DEPARTMENT OF ECONOMICS
YALE UNIVERSITY

P.O. Box 208268
New Haven, CT 06520-8268

http://www.econ.yale.edu/

Economics Department Working Paper No. 40

Cowles Foundation Discussion Paper No. 1642

Foundations of Intrinsic Habit Formation

Kareen Rozen

March 2008

Foundations of Intrinsic Habit Formation

Kareen Rozen* †
Yale University

July 2007

Abstract

We provide theoretical foundations for several common (nested) representations of intrinsic linear habit formation. These representations are dynamically consistent and additive, with geometrically decaying coefficients of habit formation. Our axiomatization introduces a revealed preference theory of *weaning* a decision-maker from her habits using the device of compensation. We characterize linear habit formation in terms of the ability to wean using uniquely determined *compensating streams*. Moreover, we distinguish between habits that are *responsive* to weaning and those that are *persistent*, develop a simple choice-theoretic measure of the rate of *habit decay*, and demonstrate how to recover the entire sequence of habit formation coefficients from observed choice behavior. We introduce novel monotonicity and separability axioms that are appropriate for time-nonseparable preferences. Our analysis suggests techniques for eliciting dynamic reference points from choice behavior and obtaining discounted utility representations on endogenously generated auxiliary spaces.

Keywords: linear habit formation, time-nonseparable preferences, compensation, weaning, compensated separability, gains monotonicity, endogenously generated auxiliary spaces.
JEL classification: C60, D11, D90

*I am indebted to Roland Benabou, Wolfgang Pesendorfer, and especially Eric Maskin for their comments on this paper. I also thank Dilip Abreu, Faruk Gul, Jonathan Parker, Michael Rothschild, Ron Siegel, and seminar participants at Brown, Carnegie Mellon, Northwestern Economics and Kellogg MEDS, Princeton, UC-Los Angeles, UC-San Diego, University of Pennsylvania, Washington University in St. Louis, and Yale for useful suggestions. This paper is based on the first chapter of my doctoral dissertation at Princeton University.
†Cowles Foundation and Department of Economics, Yale University, Box 208281, New Haven, CT 06520-8281. Email: kareen.rozen@yale.edu

1 Introduction

Is an individual's valuation for a good affected by its frequency of consumption? Will someone who is accustomed to certain levels of comfort and quality come to expect and demand the same? And is an increase in consumption always preferable, even if the increase is only temporary?

The common thread binding these questions is that not one can be properly addressed in the standard intertemporally separable model of choice. Consequently, the literature in such varied fields as macroeconomics, finance, and labor economics has seen a surge in models incorporating intertemporal nonseparability through habit formation. By presuming a correlation between an individual's prior consumption levels (her *intrinsic habit*) and her felicity from present and future consumption, such models have had success in accounting for notable phenomena that more traditional theory has been unable to explain.[1]

The literature on habit formation has, however, faced at least two difficulties. First, it has been unable to come to a consensus on a single model of habits, and in some cases the predictions of the most commonly utilized models disagree (Wendner (2003)).[2] Second, since a more flexible preference improves the ability to explain data, models of habit formation - along with other models employing exotic preferences - are vulnerable to the critique that they are "an excuse for free parameters" (Backus, Routledge & Zin (2004)). Related to these two critiques is the scarcity of theoretical work examining the underpinnings of habit forming preferences. By clarifying the implications for choice behavior, such work would help illuminate why one utility representation of habit formation might be more inherently reasonable than another; or why the commonly used incarnations of habit formation are reasonable at all. We contribute to the literature in that theoretical vein.

[1] Variations of the model of *intrinsic linear habit formation* we axiomatize have shed light on data indicating individuals are far more averse to risk than expected (e.g., Constantinides (1990) on the equity premium); suggested why consumption growth is connected strongly to income, but only weakly to interest rates (see Boldrin, Christiano & Fisher (2001) for a real business cycles model with habit formation and intersectoral inflexibilities); and explained the consumption contractions seen before exchange rate stabilization programs collapse (Uribe (2002)).

[2] While intrinsic linear habit formation is the most commonly used model, some models posit habits that are nonlinear, extrinsic (the "catching up with the Joneses" effect of Abel (1990)), or enter the discount factor (Shi & Epstein (1993)). Chen & Ludvigson (2004) use formal estimation to argue that habits are intrinsic and nonlinear. One common nonlinear model specifies a linear habit aggregator that divides consumption in the felicity (Carroll, Overland & Weil (2000)); Wendner (2003) criticized this model for its counterintuitive implications for consumption growth.

1.1 Intrinsic linear habits

In this paper we formulate a theory of history dependent intertemporal choice using the device of compensating a decision-maker (DM) for giving up her habits. In particular, we provide a revealed preference theory of *weaning* a DM from her habits using *compensating streams* that are analogous to drug cessation aids such as the nicotine patch. Our representative DM is described by a family of continuous preference relations that govern her choice behavior at every history. She is dynamically consistent given her consumption history, can be weaned from her habits using weakly decreasing streams of compensation, and satisfies novel separability and monotonicity axioms that are appropriate for time-nonseparable preferences. Though our DM is fully rational, her history dependent behavior violates the axioms of Koopmans (1960), upon which the standard theory of discounted utility rests.

Instead, our compensation-based theory lays the foundation for the model of linear habit formation, in which a DM evaluates consumption at each point in time with respect to a reference point that is generated linearly from her consumption history. We denote the set of possible infinite consumption streams of some good by C and the set of possible infinite consumption histories of that good by H.[3] A consumption stream $c \in C$ is written as $c = (c_0, c_1, c_2 \ldots)$, with c_t denoting the prescribed consumption level t periods from today; while a history $h \in H$ is written as $h = (\ldots, h_3, h_2, h_1)$, with h_k denoting the consumption level k periods ago. We refer to a consumption history as a *habit*. For the case of scalar consumption, the model of linear habit formation specifies that for each habit $h \in H$ the DM evaluates the stream $c \in C$ using the utility function $U_h : C \to \mathbf{R}$ given by

$$U_h(c) = \sum_{t=0}^{\infty} \delta^t u\big(c_t - \varphi(h, c_0, c_1, \ldots, c_{t-1})\big), \quad \delta \in (0,1)$$

The felicity $u : \mathbf{R} \to \mathbf{R}$ is continuous and is often assumed to be monotonic. The habit aggregator $\varphi : H \to \mathbf{R}$ is a strictly increasing linear functional with geometrically decaying habit formation coefficients $\{\lambda_k\}_{k \geq 1}$ that sum to no more than one. That is, for each $h \in H$, $\varphi(h) = \sum_{k=1}^{\infty} \lambda_k h_k$, where each $\lambda_k \in (0,1)$ and $\frac{\lambda_{k+1}}{\lambda_k} \leq 1 - \lambda_1$. If the DM initially has the habit $h \in H$ and chooses to consume the stream $c \in C$, then after t periods her habit will be $(h, c_0, c_1, \ldots, c_{t-1})$. The value $\varphi(h, c_0, c_1, \ldots, c_{t-1})$ serves as the *reference point* against

[3] The history and consumption spaces are formalized in Section 2.

which the DM evaluates her period-t consumption level c_t, and is often referred to as the *habit stock*.

A number of variations of this model are prevalent in the literature. We provide foundations for this general formulation, which has an infinite sequence of habit formation coefficients, as well as for some common tractable variants that impose special restrictions on the $\{\lambda_k\}_{k\geq 1}$. Our approach also generalizes to a multidimensional version that can accommodate multiple habit-forming and non-habit forming goods. While our main theorem axiomatizes the case in which the felicity $u(\cdot)$ is monotonic, as is often assumed in applied work, a corollary relaxes the monotonicity assumption to allow for the possibility that the habit-forming good is harmful. The theorems we present are explicitly stated for the case of infinite histories, but readily accommodate bounded histories as well as histories that are finite but grow unboundedly.[4] Moreover, the representation may be easily generalized to one of risky choice over consumption streams, in which the DM cares about the *correlation* in lotteries. The results we provide shed light on the defining characteristics of linear habit formation, clarify the differences among nested specifications of the model, and provide various measures of habit-forming tendencies.

1.2 Compensating the DM to induce habit-free behavior

Our main axiom of choice behavior, *Habit Compensation* (HC), attempts to capture the essence of linear habit formation. This axiom clarifies how the DM's indifference curves change for different habits and provides a means to elicit her endogenously changing reference points from choice behavior. The axiom is composed of three parts, of which the first two play central roles.

Axiom HC(i), *Weaning*, offers a revealed-preference definition of weaning that builds on the notion underlying such drug cessation aids as the nicotine patch. The nicotine patch weans a smoker from her habit by providing a fading stream of nicotine that satisfies her cravings. When wearing the patch, the smoker behaves *as if* she were a nonsmoker. We allow for "patches" that help a DM with a habit h behave *as if* she had a lower habit h'.[5] That is, the patch may reduce her habit but not necessarily eliminate it. A DM is

[4] Models with finite but unboundedly growing consumption histories may be seen as applications of our infinite history representation with a constant initial habit $(\ldots, q, q)$.

[5] By this we mean $h \geq h'$, i.e. $h_k \geq h'_k$ for every k, with at least one strict inequality.

weaned from her habit h to a lower habit h' if, when endowed with some appropriate weakly decreasing stream of consumption $d^{h',h}$ at habit h, her choice behavior at habit h is the same as her behavior at habit h'. We call the stream of consumption $d^{h',h}$ a *compensating stream*.

Axiom HC(i) asserts that the DM can be weaned in this manner. More formally, the axiom asserts the existence and uniqueness of a compensating stream $d^{h',h}$ for each pair of habits $h \geq h'$. Our representation theorem shows that the restriction that the compensating streams are unique corresponds to a mild condition on the DM's felicity that is nearly always satisfied. Moreover, our results show that the compensating streams are determined *independently* of the DM's felicity and discount factor. Therefore, Axiom HC(i) describes a choice behavior that *isolates* the effect of the habit formation coefficients. In fact, we suggest a choice experiment from which the DM's entire sequence of habit coefficients can be recovered without regard to her exact felicity, discount factor, or consumption history, thereby offering a response to the "free parameters" critique of the model of linear habit formation.

Axiom HC(ii), *Compensated Separability*, ensures that with appropriate compensation, the choice between two streams is independent of future considerations, as long as the continuation path is the same in both. To illustrate, imagine that today is Saturday and that our DM is a smoker planning to reduce her cigarette consumption as soon as the weekend is over. Suppose that she will receive a nicotine patch on Monday, which she plans to use to cut down to x cigarettes a day.[6] In the meanwhile, she has one pack of cigarettes languishing in her pocket and must decide what to do with it. One option is to smoke the entire pack today and abstain tomorrow; another is to smoke half the pack today and half tomorrow. How should she compare these two options? If the smoker is assured of receiving the appropriate patch in each case on Monday, then Axiom HC(ii) implies that her choice is independent of x (i.e., she can focus on this weekend's consumption when deciding).

This axiom is trivially satisfied by the standard model of discounted utility *a la* Koopmans (1960), since a DM with time-separable preferences does not require any compensa-

[6]Without complicating this story using our multidimensional model of habits, we cannot explain the smoker's reasons for cutting down. This exercise simply asks her to compare streams in which she smokes various amounts over the weekend and then cuts down using a patch.

tion: the "compensating streams" in that model would be identically zero. Axiom HC(ii) may be seen as an appropriate *generalization of separability* for time-nonseparable preferences. However, it is not a typical separability condition because our task is quite different than that of Koopmans, whose representative DM satisfies time-separability on the *actual* consumption space. In contrast, our DM satisfies time-separablity only on an *endogenously generated auxiliary space* composed of streams of the form $(c_0 - \varphi(h), c_1 - \varphi(h, c_0), c_2 - \varphi(h, c_0, c_1), \ldots)$.[7] The techniques we develop suggest a means to elicit subjective reference points from choice behavior and derive discounted utility representations on spaces endogenously defined by these reference points.

1.3 A new monotonicity axiom

Most theories of choice assume monotonic preferences, in which the DM is better off whenever consumption in any period is increased. We claim this is too strong an assumption if the DM's preferences are time-nonseparable. For example, suppose a content individual of modest means is briefly permitted to live luxuriously, but must then return to her humble lifestyle. Would her welfare be increased, or might the experience of luxury render the return to her former circumstances unbearable? The answer may depend on the individual in question. We propose a weakened monotonicity axiom, *Gains Monotonicity* (GM), that accommodates either possibility.

Since increasing only a finite number of elements in a consumption stream may reduce future enjoyment of that stream, we refer to a "gain" as a *uniform increase in every period's consumption* from some point forward. Our Axiom GM says that the individual's welfare is unambiguously increased if she is permitted to keep her newfound increase in consumption indefinitely. We prove that Axiom GM ensures that the felicity in the model of linear habit formation is monotonic, and argue that it is consistent with experiments indicating that individuals favor increasing streams.[8]

Gains Monotonicity is relevant to cases in which the consumption good is a desirable one, as is typically assumed in the applied literature on habit formation. We suggest

[7]Axiom HC(ii) has the flavor of a separability axiom but does not translate directly to a separability condition on the auxiliary space. We prove using HC(i) and the surjectivity of a certain mapping that HC(ii) implies the separability conditions of Gorman (1968) on the auxiliary space.

[8]Refer to Camerer& Loewenstein (2004) for a comprehensive survey of such results.

a weakening of this axiom, *Gains Sensitivity* (GS), that is applicable when the good in question can be harmful when consumed excessively relative to past consumption (e.g, as in the case of alcohol) or in our model of multidimensional habit formation (due to potential contemporaneous inseparabilities). Replacing Gains Monotonicity with Gains Sensitivity relaxes the monotonicity requirements on the DM's felicity but leaves our representation theorems otherwise unchanged.

1.4 Overview of results

Section 4 offers our main representation theorem, which characterizes linear habit formation in terms of the ability to wean a DM using uniquely determined compensating streams. The result implies that the habit formation coefficients and compensating streams are unique for nearly all applications of linear habit formation. In fact, Section 5 demonstrates how to recover the DM's entire sequence of habit formation coefficients directly from observed choice behavior, without regard to her felicity, discount factor, or consumption history.

We introduce various measures of habit formation. Section 6 considers the DM's rate of habit decay in a common version of the model where the habit aggregator satisfies an autoregressive law of motion. There we provide an additional axiom, *Immediate Equilibration* (IE), that describes a simple choice experiment which simultaneously generates the autoregressive structure and calibrates the DM's habit decay parameter in that model. Section 7 distinguishes between models of linear habit formation in which habits are *responsive* to weaning and those in which habits are *persistent*. Responsive habits are those from which the DM may be weaned *efficiently*, using a finite amount of compensation; in contrast, persistent habits require a constant level of compensation (e.g., some heroin addicts require a steady dose of methadone for the rest of their lives). We show that the distinction between these two types of habits corresponds to a simple difference in modeling that can markedly affect choice behavior.

We also consider several generalizations of our results. Section 8 extends our theory to a multidimensional model of habit formation, in which the DM forms independent linear habits over a subset of the commodity space. Such a model can be used to understand multiple addictions and addiction cycles. Section 9 extends our representation to risky consumption streams and argues that by reinterpreting the commodity space, our axiomatization also provides foundations for a model of habits in which the DM forms habits over rates of consumption growth (i.e., the argument of the felicity is *log-linear*).

1.5 Connections to the literature

This paper is related to a growing theoretical literature on intrinsic and dynamically consistent habit formation, beginning with the seminal papers by Iannaccone (1986) and Becker & Murphy (1988) on rational addiction. These papers offered the first models in which individuals might rationally choose to experiment with addictive substances, even though they know this may lead them on a path to addiction.

Rustichini & Siconolfi (2005) also offer an axiomatization of dynamically consistent habit formation over consumption streams. Unlike this paper, they axiomatize a model of (recursive) habit formation that does not offer a particular structure for the utility or form of habit aggregation; and in general, their axioms have a very different flavor from our own. Other work on dynamically consistent habit formation includes Gul & Pesendorfer (2007), who consider preferences over opportunity sets of streams of consumption, rather than over the streams themselves; their decision-makers have difficulties with self-control and may prefer to restrict their options. Other approaches to habit formation include decision-makers who are myopic (Pollak (1970)), discount rates that depend on prior consumption (Shi & Epstein (1993)), and habits that are extrinsic (e.g., Abel (1990) and Campbell & Cochrane (1999)).

The type of preference in which we are interested falls under the rubric of intrinsic reference dependence. A well-known member of this class of preferences is loss aversion, discussed in Tversky & Kahneman (1991) in the context of riskless choice. The salient feature of loss aversion is that the disutility from losses is more acute than utility from gains. While the classical representation of loss aversion is not dynamically consistent, our representation can easily accommodate a dynamically consistent version in which the felicity takes the well-known "S"-shaped form.

A dynamically inconsistent foundation for a special case of loss aversion is offered by Shalev (1997).[9] Another model incorporating loss aversion is suggested by Loewenstein & Prelec (1992); that model combines hyperbolic discounting with a specially curved utility

[9]Shalev's model is not dynamically consistent because it is based on the model of Gilboa (1989), which relaxes Savage's sure-thing principle.

function over gains and losses. Both models attempt to account for experimentally observed departures from discounted utility. These anomalies of intertemporal choice, summarized in the survey of Camerer & Loewenstein (2004), include evidence suggesting that different discount rates are applied to gains and losses; that the discount rate for losses appears lower than that for gains; and that individuals get more disutility from postponing consumption than they get utility from speeding it up. However, the time-nonseparable representation which we axiomatize can also account for such phenomena; as Camerer and Loewenstein note, "these effects are consistent with stable, uniform, time discounting once one measures discount rates with a more realistic utility function."

Finally, our analysis contributes to the choice-theoretic literature on reference dependence, particularly to the static model of Neilson (2006). That paper offers axiomatic foundations for a model in which the DM chooses a bundle, the first component of which serves as the reference point against which the other components are evaluated. Neilson approaches the problem by using an axiom that identifies the first component of the bundle as the reference point. By contrast, our approach does not assume a particular reference point but derives an infinite sequence of endogenously changing reference points that are weighted averages of historical consumption. Our reference points are extracted using a sequence of functional equations generated from the dynamics of the problem. Our method allows us to link an entire family of history dependent preferences in a dynamically consistent manner.

2 The framework

We consider a sophisticated DM who has preferences over streams of consumption. Formally, suppose that the DM faces an infinite-horizon decision problem in which a single habit forming good is consumed in every period $t \in \mathbb{N} = \{0, 1, 2, \ldots\}$ from the same set $Q = \mathbf{R}_+$.[10] The choice $q \in Q$ may be interpreted as a choice of either quantity or quality of the consumption good, with higher q corresponding to a higher level. We will use $q \in Q$ to refer to a generic one-period consumption level.

The DM's preferences are stationary (i.e., they do not depend on calendar time). However, they do depend on her consumption history, her *habit*. The set of possible habits is

[10] To streamline the presentation we leave the extension $Q = \mathbf{R}^n_+$ $(n \geq 1)$ to Section 8.

time-invariant and given by the space of bounded streams

$$H = \{h \in \times_{k=\infty}^{1} Q \mid \sup_k h_k < \infty \}.$$

Each habit $h \in H$ is an infinite stream denoting prior consumption and is written as $h = (\dots, h_3, h_2, h_1)$, where h_k denotes the consumption level of the DM k periods ago. We endow the space H with the sup metric $\rho^H(h, h') = \sup_k |h_k - h'_k|$.

At any point in time, the DM's preferences under habit $h \in H$ are given by $\succeq_h$ and are defined on the set of bounded consumption streams

$$C = \{c \in \times_{t=0}^{\infty} Q \mid \sup_t c_t < \infty \}.$$

A choice $c = (c_0, c_1, c_2, \dots) \in C$ is an infinite consumption stream; c_t is the consumption level prescribed for t periods after the date at which the DM makes her choice, which is interpreted as the current date. We consider C as a metric subspace of $\times_{t=0}^{\infty} Q$ endowed with the product metric $\rho^C(c, c') = \sum_{t=0}^{\infty} \frac{1}{2^t} \frac{|c_t - c'_t|}{1+|c_t - c'_t|}$. Since $\times_{t=0}^{\infty} Q$ endowed with ρ^C is a topologically separable metric space, so is C when viewed as a metric subspace.[11]

The DM is cognizant that her future tastes will be influenced by her consumption history. Starting from any initial habit $h \in H$, consuming the stream $c \in C$ results in the date-t habit $(h, c_0, c_1, \dots, c_{t-1})$. Therefore, DM's preferences may undergo an infinite succession of changes endogenously induced from her choice of consumption stream. Each resulting preference is a member of the family $\succeq = \{\succeq_h\}_{h \in H}$.

Our setup explicitly presumes histories are infinite because this assumption is commonly invoked in the literature. Alternatively, one could assume that the DM's preferences are affected only by her last $K \geq 3$ consumption levels.[12] The notation in our analysis would remain the same so long as current and future habits are truncated after K components; that is, (h, c_0) would denote the habit $(h_{K-1}, \dots, h_2, h_1, c_0)$.

[11] Ensuring that C is separable in this manner allows us to concentrate on the structural elements of habit formation. Alternatively we could impose separability directly as in Rustichini & Siconolfi (2005). Bleichrodt, Rohde & Wakker (2007) is representative of a literature that concentrates on relaxing assumptions about the consumption space, including separability.

[12] $K \geq 3$ is required only for the proof of time-additivity.

2.1 Useful notation

Here we collect some notation that will be used throughout the paper. We reserve the variable $k \in \mathbb{N} = \{1, 2, 3, \ldots\}$ to signify a period of previous consumption and the variable $t \in \{0, 1, 2, \ldots\}$ to signify a period of impending consumption. When the stream c is consumed under habit h, the resulting date-t habit $(h, c_0, c_1, \ldots, c_{t-1})$ is denoted by $h^{(t)}$ as long as this is unambiguous. At times it will be convenient to let hq denote the habit (h, q) which forms after consuming q under habit h. The notation $c + c'$ or $h + h'$ refers to usual vector addition. As is customary, ${}^t c$ denotes $(c_t, c_{t+1}, c_{t+2}, \ldots)$ and c^t denotes $(c_0, c_1, \ldots, c_t)$. If $c' \in C$ we write $(c^t, {}^{t+1}c + c')$ to denote $(c_0, c_1, \ldots, c_t, c_{t+1} + c'_0, c_{t+2} + c'_1, \ldots)$. For $\alpha \in \mathbf{R}$ we use the similar notation α^t to signify the t-period repetition $(\alpha, \alpha, \ldots, \alpha)$ and $(c^t, {}^{t+1}c + \alpha)$ to compactly denote $(c_0, c_1, \ldots, c_t, c_{t+1} + \alpha, c_{t+2} + \alpha, \ldots)$ whenever the resulting stream is in C. The zero habit $(\ldots, 0, 0)$ will typically be denoted by $\vec{0}$. Finally, by convention $h \geq h'$ (or $c \geq c'$) means $h_k \geq h'_k$ for all k (or $c_t \geq c'_t$ for all t), with at least one strict inequality.

3 The main axioms

This section presents axioms of choice behavior that are necessary and sufficient for representing each $\succeq_h$ by a utility function $U_h : C \to \mathbf{R}$ of the form

$$U_h(c) = \sum_{t=0}^{\infty} \delta^t u\big(c_t - \sum_{k=1}^{\infty} \lambda_k h_k^{(t)}\big), \tag{1}$$

where $\delta \in (0, 1)$; each $\lambda_k \in (0, 1)$ and $\frac{\lambda_{k+1}}{\lambda_k} \leq 1 - \lambda_1$ for $k \geq 1$; and $u : \mathbf{R} \to \mathbf{R}$ is continuous and monotonic.[13] The following axioms are imposed for all $h \in H$.

3.1 Rational choice

The first three axioms are familiar in the theory of rational choice over consumption streams. First, the DM's choices are derived from a preference relation.

Axiom PR (Preference Relation) $\succeq_h$ is a complete and transitive binary relation.

The DM's preferences are also required to be continuous. As usual, $\succ_h$ denotes the asymmetric part of $\succeq_h$.

[13] See Theorem 4.1 for a precise description of the monotonicity and "acyclicity" of $u(\cdot)$.

Axiom C (Continuity) For all $c \in C$, $\{c' : c' \succ_h c\}$ and $\{c' : c \succ_h c'\}$ are open.

These axioms ensure a continuous utility representation on our separable space.

Proposition 3.1. *There exists a continuous utility representation of* $\succeq_h$.

We further assume that the DM's preferences are dynamically consistent in a history dependent manner, in the sense that given the relevant histories, she will not change her mind tomorrow about the consumption stream she chooses today.

Axiom DC (Dynamic Consistency) For any $q \in Q$ and $c, c' \in C$, $(q, c) \succeq_h (q, c')$ if and only if $c \succeq_{hq} c'$.

Dynamic consistency helps ensure that dynamic programming techniques can be used to solve the DM's choice problem, and that the DM's welfare can be analyzed unambiguously. If dynamic consistency is violated it becomes more difficult to interpret the DM's choices for the future (her future self may wish to change plans midstream) as well as discuss the welfare implications of her choices (the welfare of her present self may come at the expense of future selves, and vice versa). Without dynamic consistency, the DM's choice must be modeled through an equilibrium concept rather than as a decision problem.[14] Moreover, as noted in Section 1.5, the axiom is able to accommodate a number of observed time-discounting anomalies.

3.2 A weakening of monotonicity

Individuals are often assumed to have monotonic preferences over consumption streams; that is, the DM is assumed to be better off whenever consumption in any period is increased. This seemingly innocuous assumption comes into question when the DM's preferences are time-nonseparable. When consuming a stream c, one may imagine that the DM experiences a feeling of gain *in a particular period* t if $c_t - c_{t-1} > 0$, or more generally, if period-t consumption is larger than some composite of prior consumption levels. However, if only a finite number of elements in a consumption stream are increased then the DM may face disappointment (or at least a reduction in enjoyment) when the increase ends. The short-

[14] This is the multi-selves approach of Strotz (1955). A related equilibrium notion is studied in Köszegi & Rabin (2006b)'s model of dynamic reference dependence (which extends the static model in Köszegi & Rabin (2006a)). There, the utility over sequences of consumption and beliefs is technically consistent but beliefs are forced to be determined rationally in a *personal equilibrium*.

term benefit from the temporary consumption increase might not suffice to overcome the long-term loss.

Therefore, this paper refers to a "gain" more specifically as a uniform increase in all consumption from some point forward. For any $\alpha \in \mathbf{R}$ and t, we refer to a stream $(c^t, {}^{t+1}c + \alpha)$ as a gain over c when $\alpha > 0$ and a loss relative to c when $\alpha < 0$. Formally, we impose the following weakened monotonicity axiom.

Axiom GM (Gains Monotonicity) If $c \in C$ and $\alpha > 0$, then $c + \alpha \succ_h c$.

GM says that the DM is happy with an increase in her future consumption so long as the increase is applied to every period; that is, raising q will raise her utility so long as it doesn't affect future gains and losses in consumption across successive periods. It is easy to see that if the preference relation satisfies GM and DC, then it satisfies the following strengthened version of Gains Monotonicity, which corresponds exactly to the notion of a gain as we have defined it.

Axiom GM* If $c, c' \in C$ and $c = (c'^t, {}^{t+1}c' + \alpha)$ for some t and $\alpha > 0$, then $c \succ_h c'$.

To provide an alternate characterization, consider the following definition.

Definition 3.2. *Let $c, c' \in C$ be two consumption streams with $c \geq c'$. We say that $c >_{GD} c'$, or c gains-dominates c', if c has larger period-to-period gains and smaller period-to-period losses: that is, $c_t - c_{t-1} \geq c'_t - c'_{t-1}\ \forall\ t \geq 1$.*

That is, increasing the gains from any point forward in a consumption stream without increasing the losses leads to a gains-dominating consumption stream. Observe that a stream will gains-dominate another if and only if the difference between the two streams is positive and increasing. The following proposition, proved in the Appendix, asserts that a continuous preference relation satisfies GM* if and only if the preference respects gains-domination.[15]

Proposition 3.3 (Respect of gains-domination)**.** *A preference relation continuous in the product topology satisfies GM* if and only if it respects gains-domination; that is, for any $c, c' \in C$, $c >_{GD} c'$ implies that $c \succ c'$.*

[15]As an aside, compare GM* with the weaker constant-tail monotonicity axiom of Shalev (1997), which says (restricted to deterministic streams) that if a stream constantly gives q from time t onwards, then raising q to some $q' > q$ from t onwards improves the stream. This is equivalent to the statement that a weakly increasing (decreasing) consumption stream is at least as good (bad) as getting its worst (best) element constantly.

While Proposition 3.3 offers a natural argument for Axiom GM, there may be cases in which the axiom lacks normative appeal. This may be because of the nature of the good itself (e.g., an individual may feel ill from consuming alcohol excessively relative to past consumption), or due to the presence of contemporaneous interactions between consumption goods, such as in the multidimensional formulation we later offer. Therefore, we offer the following weakening.

Axiom GS (Gains Sensitivity) There exist $c \in C$ and $\alpha > 0$ such that $c + \alpha \not\sim_h c$.

Replacing Gains Monotonicity with Gains Sensitivity in our representation theorems allows us to drop the monotonicity requirement on the DM's felicity.

3.3 Compensation

We now introduce our main structural axiom of habit formation, *Habit Compensation* (HC). To present the axiom we define the set of ordered pairs of consumption histories $\mathcal{H} = \{(h', h) \in H \times H \mid h' \leq h\}$ and introduce one further piece of terminology. For any k, we say that habits $(h', h) \in \mathcal{H}$ *agree on* k if $h_k = h'_k$. Similarly, we say that the habits $(h', h) \in \mathcal{H}$ agree on a subset of indices $\mathbb{K} \subset \{1, 2, \ldots\}$ if they agree on each $k \in \mathbb{K}$. Axiom HC provides a revealed preference theory of weaning a DM from her habits; formally, it says the following.

Axiom HC (Habit Compensation) There is a collection $\{d^{h',h}\}_{(h',h)\in\mathcal{H}}$ of strictly positive streams such that

(i) (Weaning). Each $d^{h',h}$ is weakly decreasing and uniquely satisfies

$$c \succeq_{h'} c' \text{ iff } c + d^{h',h} \succeq_h c' + d^{h',h} \quad \forall\, c, c' \in C.$$

(ii) (Compensated Separability). For any $c, \hat{c} \in C$, $t \geq 0$ and $h' \leq hc^t, h\hat{c}^t$,

$$(c^t, d^{h',hc^t}) \succeq_h (\hat{c}^t, d^{h',h\hat{c}^t}) \text{ iff } (c^t, \bar{c} + d^{h',hc^t}) \succeq_h (\hat{c}^t, \bar{c} + d^{h',h\hat{c}^t}) \quad \forall\, \bar{c} \in C.$$

(iii) (Independence of Irrelevant Habits). For any $\hat{k}$, $(h', h) \in \mathcal{H}$ that agree on $\hat{k}$, and $q \in Q$, if $\hat{h}'_k = \begin{cases} h'_k & k \neq \hat{k} \\ q & k = \hat{k} \end{cases}$ and $\hat{h}_k = \begin{cases} h_k & k \neq \hat{k} \\ q & k = \hat{k} \end{cases}$ then $d^{h',h} = d^{\hat{h}',\hat{h}}$.

For every ordered pair of histories $(h', h) \in \mathcal{H}$, Axiom HC(i) posits the existence of a unique compensating stream $d^{h',h}$ that induces a DM with habit h to behave *as if* she had habit h'. A compensating stream serves as an endowment of the habit-forming good that is provided only at the higher habit. It is analogous to a nicotine patch that weans a smoker from her addiction; however, we generalize the underlying notion by permitting the smoker's habit to be either reduced or eliminated, depending on the patch used. Axiom HC(ii) considers the effect of compensation applied midstream and may be viewed as a generalization of separability for time-nonseparable preferences. According to HC(ii), if the DM is appropriately compensated starting in period t, then her choice between two streams depends only on the consumption levels they provide prior to t, as long as those streams agree on their continuation path. That is, the future is "separable." Axiom HC(iii) ensures that if $(h', h) \in \mathcal{H}$ agree on some k, then the compensation required to wean the DM from h to h' is independent of the period-k habit level. In other words, elements of habits that remain unchanged are irrelevant to the weaning process.

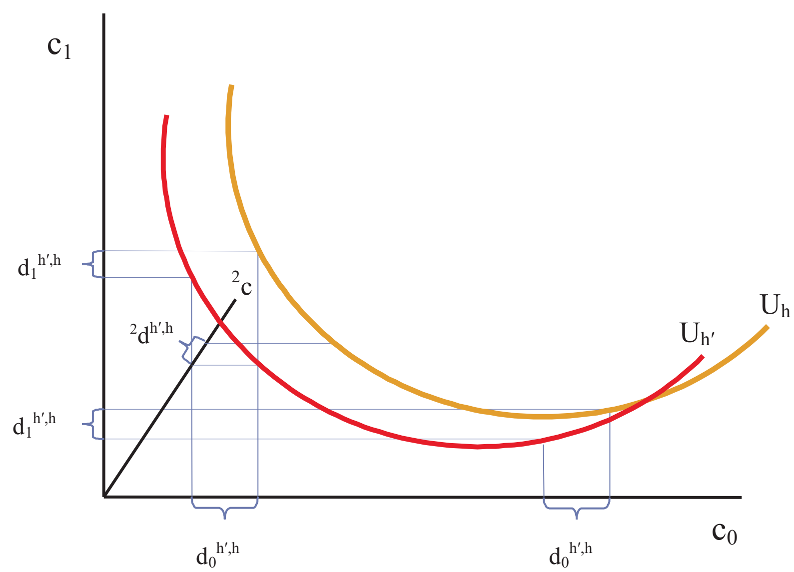

Figure 1: An h'-indifference curve over (c_0, c_1), fixing ${}^2c$, is translated under HC(i)

As illustrated in Figure 1, HC(i) establishes that the indifference curves for habit h' are translated up by the strictly positive stream $d^{h',h}$ into indifference curves for habit h.[16]

[16]There is a feature of the two indifference curves in Figure 1 that may not be immediately apparent from the axiom: they correspond to the same utility, and in fact to the same felicity in every period. The indifference curves that provide the DM with the same utility under each habit never cross; note that in

The fact that $d^{h',h}$ is weakly decreasing formalizes the manner in which higher habits lead to greater "impatience" for consumption: the DM must immediately consume the highest levels of her promised compensation in order to make the same choices that she would have made at a lower habit. This stipulation corresponds directly to the geometric decay of the habit formation coefficients.

Note that the ability to wean the DM hinges critically on the linearity of habit formation. If the argument of her felicity were nonlinear, the DM's "compensating stream" would have to be adjusted to reflect each consumption choice she makes. Hence the order of the quantifiers in Axiom HC(i) is critical for linearity.

To clarify the role of Axiom HC(iii), let us consider a variation of linear habit formation where it is violated.

Example 3.4. *Choose a felicity $u : \mathbf{R} \to \mathbf{R}$, a discount factor $\delta \in (0,1)$, and any strictly positive sequence $\{\lambda_k\}_{k\geq 1}$ satisfying $\frac{\lambda_{k+1}}{\lambda_k} \leq 1 - \lambda_1$ for all k. Endow each $h \in H$ with the sequence of habit coefficients $\{\lambda_{k,h}\}_{k\geq 1}$ given by $\lambda_{k,h} = \lambda_k \frac{\alpha + \limsup_{k'} h_{k'}}{\beta + \limsup_{k'} h_{k'}}$, where $\beta > \alpha > 0$. For each $h \in H$ define $U_h : C \to \mathbf{R}$ by*

$$U_h(c) = \sum_{t=0}^{\infty} \delta^t u\big(c_t - \sum_{k=1}^{\infty} \lambda_{k,h} h_k^{(t)}\big)$$

Then the family of utilities $\{U_h\}_{h\in H}$ satisfies HC(i) and (ii), but violates HC(iii).[17]

Example 3.4 demonstrates that Axiom HC(iii) is directly responsible for the homogeneity of the habit formation coefficients across consumption histories but is not necessary for linearity. Moreover, the habit formation coefficients in the example can only depend on the tail behavior of the consumption history; otherwise, Axioms HC(i) and DC would be violated. Similarly, note that we may also generate linearity without Axiom DC. An alternative model where the DM has preferences represented by linear habit formation with naive hyperbolic time discounting would satisfy all of our axioms except DC. Even stationarity plays a minimal role: we could replace the felicity in the representation with a non-stationary variant $u_t(\cdot)$ without violating HC or GM.

These examples correctly suggest that the compensating streams are constructed only from the habit formation coefficients, and are entirely independent of the DM's felicity and

Figure 1 one indifference curve is behind the other.

[17] The felicity should be *acyclic* (see Section 4) to ensure that compensation is unique.

discount factor. Therefore, there is an important sense in which Axioms HC(i) and HC(iii) differ from Axiom HC(ii). Axiom HC(ii) is a generalized separability axiom and does not provide a particular structure for the compensating streams. While Axioms HC(i) and HC(iii) directly generate the habit formation coefficients, Axioms HC(ii) and DC together generate the manner in which the DM aggregates her utility over time. The compensating streams would still exist and satisfy Axiom HC(iii) even in a variation of our model of linear habit formation where utility is not time-additive. However, Axiom HC(ii) would be violated.

Finally, we require two additional technical conditions on the DM's initial level of compensation. These conditions concern the strength of the DM's memory and rule out degenerate representations of the preferences we seek. First, we require that the initial compensation needed for a habit goes to zero as that habit becomes more distant in memory: i.e., for any habit $h \in H$ we have $\lim_{t\to\infty} d_0^{\vec{0},h0^t} = 0$. In counterpoint, the second condition states that for any fixed prior date of consumption, we can find two habits that differ widely enough on that date to generate any initial level of compensation: i.e., for any $q > 0$ and k, there exist $(h', h) \in \mathcal{H}$ that agree on $\mathbb{N} \setminus \{k\}$ and satisfy $d_0^{h',h} = q$. The first condition is required *only for histories of infinite length*: it rules out an undesirable term inside the utility that depends only on tail elements of the habit. The second condition rules out degenerate solutions of a critical functional equation.[18] We say the DM's memory is *non-degenerate* if these two conditions hold.

Axiom NDM (Non-Degenerate Memory) The DM's memory is non-degenerate.

4 The main representation theorem

We now present our main theorem, which offers a precise characterization of the preferences that satisfy our axioms of habit formation. The representation obtained is a dynamically consistent and additive model of intrinsic linear habit formation that has featured prominently in the applied literature. The model permits any choice of felicity, subject to a minor acyclicity condition that will soon be explained.

[18] A very weak technical assumption, which we discuss in the Appendix, would also suffice.

Theorem 4.1 (Main representation theorem)**.** *The family of preferences $\succeq$ satisfies Axioms PR, C, DC, GM, HC, and NDM if and only if each $\succeq_h$ can be represented by*

$$U_h(c) = \sum_{t=0}^{\infty} \delta^t u\Big(c_t - \sum_{k=1}^{\infty} \lambda_k h_k^{(t)}\Big) \quad \forall\, c \in C, \tag{2}$$

where $\delta \in (0,1)$; the habit formation coefficients $\{\lambda_k\}_{k\geq 1}$ are unique and satisfy

$$\lambda_k \in (0,1) \text{ and } \frac{\lambda_{k+1}}{\lambda_k} \leq 1 - \lambda_1 \text{ for all } k \geq 1; \tag{3}$$

and the felicity $u : \mathbf{R} \to \mathbf{R}$ satisfies the following conditions:

(i) Continuous, weakly increasing, unique up to positive affine transformation

(ii) If $\sum_{k=1}^{\infty} \lambda_k < 1$ then $u(\cdot)$ increases strictly on $(0, \infty)$ and is not "quasi-cyclic"[19]

(iii) If $\sum_{k=1}^{\infty} \lambda_k = 1$ then there exists $0 < a \leq \infty$ such that $u(\cdot)$ increases strictly either on $(-a, \infty)$ or $(-\infty, a)$ and is not "cyclic."

The discount factor δ, coefficients $\{\lambda_k\}$, and felicity $u(\cdot)$ are independent of h.

The proof, which is outlined in Section 4.1, is given in Appendices B (sufficiency) and C (necessity). As can be seen from the proof, the monotonicity requirements on $u(\cdot)$ can be dropped by weakening Axiom GM to GS. The representation theorem is otherwise unchanged (in particular, the acyclicity conditions remain).

Corollary 4.2 (Main representation theorem, non-monotonic felicity)**.** *The family of preference relations $\succeq$ satisfies Axioms PR, C, DC, GS, HC, and NDM if and only if each $\succeq_h$ can be represented as in Theorem 4.1 but without the monotonicity requirements on $u(\cdot)$ in assertions (i)-(iii).*

An important implication of these two results is that the compensating sequences are unique for nearly all economic applications of linear habit formation. Moreover, we will soon see that they are determined independently of the DM's exact felicity and discount factor. The proof of the theorem shows that compensation is uniquely determined if an appropriate acyclicity condition holds. Once we explain the meaning of this condition, it will be evident that it essentially never binds in practice.

[19]The definitions of the terms quasi-cyclic and cyclic follow the theorem.

Definition 4.3. *A function $u : \mathbf{R} \to \mathbf{R}$ is cyclic if there are $\alpha \in \mathbf{R}$ and $\gamma > 0$ such that $u(x+\gamma) = u(x) + \alpha$ for all $x \in \mathbf{R}$. A function u: $\mathbf{R} \to \mathbf{R}$ is quasi-cyclic if there are $\alpha \in \mathbf{R}$ and $\beta, \gamma > 0$ such that $u(x+\gamma) = \beta u(x) + \alpha$ for all $x \in \mathbf{R}$.*[20]

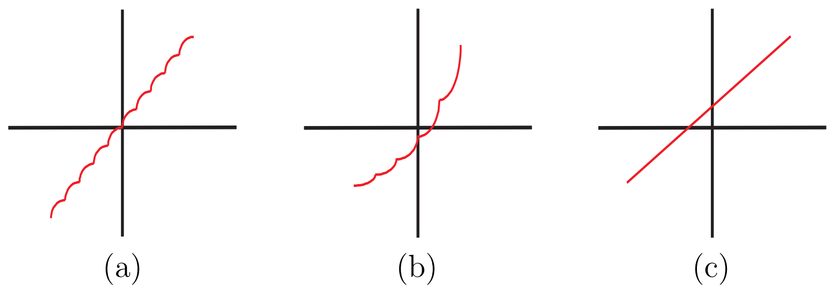

Figure 2: Violations of acyclicity. (a) $\beta = 1$; (b) $\beta > 1$; (c) $\beta = 1$ and affine.

Some cyclic and quasi-cyclic functions are illustrated in Figure 2. In light of Definition 4.3, *the class of felicities permitted by the representation contains nearly all felicities used in models of habit formation*. Unless it is affine, a quasi-cyclic function cannot be both smooth and concave. Furthermore, even a non-concave felicity such as the well-known S-shaped function of prospect theory cannot be quasi-cyclic; indeed, quasi-cyclic functions cannot have a finite (and non-zero) number of kinks. Finally, while affine functions are technically cyclic, they are not particularly well-suited to models of linear habit formation since choice behavior would be observationally equivalent to that in a model without habit formation.

4.1 Roadmap to the proof

The proof of Theorem 4.1 is constructive. In this section we offer a roadmap to a few of the main steps of the construction. Before doing so, let us gain insight into the proof by examining why Axiom HC(i) is implied by the utility representation.

Consider a DM whose family of preferences $\succeq$ may be represented as in Theorem 4.1 and choose any ordered pair of habits $(h', h) \in \mathcal{H}$. The DM's utility from a stream $c \in C$ under these two habits will differ in that her period-t felicity is given by $u\big(c_t - \varphi(h', c^{t-1})\big)$ under h', while it is given by $u\big(c_t - \varphi(h, c^{t-1})\big)$ under h. However, observe that we may

[20]Notice that a cyclic function is quasi-cyclic with $\beta = 1$.

reformulate the former by adding and subtracting $\varphi(h, c^{t-1})$:

$$u\big(c_t - \varphi(h', c^{t-1})\big) = u\big(c_t + [\varphi(h, c^{t-1}) - \varphi(h', c^{t-1})] - \varphi(h, c^{t-1})\big). \tag{4}$$

Since the habit aggregator $\varphi(\cdot)$ is a strictly increasing linear functional, the bracketed term $\varphi(h, c^{t-1}) - \varphi(h', c^{t-1})$ is strictly positive and given by $\varphi\big(h - h', 0^t\big)$.

Our observation in (4) aids in the construction of the compensating streams. Indeed, let us generate the compensating stream $d^{h',h}$. In period 0, we provide the DM with the amount $d_0^{h',h} = \varphi(h - h')$; as seen from (4), the DM's period-0 felicity from consuming $c_0 + d_0^{h',h}$ under habit h is the same as her period-0 felicity from consuming c_0 under habit h'. In the following period, we must take into account that we have compensated the DM using the habit forming good; that is, we must compensate for the compensation. The period-0 consumption level that corresponds to habit h in (4) is $c_0 + d_0^{h',h}$. Using linearity, we may then define $d_1^{h',h} = \varphi\big(h - h', \varphi(h - h')\big)$. Continuing in this manner, we find a stream $d^{h',h}$ that weans the DM from habit h to habit h' and has the recursive structure

$$d^{h',h} = \Big(\varphi(h-h'), \varphi\big(h-h', \varphi(h-h')\big), \varphi\big(h-h', \varphi(h-h'), \varphi(h-h', \varphi(h-h'))\big), \dots\Big). \tag{5}$$

In light of (5), it is evident that our constructive proof should define the habit aggregator $\varphi(h)$ by $d_0^{\vec{0},h}$. Therefore, the first task at hand is to prove that $d_0^{\vec{0},h}$ will have the desired linear structure under our axioms.

Constructing the linear habit aggregator $\varphi : H \to \mathbf{R}_+$

The utility representation given in Theorem 4.1 stipulates that the desired habit aggregator $\varphi(\cdot)$ satisfies three properties: (i) there exist functions $\varphi_k : Q \to \mathbf{R}_+$ such that for each $h \in H$, $\varphi(\cdot)$ can be written in the additive form $\varphi(h) = \sum_{k=1}^{\infty} \varphi_k(h_k)$; (ii) each $\varphi_k(\cdot)$ is strictly increasing and linear, i.e. $\varphi_k(q) = \lambda_k q$ for some $\lambda_k > 0$; and (iii) the sequence $\{\lambda_k\}_{k\geq 1}$ decays at least geometrically fast and sums to no more than one: $\frac{\lambda_{k+1}}{\lambda_k} \leq 1 - \lambda_1$.

Lemmas B.2-B.3 demonstrate that Axiom HC(i) is the key axiom generating the underlying additivity of $\varphi(\cdot)$. Recall that the compensating stream is consumed as a fixed supplement to regular consumption, akin to the manner in which a nicotine patch is administered. This patch-like structure permits each $d^{h',h}$ to be written as an infinite sum of successive one-step transitions from h' to h, where only one element is changed at a time. Axiom HC(iii) ensures that the order in which the successive one-step compensations occur is unimportant; but each one-step compensation may still depend on the entire initial habit level h'. Within the additive structure generated by Axiom HC(i), however, we prove that a one-step compensation is genuinely independent of all elements left unchanged. Thus, we may denote the compensation for changing the k-th element of a habit from 0 to q as $d^{k,0,q}$ without any loss of generality. Given the first part of Axiom NDM, which rules out an undesirable limiting term that depends on the tail elements of the habits, we may write $d^{\vec{0},h} = \sum_{k=1}^{\infty} d^{k,0,h_k}$. Hence we naturally define each $\varphi_k(\cdot)$ by $\varphi_k(q) = d_0^{k,0,q}$ for $q > 0$ and $\varphi_k(0) = 0$.

Axiom HC(i) posits the existence of compensating streams; to determine their actual form we utilize the dynamics of the problem. In Lemmas B.5-B.7 we use Axiom HC(i) in conjunction with DC and HC(iii) to manipulate the compensating sequences $d^{k,0,q}$ and show that each $\varphi_k(\cdot)$ satisfies the functional equation

$$\varphi_k(\varphi_k(q) + q') = \varphi_k(\varphi_k(q)) + \varphi_k(q') \quad \forall\, q, q' \in Q. \tag{6}$$

Equation (6) is a restricted Cauchy equation that is complicated by the fact that the domain on which it holds is endogenous; as a consequence, the solution of this functional equation has not been fully characterized.[21] One known result is that of Jarczyk (1991), which states that a continuous function $\varphi_k : [0, \infty) \to [0, \infty)$ solving (6) must take the form $\varphi_k(q) = \lambda_k q$ for some $\lambda_k > 0$. However, we may only conclude that $\varphi_k(\cdot)$ is almost everywhere continuous using Axiom HC(i). The second part of Axiom NDM ensures that the range of each $\varphi_k(\cdot)$ is the entire domain Q; reparametrization then reduces (6) to the simple Cauchy equation, for which almost everywhere continuity suffices to ensure linearity. In Lemma B.11 we suggest replacing the second component of Axiom NDM with a very weak alternative condition that can be shown to rule out discontinuities of $\varphi_k(\cdot)$ on sets of measure zero, thereby permitting us to directly apply the theorem of Jarczyk.

Let us examine the properties of these coefficients $\{\lambda_k\}_{k\geq 1}$. Recall that the compensation required to the wean the DM is weakly decreasing. Lemma B.13 shows that the

[21] A function $f : \mathbf{R} \to \mathbf{R}$ satisfies the Cauchy functional equation on some domain X if for every $x', x \in X$, $f(x' + x) = f(x') + f(x)$.

"fading" nature of compensation corresponds directly to the geometric decay of the habit formation coefficients. The critical result used to prove this is Lemma B.12, which recursively characterizes compensation by ${}^t d^{\vec{0},h} = d^{\vec{0},h d_0^{\vec{0},h} d_1^{\vec{0},h} \cdots d_{t-1}^{\vec{0},h}}$. In light of the definition $d_0^{\vec{0},h} = \varphi(h)$, this implies that $d_1^{\vec{0},h} = \varphi(h\varphi(h))$, $d_2^{\vec{0},h} = \varphi(h\varphi(h)\varphi(h\varphi(h)))$, etc. Earlier we obtained this characterization from the desired representation; Lemma B.12 obtains it directly from the axioms.

Construction of an auxiliary space and preference relation

Using the $\varphi(\cdot)$ just constructed, define the transformation $g : H \times C \to \times_{i=0}^{\infty}\mathbf{R}$ by

$$g(h, c) = (c_0 - \varphi(h), c_1 - \varphi(hc_0), c_2 - \varphi(hc_0c_1), \ldots)$$

Let $C^* = g(H \times C)$ and $C_h^* = g(\{h\}, C)$ be the image and projected images of g, respectively.[22] C^* is the set of all *habit-adjusted consumption streams* and C_h^* is the set of all *h-adjusted consumption streams*.

From here on we will be working solely in this auxiliary space C^*, rather than in the actual consumption space C. To understand the structure of C^* we must examine the properties of $g(\cdot,\cdot)$. We prove in Lemma B.17 that $g(\cdot,\cdot)$ is continuous and that $g(h,\cdot)$ is a homeomorphism into C_h^* for each $h \in H$. However, $g(\cdot,\cdot)$ itself fails to be invertible because it violates injectivity. Because each $c^* \in C^*$ is generated from some $h \in H$ and $c \in C$, there is a "natural inverse" given by the inverse of $g(h,\cdot)$,

$$g^{-1}(h, c^*) = (c_0^* + \varphi(h), c_1^* + \varphi(h, c_0^* + \varphi(h)), c_2^* + \varphi\big(h, c_0^* + \varphi(h), c_1^* + \varphi(h, c_0^* + \varphi(h))\big), \ldots) \quad (7)$$

which maps h-adjusted consumption back to regular consumption. The difficulty is that a single c^* may belong to infinitely many sets C_h^*. In fact, Lemma B.18 uses Axiom HC(i) to show that the projection sets are nested: $C_{h'}^* \subseteq C_h^*$ for all $h' \leq h$.

This failure of injectivity plays an important role in the analysis. First, it permits us to prove in Lemma B.19 that C^* is a separable space. Separability will be needed to obtain a utility representation for a continuous preference that we will construct on C^*. To that end, we define an auxiliary relation $\succeq^*$ on $C^* \times C^*$ by $c \succeq_h c'$ if and only if

[22] We endow the space $\times_{i=0}^{\infty}\mathbf{R}$ with the product topology; metrize $H \times C$ by $\rho^{H\times C}((h,c),(h',c')) = \rho^H(h,h') + \rho^C(c,c')$; and consider C^* as a metric subspace of $\times_{i=0}^{\infty}\mathbf{R}$.

$g(h,c) \succeq^* g(h,c')$. This says that the DM prefers the stream c to c' at habit h if and only if she ranks the corresponding habit-adjusted consumption streams likewise. Although $\succeq^*$ is defined through every h, our notation is independent of h. How do we know that $\succeq^*$ is well-defined, i.e., how can we rule out a situation where the DM prefers the h-adjusted consumption stream corresponding to c in one case and the h'-adjusted consumption stream corresponding to c' in the other?

The proof of this result, given in Lemma B.20, offers insight into the workings of endogenously defined auxiliary relations. In essence, $\succeq^*$ is well-defined due to the ability to (i) write C^* as a countable union of overlapping projections and (ii) use the compensating streams to maneuver between the various habits from which $\succeq^*$ may be derived without affecting the preference. To see this, observe that we may rewrite the definition of $\succeq^*$ as $c^* \succeq^* \dot{c}^*$ if and only if $c^*, \dot{c}^* \in C^*_h$ and $g^{-1}(h,c^*) \succeq_h g^{-1}(h,\dot{c}^*)$ for some $h \in H$. Suppose there are h, h' and $c^*, \dot{c}^* \in C^*_h, C^*_{h'}$ with $g^{-1}(h,c^*) \succeq_h g^{-1}(h,\dot{c}^*)$ and $g^{-1}(h',\dot{c}^*) \succ_{h'} g^{-1}(h',c^*)$. Take $\bar{h} \geq h', h$. By nestedness, $c^*, \dot{c}^* \in C^*_{\bar{h}}$. By HC(i),

$$g^{-1}(h,c^*) + d^{h,\bar{h}} \succeq_{\bar{h}} g^{-1}(h,\dot{c}^*) + d^{h,\bar{h}} \text{ and } g^{-1}(h',\dot{c}^*) + d^{h',\bar{h}} \succ_{\bar{h}} g^{-1}(h',c^*) + d^{h',\bar{h}} \qquad (8)$$

But we can show that $g^{-1}(h,c^*) + d^{h,\bar{h}}$ and $g^{-1}(h',c^*) + d^{h',\bar{h}}$ are both equal to $g^{-1}(\bar{h},c^*)$ (and similarly for $\dot{c}^*$), implying that (8) consists of the contradictory statements $a \succeq b$ and $b \succ a$. Hence $\succeq^*$ is well-defined. The other properties required to obtain a continuous utility representation for $\succeq^*$ (completeness, transitivity, and continuity) also hinge on Axiom HC(i) and are proved in Lemma B.21.

Obtaining additivity on the auxiliary space

While the DM's preferences are neither additively separable, monotonic, nor dynamically consistent in a manner independent of history, we can prove that the auxiliary preference relation $\succeq^*$ does satisfy these properties.

While additive separability conditions are normally defined on the consumption set itself, our additive separability conditions must hold for the endogenously defined preference $\succeq^*$ on the endogenously defined space C^*. Due to the transformation between the two spaces, a known separability condition in C^* would translate directly into an unintelligible condition in C. Consequently we impose a simplified condition on C, Axiom HC(ii), which has the flavor of a separability axiom. Given our other axioms, we can show that Axiom HC(ii) implies the separability conditions of Gorman (1968, Theorem 1) in the auxiliary space.[23]

Gorman reformulates the standard conditions for additive separability.[24] His conditions assert that the set of separable indices is closed under unions, intersections, and differences given certain restrictions on the preference and its domain. In combination with our prior results, Lemmas B.22 and B.27 will ensure that Gorman's restrictions are not violated. Using Gorman's theorem and the dynamic consistency of $\succeq^*$ (shown in Lemma B.23), the conditions of Koopmans (1960) will hold on the auxiliary space if $(c^{*t}, \bar{c}^*) \succeq^* (\hat{c}^{*t}, \bar{c}^*)$ iff $(c^{*t}, \bar{\bar{c}}^*) \succeq^* (\hat{c}^{*t}, \bar{\bar{c}}^*)$ for any $t \geq 0$ and $c^*, \hat{c}^*, \bar{c}^*, \bar{\bar{c}}^* \in C^*$ such that $(c^{*t}, \bar{c}^*)$, $(\hat{c}^{*t}, \bar{c}^*)$, $(c^{*t}, \bar{\bar{c}}^*)$, $(\hat{c}^{*t}, \bar{\bar{c}}^*) \in C^*$.

We must show these conditions hold on the entire space C^*. However, they do not correspond directly to Axiom HC(ii) using the "natural inverse" given in (7). Instead, to ensure that HC(ii) implies Gorman's conditions hold *on all of* C^* we must prove that the "compensated consumption" map $\xi_t : H \times C \to C^*$ given by

$$\xi_t(h, c) = g\big(h, (c^{t-1}, {}^t c + d^{h0^t, hc^{t-1}})\big)$$

is surjective. Here the failure of $g(\cdot, \cdot)$ to be injective once again plays the central role. Previously, we noted that $C^*_{h'} \subseteq C^*_h$ for $h \geq h'$. To show that the map ξ_t is surjective we must understand how the projection sets C^*_h and $C^*_{h'}$ may overlap even when h', h cannot be ordered. The requisite proofs are given in Lemmas B.24-B.26.

5 Recovering the habit formation coefficients

In this section we demonstrate how to use a single compensating stream to recover a DM's entire sequence of habit formation coefficients, regardless of her particular felicity, discount factor, or consumption history. In doing so we offer a partial response to the "free parameters" critique as it applies to the model of linear habit formation. While

[23]The only other paper of which we are aware that applies Gorman-type conditions to infinite streams in order to obtain a discounted utility representation is Bleichrodt, Rohde & Wakker (2007), which is unrelated to habit formation.

[24]See Fishburn (1970) for an excellent discussion of the standard additivity theories.

the indifference curves of a DM can determine her utility only up to a strictly increasing transformation, they uniquely determine the entire sequence $\{\lambda_k\}_{k\geq 1}$ of habit formation coefficients for a DM satisfying our axioms.

Proposition 5.1 (Independent recovery)**.** *The DM's entire sequence of habit formation coefficients $\{\lambda_k\}_{k\geq 1}$ can be recovered directly from observed choice behavior. In particular, $\{\lambda_k\}_{k\geq 1}$ may be found using a single compensating stream, without knowledge of the DM's felicity, discount factor, or consumption history.*

The habit formation coefficients may be recovered without regard to the DM's felicity or discount factor because the compensating streams are constructs only of the *arguments* of the DM's utility over streams - not of the utility itself. Moreover, the same choice experiment may be applied regardless of the DM's initial habit.

To see this, suppose that the DM has some arbitrary habit $\tilde{h} \in H$. Choose any strictly positive $q \in Q$ and let $h' = \tilde{h}0$ and $h = \tilde{h}q$.[25] Using dynamic consistency, the choice experiment required to find the compensating stream $d^{h',h}$, in which only the most recent consumption memory is affected, takes a simple form. The compensating stream $d^{h',h}$ is the unique $d \in C$ satisfying

$$(0, c) \succeq_{\tilde{h}} (0, c') \text{ if and only if } (q, c + d) \succeq_{\tilde{h}} (q, c' + d) \text{ for all } c', c \in C \tag{9}$$

Under our axioms, the compensating stream $d^{h',h}$ depends only on $h - h'$; consequently the choice experiment (9) delivers the same compensating stream regardless of the initial habit $\tilde{h}$. Recall that the proof of Theorem 4.1 demonstrates that for any $(h', h) \in \mathcal{H}$ the compensating stream $d^{h',h}$ is characterized recursively by

$$d^{h',h} = \Big(\varphi(h-h'), \varphi\big(h-h', \varphi(h-h')\big), \varphi\big(h-h', \varphi(h-h'), \varphi(h-h', \varphi(h-h'))\big), \dots\Big) \tag{10}$$

Since $d^{h',h}$ is determined independently of the felicity and discount factor, the task of recovering the entire sequence of habit formation coefficients $\{\lambda_k\}_{k\geq 1}$ may now be easily accomplished. Given the linearity of $\varphi(\cdot)$ and the recursive characterization (10), knowing the compensating stream $d^{h',h}$ means that the entire sequence $\{\lambda_k\}_{k\geq 1}$ can be determined

[25]The choice of 0 is arbitrary and could be replaced by any date-0 consumption level $q' \in Q$ with $q' < q$. In the triangular linear system that follows q must then be replaced by $(q - q')$.

straightforwardly from the triangular linear system

$$
\begin{aligned}
d_0^{h',h} &= \lambda_1 q \\
d_1^{h',h} &= \lambda_2 q + \lambda_1 d_0^{h',h} \\
d_2^{h',h} &= \lambda_3 q + \lambda_2 d_0^{h',h} + \lambda_1 d_1^{h',h} \\
d_3^{h',h} &= \lambda_4 q + \lambda_3 d_0^{h',h} + \lambda_2 d_1^{h',h} + \lambda_1 d_2^{h',h} \\
&\vdots
\end{aligned}
$$

6 The autoregressive model and habit decay

We now examine the rate at which habits fade as they become more distant in time, i.e. the rate at which the DM's habit stock decays. As seen in Appendix D, the utility representation in Theorem 4.1 contains too many parameters to offer a succinct measure of habit decay. To that end, we use a frequently invoked *autoregressive* specification of the habit aggregator to facilitate such a characterization. According to that model, there exist $\alpha, \beta > 0$ with $\alpha + \beta \leq 1$ such that the habit aggregator satisfies the autoregressive law of motion $\varphi(hq) = \alpha\varphi(h) + \beta q$ for all $h \in H$ and $q \in Q$.[26] In this section we examine the implications of this simplification for choice behavior. Specifically, we show that the autoregressive structure of the habit aggregator corresponds to an additional axiom that can calibrate the habit decay parameter α in that model.

6.1 A simple choice experiment

Let us consider a DM facing two possible scenarios, A and B, each of which determine today's consumption level only. In scenario A the DM has a very high level of consumption today, whereas in scenario B she has a very low level of consumption today. The DM's preferences tomorrow will differ based on how much she consumes today. We would like to know whether the single consumption level determined by these two scenarios has an irreversible effect on the DM's behavior. That is, if the DM were to consume very little for some time after scenario A and very much for some time after scenario B, could the

[26]This model appears in Boldrin, Christiano & Fisher (1997) in our discrete time form and in Constantinides (1990) and Schroder & Skiadas (2002) in the continuous time version.

opposing effects from consumption cancel so that her preferences following each scenario eventually coincide? The next axiom describes a choice behavior for which such equilibration is possible.

Axiom IE (Immediate Equilibration) For all $c_0, \hat{c}_0 \in Q$, there exist $c_1, \hat{c}_1 \in Q$ such that for all $\bar{c}, \bar{\bar{c}} \in C$, $(c_0, c_1, \bar{c}) \succeq_h (c_0, c_1, \bar{\bar{c}})$ if and only if $(\hat{c}_0, \hat{c}_1, \bar{c}) \succeq_h (\hat{c}_0, \hat{c}_1, \bar{\bar{c}})$.

In light of dynamic consistency, Axiom IE implies that $\succeq_{hc_0c_1}$ and $\succeq_{h\hat{c}_0\hat{c}_1}$ describe the same preferences. Therefore, the axiom says that we can undo by tomorrow the effect of a difference in consumption today.

We would like to use Axiom IE to provide a comparative measure of habit decay. Let us fix any period-0 consumption levels $\hat{c}_0 > c_0$ and consider the corresponding period-1 consumption levels $\hat{c}_1, c_1$ that are given by Axiom IE. Intuitively, if the DM's habits decay slowly then the effects of prior consumption linger strongly, so c_1 will have to be quite large and $\hat{c}_1$ will have to be quite small in order to offset the initial difference. More formally, for fixed $\hat{c}_0 > c_0$ one would expect that the difference $c_1 - \hat{c}_1$ in the period-1 consumption levels required by Axiom IE should be larger for those DM's whose habits decay more slowly. This intuition is confirmed by the following representation theorem, which reveals that Axiom IE implies habits decay at the constant rate $\frac{c_1 - \hat{c}_1}{\hat{c}_0 - c_0}$.

Theorem 6.1 (Autoregressive habit formation)**.** *The family of preference relations $\succeq$ satisfies Axioms PR, C, DC, GM, HC, NDM and IE if and only if each $\succeq_h$ can be represented by $U_h(c) = \sum_{t=0}^{\infty} \delta^t u\big(c_t - \varphi(h^{(t)})\big)$ as in Theorem 4.1 and in addition there exist $\alpha, \beta > 0$ with $\alpha + \beta \leq 1$ such that the habit aggregator $\varphi(\cdot)$ satisfies the autoregressive law of motion*

$$\varphi(hq) = \alpha\varphi(h) + \beta q \quad \forall\, h \in H,\ q \in Q. \tag{11}$$

Moreover, using the values $c_0, c_1, \hat{c}_0, \hat{c}_1$ for any h from Axiom IE, α is given by

$$\alpha = \frac{c_1 - \hat{c}_1}{\hat{c}_0 - c_0}. \tag{12}$$

Remark 6.2. *For finite histories of length $K \geq 3$, the habit aggregator cannot be written in the form (11) but the result of Theorem 6.1 is unchanged: the ratio of successive habit formation coefficients $\frac{\lambda_{k+1}}{\lambda_k}$ is constant and given by $\frac{c_1 - \hat{c}_1}{\hat{c}_0 - c_0}$.*

Theorem 6.1 reveals that the autoregressive model corresponds precisely to the choice behavior of Axiom IE, and that the choice experiment provided directly calibrates the parameter α. The proof of this theorem appears in the Appendix.[27]

Observe that the choice experiment in Axiom IE immediately recovers the single parameter λ in the *geometric coefficients* model, where the aggregator satisfies the law of motion $\varphi(hq) = (1-\lambda)\varphi(h) + \lambda q \; \forall \; h \in H$ and $q \in Q$. By the result in Theorem 6.1 we know that the DM must also satisfy Axiom IE and that

$$\lambda = 1 - \frac{c_1 - \hat{c}_1}{\hat{c}_0 - c_0}, \tag{13}$$

where the values $c_0, c_1, \hat{c}_0, \hat{c}_1$ are found from Axiom IE and are independent of h.

Finally, observe that the autoregressive model has two free parameters, α and β, and that the effect of a prior difference in consumption can be undone in one period. The proof of Theorem 2 actually suggests a more general result. *It can similarly be shown that a generalization of the autoregressive model that has n free parameters corresponds to a generalized $n-1$ period version of equilibration in which it takes $n-1$ periods to undo the effect of a single difference in consumption.*

7 Persistent versus responsive habits

We now return to the general setting of linear habit formation and suggest a universal characterization of the DM's habit-forming tendencies. We distinguish between two types of preferences that satisfy our axioms, those whose habits are *responsive* to weaning and those whose habits are *persistent*.

Recall that Axiom HC sheds light on how the preferences of the DM transform as she goes from a lower habit h' to a higher habit h: the indifference curves for the preference $\succeq_{h'}$ are translated up by $d^{h',h}$ into indifference curves for $\succeq_h$. Stated differently, $d^{h',h}$ measures the distance between the indifference curves of $\succeq_{h'}$ and $\succeq_h$. Whether the DM can be weaned easily using a small and quickly fading stream of compensation, or must be weaned using possibly high levels of consumption that fade slowly - or never at all - will determine

[27] Axiom GM may be relaxed to GS with the obvious relaxation of monotonicity. Also, consider the following alternative to IE: $\forall \; h, \exists \; q \in Q$ s.t. for all $\bar{c}, \bar{\bar{c}} \in C$, $\bar{c} \succeq_h \bar{\bar{c}}$ iff $(q, \bar{c}) \succeq_h (q, \bar{\bar{c}})$. It can be shown by that this axiom may replace Axiom IE in our axiomatization of the autoregressive model. However this alternative axiom does not provide a means to calibrate the parameter α.

how profoundly consumption affects her preferences. In particular, we can characterize the DM's habit-forming tendencies by whether or not the total amount of compensation she requires over time is finite.

Definition 7.1. *The DM is responsive to weaning if she can always be weaned using a finite amount of compensation; that is, for every $(h', h) \in \mathcal{H}$, the total amount $\sum_{t=0}^{\infty} d_t^{h',h}$ is finite. The DM has persistent habits if she can never be weaned using a finite amount of compensation; that is, for every $(h', h) \in \mathcal{H}$, $\sum_{t=0}^{\infty} d_t^{h',h} = \infty$.*

A responsive DM would respond well to a "packaged" drug cessation aid such as the nicotine patch; but if she has persistent habits, no finite number of nicotine patches will cure her of her habit. Note that our definition requires that a DM's total compensation is *always finite* or *always infinite* for every pair of habits $(h', h) \in \mathcal{H}$. But what about a DM who requires a finite amount of total compensation for some pairs of habits, and an infinite amount for others? We show in the following proposition that such a DM cannot exist. Furthermore, we show that the compensation required by a persistent DM not only sums to infinity, but remains forever constant. A similar phenomenon may be seen in heroin addicts weaned using the drug methadone: while the vast majority of methadone patients eventually taper their use, some patients receive a steady dose of the drug for the rest of their lives (Bertschy (1995)).

Proposition 7.2 (The dichotomy)**.** *Suppose the DM's preference $\succeq$ satisfies Axioms PR, C, DC, GS, HC, and NDM. Then the following statements hold.*[28]

(i) The DM's habits must be either responsive or persistent.

(ii) Her habits are responsive to weaning if and only if $\sum_{k=1}^{\infty} \lambda_k < 1$.

(iii) Her habits are persistent if and only if each compensating stream is constant.

Because her total compensation is finite, the compensating streams of a responsive DM must tend to zero. In contrast, the compensation of a persistent DM *never* decreases.

Since a DM who satisfies our axioms has persistent habits if and only if the ratio $\frac{\lambda_{k+1}}{\lambda_k} = 1 - \lambda_1$ for every k, persistent habits correspond directly to the model of linear habit formation with geometric coefficients, in which there exists $\lambda \in (0, 1)$ such that the habit

[28]The proof follows from Lemmas B.15-B.16 in the Appendix.

aggregator $\varphi(\cdot)$ satisfies the law of motion $\varphi(hq) = (1-\lambda)\varphi(h)+\lambda q$ for all $h \in H$ and $q \in Q$. In comparison, consider the more general autoregressive model of linear habit formation axiomatized in Theorem 6.1, which specifies that there exist $\alpha, \beta > 0$ with $\alpha + \beta \leq 1$ such that the habit aggregator satisfies $\varphi(hq) = \alpha\varphi(h) + \beta q$ for all $h \in H$ and $q \in Q$. While it collapses to the geometric coefficients model when $\alpha + \beta = 1$, the autoregressive model provides an additional degree of freedom that accommodates responsive habits when $\alpha + \beta < 1$. As the DM's habit-forming tendencies are quite different under responsive and persistent habits, the choice of $\alpha + \beta$ in this commonly used model should be made with care.

In order to understand how choice behavior differs under responsive and persistent habits, the following proposition examines the rate at which compensation fades when habits are responsive. In light of the translation of indifference curves depicted in Figure 1, the result demonstrates how sharply choice behavior may differ at different habits when the sum $\sum_{k=1}^{\infty} \lambda_k$ is exactly one or very slightly below one, *ceteris paribus*. In contrast to the case of persistent habits, the compensating streams of a DM with responsive habits decrease at least geometrically fast within finite time; and when $h > h'$, the decrease is immediate.

Proposition 7.3 (Geometric decay of compensation)**.** *Let the DM's preference $\succeq$ satisfy PR, C, DC, GS, HC, and NDM. Suppose her habits are responsive, i.e. for some k^* the ratio $\frac{\lambda_{k^*+1}}{\lambda_{k^*}} < 1 - \lambda_1$. Let $\varepsilon = 1 - \lambda_1 - \frac{\lambda_{k^*+1}}{\lambda_{k^*}}$. Then compensation decreases at least at a geometric rate*[29] *for all $t > k^*$: for any $(h', h) \in \mathcal{H}$,*

$$\inf_k \left\{\frac{\lambda_{k+1}}{\lambda_k}\right\} + \lambda_1 \leq \frac{d_t^{h',h}}{d_{t-1}^{h',h}} \leq 1 - \varepsilon\lambda_{k^*} < 1.$$

Moreover, if $h > h'$ then compensation immediately begins to strictly decrease (at least at a geometric rate if $\sup_k \left\{\frac{\lambda_{k+1}}{\lambda_k}\right\} < 1 - \lambda_1$*).*

Proposition 7.3 suggests that persistent habits are a knife-edge case in terms of their implied behavior. As seen from the upper bound $1-\varepsilon\lambda_{k^*}$ on the rate of decrease for $t > k^*$, if $\sum_{k=1}^{\infty} \lambda_k = 1 - \gamma$ then behavior sharply deviates from that of a persistent DM if most of the γ decrease can be attributed to a drop in $\frac{\lambda_{k^*+1}}{\lambda_{k^*}}$ for a small value of k^*. In such a

[29]When $\inf_k \left\{\frac{\lambda_{k+1}}{\lambda_k}\right\}$ is strictly positive compensation decreases at an exactly geometric rate.

case, compensation decreases rapidly very early and the translation of the indifference map between two habits $(h', h) \in \mathcal{H}$ is much milder than it would be if habits were persistent. It may be best to think of persistent habits as a limiting case of a sequence of responsive habits $\{\lambda_k^n\}_{k\geq 1}$, for each of which $\sum_{k=1}^{\infty} \lambda_k^n = 1 - \gamma_n < 1$ and where each γ_n is attributed to a drop in a more distant ratio of habit coefficients.

Finally, we examine the different welfare effects of persistent and responsive habits by considering the DM's valuation for streams that provide at least some subsistence level $q > 0$ in every period. To fix ideas, suppose that whenever the argument of the DM's felicity is negative she suffers a feeling of loss relative to her accustomed level of consumption; conversely, whenever the argument of the felicity is positive she derives some enjoyment from the stream. Can there be an initial habit so high and a consumption stream so "disappointing" that the DM suffers every period? The following proposition shows that the model of linear habit formation implies a certain "resilience of spirit" in eventually overcoming losses. However, while a DM with responsive habits will always "appreciate" a consumption stream that is bounded away from zero, the same cannot be said of a DM with persistent habits, who may be more "fastidious."[30]

Proposition 7.4. *Suppose that the DM's preference $\succeq$ satisfies Axioms PR, C, DC, GS, HC, and NDM. Then,*

(i) (Resilience). *For any $\gamma < 0$, there does not exist a stream $c \in C$ and a habit $h \in H$ such that $c_t - \varphi(hc^{t-1}) \leq \gamma$ for every t.*

(ii) (Fastidiousness). *If the DM's habits are persistent, then for any $q > 0$ there is a habit $h \in H$ and stream $c \in C$ with $c \geq (q, q, \ldots)$ such that $c_t - \varphi(hc^{t-1}) \leq 0$ for every t.*

(iii) (Appreciativeness). *If the DM's habits are responsive, then for any $q > 0$ there does not exist a habit $h \in H$ and a stream $c \in C$ with $c \geq (q, q, \ldots)$ such that $c_t - \varphi(hc^{t-1}) \leq 0$ for every t.*

Proposition 7.4 implies that a DM who satisfies our axioms will ultimately overcome any level of loss, but a consumption stream offering at least some subsistence level $q > 0$

[30]We also note a related result that has implications for the use of felicities satisfying the Inada condition $\lim_{x\to 0} u'(x) = \infty$ when habits are persistent. Indeed, when $\sum_{k=1}^{\infty} \lambda_k = 1$ there cannot exist a stream $c \in C$ and habit $h \in H$ such that habit-adjusted consumption $c_t - \varphi(hc^{t-1})$ is always strictly positive and bounded away from zero.

may not be appreciated by a DM with persistent habits. While a persistent DM may never derive any enjoyment from such a stream, iterated application of Proposition 7.4 implies that a responsive DM will derive enjoyment from it infinitely often. As seen in Appendix C.1, Proposition 7.4 is the root of the different monotonicity requirements for responsive and persistent habits.[31]

8 Multidimensional habit formation

In this section we extend our analysis to a multidimensional framework that can accommodate models of addiction. Moreover, as noted in Deaton (1987), a multidimensional model of habit formation of the type we axiomatize can explain observed patterns of consumption that are irreconcilable with the standard life-cycle model of consumer behavior (e.g., excess sensitivity of consumption to income).

The commodity space is now $Q = \mathbf{R}^n_+$, $n \geq 1$. To fix ideas, we interpret $q \in Q$ as a bundle of goods and permit the DM to form habits over only a subset of these goods. Suppose that $n \geq 2$ and that the DM only forms habits over goods $\{1, 2, \ldots, M\}$, where $1 \leq M \leq n$. For any $m \leq n$ and $q \in Q$, q_m will correspond to the consumption level of the m-th good in the bundle q. We let $Q^m = \{q \in Q \mid q_{\hat{m}} = 0 \; \forall \; \hat{m} \neq m\}$ be the set of bundles offering good m only.

In order to simplify the notation we include all goods in the DM's consumption history, with the understanding that her preferences are independent of her past consumption of the non-habit forming goods. The set of consumption histories is given by $H = \{h \in \times_{k=\infty}^{1} Q \mid \sup_{m \leq n, k \geq 1} \; h_{m.k} < \infty \}$, where $h_{m.k}$ denotes her consumption of the m-th good k periods ago. For any $h \in H$, $h_m = (\ldots, h_{m.2}, h_{m.1})$ denotes the restriction of the history to the m-th good. The set of consumption streams is given by $C = \{c \in \times_{t=0}^{\infty} Q \mid \sup_{m \leq n, t \geq 0} c_{m.t} < \infty\}$, where $c_{m.t}$ denotes the consumption of good m in period t and c_m is the stream $(c_{m.0}, c_{m.1}, \ldots)$. We define the set $C^m = \{c \in C \mid c_{\hat{m}.t} = 0 \; \forall \; \hat{m} \neq m, t \geq 0 \}$ to be the set of consumption streams offering good m only.

As before, $h^{(t)} = hc^{t-1}$ is the date-t consumption history that results from consuming the stream c given the initial habit h. The restriction of $h^{(t)}$ to the m-th good is denoted

[31] The assertions in Proposition 7.4 follow from the proofs of Lemmas B.31, B.33 and C.1. We note the result is independent of Axiom GM, though the interpretation presumes some monotonicity.

$h_m^{(t)}$. The set $\mathcal{H}$ of ordered pairs of habits is defined as before, with a slight modification of the notion of habits agreeing. We say the pair (h', h) *agree on a date* k *and a good* m if $h_{m.k} = h'_{m.k}$. We say (h', h) *agree on a date* k if for every m, (h', h) agree on k and m. Finally, for each $(h', h) \in \mathcal{H}$ it will be convenient to define the set

$$C(h', h) = \{c \in C \mid \text{for all } t \geq 0 \text{ and } m \leq n,\ c_{m.t} > 0 \text{ iff } m \leq M \text{ and } h_m \neq h'_m\}.$$

This is the set of streams where the consumption of the m-th good is strictly positive if and only if it is a habit-forming good for which the good-specific histories differ.

Let us now consider the extension of our axioms to the multidimensional case. Using the newly defined notation, only the statements of Axioms GS, HC, and NDM must be modified. In particular, we need only impose Gains Sensitivity for a single habit-forming good. That is, for some good $m \leq M$ there must be a uniform increase in consumption that leaves the DM either strictly better or strictly worse.

Axiom GS$^\diamond$ There exist $c \in C$, $m \leq M$, and $\alpha \in Q^m$ such that $c + \alpha \not\sim_h c$.

Axiom HC requires modification of parts (i) and (iii). First, a compensating stream is positive only along habit-forming dimensions from which the DM is being weaned. Moreover, we need only require uniqueness of compensating streams that wean the DM from a single habit-forming good. Next, HC(iii) takes on a dual meaning: in addition to compensation satisfying the same condition as before along each good, the axiom posits that compensation is independent across goods.

Axiom HC$^\diamond$ There is a collection $\{d^{h',h}\}_{(h',h)\in\mathcal{H}}$ of streams such that

(i) (Weaning). Each $d^{h',h} \in C(h', h)$, is weakly decreasing, and satisfies

$$c \succeq_{h'} c' \text{ iff } c + d^{h',h} \succeq_h c' + d^{h',h} \forall\, c, c' \in C; \tag{14}$$

and if $d^{h',h} \in C_m$ for some $m \leq M$ then it uniquely satisfies (14) in C_m.

(ii) (Compensated Separability). For any $c, \hat{c} \in C$, $t \geq 0$ and $h' \leq hc^t, h\hat{c}^t$,

$$(c^t, d^{h',hc^t}) \succeq_h (\hat{c}^t, d^{h',h\hat{c}^t}) \text{ iff } (c^t, \bar{c} + d^{h',hc^t}) \succeq_h (\hat{c}^t, \bar{c} + d^{h',h\hat{c}^t}) \quad \forall\, \bar{c} \in C.$$

(iii) (Independence of Irrelevant Habits). Take any $\hat{k}, \hat{m} \geq 1$ and $q_{\hat{m}}$. If the pairs of habits $(h', h), (\hat{h}', \hat{h}) \in \mathcal{H}$ agree on $\hat{k}, \hat{m}$ and satisfy $\hat{h}'_{\hat{m}.k} = \begin{cases} h'_{\hat{m}.k} & k \neq \hat{k} \\ q_{\hat{m}} & k = \hat{k} \end{cases}$ and $\hat{h}_{\hat{m}.k} = \begin{cases} h_{\hat{m}.k} & k \neq \hat{k} \\ q_{\hat{m}} & k = \hat{k} \end{cases}$ for all k, then $d^{h',h}_{\hat{m}} = d^{\hat{h}',\hat{h}}_{\hat{m}}$.

Observe that while the statement of HC(ii) has not changed, it has become a hybrid separability condition that accommodates both the habit-forming and non-habit forming dimensions of consumption.

Next, we impose Axiom NDM only along the habit-forming dimensions.

Axiom NDM$^\diamond$(i) For any $m \leq M$ and $h \in H$, $\lim_{t\to\infty} d^{\vec{0},h0^t}_{m.0} = 0$; (ii) For any $m \leq M$, $q_m > 0$ and k, there exist $(h', h) \in \mathcal{H}$ agreeing on $\mathbb{N} \setminus \{k\}$ with $d^{h',h}_{m.0} = q_m$.

Before presenting the multidimensional representation, we redefine the notions of cyclicity and quasi-cyclicity. For any function $u : \mathbf{R}^n \to \mathbf{R}$ and any $x \in \mathbf{R}^n$, define the restriction $u_{m,x} : \mathbf{R} \to \mathbf{R}$ by $u_{m,x}(\cdot) = u(x_1, \ldots, x_{m-1}, \cdot, x_{m+1}, \ldots, x_n)$. Then, u is *quasi-cyclic in the m-th good* if there exists $x \in \mathbf{R}^n$ such that the restriction $u_{m,x}$ is quasi-cyclic; similarly, u is *cyclic in the m-th good* if there is $x \in \mathbf{R}^n$ such that the restriction $u_{m,x}$ is cyclic.

Given this reformulation of the framework, we have the following extension of our main representation theorem to the multidimensional setting, in which the DM forms independent linear habits over the first M goods.

Theorem 8.1 (Multidimensional habit formation)**.** *The family of preference relations $\succeq$ satisfies Axioms PR, C, DC, GS$^\diamond$, HC$^\diamond$, and NDM$^\diamond$ if and only if each $\succeq_h$ can be represented by*

$$U_h(c) = \sum_{t=0}^{\infty} \delta^t u\Big(c_{1.t} - \varphi_1(h^{(t)}), \ldots, c_{M.t} - \varphi_M(h^{(t)}), c_{M+1.t}, c_{M+2.t}, \ldots, c_{n.t}\Big), \tag{15}$$

with $\delta \in (0, 1)$. For each good $m \leq M$ the habit aggregator $\varphi_m : H \to \mathbf{R}$ is given by

$$\varphi_m(h) = \sum_{k=1}^{\infty} \lambda_{m.k} h_{m.k},$$

where the coefficients $\{\lambda_{m.k}\}_{k\geq 1} \in \mathbf{R}$ *for the* m*-th good are unique and satisfy*

$$\lambda_{m.k} \in (0,1) \text{ and } \frac{\lambda_{m.k+1}}{\lambda_{m.k}} \leq 1 - \lambda_{m.1} \text{ for all } k \geq 1. \tag{16}$$

The felicity $u : \mathbf{R}^n \to \mathbf{R}$ *is continuous, unique up to positive affine transformation, and satisfies the following acyclicity conditions: for each* $m \leq M$, $u(\cdot)$ *is not quasi-cyclic in the* m*-th good if* $\sum_{k=1}^{\infty} \lambda_{m.k} < 1$ *and is not cyclic in the* m*-th good if* $\sum_{k=1}^{\infty} \lambda_{m.k} = 1$.[32]

If there are exactly two habit-forming goods and one regular consumption good, our model is a special linear case of that in Palacios-Huerta (2003), which demonstrates the existence of addiction cycles in a model of multiple independent habits.[33]

9 Extensions

In this paper we have provided theoretical foundations for intrinsic linear habit formation by introducing the device of compensating a DM for giving up her habits. We have also offered separability and monotonicity axioms appropriate for nonseparable preferences. In addition to accommodating consumption histories of various lengths and clarifying the differences between nested specifications of the model, our axiomatic approach is flexible enough to accommodate other incarnations of habit formation. We conclude by briefly discussing some of these extensions.

First, suppose that the DM instead cares about, and forms habits over, the *rate* of growth of her consumption over time. Formally, assume that consumption is bounded below by some small $\varepsilon > 0$ and that the DM forms habits over the *logarithms* of her past consumption levels, $(\ldots, \log h_2, \log h_1)$, rather than over her actual consumption history. Similarly, suppose that her preferences are defined over streams of logarithms of consumption $(\log c_0, \log c_1, \ldots)$. Consider our axioms in this new setting. For example, in Axiom HC(i), the DM would have to be compensated in terms of rates of consumption growth rather than using consumption itself. In terms of the implications of GM, one stream would gains-dominate another if its period-to-period growth rate is larger. Applying our main representation theorem we would obtain a utility representation of habit formation where

[32] The proof is similar to that of Theorem 4.1.

[33] In contrast, cycles in Becker & Murphy (1988) require one dimension to have two stocks.

the DM cares about the ratio of her current consumption level to a *geometric average* of her prior consumption levels. That is, she behaves as if her history dependent utility over *actual* consumption is given by $U_h(c) = \sum_{t=0}^{\infty} \delta^t u\left(\frac{c_t}{\varphi(h^{(t)})}\right)$, where $\varphi(h) = \prod_{k=1}^{\infty} h_k^{\lambda_k}$ and the habit formation coefficients $\{\lambda_k\}_{k\geq 1}$ satisfy the same conditions as before.[34]

Moreover, our general approach can potentially be extended to axiomatize other models of non-linear habit formation. If the definition of weaning is generalized so that compensation may depend on the choice set, then the critical assumption generating linearity is relaxed. One may then presumably generate other models of habit formation by placing the appropriate axioms on compensation and modifying our techniques to obtain a discounted utility representation on the relevant auxiliary space.

Finally, while this paper has addressed riskless choice, our approach can be immediately generalized to expected utility by imposing the von Neumann-Morganstern axioms on lotteries and our axioms on the degenerate lotteries. The prototypical example used to illustrate the failings of expected discounted utility is that DM would be indifferent between a lottery that offers a .5-probability of $(100, 100, \ldots)$ and a .5-probability of $(0, 0, \ldots)$, and another lottery that offers a .5-probability of getting each 0 or 100, independently in every period. Under expected linear habit formation, the DM is not indifferent between these lotteries because they affect her consumption path; she cares about the correlation in the lottery because she is maximizing the expectation of a time-nonseparable utility. Models of habit formation with expected utility offer some separation between risk aversion and the intertemporal elasticity of substitution, thereby avoiding an important pitfall of expected discounted utility.[35]

[34]Such a model is proposed by Kozicki & Tinsley (2002) and is particularly appealing in light of Wendner (2003), which shows the counterintuitive implications of a common model in which the argument of the felicity is current consumption divided by a *linear* habit stock.

[35]For example, see Constantinides (1990) or Boldrin, Christiano & Fisher (1997).

Appendix

A Proof of Proposition 3.3

Let the preference $\dot{\succeq}$ be continuous in the product topology. If $\dot{\succeq}$ respects gains-domination then it clearly satisfies GM*. To show the converse, suppose that c gains-dominates c' and that the preference satisfies GM*. By rearranging the inequality in the definition of gains-domination, we see that $c_t - c'_t \geq c_{t-1} - c'_{t-1}$ for every $t \geq 1$; i.e., the gap between c and c' is widening (since $c \geq c'$). Let $t^1 = \min\{t : c_t > c'_t\}$ and $\alpha^1 = c_{t^1} - c'_{t^1}$. Define c^1 by $c^1_t = c'_t + \alpha^1 \; \forall \; t \geq t^1$ and $c^1_t = c'_t$ otherwise. Note that $c^1 \dot{\succ} c'$ by gains-monotonicity*. For $n \geq 2$, inductively define $t^n = \min\{t : c_t > c^{n-1}_t\}$. So long as $t^n < \infty$, let $\alpha^n = c_{t^n} - c^{n-1}_{t^n}$ and define c^n by $c^n_t = c^{n-1}_t + \alpha^n \; \forall \; t \geq t^n$ and $c^n_t = c^{n-1} \; \forall \; t < t^n$. By GM*, $c^n \dot{\succeq} c^{n-1} \; \forall \; n \geq 2$ and so by transitivity, $c^n \dot{\succ} c^1 \dot{\succ} c'$. Note that gains-domination implies $t^n > t^{n-1}$, and by construction, $c^n_t = c_t \; \forall \; t \leq t^n$. Hence $\lim_{n\to\infty} c^n = c$ in the product topology. Continuity of $\dot{\succeq}$ in the product topology then guarantees the desired result $c \dot{\succeq} c^1 \dot{\succ} c'$.

B Proof of sufficiency in Theorem 4.1

In this section of the Appendix we prove that the axioms PR, C, DC, GM, HC, and NDM imply the desired representation. The axioms PR, C, DC, HC, GS and NDM are implicit in all hypotheses, and we make it clear in the statement of the lemmas when GM is invoked to clarify why Corollary 4.2 follows.

Results about the sequences $d^{h',h}$

Here we establish that each $d^{h',h}$ may be written as an infinite sum of independent one-step transitions, and begin proving that each such summand will satisfy a particular functional equation. Because we will ultimately show that $d^{h',h} = d^{\vec{0},h-h'}$, we save on notation by using d^h whenever $d^{\vec{0},h}$ is intended.

Lemma B.1. *For each h' there is no nonzero $\bar{c} \in C$ such that $c + \bar{c} \succeq_{h'} c' + \bar{c}$ iff $c \succeq_{h'} c'$ for all $c, c' \in C$. Consequently we may define $d^{h,h} = (0, 0, \ldots)$.*

Proof. If there were, then for any $h \geq h'$ both $\bar{c} + d^{h',h}$ and $d^{h',h}$ would compensate from h' to h, violating uniqueness. □

The next lemma provides a useful "triangle equality."

Lemma B.2. *(The Triangle Equality) Let $h'' \geq h' \geq h$. Then $d^{h,h''} = d^{h,h'} + d^{h',h''}$.*

Proof. By application of Weaning, $c \succeq_h c'$ iff $c + d^{h,h'} \succeq_{h'} c' + d^{h,h'}$. Using Weaning again, $c + d^{h,h'} \succeq_{h'} c' + d^{h,h'}$ iff $c + d^{h,h'} + d^{h',h''} \succeq_{h''} c' + d^{h,h'} + d^{h',h''}$. Therefore, $c \succeq_h c'$ iff $c + d^{h,h'} + d^{h',h''} \succeq_{h''} c' + d^{h,h'} + d^{h',h''}$ for arbitrary $c, c' \in C$ and so $d^{h,h'} + d^{h',h''}$ plays the role of $d^{h,h''}$.[36] The result follows by uniqueness of $d^{h,h''}$. □

The following connects compensating for one memory and for an entire habit.

Lemma B.3. *Let $h^0 = h'$ and for every k inductively define h^k by $h^k_k = h_k$ and $h^k_i = h^{k-1}_i$ for all $i \neq k$. Then, for any $(h', h) \in \mathcal{H}$, $d^{h',h} = \sum_{k=1}^{\infty} d^{h^{k-1},h^k}$. Moreover, d^{h^{k-1},h^k} depends only on the tuple (k, h'_k, h_k).*

Remark B.4. *Lemma B.3 implies that we may write d^{k,h'_k,h_k} instead of d^{h^{k-1},h^k}, since the compensation is independent of the values of h' and h on $\mathbb{N} \setminus \{k\}$. Indeed, $d^{k,q',q}$ will denote any transition between any two habit vectors in which only the k-th element is changed from q' to q. We also define one other piece of notation: for any $h \in H$, $q \in Q$ and $k \in \mathbb{N}$, the habit $h^{k,q} \in H$ is defined by $h^{k,q}_k = q$ and $h^{k,q}_t = h_t$ for every $t \neq k$.*

Proof. We prove the lemma in three steps: (1) for any $(h', h) \in \mathcal{H}$, we may write $d^{h',h} = \sum_{k=1}^{\infty} d^{h^{k-1},h^k} + \lim_{K\to\infty} d^{h^K,h}$, where the final term is $(0, 0, \ldots)$ if h', h are eventually identical; (2) each d^{h^{k-1},h^k} is independent of the values of h' and h on $\mathbb{N} \setminus \{k\}$; and (3) $\lim_{K\to\infty} d^{h^K,h} = (0, 0, \ldots)$.

(i) Using iterated application of the triangle equality from Lemma B.2, observe that for habits $(h', h) \in \mathcal{H}$ that eventually agree (WLOG, suppose they agree on $\{t, t+1, \ldots\}$) we have $d^{h',h} = \sum_{k=1}^{t} d^{h^{k-1},h^k}$. Now consider arbitrary $(h', h) \in \mathcal{H}$. For any $K \in \mathbb{N}$ and any $c, c' \in C$, $c \succeq_{h'} c'$ iff $c + \sum_{k=1}^{K} d^{h^{k-1},h^k} \succeq_{h^K} c' + \sum_{k=1}^{K} d^{h^{k-1},h^k}$. But again by Weaning in Axiom HC, $c + \sum_{k=1}^{K} d^{h^{k-1},h^k} \succeq_{h^K} c' + \sum_{k=1}^{K} d^{h^{k-1},h^k}$ iff $c + \sum_{k=1}^{K} d^{h^{k-1},h^k} + d^{h^K,h} \succeq_h c' + \sum_{k=1}^{K} d^{h^{k-1},h^k} + d^{h^K,h}$. Therefore, for arbitrary K, $d^{h',h} = \sum_{k=1}^{K} d^{h^{k-1},h^k} + d^{h^K,h}$. Taking the limit over K, $d^{h',h} = \sum_{k=1}^{\infty} d^{h^{k-1},h^k} + \lim_{K\to\infty} d^{h^K,h}$.

[36] While in the intermediate step Axiom HC does not explicitly guarantee that $d^{h',h''}$ uniquely satisfies $c + d^{h,h'} \succeq_{h'} c' + d^{h,h'}$ iff $c + d^{h,h'} + d^{h',h''} \succeq_{h''} c' + d^{h,h'} + d^{h',h''}$, it suffices that the equivalence hold.

(ii) We now show that each d^{h^{k-1},h^k} is independent of the values of h' and h on $\mathbb{N} \setminus \{k\}$. In fact, we will show that for arbitrary $q' \leq q$ and $(\underline{h}, \bar{h}) \in \mathcal{H}$,

$$d^{\underline{h}^{k,q'},\underline{h}^{k,q}} = d^{\bar{h}^{k,q'},\bar{h}^{k,q}} \text{ if and only if } d^{\underline{h}^{k,q},\bar{h}^{k,q}} = d^{\underline{h}^{k,q'},\bar{h}^{k,q'}}. \tag{17}$$

To see this, use Lemma B.2 to write $d^{\underline{h}^{k,q'},\bar{h}^{k,q}} = d^{\underline{h}^{k,q'},\bar{h}^{k,q'}} + d^{\bar{h}^{k,q'},\bar{h}^{k,q}}$ as well as $d^{\underline{h}^{k,q'},\bar{h}^{k,q}} = d^{\underline{h}^{k,q'},\underline{h}^{k,q}} + d^{\underline{h}^{k,q},\bar{h}^{k,q}}$. Combining these two expressions,

$$d^{\underline{h}^{k,q'},\bar{h}^{k,q'}} - d^{\underline{h}^{k,q},\bar{h}^{k,q}} = d^{\underline{h}^{k,q'},\underline{h}^{k,q}} - d^{\bar{h}^{k,q'},\bar{h}^{k,q}}.$$

This proves the claim (17). Note that $d^{\underline{h}^{k,q'},\underline{h}^{k,q}} = d^{\bar{h}^{k,q'},\bar{h}^{k,q}}$ holds by Axiom HC(iii). Consider arbitrary $h' \leq h$ and k, and recall the definition of h^k. Since h^k and h^{k+1} agree on $\mathbb{N} \setminus \{k\}$, (17) implies that $d^{h^k,h^{k+1}} = d^{\vec{0}^{k,h'_k},\vec{0}^{k,h_k}}$. Hence d^{h^{k-1},h^k} depends only on (k, h'_k, h_k), as claimed.

(iii) Now we show that $\lim_{K\to\infty} d^{h^K,h} = (0, 0, \ldots)$. Since the habits h^K and h agree on $\{1, 2, \ldots, K\}$, iterated application of Axiom HC(iii) implies that for each K, $d^{h^K,h} = d^{h'0^K,h0^K}$. But by the triangle equality, $d^{h',h}$ is decreasing in h'. Hence $d^{h'0^K,h0^K} \leq d^{\vec{0},h0^K}$. Therefore,

$$(0, 0, \ldots) \leq \lim_{K\to\infty} d^{h^K,h} = \lim_{K\to\infty} d^{h'0^K,h0^K} \leq \lim_{K\to\infty} d^{\vec{0},h0^K} = (0, 0, \ldots),$$

where the last equality is due to Axiom NDM and $d^{h',h}$ decreasing in h'. □

Next we use dynamic consistency in conjunction with compensation.

Lemma B.5. *For any $q, \hat{q} \in Q$ and k,*

$$d^{\vec{0}^{k,q},(\vec{0}^{k,\hat{q}},q+d_0^{\vec{0}^{k,\hat{q}}},d_1^{\vec{0}^{k,\hat{q}}},\ldots,d_{k-1}^{\vec{0}^{k,\hat{q}}})} = d^{(\vec{0}^{k,\hat{q}},d_0^{\vec{0}^{k,\hat{q}}},d_1^{\vec{0}^{k,\hat{q}}},\ldots,d_{k-1}^{\vec{0}^{k,\hat{q}}})} = {}^k d^{\vec{0}^{k,\hat{q}}}. \tag{18}$$

Proof. Consider any $c, c' \in C$ such that $(c_0, c_1, \ldots, c_{k-1})$ and $(c'_0, c'_1, \ldots, c'_{k-1})$ are both equal to $(q, 0, 0, \ldots, 0)$. According to Weaning,

$$c \succeq_{\vec{0}} c' \text{ iff } c + d^{\vec{0}^{k,\hat{q}}} \succeq_{\vec{0}^{k,\hat{q}}} c' + d^{\vec{0}^{k,\hat{q}}}. \tag{19}$$

Applying DC to the RHS of (19),

$$c + d^{\vec{0}^{k,\hat{q}}} \succeq_{\vec{0}^{k,\hat{q}}} c' + d^{\vec{0}^{k,\hat{q}}} \text{ iff } {}^k c + {}^k d^{\vec{0}^{k,\hat{q}}} \succeq_{(\vec{0}^{k,\hat{q}}, q + d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})} {}^k c' + {}^k d^{\vec{0}^{k,\hat{q}}}. \tag{20}$$

But again by DC, this time applied to the LHS of (19),

$$c \succeq_{\vec{0}} c' \text{ iff } {}^k c \succeq_{\vec{0}^{k,q}} {}^k c'. \tag{21}$$

Combining expressions (20) and (21) using (19),

$${}^k c \succeq_{\vec{0}^{k,q}} {}^k c' \text{ iff } {}^k c + {}^k d^{\vec{0}^{k,\hat{q}}} \succeq_{(\vec{0}^{k,\hat{q}}, q + d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})} {}^k c' + {}^k d^{\vec{0}^{k,\hat{q}}}. \tag{22}$$

Since both have a q in the k-th place, $\vec{0}^{k,q} \leq (\vec{0}^{k,\hat{q}}, q + d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})$. As ${}^k c$ and ${}^k c'$ are arbitrary, ${}^k d^{\vec{0}^{k,\hat{q}}} = d^{\vec{0}^{k,q},(\vec{0}^{k,\hat{q}}, q + d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})}$. In particular, the choice of c, c' (which depended on q) does not affect $d^{\vec{0}^{k,\hat{q}}}$. This means ${}^k d^{\vec{0}^{k,\hat{q}}} = d^{(\vec{0}^{k,\hat{q}}, d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})}$ as well, giving the desired conclusion. □

Lemma B.6. *For any $q, \hat{q} \in Q$ and k, we have*

$$d_0^{\vec{0}^{k,q}, \vec{0}^{k, q + d_0^{\vec{0}^{k,\hat{q}}}}} = d_0^{\vec{0}^{k, d_0^{\vec{0}^{k,\hat{q}}}}}. \tag{23}$$

Proof. By Lemma B.5,

$$d^{\vec{0}^{k,q},(\vec{0}^{k,\hat{q}}, q + d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})} = d^{(\vec{0}^{k,\hat{q}}, d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})}. \tag{24}$$

for any $q, \hat{q} \in Q$. Using Lemma B.2,

$$d_0^{\vec{0}^{k,q},(\vec{0}^{k,\hat{q}}, q + d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})} = d_0^{\vec{0}^{k,q}, \vec{0}^{k, q + d_0^{\vec{0}^{k,\hat{q}}}}} + d_0^{\vec{0}^{k, q + d_0^{\vec{0}^{k,\hat{q}}}}, (\vec{0}^{k,\hat{q}}, q + d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})}. \tag{25}$$

Similarly,

$$d_0^{(\vec{0}^{k,\hat{q}}, d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})} = d_0^{\vec{0}^{k, d_0^{\vec{0}^{k,\hat{q}}}}} + d_0^{\vec{0}^{k, d_0^{\vec{0}^{k,\hat{q}}}}, (\vec{0}^{k,\hat{q}}, d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})}. \tag{26}$$

Because of Lemma B.3 and the fact that $d^{k,q,q} = (0, 0, \ldots)$,

$$d_0^{\vec{0}^{k, q + d_0^{\vec{0}^{k,\hat{q}}}}, (\vec{0}^{k,\hat{q}}, q + d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})} = d_0^{\vec{0}^{k, d_0^{\vec{0}^{k,\hat{q}}}}, (\vec{0}^{k,\hat{q}}, d_0^{\vec{0}^{k,\hat{q}}}, d_1^{\vec{0}^{k,\hat{q}}}, \ldots, d_{k-1}^{\vec{0}^{k,\hat{q}}})}. \tag{27}$$

Setting equations (25) and (26) equal using (24) and canceling the parts corresponding to (27) gives the desired result. □

Construction of the habit aggregator

We now use the prior results to derive a sequence of functional equations and define the habit aggregator $\varphi : H \to \mathbf{R}$ as an infinite sum of linear functions. We develop properties of the aggregator, including geometric bounds on the habit formation coefficients and a recursive representation of compensation.

For each k define $\varphi_k : Q \to \mathbf{R}$ by $\varphi_k(q) = d_0^{\vec{0}^{k,q}}$ if $q > 0$ and $\varphi_k(0) = 0$. We naturally define $\varphi : H \to \mathbf{R}$ by $\varphi(h) = d_0^h = \sum_{k=1}^{\infty} \varphi_k(h_k)$.

Lemma B.7. *For each k and $q, \hat{q} \in Q$,*

$$d_0^{\vec{0}^{k,q+d_0^{\vec{0}^{k,\hat{q}}}}} = d_0^{\vec{0}^{k,q}} + d_0^{\vec{0}^{k,d_0^{\vec{0}^{k,\hat{q}}}}}. \tag{28}$$

Consequently, each $\varphi_k(\cdot)$ is additive on its image, i.e., for every k,

$$\varphi_k(\varphi_k(q) + q') = \varphi_k(\varphi_k(q)) + \varphi_k(q') \quad \forall\, q, q' \in Q. \tag{29}$$

Proof. To see (28), observe that $d_0^{\vec{0}^{k,q+d_0^{\vec{0}^{k,\hat{q}}}}} = d_0^{\vec{0}^{k,q}} + d_0^{\vec{0}^{k,q},\vec{0}^{k,q+d_0^{\vec{0}^{k,\hat{q}}}}}$ by Lemma B.2 and that $d_0^{\vec{0}^{k,q},\vec{0}^{k,q+d_0^{\vec{0}^{k,\hat{q}}}}} = d_0^{\vec{0}^{k,d_0^{\vec{0}^{k,\hat{q}}}}}$ by Lemma B.6. Then (29) follows by construction. □

Remark B.8. *Equation (29) is a complex functional equation, the solution of which has not yet been completely characterized. A theorem of W. Jarczyk (proved in Jarczyk (1991, pp. 52-61)) asserts that continuous solutions of (29) must be affine. Lemma B.11 examines a weak technical condition which ensures continuity of $\varphi_k(\cdot)$. But first we show how Axiom NDM reduces the functional equation (29) to a simple nonnegativity-restricted Cauchy equation that can be solved.*[37]

Lemma B.9. $\varphi_k(q) = \lambda_k q$ *for some* $\lambda_k > 0$ *and* $\forall\, q \in Q$*; and* $d^{h',h} = d^{\vec{0},h-h'}$ *for every* $(h', h) \in \mathcal{H}$.

[37] A function $f : \mathbf{R}_+ \to \mathbf{R}_+$ satisfies a non-negativity restricted Cauchy equation if for every $x, y \in \mathbf{R}_+$, $f(x + y) = f(x) + f(y)$.

Remark B.10. *Since this lemma proves φ_k is linear, we may reduce the number of arguments of $d^{h',h}$ and apply $\varphi(\cdot)$: i.e., $\varphi(h-h') = d^{h-h'} = d^{h',h}$.*

Proof. By Axiom NDM and construction, $\varphi_k(\cdot)$ is onto Q. Hence (29) is identical to the non-negativity restricted Cauchy equation under the reparametrization $q'' = \varphi_k(q')$. We know $\varphi_k(\cdot)$ is strictly monotone; in combination with Axiom NDM, this implies continuity. Then we know from Aczel & Dhombres (1989, Corollary 9) that φ_k is a strictly increasing and affine function. To see the second result, note for each $q', q \in Q$ and k, $d^{k,q',q} = d^{k,0,q} - d^{k,0,q'}$. But $\varphi_k(q) = d^{k,0,q}$ is linear, so $d^{k,q',q} = d^{k,0,q-q'}$. The result follows immediately from the additive characterization of $d^{h',h}$. $\square$

We now consider an alternative condition forcing continuity of $d^{k,0,q}$ given that it satisfies the functional equation (29). The condition says that the set

$$\{\varepsilon > 0 \mid \lim_{n\to\infty} d^{k,\varepsilon,q^n} \neq d^{k,\varepsilon,q}\}$$

has Lebesgue measure zero for every $q > 0$ and $\{q^n\} \in Q$ with $\lim_{n\to\infty} q^n = q$. We call this condition $(*)$. It differs from the condition that $d^{k,\varepsilon,q}$ is almost everywhere continuous in q for any fixed ε, which is implied by monotonicity in q.

Lemma B.11. *Suppose condition $(*)$ is satisfied. Let the sequence $\{q^n\}_{n\in\mathbb{N}}$ be in Q and let $\lim_{n\to\infty} q^n = q$. Then, for any k, $\lim_{n\to\infty} d^{k,0,q^n} = d^{k,0,q}$ for every $q > 0$ and $\lim_{n\to\infty} d^{k,0,q^n} = (0,0,\ldots)$ for $q = 0$.*

Proof. The triangle equality in Lemma B.2 shows that $d^{h',h}$ is monotonic in each argument and therefore almost everywhere continuous in each. We wish to rule out discontinuities on a set of measure zero. Consider some $q > 0$. We prove that if $\{\hat{q}^n\} \in Q$ is s.t. $\lim_{n\to\infty} \hat{q}^n = \hat{q}$ where $0 < \hat{q} < q$, then $\lim_{n\to\infty} d^{k,\hat{q}^n,q} = d^{k,\hat{q},q}$, since $d^{k,\hat{q}^n,q} = d^{k,0,q} - d^{k,0,\hat{q}^n}$. Suppose instead that $\lim_{n\to\infty} d^{k,\hat{q}^n,q} \neq d^{k,\hat{q},q}$ and choose $\bar{q} > q$. Then $d^{k,\hat{q}^n,\bar{q}} = d^{k,\hat{q}^n,q} + d^{q,\bar{q}}$ implies $\lim_{n\to\infty} d^{k,\hat{q}^n,\bar{q}} \neq d^{k,\hat{q},\bar{q}}$ for every $\bar{q} > q$. Take $\{\bar{q}^n\}$ with $\bar{q}^n \leq \bar{q}$ and $\lim_{n\to\infty} \bar{q}^n = \bar{q}$. For some such $\bar{q}$, large n, and $\varepsilon < \hat{q}$, $d^{k,\hat{q}^n,\bar{q}} = d^{k,\hat{q}^n,\bar{q}^n} + d^{\bar{q}^n,\bar{q}}$ so $\lim_{n\to\infty} d^{k,\hat{q}^n,\bar{q}} = \lim_{n\to\infty} d^{k,\hat{q}^n,\bar{q}^n} = \lim_{n\to\infty} d^{k,\varepsilon,\bar{q}^n} - \lim_{n\to\infty} d^{k,\varepsilon,\hat{q}^n}$. By $\lim_{n\to\infty} d^{k,\hat{q}^n,\bar{q}} \neq d^{k,\hat{q},\bar{q}} = d^{k,\varepsilon,\bar{q}} - d^{k,\varepsilon,\hat{q}}$ then $\lim_{n\to\infty} d^{k,\varepsilon,\bar{q}^n} \neq d^{k,\varepsilon,\bar{q}}$. This is so for a set of ε's of nonzero measure, violating condition $(*)$. Hence $\lim_{n\to\infty} d^{k,0,q^n} = d^{k,0,q}$ if $q > 0$. Now consider the case $\hat{q} = 0$ and $\lim_{n\to\infty} \hat{q}^n = \hat{q}$. Suppose by contradiction that $\lim_{\tilde{q}\to 0} d_0^{k,0,\tilde{q}} = \varepsilon > 0$ (if $\varepsilon = 0$ then the stream is identically zero since it is decreasing). By the functional equation (28), $d_0^{\vec{0}^{k,q+d_0^{\vec{0}^{k,\tilde{q}}}}} = d_0^{\vec{0}^{k,q}} + d_0^{\vec{0}^{k,d_0^{\vec{0}^{k,\tilde{q}}}}}$. Using our prior result for $\hat{q} > 0$ and taking $\tilde{q} \to 0$ on both sides, $d_0^{k,0,q+\varepsilon} = d_0^{k,0,q} + d_0^{k,0,\varepsilon}$. Since $\varepsilon > 0$, take $q \to 0$ on both sides, giving $\lim_{q \to 0} d_0^{k,0,q} = 0 < \varepsilon$. This contradicts the definition of ε. □

The following lemma develops a recursive representation of compensation.

Lemma B.12. *For any $t \geq 0$ and $h \in H$, ${}^t d^h = d^{hd_0^h d_1^h \cdots d_{t-1}^h}$; so in particular, $d_t^h = \varphi(hd_0^h d_1^h \cdots d_{t-1}^h)$.*

Proof. By strong induction. Clearly the lemma is true for $t = 0$: $d^h = d^h$. Assume it is true for every t smaller than some $\hat{t}$ that ${}^t d^h = d^{hd_0^h d_1^h \cdots d_{t-1}^h}$. This implies that ${}^{\hat{t}+1} d^h = {}^1 d^{hd_0^h d_1^h \cdots d_{\hat{t}-1}^h}$. By the inductive hypothesis, and using $hd_0^h d_1^h \cdots d_{\hat{t}-1}^h$ as the habit vector,

$$
{}^1 d^{hd_0^h d_1^h \cdots d_{\hat{t}-1}^h} = d^{hd_0^h d_1^h \cdots d_{\hat{t}-1}^h d_0^{hd_0^h d_1^h \cdots d_{\hat{t}-1}^h}}.
$$

Once more by the inductive hypothesis, $d_{\hat{t}}^h = d_0^{hd_0^h d_1^h \cdots d_{\hat{t}-1}^h}$. Therefore, ${}^{\hat{t}+1} d^h$ is equal to $d^{hd_0^h d_1^h \cdots d_{\hat{t}}^h}$ as desired. □

Lemma B.12 implies $d_1^h = \varphi(h\varphi(h))$, $d_2^h = \varphi(h\varphi(h)\varphi(h\varphi(h)))$, etc. We also use this lemma to prove the geometric decay of the λ's.

Lemma B.13. *Every $d^{k,q,q'}$ is decreasing if and only if $\lambda_1 \in (0,1)$ and*

$$
\frac{\lambda_{k+1}}{\lambda_k} \leq 1 - \lambda_1 \quad \forall\, k \geq 1. \tag{30}
$$

We remark that (30) clearly implies that $\sum_{k=1}^{\infty} \lambda_k \leq 1$. Note that the geometric bounds are found from each $d^{k,q,q'}$ and do not require the first part of Axiom NDM.

Proof. Lemmas B.3, B.9 and B.12 together prove that each $d_t^{k,q',q}$ may be written

$$
d_t^{k,q',q} = (q - q')\lambda_{t+k} + \sum_{i=0}^{t-1} d_i^{k,q',q} \lambda_{t-i}. \tag{31}
$$

Therefore, for $t \geq 1$,

$$
d_{t-1}^{k,q',q} - d_t^{k,q',q} = \sum_{i=0}^{t-2} d_i^{k,q',q} \lambda_{t-i-1} + \lambda_{t-1+k}(q - q') - \sum_{i=0}^{t-1} d_i^{k,q',q} \lambda_{t-i} - \lambda_{t+k}(q - q'). \tag{32}
$$

When $t = 1$, only the term $(\lambda_k - \lambda_k\lambda_1 - \lambda_{k+1})(q - q')$ remains in (32) for each k. Hence, the condition (30) holds if and only if $d_0^{k,q',q} \geq d_1^{k,q',q}$ for every k. Note that this also has the effect of implying $\lambda_1 < 1$, since $\lambda_k > 0$ for every k by Lemma B.9. Now, we show that (30) guarantees that $d_{t-1}^{k,q',q} \geq d_t^{k,q',q}$ for every t. Indeed, rearranging (32) and plugging in from (31), we obtain

$$\begin{aligned}
d_{t-1}^{k,q',q} - d_t^{k,q',q} &= \sum_{i=0}^{t-2} d_i^{k,q',q}[\lambda_{t-i-1} - \lambda_{t-i}] + (q - q')[\lambda_{t-1+k} - \lambda_{t+k}] - \lambda_1 d_{t-1}^{k,q',q} \\
&= \sum_{i=0}^{t-2} d_i^{k,q',q}[\lambda_{t-i-1}(1 - \lambda_1) - \lambda_{t-i}] + (q - q')[\lambda_{t-1+k}(1 - \lambda_1) - \lambda_{t+k}] \\
&\geq 0,
\end{aligned}$$

where the nonnegativity follows from condition (30). □

Lemma B.14. *For any $h \in H$ and t, $\left(\inf_k\{\frac{\lambda_{k+1}}{\lambda_k}\}\right)^t \leq \varphi(h0^t) \leq \left(\sup_k\{\frac{\lambda_{k+1}}{\lambda_k}\}\right)^t \varphi(h)$, with strict inequality holding for the upper bound if there exists k such that $\frac{\lambda_{k+1}}{\lambda_k} < 1 - \lambda_1$ and $h_k \neq 0$.*

Proof. This follows from observing that $\varphi(h0^t) = \sum_{k=t+1}^{\infty} \frac{\lambda_k}{\lambda_{k-t}} \lambda_{k-t} h_{k-t}$. □

We next demonstrate that the compensating sequences are either constant or tend to zero asymptotically (so fast they are infinitely summable), depending on whether $\sum_{k=1}^{\infty} \lambda_k = 1$ or $\sum_{k=1}^{\infty} \lambda_k < 1$, respectively.

Lemma B.15. *For any $h \geq 0$, $\lim_{t\to\infty} \frac{d_t^h}{d_{t-1}^h} < 1$ if $\sum_{k=1}^{\infty} \lambda_k < 1$. In particular, if for some $\varepsilon > 0$ there is k^* such that $\frac{\lambda_{k^*+1}}{\lambda_{k^*}} = 1 - \lambda_1 - \varepsilon$ then $\lim_{t\to\infty} \frac{d_t^h}{d_{t-1}^h} \leq 1 - \varepsilon\lambda_{k^*}$. Consequently, if $\sum_{k=1}^{\infty} \lambda_k < 1$ then for any h, $\sum_{t=1}^{\infty} d_t^h < \infty$ and $\lim_{t\to\infty} d_t^h = 0$.*

Proof. Let $\varepsilon = 1 - \lambda_1 - \frac{\lambda_{k^*+1}}{\lambda_{k^*}}$ and $x_{t,k^*} = \begin{cases} d_{t-1-k^*}^h & \text{if } t > k^* \\ h_{k^*+1-t} & \text{if } t \leq k^* \end{cases}$ Using the recursive construction of Lemma B.12 and the previous lemma,

$$\begin{aligned}
\frac{d_t^h}{d_{t-1}^h} = \frac{\varphi(hd_0^h \cdots d_{t-1}^h)}{d_{t-1}^h} &= \frac{\varphi(hd_0^h \cdots d_{t-2}^h 0) + \lambda_1 d_{t-1}^h}{d_{t-1}^h} \\
&\leq \frac{(1 - \lambda_1)d_{t-1}^h - \varepsilon x_{t,k^*}\lambda_{k^*} + \lambda_1 d_{t-1}^h}{d_{t-1}^h},
\end{aligned}$$

with equality if k^* is the unique k such that $\frac{\lambda_{k+1}}{\lambda_k} < 1 - \lambda_1$. Because we know that $d^h_{t-1-k^*} \geq d^h_{t-1}$ for every $t > k^*$,

$$\frac{d^h_t}{d^h_{t-1}} \leq \frac{(1-\lambda_1)d^h_{t-1} - \varepsilon d^h_{t-1-k^*}\lambda_{k^*} + \lambda_1 d^h_{t-1}}{d^h_{t-1}} = (1-\lambda_1) - \varepsilon \frac{d^h_{t-1-k^*}}{d^h_{t-1}}\lambda_{k^*} + \lambda_1 \leq 1 - \varepsilon\lambda_{k^*}.$$

□

Finally, the compensating streams are constant when $\sum_{k=1}^{\infty} \lambda_k = 1$, as claimed.

Lemma B.16. *If $\sum_{k=1}^{\infty} \lambda_k = 1$ (i.e., $\frac{\lambda_{i+1}}{\lambda_i} = 1 - \lambda_1 \ \forall \ i$), then every compensating stream $d^{h',h}$ is constant.*

Proof. It is easily checked for this case that given any $h \in H$ and $q \in Q$, we have $\varphi(hq) = (1-\lambda_1)\varphi(h) + \lambda_1 q$. Therefore $\varphi(h\varphi(h)) = \varphi(h)$. The claim easily follows from induction and the recursive characterization in Lemma B.12. □

Construction of the continuous preference relation $\succeq^*$

In this subsection we show that Axiom HC permits the construction of a continuous map g from $H \times C$ into an auxiliary space C^*. We also develops critical properties of g that allow us to construct a well-defined continuous preference relation on C^* which preserves the original preference.

We endow the space $\times_{i=0}^{\infty}\mathbf{R}$ with the product topology and define the transformation $g : H \times C \to \times_{i=0}^{\infty}\mathbf{R}$ by $g(h,c) = (c_0 - \varphi(h), c_1 - \varphi(hc_0), c_2 - \varphi(hc_0c_1), \ldots)$. Let $C^* = g(H \times C)$ and $C^*_h = g(\{h\} \times C)$, for any $h \in H$, be the image and projected image under g, respectively. We shall consider C^* to be a metric subspace of $\times_{t=0}^{\infty}\mathbf{R}$, implying that C^* is a metric space in its own right. As a reminder, the spaces H and C are metrized by the sup metric $\rho^H(h,h') = \sup_k |h_k - h'_k|$ and the product metric $\rho^C(c,c') = \sum_{t=0}^{\infty} \frac{1}{2^t}\frac{|c_t - c'_t|}{1+|c_t - c'_t|}$ respectively. We will naturally metrize $H \times C$ by $\rho^{H\times C}((h,c),(h',c')) = \rho^H(h,h') + \rho^C(c,c')$.

Using these metrics it is not difficult to see, as the ensuing lemma asserts, that $g(\cdot,\cdot)$ is jointly continuous in all arguments and that once one fixes an $h \in H$ the map $g(h,\cdot) : C \to C^*_h$ is a homeomorphism (that is, it is continuous and also has a continuous inverse). The continuity in h is a direct consequence of the use of the sup metric on the space H and would fail if the product metric were used instead.[38]

[38]To see this, we can show that $\varphi(\cdot)$ is discontinuous at $\vec{0}$ under the product metric. Take any $\varepsilon > 0$ and

Lemma B.17. *$g(\cdot,\cdot)$ is a continuous mapping; moreover, for any given $h \in H$, $g(h,\cdot)$ is a homeomorphism into C^*_h.*

Proof. The map is continuous in the topology if every component is. Linearity of φ implies that the t-th component can be written as $c_t - \varphi(h0^t) - \sum_{k=1}^{t} \lambda_k c_{t-k}$; as only there is only a finite sum of elements of c in each component, the map is continuous with respect to C. Using the sup metric it is clear that $\varphi(\cdot)$ is continuous with respect to H. The desired joint continuity is evident under the respective metrics. Finally, for any $h \in H$ we can directly exhibit the inverse $g^{-1}(h,\cdot) : C^*_h \to C$. It is the mapping given by the clearly continuous map

$$g^{-1}(h,c^*) = (c^*_0+\varphi(h), c^*_1+\varphi(h, c^*_0+\varphi(h)), c^*_2+\varphi\big(h, c^*_0+\varphi(h), c^*_1+\varphi(h, c^*_0+\varphi(h))\big), \ldots), \quad (33)$$

□

We will often take advantage of the next result characterizing the sets C^*_h.

Lemma B.18. *$C^*_{h'} \subseteq C^*_h$ for any $(h',h) \in \mathcal{H}$.*

Proof. To show that the C^*_h are an increasing sequence, take $(c_0 - \varphi(h'), c_1 - \varphi(h'c_0), c_2 - \varphi(h'c_0c_1), \ldots) \in C^*_{h'}$, so that $(c_0, c_1, c_2, \ldots) \in C$. Consider any $(h',h) \in \mathcal{H}$. It is clear that $c + d^{h',h} \in C$. By Lemma B.12 we know that $d^{h',h} = d^{h-h'} = (\varphi(h-h'), \varphi(h-h', \varphi(h'-h)), \ldots)$. Moreover, φ is affine by Lemma B.9. Using the additivity of φ,

$$\begin{aligned}
(c_0+&\varphi(h-h') - \varphi(h), c_1 + \varphi(h-h', \varphi(h-h')) - \varphi(h, c_0 + \varphi(h-h')), \ldots) \\
&= (c_0 + \varphi(h-h'-h), c_1 + \varphi(h-h'-h, \varphi(h-h') - c_0 - \varphi(h-h')), \ldots) \\
&= (c_0 - \varphi(h'), c_1 - \varphi(h'c_0), c_2 - \varphi(h'c_0c_1), \ldots) \in C^*_h.
\end{aligned}$$

□

We would like to know that C^* is separable, connected and convex. Since H and C are each connected, so is $H \times C$; so C^* is certainly connected, being the continuous image of a connected space. We cannot directly use the well-known result that the continuous image

for $j \in \mathbb{N}$ define h^j by $h^j_k = \begin{cases} 0 & k \neq j \\ \frac{\varepsilon}{\lambda_j} & k = j \end{cases}$. Each $h^j \in H$ since it is bounded, and h^j satisfies $\varphi(h^j) = \varepsilon > 0$. Yet h^j goes to $\vec{0}$ in the product metric.

of a separable space is separable, since H is not separable under the sup metric; and if we were to make H separable by endowing it with the product topology instead, then g would not be continuous with respect to h.

Lemma B.19. *C^* is separable, connected and convex.*

Proof. We have already shown connectedness. To see separability, construct the sequence $\{h^n\}_{n\in\mathbb{Z}}$ by $h^n = (\ldots, n, n, n)$. By Lemma B.18, $C^* = \bigcup_{n\in\mathbb{Z}} C^*_{h^n}$. Since each $g(h^n, \cdot)$ is continuous, each $C^*_{h^n}$ is the continuous image of a separable space and therefore separable. Letting $\tilde{C}^*_{h^n}$ denote the countable dense subset of each $C^*_{h^n}$, we offer $\bigcup_{n\in\mathbb{Z}} \tilde{C}^*_{h^n}$ as a countable dense subset of C^*. Finally, to see that C^* is convex, note that the transformation $g(\cdot,\cdot)$ is linear and that C and H are convex. □

We define a binary relation $\succeq^*$ on $C^* \times C^*$ by

$$g(h, c) \succeq^* g(h, \dot{c}) \text{ iff } c \succeq_h \dot{c}. \tag{34}$$

Lemma B.20. *The relation $\succeq^*$ is well-defined.*

Proof. To see that $\succeq^*$ is well-defined, note that we may alternatively write (34) as $c^* \succeq^* \dot{c}^*$ if and only if

$$(c^*_0 + \varphi(h), c^*_1 + \varphi(h, c^*_0 + \varphi(h)), \ldots) \succeq_h (\dot{c}^*_0 + \varphi(h), \dot{c}^*_1 + \varphi(h, \dot{c}^*_0 + \varphi(h)), \ldots) \tag{35}$$

for some $h \in H$ such that $c^*, \dot{c}^* \in C^*_h$. Suppose by contradiction that for some $h' \neq h$ such that $c^*, \dot{c}^* \in C^*_{h'}$,

$$(\dot{c}^*_0 + \varphi(h'), \dot{c}^*_1 + \varphi(h', \dot{c}^*_0 + \varphi(h')), \ldots) \succ_{h'} (c^*_0 + \varphi(h'), c^*_1 + \varphi(h', c^*_0 + \varphi(h')), \ldots). \tag{36}$$

Take any $\bar{h} \geq h, h'$. Since the $C^*_{\hat{h}}$ are nested by Lemma B.18, $c^*, \dot{c}^* \in C^*_{\bar{h}}$. By application of Weaning to the RHS of (35) and Lemma B.12,

$$\begin{gathered}(c^*_0 + \varphi(h) + \varphi(\bar{h} - h), c^*_1 + \varphi(h, c^*_0 + \varphi(h)) + \varphi(\bar{h} - h, \varphi(\bar{h} - h)), \ldots) \succeq_{\bar{h}} \\ (\dot{c}^*_0 + \varphi(h) + \varphi(\bar{h} - h), \dot{c}^*_1 + \varphi(h, \dot{c}^*_0 + \varphi(h)) + \varphi(\bar{h} - h, \varphi(\bar{h} - h)), \ldots).\end{gathered} \tag{37}$$

Using the affine nature of φ, (37) is equivalent to

$$(c_0^* + \varphi(\bar{h}), c_1^* + \varphi(\bar{h}, c_0^* + \varphi(\bar{h})), \ldots) \succeq_{\bar{h}} (\dot{c}_0^* + \varphi(\bar{h}), \dot{c}_1^* + \varphi(\bar{h}, \dot{c}_0^* + \varphi(\bar{h})), \ldots). \tag{38}$$

Similarly, by applying Weaning to (36) we see that (36) holds if and only if

$$(\dot{c}_0^* + \varphi(\bar{h}), \dot{c}_1^* + \varphi(\bar{h}, \dot{c}_0^* + \varphi(\bar{h})), \ldots) \succ_{\bar{h}} (c_0^* + \varphi(\bar{h}), c_1^* + \varphi(\bar{h}, c_0^* + \varphi(\bar{h})), \ldots),$$

contradicting (38). Therefore $\succeq^*$ is well-defined. □

Now that $\succeq^*$ is well-defined, we show that it is a continuous preference relation.

Lemma B.21. $\succeq^*$ *is a continuous preference relation.*

Proof. The $C^*_{h'}$ are nested by Lemma B.18. Thus for any $c^*, \dot{c}^*, \hat{c}^* \in C^*$, there is $h \in H$ large enough so that $c^*, \dot{c}^*, \hat{c}^* \in C^*_h$. Hence $\succeq^*$ inherits completeness and transitivity over $\{c^*, \dot{c}^*, \hat{c}^*\}$ from $\succeq_h$. As $c^*, \dot{c}^*, \hat{c}^* \in C^*$ were arbitrary, $\succeq^*$ is a preference relation.

To prove that $\succeq^*$ is continuous in the product topology, we will show that the weak upper contour sets are closed; the argument for the weak lower contour sets is identical. Consider any sequence of streams $\{c^{*n}\}_{n\in\mathbb{Z}} \in C^*$ converging to some $c^* \in C^*$ and suppose that there is some $\hat{c}^* \in C^*$ such that $c^{*n} \succeq^* \hat{c}^*$ for all n. Take any h and c such that $g(h, c) = c^*$. By Lemma B.17, g is continuous: for any ε-ball $Y \subset C^*$ around c^* there is a δ-ball $X \subset H \times C$ around (h, c) such that $g(X) \subset Y$. Because the sequence $\{c^{*n}\}$ converges to c^* there is a subsequence $\{c^{*m}\} \in Y$ also converging to c^*. By our use of the sup metric on H we know that any $(h', c') \in X$ must satisfy $h' \leq h + (\delta, \delta, \ldots)$. Then Lemma B.18 ensures that for every m, $c^{*m} \in C^*_{h+(\ldots,\delta,\delta)}$. Take $\bar{h} \geq h + (\ldots, \delta, \delta)$ and large enough that $\hat{c}^* \in C^*_{\bar{h}}$. This permits us to compare all of the corresponding streams in C under $\succeq_{\bar{h}}$. Indeed, using $g^{-1}(\bar{h}, \cdot)$ as defined in (33), take $\bar{c}^m = g^{-1}(\bar{h}, c^{*m})$ for each m, $\bar{c} = g^{-1}(\bar{h}, c^*)$, and $\hat{\bar{c}} = g^{-1}(\bar{h}, \hat{c}^*)$. Using the hypothesis and the definition of $\succeq^*$ we know that $\bar{c}^m \succeq_{\bar{h}} \hat{\bar{c}}$ for every m. Lemma B.17 asserts that $g^{-1}(\bar{h}, \cdot)$ is continuous, so $\bar{c}^m$ converges to $\bar{c}$. Since $\succeq_{\bar{h}}$ is continuous by Axiom C, we know $\bar{c} \succeq_{\bar{h}} \hat{\bar{c}}$. This proves the desired result that $c^* \succeq^* \hat{c}^*$. □

Because $\succeq^*$ is a continuous preference relation and C^* is separable, we note for future reference that the argument for Proposition 3.1 shows that $\succeq^*$ has a continuous representation $U^* : C^* \to \mathbf{R}$.

We now show that $\succeq^*$ exhibits strict preference over some pair of streams that differ only by a single element. The proof relies on Gains Sensitivity.

Lemma B.22. *There exist $q^*, \hat{q}^* \in \mathbf{R}$, $c^* \in C^*$, and $t \in \mathbb{N}$ such that $(c^{*t-1}, q^*, {}^{t+1}c^*) \succ^* (c^{*t-1}, \hat{q}^*, {}^{t+1}c^*)$.*

Proof. Let $\alpha > 0$, $h \in H$, and $c \in C$ be such that $c + \alpha \not\sim_h c$. We make the following observation: since the compensating streams are (weakly) decreasing and for each $\alpha > 0$, $d_0^{\vec{0}\alpha} < \alpha$, we can write any positive constant stream as a staggered sum of streams of the form $(\alpha, d^{\vec{0}\alpha})$. Formally, for any $\alpha > 0$ we can find a sequence of times $t^1 < t^2 < \cdots$ and positive numbers $\alpha > \alpha^1 > \alpha^2 > \cdots$ such that the stream $(\alpha, \alpha, \ldots)$ can be written as the consumption stream created by $(\alpha, d^{\vec{0}\alpha})$ starting at time 0, plus the stream $(\alpha^1, d^{\vec{0}\alpha^1})$ starting at time t^1, plus the stream $(\alpha^2, d^{\vec{0}\alpha^2})$ starting at time t^2, etc. Now, let us suppose by contradiction that $\forall\, q^*, \hat{q}^* \in \mathbf{R}$, $c^* \in C^*$, and $t \in \mathbb{N}$, $(c^{*t-1}, q^*, {}^{t+1}c^*) \sim^* (c^{*t-1}, \hat{q}^*, {}^{t+1}c^*)$. Let $g(h, c) = c^*$ where h, c are given as initally stated. Then $(c^{*t-1}, c_t^* + \alpha, {}^{t+1}c^*) \sim^* c^*$ by hypothesis. By definition, this means that $g^{-1}(h, (c^{*t-1}, c_t^* + \alpha, {}^{t+1}c^*)) \sim_h g^{-1}(h, (c^*))$, or $(c^{t-1}, c_t + \alpha, {}^{t+1}c + d^{\vec{0}\alpha}) \sim_h c$. In light of our initial observation, iterative application of the indifference for $\alpha^1, \alpha^2, \ldots$ and product continuity, this implies that $c + (\alpha, \alpha, \ldots) \sim_h c$, violating Gains Sensitivity. $\square$

Separability conditions for $\succeq^*$

We now prove a stronger version of stationarity than given in Koopmans (1960).

Lemma B.23. *For any $c^*, \dot{c}^* \in C^*$ with $c_0^* = \dot{c}_0^*$, $(c_0^*, {}^1c^*) \succeq^* (c_0^*, {}^1\dot{c}^*)$ iff ${}^1c^* \succeq^* {}^1\dot{c}^*$.*

Proof. To see this, note that $(c_0^*, {}^1c^*) \succeq^* (c_0^*, {}^1\dot{c}^*)$ iff

$$(c_0^* + \varphi(h), c_1^* + \varphi(h, c_0^* + \varphi(h)), \ldots) \succeq_h (c_0^* + \varphi(h), \dot{c}_1^* + \varphi(h, c_0^* + \varphi(h)), \ldots) \tag{39}$$

for some $h \in H$ such that $c^*, \dot{c}^* \in C_h^*$. Since $c^* \in C_h^*$, $c_0^* + \varphi(h) \in Q$. Because $\succeq_h$ satisfies the dynamic consistency axiom DC, the relation in (39) holds iff

$$\begin{aligned}&(c_1^* + \varphi(h, c_0^* + \varphi(h)), c_2^* + \varphi\big(h, c_0^* + \varphi(h), c_1^* + \varphi(h, c_0^* + \varphi(h))\big), \ldots)\\ &\quad \succeq_{h(c_0^* + \varphi(h))} (\dot{c}_1^* + \varphi(h, c_0^* + \varphi(h)), \dot{c}_2^* + \varphi\big(h, c_0^* + \varphi(h), \dot{c}_1^* + \varphi(h, c_0^* + \varphi(h))\big), \ldots).\end{aligned}$$

This means by definition that ${}^1c^* \succeq^* {}^1\dot{c}^*$. $\square$

We now prove that Compensated Separability suffices for the required additive separability conditions to hold; we must show that a certain mapping from C into C^* is surjective, so that the needed conditions apply for all elements of C^*.

Recall Lemma B.18, which showed that the $C^*_{h'}$ are nested. In this case we need an orthogonal result, that even though $\hat{h}0^{t+1} \not\geq hc^t$ it is possible that $\hat{h}$ is sufficiently large to ensure that $C^*_{hc^t} \subseteq C^*_{\hat{h}0^{t+1}}$ (we remind the reader that $c^t \in Q^{t+1}$).

Lemma B.24. *For any $h \in H$, $t \geq 0$ and $(c_0, c_1, \ldots, c_t) \in Q^{t+1}$, there exists $\hat{h} \in H$ large enough that $C^*_{hc^t} \subseteq C^*_{\hat{h}0^{t+1}}$.*

Proof. Since φ is linear and strictly increasing in each component, we may choose $\hat{h} > h$ such that

$$\varphi(\hat{h}0^{t+1}) - \varphi(hc^t) \geq \sum_{s=0}^{t}(1-\lambda_1)^{s+1}c_s. \tag{40}$$

We will now show that we can find $\hat{c} \in C$ such that the claim of the lemma is true. Indeed, choose any $c^* \in C^*_{hc^t}$. Then, there is a $\dot{c} \in C$ such that $g(hc^t, \dot{c}) = c^*$. For it to also be true that $c^* \in C^*_{\hat{h}0^{t+1}}$ it must be that for some $\hat{c} \in C$,

$$\hat{c}_s - \varphi(\hat{h}0^{t+1}\hat{c}^{s-1}) = c^*_s = \dot{c}_s - \varphi(hc^t\dot{c}^{s-1}) \quad \forall s \geq 0, \tag{41}$$

where $c^{-1}, \dot{c}^{-1}$ are ignored for the case $s = 0$. We claim that we can construct $\hat{c}$ by using (41) to recursively define $\hat{c}_s = \varphi(\hat{h}0^{t+1}\hat{c}^{s-1}) + \dot{c}_s - \varphi(hc^t\dot{c}^{s-1})$ for every $s \geq 0$.

We need only show that $\hat{c} \in C$ to complete the proof. To do this we must show two things: first, that $\hat{c}$ is nonnegative, and second, that $\hat{c}$ remains bounded.

(i) $\hat{c}$ is nonnegative: it will suffice to show that $\hat{c} \geq \dot{c}$. For $s = 0$ it is clearly true that $\hat{c}_0 \geq \dot{c}_0$, since we have chosen $\varphi(\hat{h}0^{t+1}) - \varphi(hc^t) \geq 0$ in (40). We will proceed by strong induction, assuming that $\hat{c}_{\hat{s}-1} \geq \dot{c}_{\hat{s}-1}$ for every $\hat{s} \leq s$ for some s. From (41), to prove that $\hat{c}_s \geq \dot{c}_s$ it must be shown that $\varphi(\hat{h}0^{t+1}\hat{c}^{s-1}) - \varphi(hc^t\dot{c}^{s-1}) \geq 0$. In fact, using the linearity of φ, the fact that $\hat{h} \geq h$, and the inductive hypothesis, observe that

$$
\begin{aligned}
\varphi(\hat{h}0^{t+1}\hat{c}^{s-1}) - \varphi(hc^t\dot{c}^{s-1}) &= \varphi\big((\hat{h}-h)0^{t+s+1}\big) + \varphi\big(\vec{0}(\hat{c}_1-\dot{c}_1)\cdots(\hat{c}_{s-1}-\dot{c}_{s-1})\big) + \\
&\qquad \varphi\big(\vec{0}(\hat{c}_0-\dot{c}_0)0^{s-1}\big) - \varphi(\vec{0}c^t0^s) \\
&\geq \varphi\big(\vec{0}(\hat{c}_0-\dot{c}_0)0^{s-1}\big) - \varphi(\vec{0}c^t0^s) \\
&= \varphi\Big(\vec{0}\big(\varphi(\hat{h}0^{t+1}) - \varphi(hc^t)\big)0^{s-1}\Big) - \varphi(\vec{0}c^t0^s) \\
&\geq \lambda_s \sum_{i=0}^{t}(1-\lambda_1)^{i+1}c_i - \sum_{i=0}^{t}\lambda_{s+1+i}c_i \\
&= \sum_{i=0}^{t} c_i[\lambda_s(1-\lambda_1)^{i+1} - \lambda_{s+1+i}] \\
&\geq 0,
\end{aligned}
\tag{42}
$$

where the first inequality comes from the nonnegativity of φ; the equality comes from plugging in for $\hat{c}_0 - \dot{c}_0$ from (41); the second inequality comes from (40) and the linear representation for φ in Lemma B.9; and the last inequality comes from the result in Lemma B.13 that $\frac{\lambda_{s+1+i}}{\lambda_s} \leq (1-\lambda_1)^{i+1}$.

(ii) $\hat{c}$ remains bounded: since $\dot{c} \in C$ it is bounded, so it will suffice to show that the difference between $\hat{c}$ and $\dot{c}$ is bounded. Let us denote by y the quantity $\varphi\big((\hat{h} - h)0^{t+2}\big) + \varphi\big(\vec{0}(\hat{c}_0 - \dot{c}_0)0\big)$. By construction, for every $s \geq 1$, $\hat{c}_s - \dot{c}_s$ is equal to the sum on the RHS of the first equality in (42). By the fading nature of compensation, all terms but $\varphi\big(\vec{0}(\hat{c}_1 - \dot{c}_1)\cdots(\hat{c}_s - \dot{c}_s)\big)$ converge to 0 as s tends to infinity. In fact, for any h and t, $\varphi(h0^t) \leq (1-\lambda_1)^t\varphi(h)$ by Lemma B.14. Consequently, the sum $\varphi\big((\hat{h}-h)0^{t+s+1}\big) + \varphi\big(\vec{0}(\hat{c}_0-\dot{c}_0)0^{s-1}\big)$ is no bigger than $(1-\lambda_1)^{s-1}y$ for any s. Let us drop the negative term $-\varphi(\vec{0}c^t0^s)$ in (42) to obtain an upper bound. By the definition of y, we see that $\hat{c}_1 - \dot{c}_1 \leq y$. We claim that for all $s \geq 1$, $\hat{c}_s - \dot{c}_s \leq y$. The proof proceeds by strong induction. Using the inductive hypothesis, $\hat{c}_s - \dot{c}_s \leq y(1-\lambda_1)^{s-1} + y\sum_{k=1}^{s-1}\lambda s$. But $\sum_{k=1}^{s-1}\lambda s \leq \lambda_1\sum_{k=0}^{s-2}(1-\lambda_1)^k = 1-(1-\lambda_1)^{s-1}$, so we see that $\hat{c}_s - \dot{c}_s \leq y$ as claimed.

This completes the proof of the lemma. □

For each t, define the "compensated consumption" map $\xi_t : H \times C \to C^*$ by

$$\xi_t(h, c) = g\big(h, (c^{t-1}, {}^t c + d^{h0^t, hc^{t-1}})\big). \tag{43}$$

We will wish to show that each ξ_t can "hit" every $c^* \in C^*$ with the appropriate choices of h and c; namely,

Lemma B.25. *Each ξ_t is surjective.*

Proof. Fix any $c^* \in C^*$ and $t \geq 1$. By definition, we already know that there is some $h \in H$ and $c \in C$ such that $g(h, c) = c^*$. That is, for every s, $c_s - \varphi(hc_0c_1 \dots c_{s-1}) = c^*_s$. Let us fix this h and c.

We wish to show that there exist $\hat{h} \in H$ and $\hat{c} \in C$ such that $\xi_t(\hat{h}, \hat{c}) = c^*$, i.e.

$$(\ \hat{c}_0 - \varphi(\hat{h}), \hat{c}_1 - \varphi(\hat{h}\hat{c}_0), \dots, \hat{c}_{t-1} - \varphi(\hat{h}\hat{c}_0 \dots \hat{c}_{t-2}), \hat{c}_t - \varphi(\hat{h}0^t), \hat{c}_{t+1} - \varphi(\hat{h}0^t\hat{c}_t), \dots\ ) = c^*. \tag{44}$$

Because $c^* \in C^*_h$, ${}^t c^* \in C^*_{hc^{t-1}}$. Equation (44) suggests that we must show that ${}^t c^* \in C^*_{\hat{h}0^t}$ for some $\hat{h} \in H$. This is accomplished by Lemma B.24, which provides a $\bar{c}$ and $\hat{h} > h$ s.t. $g(\hat{h}0^t, \bar{c}) = {}^t c^*$. Moreover, since $\hat{h} > h$, $c^* \in C^*_{\hat{h}}$. Therefore, there exists $\bar{\bar{c}} \in C$ such that $g(\hat{h}, \bar{\bar{c}}) = c^*$ and in particular, $g(\hat{h}, \bar{\bar{c}})^{t-1} = c^{*t-1}$. Setting $\hat{c} = (\bar{\bar{c}}^{t-1}, {}^t\bar{c})$, we have $\xi_t(\hat{h}, \hat{c}) = c^*$. □

We shall now demonstrate that additive separability is satisfied.

Lemma B.26. *$\succeq^*$ satisfies the following separability conditions:*
(i) Take any $c^*, \hat{c}^* \in C^*$ with $c^*_0 = \hat{c}^*_0$. Then, for any $\bar{c}^*_0$ s.t. $(\bar{c}^*_0, {}^1c^*), (\bar{c}^*_0, {}^1\hat{c}^*) \in C^*$,

$$(c^*_0, {}^1c^*) \succeq^* (c^*_0, {}^1\hat{c}^*) \text{ iff } (\bar{c}^*_0, {}^1c^*) \succeq^* (\bar{c}^*_0, {}^1\hat{c}^*). \tag{45}$$

(ii) For any $t \geq 0$ and $c^*, \hat{c}^*, \bar{c}^*, \bar{\bar{c}}^* \in C^*$ s.t. $(c^{*t}, \bar{c}^*)$, $(\hat{c}^{*t}, \bar{c}^*)$, $(c^{*t}, \bar{\bar{c}}^*)$, $(\hat{c}^{*t}, \bar{\bar{c}}^*) \in C^*$,

$$(c^{*t}, \bar{c}^*) \succeq^* (\hat{c}^{*t}, \bar{c}^*) \text{ iff } (c^{*t}, \bar{\bar{c}}^*) \succeq^* (\hat{c}^{*t}, \bar{\bar{c}}^*). \tag{46}$$

Proof. The proof of condition (i) follows directly from stationarity of $\succeq^*$, proved in Lemma B.23. The remainder of the proof deals with condition (ii).

Find h large enough so that $(c^{*t}, \bar{c}^*)$, $(\hat{c}^{*t}, \bar{c}^*)$, $(c^{*t}, \bar{\bar{c}}^*)$, $(\hat{c}^{*t}, \bar{\bar{c}}^*) \in C^*_h$. Hence, there exist $\tilde{c}, \tilde{\tilde{c}}, \dot{c}, \ddot{c}$ such that $g(h, \tilde{c}) = (c^{*t}, \bar{c}^*)$, $g(h, \tilde{\tilde{c}}) = (\hat{c}^{*t}, \bar{c}^*)$, $g(h, \dot{c}) = (c^{*t}, \bar{\bar{c}}^*)$, and $g(h, \ddot{c}) = (\hat{c}^{*t}, \bar{\bar{c}}^*)$. Moreover, we must have $\tilde{c}^t = \dot{c}^t$ and $\tilde{\tilde{c}}^t = \ddot{c}^t$.

By Lemma B.25, ξ_t is surjective. We claim that there are $\hat{h}$ and $c, \hat{c}, \bar{c}, \bar{\bar{c}} \in C$ so that

$$\begin{aligned} &\xi_t(\hat{h}, (c^t, \bar{c})) = (c^{*t}, \bar{c}^*),\ \xi_t(\hat{h}, (\hat{c}^t, \bar{c})) = (\hat{c}^{*t}, \bar{c}^*), \\ &\xi_t(\hat{h}, (c^t, \bar{\bar{c}})) = (c^{*t}, \bar{\bar{c}}^*),\ \xi_t(\hat{h}, (\hat{c}^t, \bar{\bar{c}})) = (\hat{c}^{*t}, \bar{\bar{c}}^*). \end{aligned} \tag{47}$$

Recalling the construction in Lemma B.24, choose $\hat{h} > h$ large enough so that

$$\varphi(\hat{h}0^{t+1}) \geq \max\Big\{ \sum_{s=0}^{t} (1-\lambda_1)^{s+1} \tilde{c}_s + \varphi(h\tilde{c}^t), \sum_{s=0}^{t} (1-\lambda_1)^{s+1} \tilde{\tilde{c}}_s + \varphi(h\tilde{\tilde{c}}^t) \Big\}.$$

Now that we have an $\hat{h}$ that will work uniformly for these four streams in C^*, note again from the construction in Lemma B.24 that the required continuation streams depend only on $\tilde{c}^t = \dot{c}^t$ and $\tilde{\tilde{c}}^t = \ddot{c}^t$. Therefore, $\bar{c}$ and $\bar{\bar{c}}$ may be constructed as desired in (47). From the construction at the end of Lemma B.25 and the fact that $\hat{h}$ has been chosen to work uniformly, c and $\hat{c}$ may be chosen to satisfy (47).

Consequently, using (47), the desired result (46) holds if and only if

$$\xi_t(\hat{h}, (c^t, \bar{c})) \succeq^* \xi_t(\hat{h}, (\hat{c}^t, \bar{c})) \text{ iff } \xi_t(\hat{h}, (c^t, \bar{\bar{c}})) \succeq^* \xi_t(\hat{h}, (\hat{c}^t, \bar{\bar{c}})),$$

which, using the definitions of ξ_t in (43) and $\succeq^*$, holds if and only if

$$\begin{aligned} &(c^{t-1}, \bar{c} + d^{h0^t, hc^{t-1}}) \succeq_{\hat{h}} (\hat{c}^{t-1}, \bar{c} + d^{h0^t, h\hat{c}^{t-1}}) \text{ if and only if} \\ &(c^{t-1}, \bar{\bar{c}} + d^{h0^t, hc^{t-1}}) \succeq_{\hat{h}} (\hat{c}^{t-1}, \bar{\bar{c}} + d^{h0^t, h\hat{c}^{t-1}}). \end{aligned}$$

But this is immediately true by Compensated Separability. □

For each subset of indices $\mathbb{K} \subset \mathbb{N}$, we will define the projection correspondences $\iota_{\mathbb{K}} : C^* \rightsquigarrow \times_{i\in\mathbb{K}} \mathbf{R}$ by $\iota_{\mathbb{K}}(\hat{C}^*) = \{x \times_{i\in\mathbb{K}} \mathbf{R} \mid \exists\ c^* \in \hat{C}^* \text{ s.t. } c^*|_{\mathbb{K}} = x\ \}$, where $c^*|_{\mathbb{K}}$ denotes the restriction of the stream c^* to the indices in $\mathbb{K}$ (e.g., $c^*|_{\{3,4\}} = (c^*_3, c^*_4)$). For any $t \geq 0$ we will use C^*_t and ${}^tC^*$ to denote the projected spaces $\iota_{\{t\}}(C^*)$ and $\iota_{\{t,t+1,\ldots\}}(C^*)$, respectively. Since $g(\cdot,\cdot)$ is continuous the projected image C^*_t is connected for every t. Moreover each

C_t^* is separable. It is evident by the arbitrariness of histories used to construct these spaces that for any t, ${}^tC^* = C^*$.

Lemma B.27. *Choose some t and $\hat{c}^* \in {}^tC^*$, and take $c_s^* \in C_s^*$ for every $0 \leq s \leq t$. Then $(c_0^*, c_1^*, \ldots, c_t^*, \hat{c}^*) \in C^*$.*

Proof. Pick $\hat{h} \in H$ and $\hat{c} \in C$ such that $\hat{c}^* \in C^*_{\hat{h}\hat{c}^t}$, and let $\tilde{c}*^t = g(\hat{h}, \hat{c})|_{\{0,1,\ldots,t\}}$. Choose any $\varepsilon \geq \max\{0, \max_{0 \leq i \leq t} \frac{\tilde{c}_i^* - c_i^*}{\sum_{k=i+1}^{\infty} \lambda_k}\}$ and set $h = \hat{h} + (\ldots, \varepsilon, \varepsilon)$. Recall the inverse function $g^{-1}(h, \cdot)$, which takes an element of C^* and returns an element of C. We do not yet know that $(c_0^*, c_1^*, \ldots, c_t^*, \hat{c}^*) \in C^*$, but we demonstrate that applying the transformation used in $g^{-1}(h, \cdot)$ to $(c_0^*, c_1^*, \ldots, c_t^*, \hat{c}^*)$ returns a nonnegative stream. Let us take $c^t = g^{-1}(h, (c_0^*, c_1^*, \ldots, c_t^*, \hat{c}^*))|_{\{0,1,\ldots,t\}}$. Since the $C^*_{h'}$ are nested and $h \geq \hat{h}$, it will suffice to prove that $c^t \geq \hat{c}^t$, for then $\hat{c}^* \in C^*_{hc^t}$ and there is a $\bar{c} \in C$ such that $g(h, (c^t, \bar{c})) = (c_0^*, c_1^*, \ldots, c_t^*, \hat{c}^*)$. Using the transformation, $c^t \geq \hat{c}^t$ if and only if $c_0^* + \varphi(h) \geq \tilde{c}_0^* + \varphi(\hat{h})$, $c_1^* + \varphi(hc_0) \geq \tilde{c}_1^* + \varphi(\hat{h}\hat{c}_0)$, up through $c_t^* + \varphi(hc_0 \ldots c_{t-1}) \geq \tilde{c}_t^* + \varphi(\hat{h}\hat{c}_0 \ldots \hat{c}_{t-1})$. But this can be seen through forward induction by using the choice of ε, definition of h, and the fact that φ is linear and increasing in each component. $\square$

The previous discussion has proved that $C^* = C_1^* \times C_2^* \times C_3^* \times C^*$ and that $\succeq^*$ is continuous and sensitive (stationarity then implies essentiality of all periods). Hence C^* is a product of separable and connected topological spaces. Suppose we could also show that for any $c^*, \hat{c}^* \in C^*$,

$$\begin{aligned}
&(c_0^*, c_1^*, c_2^*, {}^3c^*) \succeq^* (c_0^*, \hat{c}_1^*, \hat{c}_2^*, {}^3\hat{c}^*) \text{ iff } (\hat{c}_0^*, c_1^*, c_2^*, {}^3c^*) \succeq^* (\hat{c}_0^*, \hat{c}_1^*, \hat{c}_2^*, {}^3\hat{c}^*),\\
&(c_0^*, c_1^*, c_2^*, {}^3c^*) \succeq^* (\hat{c}_0^*, c_1^*, \hat{c}_2^*, {}^3\hat{c}^*) \text{ iff } (c_0^*, \hat{c}_1^*, c_2^*, {}^3c^*) \succeq^* (\hat{c}_0^*, \hat{c}_1^*, \hat{c}_2^*, {}^3\hat{c}^*),\\
&(c_0^*, c_1^*, c_2^*, {}^3c^*) \succeq^* (\hat{c}_0^*, \hat{c}_1^*, c_2^*, {}^3\hat{c}^*) \text{ iff } (c_0^*, c_1^*, \hat{c}_2^*, {}^3c^*) \succeq^* (\hat{c}_0^*, \hat{c}_1^*, \hat{c}_2^*, {}^3\hat{c}^*), \text{ and}\\
&(c_0^*, c_1^*, c_2^*, {}^3c^*) \succeq^* (\hat{c}_0^*, \hat{c}_1^*, \hat{c}_2^*, {}^3c^*) \text{ iff } (c_0^*, c_1^*, c_2^*, {}^3\hat{c}^*) \succeq^* (\hat{c}_0^*, \hat{c}_1^*, \hat{c}_2^*, {}^3\hat{c}^*).
\end{aligned} \tag{48}$$

Then, for the case $n = 4$, we would be able to apply Debreu's well-known theorem on additive separability for $n \geq 3$ (see Fishburn (1970, Theorem 5.5) to conclude that there exist $u_0, u_1, u_2 : \mathbf{R} \to \mathbf{R}$ and $U_3 : C^* \to \mathbf{R}$ (all continuous and unique up to a similar positive linear transformation) such that $c^* \succeq^* \hat{c}^*$ iff $u_0(c_0^*) + u_1(c_1^*) + u_2(c_2^*) + U_3({}^3c^*) \geq u_0(\hat{c}_0^*) + u_1(\hat{c}_1^*) + u_2(\hat{c}_2^*) + U_3({}^3\hat{c}^*)$.

Instead of (48), Lemma B.26 proves the statement (46). We now utilize the result of

Gorman (1968, Theorem 1), which requires that each of C_1^*, C_2^*, C_3^* and C^* be arc-connected and separable. We have already shown separability; and arc-connectedness follows from being a path-connected Hausdorff space (a convex space is path-connected, and a metric space is Hausdorff). Then, Gorman's Theorem 1 asserts that the set of separable indices is closed under unions, intersections, and differences. Condition (46) implies separability of $\{(1),(2)\}$ and stationarity implies separability of $\{(2,3,4,\ldots)\}$ and $\{(3,4,5,\ldots)\}$.[39] Iterated application of Gorman's theorem implies (48).

$\succeq_h$ can be represented as in (1)

We have shown that $\succeq^*$ satisfies axioms similar to those of Koopmans (1960). To prove $\succeq$ satisfies (1) we first obtain an discounted utility representation for $\succeq^*$ using a related approach and then use the appropriate transformation.

Lemma B.28. *$U_h(c) = \sum_{t=0}^{\infty} \delta^t u(c_t - \sum_{k=1}^{\infty} \lambda_k h_k^t)$ (where h^t is the evolving habit vector) for some continuous u and $\delta \in (0,1)$, with the given restrictions on $\{\lambda_k\}$.*

Proof. $\succeq^*$ is a continuous, stationary, and sensitive preference relation; and satisfies (48), so it can be represented in the form $u_0(\cdot) + u_1(\cdot) + u_2(\cdot) + U_3(\cdot)$ on the space $C^* = C_0^* \times C_1^* \times C_2^* \times C^*$, with the additive components continuous and unique up to a similar positive affine transformation. Condition (48) clearly also implies additive representability on $C^* = C_0^* \times C_1^* \times C^*$, with the additive components again unique up to a similar positive linear transformation. By stationarity, $u_0(\cdot) + u_1(\cdot) + [u_2(\cdot) + U_3(\cdot)]$ and $u_1(\cdot) + u_2(\cdot) + U_3(\cdot)$ are both additive representations on $C^* = C_0^* \times C_1^* \times C^*$. Thus, $\exists\ \delta > 0$ and $\beta_1, \beta_2, \beta_3 \in \mathbf{R}$ s.t. $u_1(\cdot) = \delta u_0(\cdot) + \beta_1$, $u_2(\cdot) = \delta u_1(\cdot) + \beta_2 = \delta^2 u_0(\cdot) + \delta\beta_1 + \beta_2$, and for any $c^* \in C^*$, $U_3(c^*) = \delta[u_2(c_0^*) + U_3({}^1c^*)] + \beta_3 = \delta^3 u_0(c_0^*) + \delta U_3({}^1c^*) + \beta_3 + \delta\beta_2 + \delta^2\beta_1$. Each $c \in C$ and $h \in H$ is bounded and $\sum_{k=1}^{\infty} \lambda_k \leq 1$, so for each $c^* \in C^*\ \exists\ \bar{x}, \underline{x} \in \mathbf{R}$ such that $\underline{x} \leq c_t^* \leq \bar{x}\ \forall\, t$. By Tychonoff's theorem $[\underline{x}, \bar{x}]^\infty$ is compact in $\times_{i=0}^{\infty}\mathbf{R}$ and therefore $[\underline{x}, \bar{x}]^\infty \cap C^*$ is compact in C^*. Given $\underline{x}$ and $\bar{x}$, continuity of $u_0(\cdot)$ and $U_3(\cdot)$ ensures they remain uniformly bounded on $[\underline{x}, \bar{x}]$ and $[\underline{x}, \bar{x}]^\infty \cap C^*$, respectively. Using iterative substitution $U^*(c^*) = \sum_{t=0}^{\infty} \delta^t u(c_t^*)$, where $u(\cdot) = u_0(\cdot)$ is continuous and $\delta \in (0,1)$ by product continuity. To represent $\succeq_h$ as in (1) we then transform each $c \in C$ by $g(h,\cdot)$ into an argument of U^*. □

[39] Because (46) hold for all t it is an even stronger hypothesis than necessary; also, for any t, $\{(t, t+1, t+2, \ldots)\}$ is strictly sensitive by dynamic consistency.

The felicity u is increasing

We have shown that $\succeq^*$ is represented by $U^*(c^*) = \sum_{t=0}^{\infty} \delta^t u(c_t^*)$ for some continuous $u : \mathbf{R} \to \mathbf{R}$ and $\delta \in (0,1)$; and that we may derive from this the desired representation (1) for $\succeq_h$. We now use Axiom GM to put more structure on u.

Lemma B.29. *Under Axiom GM the felicity u is an increasing function.*

Before proving this we will prove the following lemma.

Lemma B.30. $\forall\, T \in \mathbb{N}$, $\vec{x} = (x_0, x_1, \ldots, x_T) \in \mathbf{R}^{T+1}$, $\exists\, h_{\vec{x},T} \in H$ *such that*

$$(x_0, x_1, \ldots, x_T, 0, 0, \ldots) \in C^*_{h_{\vec{x},T+1}}.$$

Proof. For arbitrary h, define c^h by $c_0^h = x_0 + \varphi(h)$, $c_t^h = x_t + \varphi(hc_0^h c_1^h \cdots c_{t-1}^h)$ for all $1 \leq t \leq T$, and $c_t^h = \varphi(hc_0^h c_1^h \cdots c_{t-1}^h)$ for $t > T$. φ is strictly increasing, so we may choose $h_{\vec{x},T} \in H$ sufficiently large so that $(c_0^{h_{\vec{x},T}}, c_1^{h_{\vec{x},T}}, \ldots c_T^{h_{\vec{x},T}})$ is nonnegative. But if $(c_0^{h_{\vec{x},T}}, c_1^{h_{\vec{x},T}}, \ldots c_T^{h_{\vec{x},T}})$ is nonnegative, then so is ${}^{T+1}c^{h_{\vec{x},T}}$. Moreover, the stream is ultimately weakly decreasing. Therefore $c^{h_{\vec{x},T}} \in C$. □

We may now prove Lemma B.29.

Proof. Suppose that u is not increasing. Because u is a continuous function, there must then exist some $x \in \mathbf{R}$ and $\alpha > 0$ such that $\forall\, \alpha' \in (0, \alpha]$, $u(x + \alpha') < u(x)$.

Let T be arbitrary for the moment. Note that by Lemma B.30 there is h' such that $(x, x, \ldots, x, 0, 0, \ldots) \in C^*_{h'}$ (where x is repeated $T+1$ times). Again by Lemma B.30 there is h'' such that $(x + \alpha, x, x, \ldots, x, 0, 0, \ldots) \in C^*_{h''}$ (where x by itself is repeated T times). Let $h \geq h', h''$, and recall that the $C^*_{\hat{h}}$ are nested. Using the representation for $\succeq^*$ and the fact that $u(x + \alpha) < u(x)$,

$$u(x) + \sum_{t=1}^{T} \delta^t u(x) + \sum_{t=T+1}^{\infty} \delta^t u(0) > u(x + \alpha) + \sum_{t=1}^{T} \delta^t u(x) + \sum_{t=T+1}^{\infty} \delta^t u(0). \tag{49}$$

Since $(x, x, \ldots, x, 0, 0, \ldots) \in C^*_h$, there is $c \in C$ with $g(h, c) = (x, x, \ldots, x, 0, 0, \ldots)$. Clearly $c + \alpha \in C$, and by GM we know $c + \alpha \succ_h c$. Moreover, $g(h, c + \alpha)$ is

$$(x + \alpha, x + \alpha(1 - \lambda_1), \ldots, x + \alpha(1 - \sum_{k=1}^{T} \lambda_k), \alpha(1 - \sum_{k=1}^{T+1} \lambda_k), \alpha(1 - \sum_{k=1}^{T+2} \lambda_k), \ldots), \tag{50}$$

where x appears $T+1$ times. Therefore, by the representation theorem for $\succeq^*$,

$$u(x+\alpha)+\sum_{t=1}^{T}\delta^t u\big(x+\alpha(1-\sum_{k=1}^{t}\lambda_k)\big)+\sum_{t=T+1}^{\infty}\delta^t u\big(\alpha(1-\sum_{k=1}^{t}\lambda_k)\big)>\sum_{t=0}^{T}\delta^t u(x)+\sum_{t=T+1}^{\infty}\delta^t u(0). \tag{51}$$

Combine the RHS of (49) and the LHS of (51); and rearrange by subtracting the RHS of (49) from all sides of the inequalities. This obtains

$$\sum_{t=1}^{T}\delta^t\big[u\big(x+\alpha(1-\sum_{k=1}^{t}\lambda_k)\big)-u(x)\big]+\sum_{t=T+1}^{\infty}\delta^t\big[u\big(\alpha(1-\sum_{k=1}^{t}\lambda_k)\big)-u(0)\big]>u(x)-u(x+\alpha), \tag{52}$$

which is strictly positive. Since each $\lambda_k>0$ and $\sum_{k=1}^{\infty}\lambda_k\le 1$, we know that the value $\alpha(1-\sum_{k=1}^{t}\lambda_k)\in[0,\alpha)$ for every t, and is in fact strictly positive as $t<\infty$. The assumption that u dips below $u(x)$ just to the right of x implies

$$\sum_{t=1}^{T}\delta^t\big[u\big(x+\alpha(1-\sum_{k=1}^{t}\lambda_k)\big)-u(x)\big]<0.$$

This sum decreases in T. By continuity, u is bounded on $[0,\alpha]$. Choose T large enough so that $\sum_{t=T+1}^{\infty}\delta^t\big[u\big(\alpha(1-\sum_{k=1}^{t}\lambda_k)\big)-u(0)\big]$ is small enough to bring about the contradiction $0>0$ from (52). This is possible because Lemma B.30 permits us to find h large enough so that the constructed streams are in C^*. □

Finally, we prove the strict monotonicity of $u(\cdot)$ on the relevant ranges.

Lemma B.31. *Assume Axiom GM. If $\sum_{k=1}^{\infty}\lambda_k<1$ then $u(\cdot)$ is strictly increasing on $(0,\infty)$; and if $\sum_{k=1}^{\infty}\lambda_k=1$ then there is a with $0<a\le\infty$ such that $u(\cdot)$ is strictly increasing either on $(-a,\infty)$ or on $(-\infty,a)$.*

Proof. By Lemma B.29 we know that $u(\cdot)$ is an increasing function. To prove it is strictly increasing on the relevant ranges we will consider the two cases separately.

(i) $\sum_{k=1}^{\infty}\lambda_k=1$: First we will show that $u(\cdot)$ is strictly increasing in some interval around 0. To complete the proof, we will show that there cannot exist $x>0>y$ such that $u(\cdot)$ does not increase strictly at both x and y. To see the first point, take any $q>0$ and let $h=(\ldots,q,q)$ and $c=(q,q,\ldots)$. Then $g(h,c)=(0,0,\ldots)$ and for small α, both $c+\alpha \succ_h c$ and $c \succ_h c-\alpha$ by Axiom GM. Using the representation for $\succeq^*$,

$$\sum_{t=0}^{\infty} \delta^t u\big(\alpha(1-\sum_{k=1}^{t}\lambda_k)\big) > \sum_{t=0}^{\infty}\delta^t u(0) > \sum_{t=0}^{\infty}\delta^t u\big(-\alpha(1-\sum_{k=1}^{t}\lambda_k)\big).$$

By monotonicity of $u(\cdot)$ it must be that $u(\cdot)$ increases strictly in a neighborhood of 0. For the second point, suppose by contradiction that there exist $x > 0 > y$ such that $u(\cdot)$ does not increase strictly at both x and y. By continuity and monotonicity of $u(\cdot)$ there is $\alpha > 0$ such that $u(\cdot)$ is constant on $(x, x+\alpha)$ and on $(y, y+\alpha)$. Without loss of generality suppose that x, y are rational (else take some rational x, y inside the interval). Since x, y are rational there exist m, n such that $mx = -ny$. Let $c^* = (x^m, y^n, x^m, y^n, \ldots)$ (i.e., x is repeated m times, then y is repeated n times, etc). Because the compensating streams are constant, we may use the characterization (53) in Lemma B.32 to find $h \in H$ large enough so that there is $c \in C$ satisfying $g(h, c) = c^*$. Observe by GM that $c + \frac{\alpha}{2} \succ_h c$, a contradiction to the assumption that $u(\cdot)$ is constant on $(x, x+\alpha)$ and $(y, y+\alpha)$.

(ii) $\sum_{k=1}^{\infty}\lambda_k < 1$: Observe that in this case, for any $q \in Q$, if we set $h = (\ldots, q, q)$ and $c = (q, q, \ldots)$, then $g(h, q) = (q[1-\sum_{k=1}^{\infty}\lambda_k], q[1-\sum_{k=1}^{\infty}\lambda_k], \ldots)$. As q is arbitrary, we may conclude that for any $x \geq 0$, $(x, x, x, \ldots) \in C^*$. Suppose to the contrary that $u(\cdot)$ is not increasing from the right at x. Since $u(\cdot)$ is continuous and weakly increasing, this implies that there exists some $\beta^+ > 0$ such that for every $0 < \beta \leq \beta^+$, $u(x+\beta) = u(x)$. Let c and h be such that $g(h, c) = (x, x, x, \ldots)$. By GM, $c + \beta \succ_h c$, so using the representation for $\succeq^*$, $\sum_{t=0}^{\infty}\delta^t u\big(x + \beta(1-\sum_{k=1}^{t}\lambda_k)\big) > \sum_{t=0}^{\infty}\delta^t u(x)$. However, since $0 < \beta \leq \beta^+$ and $\sum_{k=1}^{t}\lambda_k < 1$, $u\big(x+\beta(1-\sum_{k=1}^{t}\lambda_k)\big) = u(x)$ for every $t \geq 0$. This is a contradiction. □

The felicity u is not (quasi-)cyclic

Recall the definitions of a cyclic and quasi-cyclic function. To complete the proof of sufficiency we demonstrate that u must satisfy the desired acyclicity properties.

Lemma B.32. *u is not quasi-cyclic if $\sum_{k=1}^{\infty}\lambda_k < 1$ and not cyclic if $\sum_{k=1}^{\infty}\lambda_k = 1$.*

To prove this, we first prove an auxiliary result. For technical convenience the following lemma allows an extension of the definition of compensation to negative "histories;" hence if $\gamma < 0$ then $d^{(\vec{0},\gamma)} = -d^{(\vec{0},-\gamma)}$.

Lemma B.33. *Consider any sequence* $\{\gamma_t\}_{t\in\mathbb{N}}$ *and* $h \in H$. *If* $\bar{c} \in \times_{t=0}^{\infty}\mathbf{R}$ *satisfies* $\bar{c}_t = \varphi(h\bar{c}^{t-1}) + \gamma_t$ *for every* t *then each* $\bar{c}_t$ *may be alternatively written as*

$$\bar{c}_t = \gamma_t + d_t^h + \sum_{s=0}^{t-1} d_s^{\vec{0}\gamma_{t-s-1}}. \tag{53}$$

Proof. The lemma is clearly true for $t = 0$. Suppose that Equation (53) holds for every $t \leq T-1$. Then

$$\begin{aligned}
\bar{c}_T &= \gamma_T + \varphi(h\bar{c}^{T-1}) \\
&= \gamma_T + \varphi(h, \gamma_0 + d_0^h, \gamma_1 + d_1^h + d_0^{\vec{0}\gamma_0}, \ldots, \gamma_{T-1} + d_{T-1}^h + \sum_{s=0}^{T-2} d_s^{\vec{0}\gamma_{T-s-2}}) \\
&= \gamma_T + \varphi(hd_0^h \cdots d_{T-1}^h) + \sum_{s=0}^{T-1} \varphi(\vec{0}\gamma_s d_0^{\vec{0}\gamma_s} \cdots d_{T-2-s}^{\vec{0}\gamma_s}) \\
&= \gamma_T + d_T^h + \sum_{s=0}^{T-1} d_s^{\vec{0}\gamma_{T-s-1}},
\end{aligned}$$

where the second-to-last equality follows from using the recursive characterization given in Lemma B.12 and reversing the order of summation. □

Proof. (*Lemma B.32*) The two cases are examined separately.

(i) $\sum_{k=1}^{\infty} \lambda_k < 1$. Suppose that u is quasi-cyclic, so there exists $\gamma, \beta > 0$ and $\alpha \in \mathbf{R}$ such that $u(x+\gamma) = \beta u(x) + \alpha$ for every $x \in \mathbf{R}$. Apply Lemma B.33 with $\gamma_t = \gamma$ for every t and recall the summability of per-period compensation from Lemma B.15. These results imply that $\bar{c}$ as defined in Lemma B.33 remains bounded, i.e. $\bar{c} \in C$. Moreover $\bar{c}_0 = \gamma$, so c is nonzero. We claim that this $\bar{c}$ is exactly the consumption stream $\bar{c}$ ruled out in Lemma B.1, a contradiction. Indeed, by our additive representation $c + \bar{c} \succeq_h c' + \bar{c}$ if and only if

$$\sum_{t=0}^{\infty} \delta^t u\big(c_t + \bar{c}_t - \varphi(hc^{t-1}) - \varphi(\vec{0}\bar{c}^{t-1})\big) \geq \sum_{t=0}^{\infty} \delta^t u\big(c'_t + \bar{c}_t - \varphi(hc'^{t-1}) - \varphi(\vec{0}\bar{c}^{t-1})\big).$$

Consider the t-th term $u\big(c_t + \bar{c}_t - \varphi(hc^{t-1}) - \varphi(\vec{0}\bar{c}^{t-1})\big)$. By construction of $\bar{c}$, this term is equal to $u\big(c_t - \varphi(hc^{t-1}) + \gamma\big) = \beta u(c_t - \varphi(hc^{t-1})) + \alpha$. Since $\beta > 0$, it becomes evident that $c + \bar{c} \succeq_h c' + \bar{c}$ iff $c \succeq_h c'$ for any $c, c' \in C$.

(ii) $\sum_{k=1}^{\infty} \lambda_k = 1$. Suppose that u is cyclic. Then there exists $\gamma > 0$ and $\alpha \in \mathbf{R}$ such that $u(x + \gamma) = u(x) + \alpha$ for every $x \in \mathbf{R}$. In this case, simply choose $\bar{c}_0 = \gamma$ and $\bar{c}_t = \varphi(\vec{0}\bar{c}^{t-1})$ for every $t \geq 1$. Clearly $\bar{c} \in C$. It is easy to check that $c + \bar{c} \succeq_h c' + \bar{c}$ iff $c \succeq_h c'$ for any $c, c' \in C$, violating Lemma B.1. □

C Proof of necessity in Theorem 4.1

The representation clearly implies Axioms PR, C, DC, and NDM. Given the constructive proof of sufficiency, the necessity of HC may also be seen, except perhaps for the uniqueness of compensation. It may also not be obvious why the felicity need not be strictly monotonic. We resolve these matters here.

Why u need not be strictly increasing everywhere

Lemma C.1. *Suppose that $\sum_{k=1}^{\infty} \lambda_k < 1$. Then,*

(i) For any $\gamma > 0$, there does not exist a stream $c \in C$ and history $h \in H$ such that $c \geq (\gamma, \gamma, \ldots)$ and $g(h, c) \leq (0, 0, \ldots)$.

(ii) For any $\gamma < 0$, there does not exist a stream $c \in C$ and history $h \in H$ such that $g(h, c) \leq (\gamma, \gamma, \ldots)$.

Proof. To see (i), we first note that if $g(h, c) \leq (0, 0, \ldots)$ then $c_0 \leq \varphi(h)$, $c_1 \leq \varphi(hc_0)$, $c_2 \leq \varphi(hc_0c_1)$, etc. But using the monotonicity of φ and substituting in recursively, we see that $c_1 \leq \varphi(h\varphi(h))$, $c_2 \leq \varphi(h\varphi(h)\varphi(h\varphi(h)))$, etc. Therefore, it suffices to show that the compensating streams $(\varphi(h), \varphi(h\varphi(h)), \varphi(h\varphi(h)\varphi(h\varphi(h))), \ldots)$ tend to zero asymptotically. But this was accomplished in Lemma B.15.

Similarly, to see (ii), note that if $g(h, c) \leq (\gamma, \gamma, \ldots)$ then $c_0 \leq \varphi(h) + \gamma$, $c_1 \leq \varphi(hc_0) + \gamma \leq \varphi(h\varphi(h)) + \lambda_1\gamma + \gamma$. But since $\gamma < 0$, we may drop the term $\lambda_1\gamma$ to obtain $c_1 \leq \varphi(h\varphi(h)) + \gamma$. In this manner, $c_2 \leq \varphi(h\varphi(h)\varphi(h\varphi(h))) + \gamma$, and so on. But the compensating streams

$(\varphi(h), \varphi(h\varphi(h)), \varphi(h\varphi(h)\varphi(h\varphi(h))), \ldots)$ tend to zero asymptotically, and yet $\gamma < 0$ is fixed, implying that c must eventually become negative, a contradiction. $\square$

When $\sum_{k=0}^{\infty} \lambda_k < 1$, the fact that there does not exist a stream $c \in C$ and history $h \in H$ such that $c \geq (\gamma, \gamma, \ldots)$ for some $\gamma > 0$ and $g(h, c) \leq (0, 0, \ldots)$ means that the argument of the felicity in the representation cannot always be strictly negative when the consumption stream is bounded from zero. Moreover, the fact that there is no stream $c \in C$ and history $h \in H$ and $\gamma < 0$ such that $g(h, c) \leq (\gamma, \gamma, \ldots)$ means that the argument of the felicity in the representation cannot be bounded below zero. The former implies that we cannot shift a stream down using GM to conclude the felicity is increasing in the negative range, and the latter implies that there is no stream which we can shift up using GM to conclude that felicity is increasing in the negative range. Unlike in the case $\sum_{k=1}^{\infty} \lambda_k = 1$, shifting up (down) a stream starting from point t using GM leaves residual increases (decreases) in every the argument of the felicity starting from point t onward. Even though shifting a stream down (up) using the GM axiom will cause a decrease (increase) in utility, it suffices that the felicity is sensitive on the nonnegative domain for the utility to be sensitive to the GM induced shift due to Lemma C.1. To understand why it suffices that for some $0 < a \leq \infty$, $u(\cdot)$ is only strictly increasing either on $(-\infty, a)$ or $(-a, \infty)$ when $\sum_{k=0}^{\infty} \lambda_k = 1$, we may use Lemma B.33. By the characterization in (53), there cannot exist h and c such that $g(h, c)$ is always positive and bounded from zero, or always negative and bounded from zero (the stream would grow unboundedly in the first case and violate nonnegativity in the second).

On the uniqueness of compensation

Lemma C.2. *Assume the representation holds. Then for every $(h', h) \in \mathcal{H}$ there is a unique $d \in C$ satisfying $c + d \succeq_h c' + d$ iff $c \succeq_{h'} c'$ for every $c', c \in C$.*

Proof. Clearly $d^{h',h}$ as constructed earlier satisfies this; suppose that some $d \in C$, $d \neq d^{h',h}$ also satisfies the condition. According to the representation for $\succeq_{h'}$, both the utility functions $\sum_{t=0}^{\infty} \delta^t u\big(c_t - \varphi(h'c^{t-1}) + d_t - \varphi((h - h')d^{t-1})\big)$ and $\sum_{t=0}^{\infty} \delta^t u\big(c_t - \varphi(h'c^{t-1})\big)$ represent $\succeq_{h'}$. Using the uniqueness of the additive representation, there exist $\beta > 0$ and a sequence $\{\alpha_t\}_{t\geq 0}$ such that for any $c \in C$,

$$u\big(c_t - \varphi(h'c^{t-1}) + d_t - \varphi((h-h')d^{t-1})\big) = \beta u\big(c_t - \varphi(h'c^{t-1})\big) + \alpha_t.$$

Let $\gamma_t = d_t - \varphi((h-h')d^{t-1})$ for every t; we must show $\gamma_t = 0$ for all t.

Observe that for any $x \in \mathbf{R}$ and any t we may find $c \in C$ such that $c_t - \varphi(h'c^{t-1}) = x$. Indeed, if $x \geq 0$ choose $c_s = 0$ for every $s < t$ and $c_t = \varphi(h'0^t) + x$; if $x < 0$ choose $c_s = 0$ for every $s < t-1$, $c_{t-1} = \frac{x}{\lambda_1}$, and $c_t = \varphi(h'0^t)$. Therefore, for any x and t, $u(x + \gamma_t) = \beta u(x) + \alpha_t$.

Suppose that $\sum_{k=1}^{\infty} \lambda_k < 1$. Consider the first nonzero γ_t. If it is positive then u is quasi-cyclic, a contradiction. If $\gamma_t < 0$, then rearranging and changing variables gives $u(x + |\gamma_t|) = \frac{1}{\beta}u(x) - \frac{\alpha_t}{\beta}$. Hence u is quasi-cyclic, a contradiction.

Now consider the case $\sum_{k=1}^{\infty} \lambda_k = 1$. If some $\gamma_t = 0$ then $u(x)(1-\beta) = \alpha_t$ for all x, implying that $\beta = 1$ and u is cyclic, a contradiction. Hence $\gamma_t \neq 0$ for every t. We aim to show there exist $t, \hat{t}$ such that $\gamma_t \neq \gamma_{\hat{t}}$. If instead $\gamma_t = \gamma$ for every t, then we know that $\gamma > 0$ as a consequence of Lemma C.1, which says that for any $\gamma < 0$, there does not exist a stream $c \in C$ and history $\hat{h} \in H$ such that $g(\hat{h}, c) \leq (\gamma, \gamma, \ldots)$ (apply the lemma with $\hat{h} = h - h'$ and $c = d$). But if $\gamma > 0$, then $d_t = \varphi((h-h')d^{t-1}) + \gamma$ cannot be in C, a contradiction. To see this, first observe by Lemma B.16 that $d_{t-1}^{\vec{0}\gamma} = \lambda_1 \gamma > 0$ when $\sum_{k=1}^{\infty} \lambda_k = 1$; then apply Lemma B.33. Since d grows unboundedly it cannot be in C.

Therefore, we conclude that there exist nonzero $\gamma_t \neq \gamma_{\hat{t}}$ such that $u(x+\gamma_t) = \beta u(x) + \alpha_t$ and $u(x + \gamma_{\hat{t}}) = \beta u(x) + \alpha_{\hat{t}}$ for all x. Plug $x + \gamma_{\hat{t}}$ into the first equation and $x + \gamma_t$ into the second. This implies that for all x,

$$\beta u(x + \gamma_t) + \alpha_{\hat{t}} = u(x + \gamma_t + \gamma_{\hat{t}}) = \beta u(x + \gamma_{\hat{t}}) + \alpha_t.$$

Suppose WLOG that $\gamma_t > \gamma_{\hat{t}}$. By changing variables we see that for all x $u(x+\tilde{\gamma}) = u(x) + \tilde{\alpha}$, where $\tilde{\gamma} = \gamma_t - \gamma_{\hat{t}}$ and $\tilde{\alpha} = \frac{\alpha_t - \alpha_{\hat{t}}}{\beta}$. But then u is cyclic, a contradiction. □

D Proof of Theorem 6.1 and habit decay

Proof of Theorem 6.1

Proof. Note that the claim (11) is obvious when $\sum_{k=1}^{\infty}\lambda_k = 1$, for then $\frac{\lambda_{k+1}}{\lambda_k} = 1-\lambda_1$ for every k and consequently $\varphi(hq) = (1-\lambda_1)\varphi(h) + \lambda_1 q$.

For the particular h and $c_0, \hat{c}_0 \in Q$ from Axiom IE find the corresponding $c_1, \hat{c}_1$. Axioms IE and DC together imply that $\succeq_{hc_0c_1}$ and $\succeq_{h\hat{c}_0\hat{c}_1}$ are equivalent preferences, both additively representable as in (2) according to Theorem 4.1. By the uniqueness of additive representations up to positive affine transformation, there exist a $\rho > 0$ and a σ_i for every $i \geq 0$ such that for each $\bar{c} \in C$,

$$u\big(\bar{c} - \varphi(h00\bar{c}^{i-1}) - \lambda_{i+1}c_1 - \lambda_{i+2}c_0\big) = \rho u\big(\bar{c} - \varphi(h00\bar{c}^{i-1}) - \lambda_{i+1}\hat{c}_1 - \lambda_{i+2}\hat{c}_0\big) + \sigma_i. \quad (54)$$

For each i, let $\gamma_i = \lambda_{i+1}c_1 + \lambda_{i+2}c_0 - \lambda_{i+1}\hat{c}_1 - \lambda_{i+2}\hat{c}_0$.

If $\sum_{k=1}^{\infty}\lambda_k < 1$, then $\gamma_i = 0$ for every i since u cannot be quasi-cyclic. But we also wish to examine the case $\sum_{k=1}^{\infty}\lambda_k = 1$. Therefore, we note that $\rho = 1$ must hold. Indeed, since $\frac{\lambda_{i+1}}{\lambda_i} \leq 1-\lambda_1 \in (0,1)$, both $|\lambda_{i+1}\hat{c}_1 + \lambda_{i+2}c_0|$ and $|\lambda_{i+1}\hat{c}_1 + \lambda_{i+2}\hat{c}_0|$ tend to zero as i goes to infinity. As we have previously noted, for any i and $x \in \mathbf{R}$ we may find a $\bar{c} \in C$ such that $x = \bar{c} - \varphi(h00\bar{c}^{i-1})$. Then, by (54) and continuity of $u(\cdot)$, $\lim_{i\to\infty}\sigma_i = (1-\rho)u\big(x\big)$ for any $x \in \mathbf{R}$. Since the RHS depends on x while the LHS does not, we must have $\rho = 1$. Since u cannot be cyclic when $\sum_{k=1}^{\infty}\lambda_k = 1$, we have $\gamma_i = 0$ for every i in that case too.

Since $\gamma_i = 0$ for every i, we have $\frac{\lambda_{i+1}}{\lambda_i} = \frac{c_1-\hat{c}_1}{\hat{c}_0-c_0}$ for all $i \geq 1$. Then

$$\varphi(hq) = \sum_{k=2}^{\infty}\lambda_k h_{k-1} + \lambda_1 q = \sum_{k=2}^{\infty}\frac{\lambda_k}{\lambda_{k-1}}\lambda_{k-1}h_{k-1} + \lambda_1 q = \frac{c_1-\hat{c}_1}{\hat{c}_0-c_0}\varphi(h) + \lambda_1 q.$$

Now define $\alpha = \frac{c_1-\hat{c}_1}{\hat{c}_0-c_0}$ and $\beta = \lambda_1$. Clearly $\alpha + \beta \leq 1$ since $\frac{\lambda_{i+1}}{\lambda_i} \leq 1-\lambda_1$. □

Additional results on habit decay

To isolate habit decay, we examine how the compensating stream changes $d^{h',h}$ if the DM only begins to wean herself from her current habit h tomorrow. If she abstains today, the DM can be weaned starting tomorrow using $d^{h'0,h0}$ (the assumption of abstention is

WLOG). We call the ratio $\frac{d_t^{h'0,h0}}{d_t^{h',h}}$ the period-t rate of habit decay; the lower it is, the faster the DM's habit decays in period t. We would like to use this measure to compare DM's.

Definition D.1. *For each $(h',h) \in \mathcal{H}$, let $d_{i.}^{h',h}$ denote the compensating streams of DM_i and $d_{j.}^{h',h}$ denote the compensating streams of DM_j. We say that DM_i's habits decay faster than DM_j's if $\frac{d_{i.t}^{h'0,h0}}{d_{i.t}^{h',h}} \leq \frac{d_{j.t}^{h'0,h0}}{d_{j.t}^{h',h}}$ for every t and $h \geq h' \in H$. We express this relation by $DM_i \geq_{\text{HD}} DM_j$.*

While the relation $\geq_{\text{HD}}$ is transitive, it is not complete. The following example illustrates the difficulty in ranking DM's according to their rates of habit decay.

Example D.2. *Suppose that the agents DM_1 and DM_2 satisfy our axioms and that each DM's habit formation coefficients are determined by three parameters:*

$$(\lambda_{i.1}, \lambda_{i.2}, \lambda_{i.3}, \lambda_{i.4}, \lambda_{i.5}, \ldots) = (\alpha_i, \beta_i, \gamma_i, \gamma_i(1-\alpha_i), \gamma_i(1-\alpha_i)^2, \ldots) \text{ for } DM_i,\ i = 1, 2,$$

where $\frac{\beta_i}{\alpha_i} \leq 1-\alpha_i$ and $\frac{\gamma_i}{\beta_i} \leq 1-\alpha_i$ for each i. In particular, we specify that $\alpha_1 = \alpha_2 = \alpha = \frac{1}{2}$, $\beta_1 = \frac{1}{6}$, $\beta_2 = \frac{1}{5}$, $\gamma_1 = \frac{1}{17}$, and $\gamma_2 = \frac{1}{25}$. For simplicity we compare only the first two values of $\frac{d_t^{h'0,h0}}{d_t^{h',h}}$ for each DM_i given the habits satisfy $h - h' = (\vec{0}, 1)$. Using the recursive formula, we calculate that for DM_1

$$\frac{d_0^{h'0,h0}}{d_0^{h',h}} = \frac{\varphi(\vec{0},1,0)}{\varphi(\vec{0},1)} = \frac{\beta_1 \cdot 1}{\alpha \cdot 1} = \frac{1}{3} \quad \textit{and}$$

$$\frac{d_1^{h'0,h0}}{d_1^{h',h}} = \frac{\varphi(\vec{0},1,0,\varphi(\vec{0},1,0))}{\varphi(\vec{0},1,\varphi(\vec{0},1))} = \frac{\alpha \cdot \frac{1}{6} + \gamma_1 \cdot 1}{\alpha \cdot \frac{1}{2} + \beta_1 \cdot 1} \approx .34$$

Similarly, for DM_2 we calculate that

$$\frac{d_0^{h'0,h0}}{d_0^{h',h}} = \frac{\varphi(\vec{0},1,0)}{\varphi(\vec{0},1)} = \frac{\beta_2 \cdot 1}{\alpha \cdot 1} = \frac{\frac{1}{5}}{\frac{1}{2}} = \frac{2}{5} \quad \textit{and}$$

$$\frac{d_1^{h'0,h0}}{d_1^{h',h}} = \frac{\varphi(\vec{0},1,0,\varphi(\vec{0},1,0))}{\varphi(\vec{0},1,\varphi(\vec{0},1))} = \frac{\alpha \cdot \frac{1}{5} + \gamma_2 \cdot 1}{\alpha \cdot \frac{1}{2} + \beta_2 \cdot 1} \approx .31$$

Here is the difficulty: while the time-0 rate of habit decay is faster for DM_1 ($\frac{1}{3} < \frac{2}{5}$), the time-1 rate of habit decay is faster for DM_2 ($.31 < .34$). Confounding matters, DM_2's *time-1 rate of decay is faster than DM_1's time-0 rate of decay ($.31 < .\overline{33}$), but DM_1's time-1 rate of decay is faster than DM_2's time-0 rate of decay ($.34 < .4$).*

The following proves $\geq_{\text{HD}}$ is complete on the set of DM's satisfying Axiom IE and incomplete if any DM not satisfying IE is added to that set; and that under IE the notion of habit decay in Theorem 6.1 coincides with that in Definition D.1.

Proposition D.3. *Assume Axioms PR, C, DC, GS, HC, and NDM. (i) If the DM satisfies IE then $\frac{d_t^{h0}}{d_t^h} = \frac{\lambda_{k+1}}{\lambda_k} = \alpha$ for some $\alpha \in (0,1)$; (ii) for any DM_1 not satisfying IE, we may find a DM_2 who satisfies IE such that $\geq_{\text{HD}}$ cannot rank DM_1 and DM_2.*

Proof. To see (i), note by Theorem 6.1 that under Axiom IE there is $\alpha \in (0,1)$ such that $\frac{\lambda_{k+1}}{\lambda_k} = \alpha$ for all k. Since $\varphi(h0^t) = \alpha^t \varphi(h)$ by Lemma B.14, $\frac{d_0^{h0}}{d_0^h} = \alpha$. Assume $\frac{d_{t-1}^{h0}}{d_{t-1}^h} = \alpha$ for some $t-1$. By the inductive hypothesis and Lemma B.12, $\frac{d_t^{h0}}{d_t^h}$ equals

$$\frac{\varphi(h0d_0^{h0}\dots d_{t-1}^{h0})}{\varphi(hd_0^h\dots d_{t-1}^h)} = \frac{\varphi(h0d_0^{h0}\dots d_{t-2}^{h0}0) + \lambda_1 d_{t-1}^{h0}}{\varphi(hd_0^h\dots d_{t-2}^h0) + \lambda_1 d_{t-1}^h} = \frac{\alpha d_{t-1}^{h0} + \lambda_1 d_{t-1}^{h0}}{\alpha d_{t-1}^h + \lambda_1 d_{t-1}^h} = \frac{\alpha^2 d_{t-1}^h + \alpha\lambda_1 d_{t-1}^h}{\alpha d_{t-1}^h + \lambda_1 d_{t-1}^h},$$

which equals α. To see part (ii), observe that Theorem 6.1 implies that DM_1 does not satisfy Axiom IE if and only if for some $k, \hat{k}$ with $k \neq \hat{k}$, $\frac{\lambda_{k+1}}{\lambda_k} \neq \frac{\lambda_{\hat{k}+1}}{\lambda_{\hat{k}}}$. Let $k^* = \min\{k \geq 2 \mid \frac{\lambda_{k+1}}{\lambda_k} \neq \frac{\lambda_k}{\lambda_{k-1}}\}$ and define $h^* \in H$ by $h_j^* = \begin{cases} 0 & j \neq k^*-1 \\ 1 & j = k^*-1 \end{cases}$ for all $j \geq 1$. Observe that $d_0^{h^*} = \lambda_{k^*-1}$, $d_0^{h^*0} = \lambda_{k^*}$, $d_1^{h^*} = \lambda_{k^*} + \lambda_1\lambda_{k^*-1}$, and $d_1^{h^*0} = \lambda_{k^*+1} + \lambda_1\lambda_{k^*}$. Then $\frac{d_0^{h^*0}}{d_0^{h^*}} = \frac{\lambda_{k^*}}{\lambda_{k^*-1}}$, but

$$\frac{d_1^{h^*0}}{d_1^{h^*}} = \frac{\lambda_{k^*+1} + \lambda_1\lambda_{k^*}}{\lambda_{k^*} + \lambda_1\lambda_{k^*-1}} = \frac{\lambda_{k^*}}{\lambda_{k^*-1}} \cdot \left\{ \frac{\frac{\lambda_{k^*+1}\lambda_{k^*-1}}{\lambda_{k^*}} + \lambda_1\lambda_{k^*-1}}{\lambda_{k^*} + \lambda_1\lambda_{k^*-1}} \right\}$$

Therefore, $\frac{d_1^{h^*0}}{d_1^{h^*}} = \frac{d_0^{h^*0}}{d_0^{h^*}} = \frac{\lambda_{k^*}}{\lambda_{k^*-1}}$ if and only if $\frac{\lambda_{k^*+1}\lambda_{k^*-1}}{\lambda_{k^*}} = \lambda_{k^*}$. But that implies $\frac{\lambda_{k^*+1}}{\lambda_{k^*}} = \frac{\lambda_{k^*}}{\lambda_{k^*-1}}$, contradicting the definition of k^*. WLOG suppose $\frac{d_1^{h^*0}}{d_1^{h^*}} > \frac{d_0^{h^*0}}{d_0^{h^*}}$. Let $\alpha \in (\frac{d_0^{h^*0}}{d_0^{h^*}}, \frac{d_1^{h^*0}}{d_1^{h^*}})$ and take DM_2 satisfying Axiom IE and for whom $\frac{\hat{\lambda}_{k+1}}{\hat{\lambda}_k} = \alpha$ for all k. Then DM_1 and DM_2 cannot be ranked by $\geq_{\text{HD}}$ using part (i) of the result. □

References

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

Aczel, J., and J. Dhombres (1989): *Functional Equations Containing Several Variables.* Cambridge University Press, Cambridge, UK.

Backus, D., B. Routledge, and S. Zin (2004): "Exotic Preferences for Macroeconomists," *NBER Macroeconomics Annual.*

Becker, G., and K. Murphy (1988): "A Theory of Rational Addiction," *Journal of Political Economy*, 96, 675–700.

Bertschy, G. (1995): "Methadone Maintenance Treatment: An Update," *The European Archives of Psychiatry and Clinical Neuroscience*, pp. 114–124.

Bleichrodt, H., K. Rohde, and P. Wakker (2007): "Koopmans' Constant Discounting: A Simplification and an Extension to Incorporate General Economic Growth," *Mimeo.*

Boldrin, M., L. Christiano, and J. Fisher (1997): "Habit Persistence and Asset Returns in an Exchange Economy," *Macroeconomic Dynamics*, 1(2), 312–332.

——— (2001): "Habit Persistence, Asset Returns, and the Business Cycle," *The American Economic Review*, 91, 149–166.

Camerer, C., and G. Loewenstein (2004): "Behavorial Economics: Past, Present, and Future," in *Advances in Behavioral Economics*, ed. by C. Camerer, G. Loewenstein, and M. Rabin. Princeton University Press.

Campbell, J., and J. Cochrane (1999): "By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Carroll, C., J. Overland, and D. Weil (2000): "Saving and Growth with Habit Formation," *American Economic Review*, 90, 341–355.

Chen, X., and S. Ludvigson (2004): "Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Behavior," *NBER Working Paper.*

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

Deaton, A. (1987): "Life Cycle Models of Consumption: Is the Evidence Consistent With Facts?," in *Advances in Econometrics*, ed. by T. Bewley. Cambridge University Press.

Fishburn, P. (1970): *Utility Theory for Decisionmaking.* John Wiley & Sons, Inc., New York, NY.

Gilboa, I. (1989): "Expectation and Variation in Multi-Period Decisions," *Econometrica*, 57, 1153–1160.

Gorman, W. (1968): "The Structure of Utility Functions," *Review of Economic Studies*, 35, 367–390.

Gul, F., and W. Pesendorfer (2007): "Harmful Addiction," *Review of Economic Studies*, 74, 147–172.

Iannaccone, L. (1986): "Addiction and Satiation," *Economics Letters*, 21, 95–99.

Jarczyk, W. (1991): "A Recurrent Method of Solving Iterative Functional Equations," *Prace Naukowe Uniwersytetu Slaskiego w Katowicach 1206*.

Koopmans, T. (1960): "Stationary Ordinal Utility and Impatience," *Econometrica*, 28, 287–309.

Kőszegi, B., and M. Rabin (2006a): "Dynamic Reference-Dependent Preferences," *Mimeo*.

——— (2006b): "A Model of Reference Dependent Preferences," *Quarterly Journal of Economics*, pp. 1133–1166.

Kozicki, S., and P. Tinsley (2002): "Dynamic Specifications in Optimizing Trend-Deviation Macro Models," *Journal of Economic Dynamics and Control*, 26, 1585–1611.

Loewenstein, G., and D. Prelec (1992): "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics*, pp. 573–597.

Neilson, W. (2006): "Axiomatic Reference-Dependence in Behavior Towards Others and Toward Risk," *Economic Theory*, 28, 681–692.

Palacios-Huerta, I. (2003): "Multiple Addictions," *Working Paper*.

Pollak, R. (1970): "Habit Formation and Dynamic Demand Functions," *Journal of Political Economy*, 78, 745–763.

Rustichini, A., and P. Siconolfi (2005): "Dynamic Theory of Preferences: Taste for Variety and Habit Formation," *Mimeo*.

Schroder, M., and C. Skiadas (2002): "An Isomorphism Between Asset Pricing Models With and Without Linear Habit Formation," *The Review of Financial Studies*, 15, 1189–1221.

Shalev, J. (1997): "Loss Aversion in a Multi-Period Model," *Mathematical Social Sciences*, 33, 203–226.

Shi, S., and L. Epstein (1993): "Habits and Time Preference," *International Economic Review*, 34, 61–84.

Strotz, R. (1955): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23, 165–180.

Tversky, A., and D. Kahneman (1991): "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *The Quarterly Journal of Economics*, 106, 1039–1061.

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

Wendner, R. (2003): "Do Habits Raise Consumption Growth?," *Research in Economics*, 57, 151–163.