', \mu'') - \frac{\varepsilon}{2}$$

and,

$$-\gamma_n^1(0, g', \mu') > -\overline{\gamma}(0, g', \mu') - \frac{\varepsilon}{2}$$

Combining these two, we get

$$\begin{aligned}
\gamma_n^1(0, g'', \mu'') - \gamma_n^1(0, g', \mu') &> \overline{\gamma}^1(0, g'', \mu'') - \overline{\gamma}^1(0, g', \mu') - \varepsilon \\
&= \widetilde{\overline{\gamma}}^1(h'', a'', \mu) - \widetilde{\overline{\gamma}}^1(h', a', \mu) - \varepsilon \\
&= 0
\end{aligned}$$

which implies that

$$\widetilde{\gamma}_n^1(h'', a'', \mu) - \widetilde{\gamma}_n^1(h', a', \mu) > 0$$

This contradicts that $\gamma_n^1 \in S^{1,d}$.

To complete the proof of the desired result, we can apply this argument repeatedly to get successive subsets $S^d \supseteq S^{1,d} \supseteq S^{2,d}, ...$ each of which is closed. Countable intersection of closed sets are closed and hence $\cap_n S^{n,d} \equiv S^*$ (the intersection of $\{S^{n,d}\}_n$) is closed.