Compulsive and Addictive Consumption

Subir Bose[*] Miltiadis Makris[†]

May 2010

Abstract

We present a new theory of rational addiction. Addictive consumption is compulsive, influenced by subtle stochastic cravings. Crucially, the consumer lacks complete information regarding the stochastic process generating the cravings whose likelihood may be affected by past consumption. The model generates behavior, e.g., recurring recividism, associated with boundedly-rational agents. Welfare enhancing policy involves providing information about the true process. Our theory provides micro-foundation for non-linear intrinsic habit formation. There is also the possibility of extrinsic habit-forming behavior with contrarian peer group effect due to informational cascades. We derive some new results on monotonicity properties of the value function.

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

In this paper we develop a new theory of rational addiction.[1] Its importance lies in it (a) offerring novel insights regarding connections between social environment, experiences, usage, costs of self-control and consumption of addictive substances, scope of information campaigns, and the role informational peer effects for consumption of addictive substances, and (b) generating behavior often viewed as incompatible with the model of a dynamically consistent decision-maker maximizing expected utility and making Bayesian inferences whenever this is possible. The latter is very important since the most often-mentioned criticism of the rational addiction literature seems to be its perceived failure in explaining much of observed behavior of addicted individuals. In our work, we extend the basic rational addiction model in a very natural and intuitive way (see details below). It is interesting that this is enough to generate behavior that is often thought to require models of bounded rational behavior for a proper explanation.

At a more technical level, our model provides a micro-foundation for nonlinear intertemporal consumption complementarities.[2] In addition, our model gives rise to a hitherto unexplored decision-making problem where the study of the decision-maker's choices requires us proving certain monotonicity properties for the value function of the associated stochastic dynamic programming problem that do not have a counterpart in the literature on rational addiction or in the received dynamic programming literature.

Our theory is motivated by two observations. First, often consumption is addictive and (hence) compulsive. Second, the future effects of current and past consumption, though obvious at times, can also be more subtle, and without experimenting the consumer may not be able to know the extent of the addictive properties (vis-a-vis themselves) of certain consumption goods. Consider smoking as a specific example (we will continue to use this as a sort of "canonical" example throughout the paper). While smokers often feel an *obvious* desire to smoke, there are also more *subtle* ones (for example, for its supposedly relaxing role, or as an aid to help concentrate.) In fact, it is not uncommon to hear smokers claim (often as evidence of their attempt to control their smoking) that they restrict this activity only to those situations when there seems to be a dire need for it. For example, an academic who is a smoker

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

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

and under pressure to finish a paper to meet some (exogenous) deadline may feel that during those times smoking helps him/her to concentrate. While smoking may indeed help to concentrate (in what we call the craving state in the model), it is also possible that past smoking itself may cause the smoker's "need" for (or "benefit" from) smoking in order to concentrate. That is, this seeming need or benefit may in fact be illusory, and instead, a manifestation of addictiveness. However, since the situations where the smoker feels the need to smoke can be due to exogenous factors ("cue-conditioned" to borrow a term from the literature), it might not be obvious or apparent to the smoker right away that a change in his total consumption of smoking may reduce the *chances* of his *having the need* to smoke in similar future situations.

In more detail, we consider the situation where (maybe as a result of extensive public information campaign in the past) the health and monetary costs of addictive consumption are fully known by consumers. Our theory of rational addiction then is based on four central premises. First, as mentioned, addictive consumption is compulsive, influenced by the presence of temptations. These temptations are modeled as (possibly cue-triggered) taste-shocks/cravings. Second, the occurrence of temptations may depend on past behavior. Third, (addicted) individuals do understand their susceptibility to cravings and try to rationally manage the process through their consumption. In particular, and as in other models involving rational decision-makers, in the presence of an urge the consumer 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, and central, premise of our theory, and what differentiates substantially our work from the existing literature, is that decision-makers often lack precise information about the addictive properties of the substance. We should clarify, before we proceed, that while we pose the question as one where the decision-maker does not know *whether* a certain substance is addictive or not, an alternative interpretation is that even for substances known to be addictive, *individual physiological responses* to it - for example frequency of personal cravings - do vary and it is this that the decision-maker has uncertainty about. We should also emphasize that while nothing prevents interpreting this behavior as due to factors like mistakes or overconfidence etc., our theory is about a decision-maker who is a Bayesian expected utility maximizer with standard discounting facing uncertainty over the likelihood of future temptations. And the uncertainty regarding the addictive properties of the substance, or more formally, about the exact nature of the stochastic process generating these needs/craving does not have to be due to any bounded-rationalilty factors. It is perfectly plausible that the decision-maker may not have access to enough data - especially if the uncertainty is about individual physiological reaction to the addictive substance - to form a subjective prior that coincides with the true process. As a result, he/she can only learn through experimentation via consumption. The resulting behavior, though rational, may indeed seem to be deluded with respect to the true process.[3]

Obviously, 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, arguably, there are also drugs characterized by the existence of cue-conditioned cravings that override cognitive control. 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 mistakes" are not common[4] (smoking, sex, shopping, gambling, food-related addictions and kleptomania are some examples). Therefore, we view our work as an important complement to the existing literature in understanding (rational) addiction.

The four central premises of our theory differentiate it substantially from the received literature. Some of its predictions are new and offer some very interesting insights for addictive consumption. In our model there is scope for policy intervention in the form of campaigns informing consumers about the addictive properties of the various substances, but crucially not about the health and monetary costs of addictive consumption[5]. And similar to other rational addiction models, policies to force a change of behavior is not welfare enhancing (in the absence of externalities). Furthermore, our model can feature failed attempts to quit and occasional use. Such behavior occurs as a process of experimentation in the face of subjective uncertainty. Importantly, and in contrast to common conjectures and beliefs (see also below the discussion of related literature), such behavior does not require the use of any behavioral assumptions in our set up. Our theory also provides micro-foundations for nonlinear intrinsic habit-forming behavior but by starting from a standard model of fully rational decision-maker with intertemporally separable preferences. The behavior is generated due to the subjective uncertainty regarding the likelihood of future temptations, which in turn depends on past behavior and the history of past and current temptations. Our model may also give rise to extrinsic habit-forming behavior.[6] Interestingly, while there may be a *standard* peer effect (where an agent has a

[3]It is common for addicts to be overconfident about their ability to quit.

[4]On this see also footnote 18 in Bernheim and Rangel (2004).

[5]A common frustration amongst anti-smoking campaigners is that "surely by this time, people do know that smoking is bad for health".

[6]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

higher chance of consuming if more people around him do so), there could also be a *contrarian* informational peer effect. This latter happens because past and current consumption of peers may provide valuable information about the addictive properties of the substance. When only actions, but not the private information, of peers is observed, the situation is similar to ones giving rise to informational cascades as in the herding literature. An agent who would otherwise consume may refrain from doing so if many of his peers show behavior that lead the agent to be fearful about the substance's addictive property; conversely an agent who might otherwise not consume may infer from behavior of his peers that consumption is not very likely to be addictive and hence decide to consume.

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 latter is due to the fact that the environment may have a big impact on the prior of the decision-maker regarding the addictive properties of the substance, which in turn will influence the first and subsequent consumption decisions.

Second, drugs that are (perceived to be) more addictive may be associated with lower consumption among more experienced users and higher consumption among new users, in contrast to the predictions of Bernheim and Rangel (2004). In our model, rational consumers lower their consumption in the face of higher perceived degree of addiction resulting from consumption. Hence, if the increase in the perceived degree of addictiveness is increasing in past consumption then we may observe the phenomenon described. 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).[7] Similar to our model, the central premise of this method is 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 reducing both health and future monetary costs.

We discuss next the literature on addiction. The seminal paper on rational addiction, Becker and Murphy (1988) (BM hereafter), investigates a nonlinear intrinsic

reference point own past consumption.

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

habit formation model where past consumption increases with certainty the value of future consumption. In that model, the decision-maker is rational and, in contrast to our model, fully aware of the implications of her consumption decisions when maximizing her intertemporal payoff. Policy intervention is thus welfare improving only when externalities are imposed by the addicts. Orphanides and Zervos (1995) (OZ hereafter) investigate the implications of the decision-maker facing subjective uncertainty about the 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. In our model, this is not an issue as the decision-maker is fully aware of the health and monetary consequences of consumption, but not of the implications in terms of future occurrence of cravings. Gruber and Köszegi (2001) (GK hereafter) introduce dynamic inconsistency in BM by means of hyperbolic discounting. Our decision-maker's preferences, however, are not dynamically inconsistent. GK is motivated by evidence of unrealized intentions to quit at some time in the future and the search for self-control devices, such as avoiding cues or entering rehabilitation, to help quit, which is behavior that does not arise in the BM model.[8] 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. These papers do not provide any axiomatic or micro-foundation of habit forming preferences.[9] Gul and Pesendorfer (2007) (GP hereafter) in turn characterize axiomatically dynamically consistent preferences over menus of streams of consumption (rather than on streams themselves) and investigate the implications for rational addiction when the decision-maker is fully aware of the environment. In their context, past consumption affects the cost of current self-control and the decision-maker can choose consumption and future options that go against temptations. In contrast to our work, the latter are defined in GP as consumption bundles and future options that are costly to ignore (i.e. to not choose). In their setup rehab is desirable as in GK, but as a commitment device to reduce temptation; therefore, prohibitive policies are beneficial.[10] Yet, failed attempts or unrealized promises to quit are not predicted by that model. This is also true for

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

[9]For an axiomatization of hyperbolic discounting, see Harvey(1986). See also Halevy (2008) for an alternative micro-foundation of hyperbolic discounting in terms of *uncertainty* about future consumption.

[10]Moreover, the demand for temporary commitment generates rehab cycles.

the model of Bernheim and Rangel (2004) (BR hereafter). These authors also study addictive consumption in a framework that allows for micro-founded intertemporal consumption complementarities,[11] but under "cue–induced mistakes". Specifically, the decision-maker operates in a stochastic environment where depending on past consumption she may be 'hit by temptations' that carry very high physiological and psychological costs of being ignored. Thus, in contrast to our work, cue-conditioned cravings override cognitive control and the decision-maker consumes the addictive substance whenever she faces a temptation.[12] 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.[13]

In relation to the received literature, we should also point out that one can interpret the models of BM, 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, the decision-maker considers abstention to be always inferior to consumption. In these models, it is the "degree of the temptations" that depends on past consumption: in BM, 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. Unlike these, 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 (for example, due to the presence of high health costs). Furthermore, in contrast to all the above papers, temptations here are preference-shocks. However, their occurrence depends on past behavior in contrast to the cue-triggered taste-shocks in Laibson

[11] 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).

[12] 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 she 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 welfare 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.

[13] Thus, in BR, rehab cycles arise due to a form of "consumption-smoothing".

(2001) (L hereafter).[14] Our decision-maker would pay to not have a craving.[15] 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.

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 $t = 1, 2$. In Section 4, we discuss the extension of our model with $T > 2$ periods including the case of infinite horizon. 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 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 also as we have mentioned in the Introduction. An important point to note is that the consumer cannot commit to future actions: she chooses action a_t in period t.

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 in period t is to choose $a_t = x_t$ in state x_t.

[14] 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.

[15] 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). While we recognize that individuals may manage their addiction through lifestyle activities and we do not deny the importance of such activities, we abstract from this in our model to focus on the novel aspect of our theory, namely, management of addiction through changes in consumption in the face of uncertainty over the addictive properties of the substance.

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 BM, OZ, GK and 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), which 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.

The above assumptions imply that the welfare costs of the addictive consumption are known and well understood by the decision maker. In particular, note that $u(0,0) - u(1,0)$ represents the welfare cost due to health and monetary costs of compulsive consumption.[16] 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 unimportant, but only to emphasize that our results do not rely on any cumulative welfare costs. 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

[16] For some substances, health costs from consumption may depend on the craving state. For ease of exposition, we have defined in the main text the health (and monetary) costs only in terms of the no-craving state. None of the results are affected if we did otherwise.

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 the continuous function $f(h_1, a_1)$, 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 consumption (with $h_1 = 0$ representing no consumption in the past and $h_1 > 0$ representing past consumption). Note that a positive h_1 may represent past *passive* consumption. We assume the following:

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

For the formal derivation of the optimal consumption rule, it suffices to assume only that f is increasing in a; however, monotonicity with respect to h is used for the various comparative statics results. Denote by f_0 the probability that $x_1 = 1$. This probability depends positively on consumption up to and including (some artificial) period zero.[17] 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. Let μ_0 denote the subjective prior belief of the decision maker that the process is addictive.

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

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 $f \geq p$; further, 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.

The next assumption, while reasonable, also serves the technical purpose of guaranteeing that Bayes rule can be applied when $x_t = 0$ (even when $\mu_{t-1} = 1$).

Assumption 6 $f(h, 1) < 1$ *for any* h

[17] 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 avoid notational cluttering.

The above three assumptions on $f(h,a)$ and p imply that that $0 < \mu_0 f_0 + (1 - \mu_0)p < 1$.

Next we assume a supermodularity type property on f.

Assumption 7 *$f(h,1) - f(h,0)$ is increasing in h for all h.*

Assumption (7) is consistent with the literature. Further, it is used only for two results: (i) showing possibility of recidivist behavior, and (ii) exhibiting pattern of use of more or less addictive substances amongst old and young users.

We have already mentioned this in the Introduction but let us emphasize again that even though we have described the model as depicting the story where a decision-maker does not know *whether* a substance is addictive or not, it is possible (by having a slightly different specification) to capture an alternative story in which even though it might be *generally known* that the substance is addictive, *individual physiological responses* to consumption of the substance do vary, and what the decision maker is not fully sure about is *his/her own responses* to the substance. More formally, suppose the decision-maker knows that the probability of having a craving is given by the function $f(h,a,\theta)$ but des not know her addictive type $\theta \in \{\theta_l, \theta_h\}$. Under the assumptions that $f(0,0,\theta_l) = f(0,0,\theta_h) \equiv p$, $f(h,a,\theta_h) > f(h,a,\theta_l)$ except when $h = a = 0$, and that $f(h,a,\theta_h)/f(h,a,\theta_l)$ is increasing in h and a, this alternative model would give qualitatively similar results as the one we consider.[18]

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, μ_t. Let $m(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})$ be the function (of the variables, x_t, μ_{t-1}, h_{t-1},and a_{t-1}) that gives the posterior; i.e. $\mu_t = m(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})$.[19] In each period, the consumer also chooses action a_t in order to maximize her intertemporal welfare. Being a rational consumer, she takes

[18] Further, this alternative version fits our model exactly under the assumption that $f(h,a,\theta_l)$ is a constant and equal to p for all h and a. In general, one could think of $f(h,a,\theta)$ increasing in h and a for both values of θ (as long as the likelihood ratio assumption is maintained).

[19] $m(x_t, \mu_{t-1}, h_{t-1}, a_{t-1})$ is simply the application of Bayes rule. Formally,

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

and

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

into account the possible implications of her current choice of action on the likelihood of the future occurrences of cravings.

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 the (long run) benefit from not having the urge. B on the other hand represents the short run benefit from indulging to the urge net of health and monetary costs. For simplicity, we refer hereafter to B simply as the benefit/kick from consuming the substance under a craving.

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.[20] 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. 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

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

cost when the craving occurs next period, which occurs with (perceived) probability $(1-\mu_1)p+\mu_1 f(h_1,a_1)$.

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 consume, the rational consumer balances the current "kick" from satisfying the urge, B, versus the decrease in expected discounted future utility, $\delta D\mu_1\left[f(h_1,1)-f(h_1,0)\right]$, from exposing herself to a greater risk of having a craving next period.

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 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 refraining from 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.*

Inequality (2) confirms certain results of the received literature. First, suppose there is a reduction in $u(1,1)$ as a result of which B falls and D increases. In other words, the benefit to giving in to the urge is lower, while the benefit from not having an urge next period is higher. In that case, current consumption is indeed lower. Second, for any given beliefs, more addictive drugs (consider an increase in $f(h_1,1)-f(h_1,0)$ for any given h_1) are associated with lower consumption. Third, direct peer effects can reduce self-control (B increases); peer effects are discussed in more detail later. Finally, reducing the occurrence of cravings and exposure to cues (through change in habits/environment for example) is beneficial for any given priors due to a decrease in the true probability of $x_1 = 1$ (think of a reduction if p and f that does not reduce the terms $f(h,1)-f(h,0)$ for any h. There is thus no change in consumption decisions; however, the change in the parameter values are ex ante beneficial since the (true) chances of craving states are lower for any history).

It is important to note that similar to the standard rational addiction model, policy interventions such as forcing the consumer not to consume,[21] or raising taxes on the addictive good are not welfare enhancing in our setup in the absence of any consumption externalities and/or merit good arguments.

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 regarding the addictive properties of the compulsive consumption. For example, 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 $B < \delta D\left[f(h_1, 1) - f(h_1, 0)\right]$ 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.

Second, the prior μ_0 can be thought of arising as a result of the environment in which the decision-maker has been raised (representing 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 experience. Thus, for a 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 likelihood of smoking is relatively high. Recalling the alternative interpretation of our model as involving a decision-maker who does not know her individual physiological responses to consumption of an (known) addictive substance, the prior μ_0 can also encompass personal psychological traits. For example, many addicts seem to suffer from what appears to be an obvious overconfidence with respect to their ability to "kick the habit". A low μ_0 can reflect this overconfidence that then leads to experimenting (and possibly eventual addiction) with substances that are "generally" known to be addictive.

Third, the decision-maker in our model can exhibit behaviour that can be interpreted as failed attempt to quit. Suppose the prior μ_0 and (positive) past consumption is such that $[B/\delta D]/[f(h_1, 1) - f(h_1, 0)] \leq \mu_0 < \frac{\mu_0 f_0}{\mu_0 f_0 + (1-\mu_0)p}$. Then, the consumer chooses $a_1 = 0$ in period $t = 1$ but would choose $a_2 = 1$ in $t = 2$ if $x_2 = 1$. Of course, one might argue that this is an artifact of the two-period horizon of our basic model; however, as we show in more detail in the next section, the same insights carry over in a more general multi-period model. We should emphasize that the observed behaviour in this case is not due to dynamic inconsistency of preference or (cue-induced) mistake but as a result of experimentation and information acquisition as consumption and preference shocks vary over time.

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

Fourth, one needs to be careful about the role of quit aids and their welfare effects. Quit aids such as patches and inhalers are usually promoted as beneficial since they are supposed to reduce health costs - associated with inhaling tar and other additives in cigarettes - while providing smokers with the necessary nicotine intake to reduce withdrawal effects from not smoking. In addition, it is argued that this intake could be decreasing over time thus enabling smokers to get out of the habit eventually. In terms of our model, one can think of quit aids as a different substance with the following properties. First it gives a higher benefit from nicotine intake because it provides the kick without the health costs from using tobacco products. That is, quit aids are characterized by a higher B[22]. Second, the use of quit aids does not increase the value of h. That is, unlike smoking, using quit aid in period $t = 1$ is supposed to result in $h_2 \leq h_1$. However, it is not clear whether quit aids do indeed help kick the addiction by lowering the value of h. In the case when they do not help, not only is it possible that the consumer may start smoking again but also that she might simply get addicted to the quit aid product.

Fifth, substances with stronger withdrawal syndromes are associated with lower consumption. To see this, note first that self-control costs are represented by $u(0,0) - u(0,1)$ as this measures the welfare cost from having a craving given abstention. Note also that we can write $D = u(0,0) - u(0,1) - B$. Hence, an increase in $u(0,0) - u(0,1)$, other things remaining the same, will lower the chances of smoking.

Finally, our model can exhibit the phenomenon that more addictive drugs are associated with lower, rather than higher, consumption amongst the more experienced users. As mentioned before, a more addictive substance is characterized by a higher value of $f(h,1) - f(h,0)$ for any h. However, for any $h' > h''$, while both $f(h',1) - f(h',0)$ and $f(h'',1) - f(h'',0)$ are higher for a more addictive substance than for a less addictive one, it follows from assumption (7) that $f(h',1) - f(h',0) > f(h'',1) - f(h'',0)$. Hence the more experienced users (characterized by history h') may stop when the less experienced do not.

Note that even though our model involves experimentation and learning when μ_0 is different from 0 or 1, absence of learning does not change the essence of our main message insofar $\mu_0 < 1$. In fact, the case for information campaigns would be stronger if the true process was indeed the addictive one but the consumer's subjective prior puts $\mu_0 = 0$ (and hence consumes in the first period).

[22] Note thus that in a model with cumulative health costs as the one we discuss in Section 4, the use of quit aids would also lead to lower health costs in the future.

4 Extensions

In the previous sections, we have considered the 2-period model that illustrates most of the insights in the simplest setting. However, it is important to know whether these insights remain valid in a more general multi-period model. Furthermore, as mentioned earlier, certain properties of the model, for example, depicting behavior that seems like failed attempts to quit can be addressed more satisfactorily in a truly multi-period model. Finally, a general multi-period model may generate additional insights not present in the simpler 2-period one. Therefore, in this section, we extend the analysis to arbitrary $T > 2$ periods (where T can be finite or infinite). We also discuss the issue of cumulative welfare effects and end this section by showing how the present work relates to the intrinsic habit formation models. Before we do so, we discuss next some new implications of (a simple extension of) our model for peer effects.

4.1 Peer Effects

Our model features two kinds of peer effects:

(1) **Standard peer effect:** This can be modelled as follows. Whenever there are more people around the decision-maker that consume the substance, either there is an increase in $u(1, 1)$, resulting in an increase in B (and a decrease in D), or, there is an increase in the likelihoods $f(h_{t-1}, a_{t-1})$ and p (for any given history). The first depicts a direct increase in the cost of not consuming in the presence of the urge, the second reflects an increase in the desire to consume (by an increase in the likelihood of experiencing an urge). In either case, this is the more standard peer effect where an agent has a higher tendency to consume the addictive substance if more people around him also consume.

(2) **The informational peer effect:** There is a less obvious peer effect that can arise out of the process of information acquisition as a result of observing others' consumption practices. In particular, suppose there is a population of individuals and the true process is common for all. In such an environment observing others' consumption provides further information to the individual about addictiveness of the process[23]. Given our assumptions, Proposition 1 still holds. Hence for every individual, $a_t = 1$ implies that the individual has observed $x_t = 1$. However, when $a_t = 0$ it may still be the case that $x_t = 1$. Hence as in herd behavior models(e.g.

[23] To simplify the exposition, we assume that the processes of the individuals are in effect, *perfectly correlated*. It is clear that we do not need perfect correlation; rather what is needed is that others' consumption provides some valuable information regarding one's *own* process.

Bikhchandani et al. 1992, and Banerjee 1992), actions will not perfectly reveal the state/signal x, and in particular an informational cascade can arise for consumers who face $x_t = 1$. 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 informational peer group effect, however, we can get *contrarian* behavior due to informational cascades. Our model can be used to provide some micro-foundations for extrinsic habit formation (ie. for consumption history of peers to affect one's own tendency to consume); however, as the previous discussion showed, the exact nature of the extrinsic habit formation model can be (subtly) different from what one might expect at first.

4.2 Longer Horizon

We start by describing direct extensions of corresponding terms and definitions of the two-period model; naturally, we also need to introduce some new ones. Throughout, we suppress the explicit dependence on the horizon T in our notation to simplify exposition.

For any period $t \geq 1$, 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.[24] We also assume that there exists some $H > 0$ such that $h_t \in [0, H]$, for all t.[25]

For all the variables, we will sometimes use them without the time subscript (for example, write h, or a) to denote values. For example, h refers to some scalar in $[0, H]$, a refers to the number 0 or 1 etc.

$f(h_t, a_t)$ denotes the probability that $x_{t+1} = 1$ under the addictive process. We retain all the assumptions from the 2-period model about the function f. As before, μ_{t-1} and μ_t denote period-t prior and posterior beliefs, respectively, that the true

[24] The fact that g is increasing in h also differentiates our model from that in GP.

[25] This means that the function g must be such that for any h_{t-1} and a_{t-1}, $g(h_{t-1}, a_{t-1}) \in [0, H]$.

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; given that there is no additional information that is relevant for inferences regarding the true addictive process, the period $t+1$ prior is simply the period t posterior. We often write $m_t(x_t)$ to denote the period$-t$ posterior for given craving state x_t (and suppress the variables μ_{t-1}, h_{t-1} and a_{t-1} when it is not important to mention the dependence of μ_t on these).

Note that as long as $1 > \mu_{t-1} > 0$ and it is not true that $h_{t-1} = a_{t-1} = 0$, Assumption (5) implies that the period-t posterior belief μ_t is increasing in x (otherwise is independent of x_t) and[26] that $m(1, \mu_{t-1}, h_{t-1}, a_{t-1})$ is increasing, while $m(0, \mu_{t-1}, h_{t-1}, a_{t-1})$ is decreasing, in h_{t-1} a_{t-1}. Finally, as long as $h_{t-1} = a_{t-1} = 0$ is not true, the period$-t$ posterior is increasing in the prior μ_{t-1}.

It is helpful to define here the posterior probability of observing $x_{t+1} = 1$ given h_t, a_t and prior μ_t. Let this posterior probability be denoted by $\pi(a_t, h_t, \mu_t)$ 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, denoted by $a(h_t, x_t, \mu_t)$, depends on period-t history h_t, posterior μ_t and observed state x_t. When T is finite, it also depends on the horizon T. A finite period problem is essentially nonstationary; however, to avoid cluttering we drop the time subscript t and write $a(.)$ rather than $a_t(.)$. And the same applies for the functions Z and V described later. We show later stationarity of these functions for the infinite-horizon case.

The consumer's objective is to choose the consumption plan to maximize

$$\mathbb{E}\left\{\sum_{t=1}^{T} \delta^{t-1} u(a_t, x_t)\right\}$$

where $\mathbb{E}$ denotes the expectations operator conditional on x_1, h_1 and $\mu_1(x_1)$.

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,

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

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

The value function V is given by:

$$V(x_t, h_t, \mu_t) \equiv Z(a(h_t, x_t, \mu_t), x_t, h_t, \mu_t)$$

In the finite–horizon case, this can be derived by using backward induction. We can not rely on backward induction when dealing with the infinite-horizon case. Nevertheless, using dynamic programming techniques we can show that the value function exists and is well-defined even when $T \to \infty$. This is done in Section 4.2.2. below.

Bearing the above in mind, we first show, similar to what we had earlier, that under certain conditions the consumer can choose $a_t = 0$ even if $x_t = 1$, for any $T > 2$. Towards that end, we define the following terms first.

Let $\widehat{D}_{t+1}(h_t, \mu_t)$ be given by

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

That is, for any h_t and μ_t, and given that the consumer had decided to smoke in period t (i.e., $a_t = 1$) the term $\widehat{D}_{t+1}$ measures the gain - evaluated at $t+1$ - from not having an urge in period $t+1$.

Similarly, we define the terms Γ_{t+1} and Ψ_{t+1} as:

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

and

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

$\Gamma_{t+1}(h_t, \mu_t, x)$ represents the gain - evaluated at $t+1$ - from not smoking in the current period and optimally responding to next period's craving-state conditional on the latter being x. The interpretation of the term $\Psi_{t+1}(h_t, \mu_t)$ is more subtle. State in period $t+1$ is not known in period t. Abstaining from smoking in period t has two effects: it changes the likelihood of a craving next period, and it changes the next-period's value function conditional on next-period's craving state. Maintaining the likelihood of craving at the level corresponding to non-consumption in period t, $\Psi_{t+1}(h_t, \mu_t)$ measures the ex ante - i.e., evaluated at t - expected gain from the change in the next-period's value function.

We are now ready to show:

Proposition 2 *Suppose $x_t = 1$. Then, $a_t = 0$ is the optimal action in period t if and only if the following inequality holds:*

$$B \le \delta\{m_t(1)\,[f(h_t, 1) - f(h_t, 0)]\,\widehat{D}_{t+1}(h_t, \mu_t) + \Psi_{t+1}(h_t, \mu_t)\} \tag{3}$$

Proof. Payoff from action a_t in period t (when $x_t = 1$) is given by

$$\begin{aligned}
&u(a_t, 1) + \delta V(0, g(h_t, a_t), m(0, \mu_t, h_t, a_t)) \\
&-\delta\pi(a_t, h_t, \mu_t)\left[V(0, g(h_t, a_t), m(0, \mu_t, h_t, a_t)) - V(1, g(h_t, a_t), m(1, \mu_t, h_t, a_t))\right]
\end{aligned} \tag{4}$$

Using expression (4), we see that when $x_t = 1$, action $a_t = 0$ is preferred to $a_t = 1$ when

$$\begin{aligned}
&u(0, 1) + \delta V(0, g(h_t, 0), m(0, \mu_t, h_t, 0)) \\
&-\delta\pi(0, h_t, \mu_t)\left[V(0, g(h_t, 0), m(0, \mu_t, h_t, 0)) - V(1, g(h_t, 0), m(1, \mu_t, h_t, 0))\right] \\
\geq\; \\
&u(1, 1) + \delta V(0, g(h_t, 1), m(0, \mu_t, h_t, 1)) \\
&-\delta\pi(1, h_t, \mu_t)\left[V(0, g(h_t, 1), m(0, \mu_t, h_t, 1)) - V(1, g(h_t, 1), m(1, \mu_t, h_t, 1))\right]
\end{aligned}$$

Some straightforward algebraic manipulations, using the definitions of $\widehat{D}_{t+1}(h_t, \mu_t)$ and $\Psi_{t+1}(h_t, \mu_t)$, and noting that $\pi(1, h_t, \mu_t) - \pi(0, h_t, \mu_t) = m_t(1)\left[f(h_t, 1) - f(h_t, 0)\right]$ gives the desired inequality. ■

The above result follows simply from the definitions of the various terms and it the natural extension of the decision rule (2) in the current multi-period model. Since a longer horizon implies that there are more periods in which undesirable cravings can occur, it is natural to expect that, other things remaining the same, a longer horizon would reduce the temptation to smoke. In other words, we would expect $\Psi_{t+1}(h_t, \mu_t) \geq 0$ and $\widehat{D}_{t+1}(h_t, \mu_t) \geq D$, given that in the 2-period model we have $\Gamma_2(h_1, \mu_1, 1) = \Gamma_2(h_1, \mu_1, 1) = \Psi_2(h_1, \mu_1) = 0$ and $\widehat{D}_2(h_1, \mu_1) = D$. However, this is not necessarily true in the absence of further structure. In addition, there are certain properties of the two-period model, described below, that are intuitive and yet cannot extend directly to the multi-period set up without further structure.

- Since future cravings depend on past and current consumption, and since craving states are undesirable, one expects welfare (as reflected in the value function) to be adversely affected by the presence of future craving, or higher past and current consumption, or higher posterior. For the 2-period model, note that since $V(x_2, h_2, \mu_2) = u(x_2, x_2)$, and given Assumption (3), the future value function $V(x_2, h_2, \mu_2)$ is nonincreasing in the craving-state, x_2, history, h_2, and posterior, μ_2.

- Similarly, we would expect higher current and past consumption to be detrimental for future welfare even when there is no craving in the future (and the decision-maker responding optimally). Indeed, in the 2-period model we have that $V(0, g(h_1, a_1), m(0, \mu_1, h_1, a_1))$ is nonincreasing in h_1 and a_1 since $V(0, g(h_1, a_1), m(0, \mu_1, h_1, a_1)) = u(0, 0)$.

- At the optimum, the cost of having a craving state next period should be at least as large as equal to D, since cravings make consumption more likely and because higher consumption makes cravings in future periods more likely. Note that in the 2-period model, and recalling the definition of D, we have $V(0, g(h_1, a_1), m(0, \mu_1, h_1, a_1)) - V(1, g(h_1, a_1), m(1, \mu_1, h_1, a_1)) \geq D$ because of $V(x, g(h_1, a_1), m(x, \mu_1, h_1, a_1)) = u(x, x)$.

- We expect the optimal action to be abstaining from consumption in the absence of any craving; recall Proposition (1).

The natural extensions of the above properties in the model with $T > 2$ are that for any $t < T$:
(A) $V(x_{t+1}, h_{t+1}, \mu_{t+1})$ is nonincreasing,
(B) $V(0, g(h_t, a_t), m(0, \mu_t, h_t, a_t))$ is nonincreasing in h_t and a_t,
(C) $V(0, g(h_t, a_t), m(0, \mu_t, h_t, a_t)) - V(1, g(h_t, a_t), m(1, \mu_t, h_t, a_t)) \geq D$,
(D) $a(h_{t+1}, 0, m_{t+1}(0)) = 0$.
Note that when $T > 2$ Property (B) is not implied by Property (A) due to $m(0, \mu_t, h_t, a_t)$ being decreasing in h_t and a_t if $1 > \mu_t > 0$. Moreover, Property (C) does not follow from Property (A) directly despite the fact that $m(0, \mu_t, h_t, a_t) < m(1, \mu_t, h_t, a_t)$. Note that the four properties (A)-(D) are all true for $t = T - 1$, when $T > 2$ and finite, for similar reasons with the 2-period model. Nevertheless, even though the properties mentioned above are all intuitive, we require additional assumptions on the f and g function for them to hold in the general multi-period model. Since the proofs are somewhat lengthy we have relegated them to appendices. Here, we provide an informal discussion on how we obtain our results. We also discuss the basic intuition as to why these additional structures on the functions f and g are needed.

Proving Properties (A)-(D) under additional assumptions on fundamentals follows an inductive argument. First, we show in Appendix B the following Lemma:

Lemma 3 *If properties (A) and (B) are true for $t \geq n$, $n \geq 2$, then properties (C) and (D) are true for $t = n - 1$.*

That is, if the value function in all periods from n (inclusive) and onwards, with $n \geq 2$, satisfies Properties (A) and (B), then Property (C) holds for period $n - 1$. Furthermore, the optimal action in period n is to abstain from consumption in the absence of any craving. Thus, it suffices to show Properties (A) and (B) for any $t \geq 0$.

One can show that property (A) is true for period $t = T - 2$ with $T > 2$ and finite; the reader is referred to Appendix A for details. It can also be shown that, with $T > 2$ and finite, property (B) holds for period $t = T-2$ if $\pi(0, g(h,a), m(0,\mu,h,a))$ is nondecreasing in h and a.[27] However, monotonicity of $\pi(0, g(h,a), m(0,\mu,h,a))$ with respect to h and a depends on the properties of the g and f functions and is not guaranteed by the assumptions made so far. Furthermore, conditional on $\pi(0, g(h,a), m(0,\mu,h,a))$ being nondecreasing in h and a, property (A) is guaranteed to hold for any period $t \leq T-3$ when $T > 3$ and finite, if $m(1, m(0,\mu,h,a), g(h,a), 0)$ is nondecreasing in h and a. However, the latter can not be guaranteed with the assumptions made so far: further restrictions on f and g would be required. Similarly, for Properties (A) and (B) with $1 \leq t < T$ when $T \to \infty$.

Before we go on to state the restrictions we make on the f and g functions, it is useful to note first the reason why these are needed for the general multiperiod model. Here we provide the rough intuition, for details please refer to the Appendices. In the 2-period model, since the second period is the last period, the value function in the second period does not depend on the posterior or the history. This of course is not true for arbitrary periods of the general multiperiod model. In what follows, consider a period t other than the last or the second last period. The current value function depends on the current posterior through its effect on the probability of craving next period as well as through its effect on the value function next period. Now, the current posterior, in the face of a current craving state, is non-increasing in past consumption; further the likelihood of a craving next period is also higher in current and past consumption. However, while next period's posterior is higher in past consumption if next period's state is the craving state, it is actually lower if the state happens to be the non-craving state. Hence to obtain the desirable properties on the value function - Properties (A) and (B) (and thereby Properties (C) and (D)) - some additional structure in the form of restrictions on the fuctions f and g are needed.

To describe the restrictions we put on the f and g functions, it is useful to define the operator L. Given functions $g(h,a)$ and $m(0,\mu,h,a)$, the operator L is given as

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

[27] To see this, note, similar to the 2-period model, that $V(0, g(h_{T-2}, a_{T-2}), m(0, \mu_{T-2}, h_{T-2}, a_{T-2})) = u(0,0) + \delta u(0,0) - \delta\pi(0, g(h_{T-2}, a_{T-2}), m(0, \mu_{T-2}, h_{T-2}, a_{T-2}))D$.

Further, we use the convention $L^0 g(h,a) = g(h,a)$, $L^0 m(0,\mu,h,a) = m(0,\mu,h,a)$, and let $L^t g(h,a) = g(L^{t-1}g(h,a),0)$ and $L^t m(0,\mu,h,a) = m(0, L^{t-1}m(0,\mu,h,a), L^{t-1}g(h,a),0)$. Thus, L maps posterior under an urge, $m(0,\mu,h,a)$, history $g(h,a)$ and current abstention into next period's history and posterior under no craving.

We then make the following assumption (by following the convention that $\prod_{i=2}^{1} y \equiv 1$):

Assumption 8 *For any $t \geq 1$, the functions f and g are such that if $h'' \geq h'$ and $a'' \geq a'$ (with at least one inequality being strict) then*

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

We show in Appendix C that the set of functions g and f that satisfy Assumption (8) is non-empty. To get a feel for how strong the assumption is we also note that it is consistent with the functions used in the literature (BM, OZ, GK, L and BR).

We also show in Appendix C that the above assumption implies that for any $t \geq 1$

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

(ii) $m(1, L^{t-1}m(0,\mu,h,a), L^{t-1}g(h,a),0)$ is nondecreasing in h and a.

To understand these, consider the case when there is no craving, posterior is $m(0,\mu,h,a)$, history is $g(h,a)$ and consumption does not take place. Conclusion (i) then says that the likelihood of future urges is nondecreasing in h and a even in states when the urge is not present and consumption does not take place. Conclusion (ii) guarantees that after a string of no cravings and zero consumptions followed eventually by a craving, higher past consumption, h and a, will lead to (weakly) higher posterior. Note that these two Conclusion imply (just set $t = 1$) the monotonicity properties discussed above.

We can now discuss how our insights from the two-period model carry over to the extension with $T > 2$ periods. Note here that despite the fact that properties (A)-(C) and Assumption (8) are important in the forthcoming discussion *for any $T > 2$*, finite or infinite, the steps we will follow depend on whether T is finite or infinite. The reason is that in the former case we will be using backward induction, while in the latter we will be using dynamic programming techniques to show existence of the value function and certain limiting arguments to determine its properties.

4.2.1 The Finite Horizon Problem

We first consider the case when T is finite. We can then get the following result:

Lemma 4 *Assume 8. Then properties (A)-(D) hold for any* $0 \leq t \leq T-1$.

Proof. We show in Appendix B that under Assumption (8), if properties (A) and (B) are true for any $t \geq n$, then they are also true for $t = n-1$, for any given $n \geq 2$. The result is then obtained by using Lemma 3 and using an induction argument after recalling that properties (A)-(D) hold in the last period. ■

Before we go on to the infinite horizon model, we end this (sub)section by discussing how our model can feature failed attempt to quit.

Behaviour exhibiting Recidivism:

Consider the case where the outside observer observes the value of x_t (for example, because x_t represents situations or external cues, such that it is "known" that the decision-maker becomes more tempted to consume in those situations). In this case, certain behavior on the part of the agent will look like an attempt to resist temptation for some periods before "giving in" to it later.[28] To illustrate this in the simplest possible way, suppose that we are in period $T-3$ and $x_{T-3} = 1$. Suppose also that the decision-maker has consumed the substance in the past and, more specifically, $h_{T-3} > 0$ is such that

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

Hence,[29] the decision maker chooses $a_{T-3} = 0$. The outside observer may naturally then interpret this as the decision-maker's attempt to quit.

Let's now consider the path (of realizations of x_t) such that $x_{T-2} = 0$ and $x_{T-1} = 1$. We know from the above Lemma that under Assumption (8) we have $a_{T-2} = 0$, but what about consumption in period $T-1$? Since T is the terminal period, by using reasoning similar to the one in the two-period model we have that the decision-maker chooses $a_{T-1} = 1$ if (and only if)

[28] The case where the outside observer does not observe the craving state of the decision-maker is not interesting. The reason is that even if all the decision maker does is to choose $a_t = x_t$ (something she would do, for example, if she believed that the process was certainly non-addictive ($\mu_0 = 0$)), to the outside observer however, since x_t is stochastic, this might seem like the the decision-maker relapsing into consuming the substance after stopping for some periods.

[29] Note here that the particular rule for resolving indifference we have deployed, namely that an indifferent decision-maker does not smoke, is not important for our argument here though it simplifies exposition. By continuity, we could develop the same type of argument even if B was ε less that the right-hand side of the above equation with ε very small but positive scalar.

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

If $a_{t-1} = 1$, then the outside observer may naturally interpret this as if the decision-maker relapses into consuming the substance after stopping for one (or more) periods, that is, as a "failed attempt to quit".

To see under what conditions this can happen, focus on an environment with (weak) depreciation of the stock of addiction: $g(h,0) \leq h$.Note thus that if $a_{T-3} = a_{T-2} = 0$, then $h_{T-1} \leq h_{T-3}$ and $f(h_{T-1},1) - f(h_{T-1},0) \leq f(h_{T-3},1) - f(h_{T-3},0)$, where the last inequality follows from Assumption (7)[30]. 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 $a_{T-1} = 1$ is therefore that $m_{T-1}(1)$ is strictly less than $m_{T-3}(1)$. Let $f_{T-2} = f(h_{T-3},0)$ and $f_{T-1} = f(g(h_{T-3},0),0)$, to ease notation, and note that

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

the sign of which depends on the sign of $(1-f_{T-2})f_{T-1} - (1-p)p$. Now since $f_{T-1} \leq f_{T-2}$, we have

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

Thus, for $m_{T-1}(1) < m_{T-3}(1)$, it is sufficient to have $(1-f_{T-2})f_{T-2} - (1-p)p < 0$. For example, suppose $p = 1/2$, 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 $m_{T-1}(1) < m_{T-3}(1)$.

[30] For the parametric example considered in Appendix C 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.

Accordingly, if $p = 1/2$ and $f(h, a)$ is supermodular then we can have rational and dynamically consistent behavior that looks to an outside observer like a failed attempt to quit, while it is only part of experimentation via consumption when the craving state varies over time.

4.2.2 Infinite Horizon

Turning to the case of $T \to \infty$, note first that taking x_t and h_t to be the two state variables, the (subjective) transition probability function (of x_{t+1}) is non-stationary. This can be "handled" by treating μ_t to be a state variable also, in which case, the transition probability π_t becomes stationary. However, this comes at the "cost" of having a state variable, μ_t, with no clear-cut monotonicity property regarding its law of motion.

The first issue that arises then is whether the value function $V(x, h, \mu)$ is well-defined. The second is whether the value function satisfies properties (A) and (B) - in which case, by Lemma 3, it will also satisfy properties (C) and (D): that is, Lemma 4 will hold in this case also. If these are still true, our results will thus be robust to allowing for infinite horizon.

While one can show, using standard dynamic programming techniques, that $V(x, h, \mu)$ is indeed well-defined, continuous and bounded, the rest of the task requires somewhat more involved analysis. In fact, a technical contribution of this paper is to derive under Assumption (8) some new results on the monotonicity properties of the value function of a stochastic dynamic programming problem. The details are in Appendix D, here we provide a brief intuitive discussion.

The basic idea behind our proof is as follows. We first define a non-empty space of functions, S^*, with the domain of the functions being the state variables, x, h and μ, and where each function $\gamma(x, h, \mu) \in S^*$, satisfies the desirable monotonicity properties (A) and (B).[31] Now, (similar to standard dynamic programming techniques), we construct a contraction mapping that maps functions from S^* to S^* such that the fixed point of this mapping, if it exists, satisfies the Bellman equation. Thus, if S^* is a closed set, we know (see for instance Stockey and Lucas, 1989) that the fixed point of this mapping exists and is unique. Accordingly, it defines a value function with the desired properties. To show that S^* is a closed set we follow an iterative procedure. We start with S, the space of all continuous bounded functions (over the domain x, h and μ) and consider first the subset, $S^{0,d}$, such that a function $\gamma \in S^{0,d}$ if it is also nonincreasing in the variables x, h and μ. We start by noting that $S^{0,d}$ is a closed subset of S. We then apply the first iteration of the operator L on the func-

[31] S^* is non-empty as clearly the constant function belongs to this set.

tions in $S^{0,d}$ (for more details on the exact procedure, see Appendix D) to generate a subset $S^{1,d}$ of $S^{0,d}$ such that these (once) iterated functions are also all nonincreasing. The set of functions $S^{1,d}$ is now shown to be closed. Repeated application of this procedure produces a sequence of sets of functions $S^{k,d}$, for $k = 1, 2, ...$, such that each $S^{k,d}$ is a subset of $S^{k-1,d}$, is a closed set, and $S^* = \cap_k S^{k,d}$. The proof is then concluded by noting that countable intersection of closed sets is closed.

4.3 Cumulative Welfare Effects and (Intrinsic) Habit Formation

As noted before, we have deliberately excluded any cumulative welfare effects in our model so far to highlight the fact that our results do not depend on any such effect. However, in many applications, such effects are present and are important (for example long term health effects of smoking or long term wealth effects of gambling etc). Here, we show how we can modify our basic model to allow for such environments; we also show that the presence of such cumulative effects does not affect our results qualitatively. Finally, we also explore the link between our model and those in the intrinsic habit formation models.

To handle the presence of cumulative welfare effect, we change the utility function from being $u(a_t, x_t)$ to being $v(a_t, x_t, h_t)$. The (harmful) cumulative effect is captured by the assumption that $v(a_t, x_t, h_t)$ is nonincreasing in h_t (and to rule out the trivial case, is strictly decreasing for at least some values of h). To maintain the property of compulsive consumption, we also assume that $v(1, 1, h_t) > v(0, 1, h_t)$, $v(0, 0, h_t) > v(1, 0, h_t)$ and that $v(0, 0, h_t) > v(1, 1, h_t)$ and $v(a_t, x_t, h_t)$ being nonincreasing in h_t. Let us also define $D(h) = v(0, 0, h) - v(1, 1, h)$ and $B(h) = v(1, 1, h) - v(0, 1, h)$, to replace D and B from before.

By repeating the steps in Appendix B, one can easily show that Lemma 3 is still valid. As we also show in Appendix D, the value function is still well-defined[32]. Moreover, most of our results,[33] and in particular Proposition (2) and Lemma 4 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 habit formation models assume, in general, an intertemporal utility

[32] We focus on the infinite horizon case to facilitate comparison with most of the received literature on rational addiction; however, our discussion carries over to the case of finite horizon.

[33] The only results that would require, in order to still hold in this case, further assumptions are the ones that have to do with the comparison of the behavior of decision-makers with different histories. Specifically, these additional assumption will have to be on the monotonicity with respect to history of $B(h)/D(h)$.

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 are used to explain phenomena that standard intertemporally separable preferences seemingly cannot. Some such examples are: 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 strongly connected 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 and in fact until recently there have been no theoretical underpinnings of the habit formation preferences.[34] Rozen (2009), axiomatizes the so-called linear habit formation model used in some of the papers above. Rustichini and Siconolfi (2005), axiomatize dynamically consistent habit formation over consumption streams, (but do not offer a particular structure for the utility or form of habit aggregation). Gul and Pesendorfer (2007), who also axiomatize a dynamically consistent non-linear habit formation model but by considering preferences on menus of streams of consumption rather than on streams themselves. In contrast 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 while our decision-maker has preferences characterized by standard utility function, she lacks information about the determination of the state of the world (the "urge" to consume in the future), with the (perceived) stochastic process generating the states depending on an endogenous variable (past consumption). To be more specific, revert to the model with no cumulative welfare effects for simplicity of exposition (so that utility is given by $u(a,x)$). Note then that at time t the payoff is given by $u(a_t,x_t)+\sum_{i=1}^{\infty}\delta^i E_t[u(a_{t+i},X_{t+i}) \mid h_{t-1},a_{t-1},\mu_{t-1},a^{t+i-1},x_t]$, 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}, for any $i\geq 1$, given the $t-$period prior μ_{t-1}, past history h_{t-1} and consumption a_{t-1} (and hence h_t) and observed state x_t (and hence μ_t), and the consumption stream a^{t+i-1} (which will determine the history stream $(h_{t+1},h_{t+2},...,h_{t+i})$. Importantly, the perceived probability distribution over X_{t+i}, for all $i=1,...,T-t$) is $p+m(x_t,\mu_{t-1},h_{t-1},a_{t-1})[f(h_{t+i-1},a_{t+i-1})-p]$, which is dependent on history of past consumption. Hence, the *expected* utility, in our framework, falls under the rubric of *nonlinear* habit formation models.

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

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 in that it is influenced by stochastic urges. Second, cravings depend 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, the consumer is not fully aware of how easy or difficult it might be to quit since 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 micro-foundations for habit-forming behavior by starting from a standard model of a fully rational decision-maker with intertemporally separable preferences, but with subjective uncertainty over the likelihood of future temptations that depends on past consumption behavior in tempting circumstances. In our model, 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.

HALEVY, Y. (2008), "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect", American Economic Review, 98:3, 1145-1162

HARVEY, C.M., (1986) "Value Functions for Infinite-period Planning", Management Science, 32(9), 1123-1139

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 $t+1 = T-1$

Proof. Let $(x''_{T-1}, h''_{T-1}, \mu''_{T-1}) \geq (x'_{T-1}, h'_{T-1}, \mu'_{T-1})$ with at least one inequality being strict, and $a'_{T-1} = a(h'_{T-1}, x'_{T-1}, \mu'_{T-1})$ and $a''_{T-1} = a(h''_{T-1}, x''_{T-1}, \mu''_{T-1})$. By the definition of optimum, 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 similar to the two-period model that

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

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

8 Appendix B

Here we prove Lemma 3 and the inductive part of Lemma 4
Proof of Lemma 3. **We start with proving property (D) for** $t = n - 1$

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

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

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

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

where $h'' = g(h_n, 1)$. Note that $\pi(0, h_n, \mu_n(0)) \leq \pi(1, h_n, \mu_n(0))$.

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

$$\begin{aligned} & V(0, h_n, m_n(0)) \\ = \ & u(0,0) + \delta E_{\pi(0,h_n,m_n(0))} V(X, h', m(X, m_n(0), h_n, 0)) \end{aligned}$$

and similarly

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

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 $E_{\pi(1,h_n,m_n(1))} V(X, h'', m(X, m_n(1), h_n, 1))$

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

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

■

Proof of inductive part of Lemma 4. Assume that properties (A) and (B) hold for $t = n$.

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

Let $(x''_n, h''_n, \mu''_n) \geq (x'_n, h'_n, \mu'_n)$ (with at least one inequality being strict) 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 $x''_n = x'_n \equiv x_n$

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

Note now that $V(x_n, h_n'', \mu_n'') = u(a_n'', x_n) + \delta E_{\pi(a_n'', h_n'', \mu_n'')} V(X, g(h_n'', a_n''), m(X, \mu_n'', h_n'', a_n''))$. Thus, $V(x_n, h_n'', \mu_n'') \leq V(x_n, h_n', \mu_n')$ follows directly if

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

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

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

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 3 that

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

Note now that $V(1, h_n, \mu_n) = u(a_n'', 1) + E_{\pi(a_n'', h_n, \mu_n)} V(X, g(h_n, a_n''), m(X, \mu_n, h_n, a_n''))$ 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 $E_{\pi(a_n'', h_n, \mu_n)} V(X, g(h_n, a_n''), m(X, \mu_n, h_n, a_n'')) \leq$

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

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:

$V(0, L^j g(h, a), L^j m(0, \mu_{n-1}, h, a))$ 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 $V(0, L^{T-n}g(h,a), L^{T-n}m(0,\mu_{n-1},h,a)) = u(0,0)$ (recall 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 3, that

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

Recall that $L^i g(h,a) = g(L^{i-1}g(h,a),0)$ and note that $L^{i-1}g(h,a)$ is increasing in h and a. Recall also that $m(0, L^{i-1}m(0,\mu_{n-1},h,a), L^{i-1}g(h,a),0) = L^i m(0,\mu_{n-1},h,a)$. Note that $L^{i-1}g(h,a)$ and $L^{i-1}m(0,\mu_{n-1},h,a)$ 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 that (as we show in Appendix C) Assumption (8) implies that

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

(ii) $m(1, L^{i-1}m(0,\mu_{n-1},h,a), L^{i-1}g(h,a),0)$ 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 following the convention that $\prod_{i=2}^{1} y \equiv 1$):

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

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

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

has the sign of

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

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

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

which can be rewritten as

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

and hence, by iterating backwards, as

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

This is true of $\mu=0$ 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 $L^{t-1}m(0,\mu,h,a)\geq L^{t}m(0,\mu,h,a)$. Note also that $L^{t}m(0,\mu,h,a)$ is nonincreasing in h and a. We have, for any $t\leq T-j-1:$

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

Thus,

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

has the sign of

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

The sign of this is non-positive if

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

or, equivalently, if

$$\frac{L^{t-2}m(0,\mu,h',a')}{(1-L^{t-2}m(0,\mu,h',a'))}\frac{(1-f(L^{t-2}g(h',a'),0))}{(1-f(L^{t-2}g(h'',a''),0))}\frac{f(L^{t-1}g(h',a'),0)}{f(L^{t-1}g(h'',a''),0)}$$
$$\leq \frac{L^{t-2}m(0,\mu,h'',a'')}{(1-L^{t-2}m(0,\mu,h'',a''))}$$

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

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

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 is (6).

In what follows we consider in more detail how restrictive this is and how all this relates to the functions that appear in the existing literature. The reader more interested in pursuing the main arguments behind our results could skip the rest of this section.

To see how restrictive (6) is and how it relates to functions used in 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).[35] This encompasses the law of motions (conditional on history h be bounded and $\rho(h)$ being increasing) in BM and OZ (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

[35] 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 $T > 2$.

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 $\rho_2 = 1$).

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").[36] 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, OZ, 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, $L^j g(h,a) = c(0)\sum\limits_{i=0}^{j-1}\rho_1^i + c(a)\rho_1^j + \rho_1^{j+1}h + \rho_1^j\rho_2 a$. Thus, $f(L^j g(h,a),0) = p + c(0)\sum\limits_{i=0}^{j}\rho_1^i + c(a)\rho_1^{j+1} + \rho_1^{j+2}h + \rho_1^{j+1}\rho_2 a$.

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

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

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

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 if

[36] Recall that in BM, OZ, 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.

$a' < a''$ 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

Before we proceed to the proofs, we should emphasize that, as noted in Section (4.3), we can easily accommodate here the extension where temporal payoff depends also on past consumption, by letting the per period utility function to be $v(a_t, x_t, h_t)$, $t = 1, ..., \infty$. In that case, we assume that v is a bounded and continuous function on $\{0,1\} \times \{0,1\} \times [0, H]$ and have the properties: $v(0,0,h) > v(1,0,h)$, $v(0,0,h) > v(1,1,h) > v(0,1,h)$ and $v(a,x,h)$ being nonincreasing in h.

Next we provide a brief intuitive discussion, in terms of dynamic programming theory, as to why we cannot use established results on properties of value functions for what we need to show. The reader more interested in pursuing the main arguments behind our results could skip the rest and go directly to the next (sub)section.

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 increasing.

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

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

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

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

Despite the seemingly nonstationary nature of the probability measure over the stochastic state x, one can re-write it in such 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

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

Hence, by including the posterior, μ_t, in the set of state/predetermined variables the above probability becomes stationary. However this comes at the expense of having a state variable, μ_t, with a law of motion that has no clear-cut monotonicity properties. 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 nonincreasing 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

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

Clearly, M is nondecreasing in μ_{t-1}. 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 M requires more structure into the problem (ie. into 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 (8).

We are now ready to proceed with the rest of the proof.

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

The following proof uses standard dynamic programming techniques (see, for instance, Stockey and Lucas (1989)) to show that the Bellman equation is well-defined and that the value function V is given by the unique mapping N that is uniformly contracting on the set S of continuous and bounded functions of x, h and μ, with $x \in \{0,1\}$, $h \in [0,H]$ and $\mu \in [0,1]$. Readers who are familiar with such methods can skip the rest of this section and go directly to the next (sub)section.

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 $\{x_t, h_t, \mu_t\}$, the Bellman equation is

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

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

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

Note that if $\gamma \in S$, then $N\gamma$ is also bounded and continuous by Berge's theorem of maximum. Thus, the above defines a mapping N from the set S into itself. Moreover the set S with the sup norm, $\|\gamma\| = \sup_{x,h,\mu} |\gamma(x,h,\mu)|$ 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 $N : S \to S$ is a contraction mapping with modulus

β, then (a) N has exactly one fixed point, call it V, in S, and (b) for any $V_0 \in S$, $\lambda(N^n V_0, V) \leq \beta^n \lambda(V_0, V)$, $n = 0, 1, 2, ...$ To show then that the Bellman equation above is uniquely defined, we only have to show that N is a contraction mapping with modulus β (i.e. that for some $\beta \in (0,1)$, $\lambda(Nz, Ny) \leq \beta\lambda(z, y)$ 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 $Q \subseteq R^3$, and let $B(Q)$ be a space of bounded functions $V : Q \to R$ with the sup norm. Let $N : B(Q) \to B(Q)$ be an operator satisfying (a) monotonicity: $z, y \in B(Q)$ and $z(q) \leq y(q)$, for all $q \in Q$, implies $(Nz)(q) \leq (Ny)(q)$, for all $q \in Q$, (b) discounting: there exists some $\beta \in (0,1)$ such that $(N(f + d))(q) \leq (Nf)(q) + \beta d$, for all $f \in B(X)$, $d \geq 0$, $q \in Q$, where $(f + d)(q)$ is the function defined by $(f + d)(q) = f(q) + d$. Then, N is a contraction mapping with modulus β.

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

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

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 $(Ny)(x, h, \mu) \geq w(a_z, x, h, \mu; y) \geq w(a_z, x, h, \mu; z) = (Nz)(x, h, \mu)$. The discounting requirement is also trivially satisfied as $(N(V + d))(x, h, \mu) = (NV)(x, h, \mu) + \delta d$. Therefore, the mapping $N : S \to S$ is a contraction mapping with modulus δ. Hence, the Bellman equation and the value function it defines implicitly are well-defined, and N is uniformly contracting.

10.2 Properties of V

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

To this end, and recalling the definition of the operator L on functions g and μ, let S^* be the subset of S such that all functions γ in $S^*$are nonincreasing in x, h, μ and satisfy the following property:

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

S^* is nonempty (as it includes all constant functions) and notably all functions in S^* satisfy properties (A) and (B) (the latter follows by setting $j = 0$). Thus, it suffices to show that $V \in S^*$.

Given that N is uniformly contracting on the complete space S (endowed with the sup norm), we have that if $N : S^* \to S^*$ and S^* is closed, then the unique value function defined by the Bellman equation lies in S^*; see for instance Stockey and Lucas (1989). We show these two properties in turn.

10.2.1 $N : S^* \to S^*$

Here we show that $N : S^* \to S^*$.

Consider $\gamma \in S^*$ and let $(x'', h'', \mu'') \geq (x', h', \mu')$ with at least one inequality being strict. Note that the (assumed) monotonicity properties of $\pi(a, h, \mu)$, $m(x, \mu, h, a)$, $\pi(0, L^j g(h, a), L^j m(0, \mu_{n-1}, h, a))$, $m(1, L^j m(0, \mu_{n-1}, h, a), L^j g(h, a), 0)$ and $L^j m(0, \mu_{n-1}, h, a), L^j g(h, a)$, for all $j \geq 0$ and $n \geq 1$, imply that

$$\begin{aligned}
E_{\pi(a,h,\mu')}[\gamma(X_n, g(h, a), m(X_n, \mu', h, a)] &\geq E_{\pi(a,h,\mu'')}[\gamma(X_n, g(h, a), m(X_n, \mu'', h, a))] \\
E_{\pi(a,h',\mu)}[\gamma(X_n, g(h', a), m(X_n, \mu, h', a)] &\geq E_{\pi(a,h'',\mu)}[\gamma(X_n, g(h'', a), m(X_n, \mu, h'', a))] \\
E_{\pi(0,h,\mu)}[\gamma(X_n, g(h, 0), m(X_n, \mu, h, 0))] &\geq E_{\pi(1,h,\mu)}[\gamma(X_n, g(h, 1), m(X_n, \mu, h, 1))]
\end{aligned}$$

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

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

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

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

Similarly,

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

Furthermore,

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

with the last inequality following from $v(0,0,h) \geq v(a,1,h)$ for any $a \in \{0,1\}$, that $a'' \in \{0,1\}$ and the above properties of expectations.

The above monotonicity properties of $(N\gamma)(x,h,\mu)$ imply in turn that $a(0,h,\mu) = 0$ (as in the proof of Lemma 3). Thus, we have (by using again the definition of optimum and the above properties of expectations) that:

$$
\begin{gathered}
(N\gamma)(0, L^j g(h', a'), L^j m_n'(0)) \\
= v(0,0,L^j g(h',a')) \\
+\delta \left\{ \begin{array}{c} \pi(0, L^j g(h',a'), L^j m_n'(0))\gamma(1, L^{j+1} g(h',a'), m(1, L^j m_n'(0), L^j g(h',a'), 0)) \\ + (1 - \pi(0, L^j g(h',a'), L^j m_n'(0)))\, \gamma(0, L^{j+1} g(h',a'), L^{j+1} m'(0)) \end{array} \right\}\} \\
\geq v(0,0,L^j g(h'',a'')) \\
+\delta \left\{ \begin{array}{c} \pi(0, L^j g(h'',a''), L^j m_n''(0))\gamma(1, L^{j+1} g(h'',a''), m(1, L^j m_n''(0), L^j g(h'',a''), 0)) \\ + (1 - \pi(0, L^j g(h'',a''), L^j m_n''(0)))\, \gamma(0, L^{j+1} g(h'',a''), L^{j+1} m_n''(0)) \end{array} \right\} \\
= (N\gamma)(0, L^j g(h'',a''), L^j m_n''(0))
\end{gathered}
$$

Thus, we have that N maps S^* into itself. It remains to show that S^* is closed. We show this next.

10.2.2 Proof that S^* is closed

In what follows we endow all the function spaces with the sup norm.

Let S denotes the subset of all continuous and bounded functions $\gamma(x,h,\mu)$ from $\{0,1\} \times [0,H] \times [0,1]$ to $\mathbb{R}$ and let the subset of S consisting of all weakly nonincreasing functions be denoted by $S^{0,d}$. We start by noting that $S^{0,d}$ is a closed subset of S.[37]

We now consider a subset $S^{1,d}$ of $S^{0,d}$ that has some further properties and show that it is a closed set as well.

$S^{1,d}$ is a closed subset of S^d Specifically, take any function $\gamma \in S^{0,d}$. Consider any point in the domain of the form $(0,h,\mu)$. For any admissible (according to Assumption (8)) g and f, consider the function $\widetilde{\gamma}$ which maps h, a, μ to $\mathbb{R}$ as:

$$\widetilde{\gamma}(h,a,\mu) \equiv \gamma(0, g(h,a), m(0,h,\mu,a)) \tag{7}$$

[37] See Stockey and Lucas (1989). The proof requires showing that the limit function, say $\overline{\gamma}$, of a sequence of functions γ_n in $S^{0,d}$ belongs to the set $S^{0,d}$. It is straightforward that $\overline{\gamma}$ is continuous and bounded, and that it is nonincreasing (and hence belongs to $S^{0,d}$) can be shown by mimicking the steps in the next subsection.

Define then the set of functions, $S^{1,d} \subseteq S^{0,d}$ such that $\gamma \in S^{1,d}$ if and only if $\widetilde{\gamma}$ is nonincreasing in h and a.

We need to show that $S^{1,d}$ is closed. Towards that end, let us introduce some notation. We denote with γ^1 a generic element of $S^{1,d}$. Moreover, we use an upperbar to denote a limit function. Thus, for instance, $\overline{\gamma}$ denotes the limit function of a sequence of functions $\{\gamma_n\}_n$.

Consider now 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 $\overline{\gamma}^1 \in S^{1,d}$.

Since $S^{1,d}$ is a subset of $S^{0,d}$ and since $S^{0,d}$ is closed, the limit function $\overline{\gamma}^1$ must be in $S^{0,d}$ so the only way $\overline{\gamma}^1$ can not be in $S^{1,d}$ is because its corresponding $\overline{\widetilde{\gamma}}^1$ is increasing 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, $\mu' = m(0, \mu, h', a')$ and $\mu'' = m(0, \mu, h'', a'')$. From (7), 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 $\bar{\gamma}^1(0, g', \mu') = \overline{\widetilde{\gamma}}^1(h', a', \mu)$ and $\bar{\gamma}^1(0, g'', \mu'') = \overline{\widetilde{\gamma}}^1(h'', a'', \mu)$.

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,

$$\gamma_n^1(0, g'', \mu'') > \overline{\gamma}^1(0, g'', \mu'') - \frac{\varepsilon}{2} \tag{8}$$

and,

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

The latter inequality can be written as

$$-\gamma_n^1(0, g', \mu') > -\overline{\gamma}(0, g', \mu') - \frac{\varepsilon}{2} \tag{9}$$

Combining the two inequalities (8) and (9), we get

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

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 apply this argument repeatedly to get successive subsets $S^{0,d} \supseteq S^{1,d} \supseteq S^{2,d}, ...$ each of which is closed. Countable intersections of closed sets are closed and hence $\cap_k S^{k,d} \equiv S^*$ (the intersection of $\{S^{k,d}\}_k$) is closed.