When Do Habits Matter?
The Empirical Content of Dynamic Hedonic Models *

Josephine Auer†

MIT

March 3, 2026

Abstract

Hedonic models value goods through their characteristics but are typically interpreted under time-separable preferences. This assumption is restrictive: when some attributes are habit forming, observed prices reflect both contemporaneous utility and continuation values from past consumption. I develop a nonparametric revealed preference framework for dynamic hedonic valuation, deriving necessary and sufficient conditions for rationalisability over characteristics. The framework separates restrictions imposed by the hedonic price system from those imposed by intertemporal choice and provides diagnostics that quantify the severity of violations along each margin. Applied to household scanner data, I show that most failures of static hedonic valuation reflect violations of the hedonic price structure; conditional on satisfying this structure, allowing for habit formation improves behavioural fit. This alters the mapping from prices to willingness-to-pay and the implied welfare interpretation.

Keywords: Hedonic valuation, habit formation, identification, revealed preference, intertemporal choice, nonparametric analysis
JEL Classification: D11, D12, D90, C61, C14, L66

*I am especially grateful to Ian Crawford for his guidance and encouragement throughout this project. I also thank Nikhil Agarwal, Miguel Ballester, Nano Barahona, Colin Camerer, Glenn Ellison, Drew Fudenberg, Matt Polisson, Augustus Smith, and Chen Zhen for helpful comments and discussions.

†E-mail: joauer@mit.edu

1 Introduction

Many product attributes—such as sugar, nicotine, ethanol, and caffeine—are habit forming. Yet empirical hedonic valuation, following Gorman (1956) and Lancaster (1966), almost universally treats preferences as time separable.[1] When some characteristics are habit forming, time separability is no longer an innocuous normalisation but a restriction on the economic meaning of hedonic prices and choices. This raises two questions: when do observed prices and choices admit a coherent characteristics-based interpretation at all, and when does incorporating habits materially alter the implications of hedonic valuation?

Whether hedonic willingness-to-pay (WTP) is well-defined in the presence of habit formation matters because it is routinely used for welfare analysis and policy evaluation. In settings such as sugar taxation, alcohol regulation, tobacco policy, and environmental product standards, researchers recover marginal valuations for attributes under time-separable preferences and use them to rank counterfactuals.[2] If some attributes are habit forming, however, observed prices may reflect both contemporaneous utility and continuation values from past consumption. In that case, static hedonic WTP need not measure the object it is interpreted as measuring.

Existing models of habit formation are typically formulated over *goods*, not characteristics. A large literature studies intertemporal dependence in the consumption of cigarettes, alcohol, and digital products by modelling utility as a function of past and current quantities (Becker and Murphy, 1988; Gruber and Kőszegi, 2004; Demuynck and Verriest, 2013; Crawford, 2010; Allcott et al., 2022). These models characterise dynamic dependence in product quantities, but they do not determine whether prices admit a coherent interpretation as valuations of underlying attributes. When valuation and policy analysis are conducted in characteristics space, the dynamic structure must be specified at the level of attributes rather than goods. This is particularly important when reinforcement operates at the level of sugar, caffeine, or nicotine rather than at the level of the composite product itself. Understanding when a dynamic hedonic interpretation is coherent is therefore necessary to discipline applied valuation.

This paper provides a nonparametric framework for evaluating when dynamic hedonic interpretation is economically coherent. I make three contributions. First, I derive necessary and sufficient revealed-preference (RP) conditions—in the spirit of Samuelson (1948), Houthakker (1950), Afriat (1967), and Browning (1989)—under which observed prices and choices admit a coherent dynamic characteristics-based interpretation when some attributes are habit forming. Second, I decompose empirical failure into two conceptually distinct margins: (i) structural feasibility of the hedonic price system given the maintained goods-to-characteristics technology, and (ii) behavioural consistency of intertemporal choice conditional on that technology. Third, I develop computationally tractable, distance-based diagnostics that quantify how close the data are to satisfying each margin, thereby disciplining when static hedonic WTP is economically defensible.

The key insight is that dynamic hedonic rationalisability has a two-stage structure. First, observed prices must lie in the low-dimensional span implied by the goods-to-characteristics mapping; this is a structural restriction on the price system, independent of intertemporal optimisation. Second, conditional on such a representation existing, the implied characteristic shadow prices must rationalise observed choices over time; this is a behavioural restriction. Separating these margins clarifies whether a model fails because the maintained characteristics technology cannot sustain a hedonic interpretation or because intertemporal optimisation is violated conditional on that technology.

To formalise these ideas, I study the RP implications of a dynamic hedonic environment in which a single

[1] Hedonic models are central in empirical industrial organisation and applied micro; see, e.g., Smith and Desvousges (1986), Heckman and Scheinkman (1987), Berry et al. (1995), Nevo (2001), Gibbons and Machin (2003), Bajari and Kahn (2005), and Greenstone and Gallagher (2008).

[2] For examples valuing potentially habit-forming food attributes under static hedonic preferences, see Dubois et al. (2020), Haeck et al. (2022), and Le Fur and Outreville (2022).

consumer purchases goods over time at some observed prices. These observed goods map into characteristics through a maintained technology, and utility depends on both contemporaneous and lagged levels of selected characteristics. Some attributes are habit forming, so current consumption affects future marginal utility. The analysis characterises the observable restrictions on prices and choices required for a coherent dynamic characteristics-based interpretation.

I derive a dynamic Afriat-type RP characterisation that extends static characteristics-based rationalisability (e.g., Blow et al. (2008)) to accommodate habit formation. The characterisation makes the structural and behavioural margins operational, decomposes model rejection into economically distinct sources, and moves beyond binary pass-fail tests by providing distance-based diagnostics that quantify the severity of each type of violation. Together, these elements deliver a new taxonomy of empirical failure and a quantitative measure of how far the data lie from satisfying each component.

Empirically, I implement the RP test in a scanner panel of cereal purchases and compare characteristics-based and goods-based representations. Moving to characteristics delivers dimensional parsimony but imposes demanding structural price restrictions that are often taken for granted in applied work. Characteristics models therefore pass binary rationalisability tests less often than goods-based benchmarks. This lower pass rate reflects the stronger structural discipline imposed by the maintained characteristics technology rather than superior empirical performance of goods-based models. Moreover, the distance-based diagnostics show that most structural violations are modest in magnitude. Conditional on satisfying the hedonic price restrictions, allowing for habit formation systematically improves intertemporal coherence relative to static hedonic models, requiring smaller perturbations to observed budgets to rationalise behaviour over time.

My results clarify when dynamics matter for hedonic valuation. If the price system cannot sustain a characteristics-based interpretation under the maintained technology, hedonic WTP is not economically well-defined. If the structural restrictions hold but intertemporal separability fails, hedonic WTP is well-defined yet systematically misinterpreted when dynamics are ignored. The framework therefore provides a disciplined diagnostic for applied work: it distinguishes structural misspecification from behavioural misspecification and clarifies when static hedonic valuation is economically defensible.

The remainder of the paper is organised as follows. After situating the paper within the related literature, Section 2 introduces the dynamic hedonic model and derives the necessary and sufficient RP characterisation. All proofs are in Appendix A. Section 3 presents corollaries showing that static hedonic and dynamic goods-based models arise as special cases of my framework. Section 4 applies the theory to household scanner data on cereal purchases.

Related work

This paper contributes to the RP literature on dynamic choice and to the applied literatures on hedonic valuation and rational addiction. Its main theoretical contribution is to unify two strands of nonparametric analysis that have largely developed in isolation: tests of preferences over characteristics (Blow et al., 2008) and RP analyses of intertemporal dependence (Crawford, 2010), both rooted in the foundational work of Afriat (1967), Diewert (1973), and Varian (1982). Whereas Blow et al. (2008) study the feasibility of a static hedonic price system absent dynamics, and Crawford (2010) characterise dynamic rationalisability over goods without dimensionality reduction, my framework allows dimensionality reduction and intertemporal dependence simultaneously and shows how they interact. Static hedonic RP tests and dynamic goods-based RP tests arise as special cases. The resulting characterisation yields a taxonomy of empirical failure not present in either setting and provides the first nonparametric RP characterisation of rational addiction operating at the level of product characteristics rather than goods. Despite incorporating both dimensions, the conditions remain conditionally linear in observables and can be implemented using standard linear programming techniques. The framework also yields distance-based diagnostics that quantify the severity of empirical violations.

On the empirical side, the paper relates to the hedonic valuation literature following Rosen (1974), which models prices as implicit functions of product characteristics and recovers WTP from the gradient of the hedonic price schedule. This approach has been widely applied across housing, environmental economics, health, education, and IO (e.g. Smith and Desvousges, 1986; Gibbons and Machin, 2003; Bajari and Kahn, 2005; Greenstone and Gallagher, 2008). A maintained assumption in this literature is that preferences over characteristics are time-separable. When preferences exhibit intertemporal dependence, however, observed price gradients need not reflect contemporaneous marginal valuations, so standard hedonic identification arguments can break down. My RP approach complements the structural hedonic literature by testing whether observed behaviour is consistent with a dynamic hedonic model, rather than inferring preferences from market-clearing prices under strong behavioural assumptions. The analysis therefore complements structural approaches such as Bajari and Benkard (2005), which recover preference parameters from equilibrium price schedules. Here the focus instead is on whether a coherent hedonic interpretation of observed prices is empirically defensible in the first place.

The paper also contributes to the rational addiction literature initiated by Becker and Murphy (1988), which models addictive consumption as forward-looking and utility maximising. Empirical applications focus on goods such as alcohol, cigarettes, caffeine, and illicit drugs (e.g. Becker et al., 1994; Grossman et al., 1998; Gruber and Kőszegi, 2004; Demuynck and Verriest, 2013). These models impose dynamics at the level of goods rather than attributes, even though evidence from neuroscience suggests that persistence often attaches to specific chemical or sensory components rather than to goods as undifferentiated bundles (cf. Koob and Moal, 1997). A small empirical literature allows for multivariate addiction over nutrient profiles (e.g. Richards and Patterson, 2006), but relies on strong functional form assumptions. By contrast, my nonparametric RP characterisation allows addiction to operate at the level of characteristics and provides necessary and sufficient conditions for rationalisability without imposing functional form restrictions.

The framework is also distinct from models of inventory behaviour and stockpiling in consumer demand (Hendel and Nevo, 2006). In such models, observed persistence arises from intertemporal substitution and storage under time-separable preferences. By contrast, habit formation in my setting operates through state dependence in utility over characteristics, so past consumption directly shifts the marginal valuation of current and future consumption. The RP tests therefore distinguish preference-based persistence from inventory-driven dynamics.

Taken together, the paper clarifies how intertemporal dependence, dimensionality reduction, and behavioural discipline jointly determine whether hedonic valuation is economically meaningful. It shows that while moving to characteristics space imposes demanding structural restrictions, incorporating dynamics materially improves the model's ability to capture systematic patterns in behaviour conditional on satisfying those restrictions.

2 Model

I begin by formalising a dynamic hedonic environment in which goods map into characteristics and habits attach to a subset of those characteristics. I observe a single consumer for $t = 1, \dots, T$, with purchases $\boldsymbol{x}_t \in \mathbb{R}^K_+$ and present-value prices $\boldsymbol{\rho}_t \in \mathbb{R}^K_+$. Goods map into J measured characteristics via a time-invariant linear technology $\boldsymbol{z}_t = \boldsymbol{A}\boldsymbol{x}_t$ following Gorman (1956), where $\boldsymbol{A}$ is a $J \times K$ matrix (typically $J < K$).[3] I treat $\boldsymbol{A}$ as known and stable over time; it captures objectively defined product attributes, while valuation is encoded in preferences.

[3] I adopt a linear transformation from goods to characteristics space as it is the most widely used specification. Most results extend to a non-linear setting where $\boldsymbol{z} = \boldsymbol{F}(\boldsymbol{x})$ is increasing and strictly concave; see Appendix B for an analogue of the consistency definition. The key difference to the linear case arises in the marginal product: whereas $\partial \boldsymbol{z}/\partial \boldsymbol{x} = \boldsymbol{A}'$ is constant in the linear model, the marginal product varies with demand under a non-linear transformation.

I partition characteristics as $\boldsymbol{z}_t = ((\boldsymbol{z}_t^c)', (\boldsymbol{z}_t^a)')'$, where $\boldsymbol{z}_t^c \in \mathbb{R}^{J_1}$ are non-habit-forming characteristics and $\boldsymbol{z}_t^a \in \mathbb{R}^{J_2}$ are habit-forming characteristics, with $J_1 + J_2 = J$. The analyst defines this partition. Crucially, this formulation allows intertemporal dependence to operate at the level of attributes rather than goods, so persistence in behaviour need not be attributed to non-habit-forming components bundled within a product. Although preferences are defined over characteristics, choice and budget constraints remain in goods space.

To clarify the economic content of the model, it is useful to fix ideas with a simple example. Suppose the consumer chooses between two cereal products, where each good bundles two measurable attributes: a contemporaneous "taste" characteristic (e.g., nutty-ness) and a habit-forming "sensory" characteristic (e.g., salt or sugar intensity). The key modelling choice is that habits attach to the latter attribute rather than to the cereal good itself: consuming a high-intensity product today can change tomorrow's marginal value of intensity (due to sensory fatigue or craving), even if the consumer switches cereal product. In this example, the model's structural content is that observed goods prices must be representable as shadow values on attributes given $\boldsymbol{z}_t = \boldsymbol{A}\boldsymbol{x}_t$, while its behavioural content is that those shadow values must admit a concave, dynamically consistent utility representation.

Preferences are represented by a felicity function $u : \mathbb{R}^{J+J_2} \to \mathbb{R}$ that depends on current characteristics and one lag of the habit-forming subset, $u(\boldsymbol{z}_t^c, \boldsymbol{z}_t^a, \boldsymbol{z}_{t-1}^a)$, as in the one-lag "short memory habits" specification of Boyer (1978, 1983) and Becker et al. (1994). The multi-lag extension is straightforward and deferred to Online Appendix E. I assume quasi-linearity in an outside good y_t with unit price, as is standard in empirical IO and hedonic demand models for narrow product categories (Berry et al., 1995; Nevo, 2001). I also maintain local non-satiation, concavity, and superdifferentiability of u; I impose no further sign or monotonicity restrictions on the habit-forming components. I refer to this environment as the *habits-over-characteristics* model.

The consumer chooses $\{(\boldsymbol{x}_t, y_t)\}_{t=1}^T$ to solve

$$\max_{\{(\boldsymbol{x}_t, y_t)\}_{t=1}^T} \sum_{t=1}^T \beta^{t-1}\big(u(\tilde{\boldsymbol{z}}_t) + y_t\big) \quad \text{s.t.} \quad \sum_{t=1}^T \boldsymbol{\rho}_t'\boldsymbol{x}_t + \sum_{t=1}^T \beta^{t-1} y_t = W, \qquad \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{A}}\,\tilde{\boldsymbol{x}}_t, \tag{1}$$

where $\beta \in (0, 1]$ is a discount factor and W is present-value lifetime wealth. I use the augmented notation

$$\tilde{\boldsymbol{z}}_t := \big((\boldsymbol{z}_t^c)', (\boldsymbol{z}_t^a)', (\boldsymbol{z}_{t-1}^a)'\big)', \quad \tilde{\boldsymbol{x}}_t := (\boldsymbol{x}_t', \boldsymbol{x}_{t-1}')', \quad \tilde{\boldsymbol{A}} := \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0}_{J\times K} \\ \boldsymbol{0}_{J_2\times K} & \boldsymbol{A}^a \end{pmatrix}, \tag{2}$$

so $\tilde{\boldsymbol{z}}_t \in \mathbb{R}^{J+J_2}$, $\tilde{\boldsymbol{x}}_t \in \mathbb{R}^{2K}$, and $\tilde{\boldsymbol{A}}$ is a $(J + J_2) \times 2K$ block matrix. The key question is whether the observables $\{(\boldsymbol{\rho}_t, \boldsymbol{x}_t)\}_{t=1}^T$ can be rationalised by (1), and, if so, what testable restrictions this imposes on prices and choices.

2.1 Consistency

I now ask whether the observed data can be rationalised by optimising behaviour under the habits-over-characteristics model. This notion of consistency encompasses both the existence of a hedonic price representation (a structural requirement) and the coherence of intertemporal behaviour conditional on that representation.

Definition 2.1 (**Consistency**)**.** The data $\{(\boldsymbol{\rho}_t, \boldsymbol{x}_t)\}_{t=1}^T$ are *consistent* with the one-lag habits-over-characteristics model for given $\boldsymbol{A}$ if there exist $\beta \in (0, 1]$ and a locally non-satiated, concave, superdifferentiable felicity function u such that $\{\boldsymbol{x}_t\}_{t=1}^T$ solves (1).

The following lemma provides necessary and sufficient conditions for consistency.

Lemma 2.1 (**Consistency Conditions**). The data $\{\boldsymbol{\rho}_t, \boldsymbol{x}_t\}_{t=1}^T$ are consistent with the one-lag habits-over-characteristics model for a given technology matrix $\boldsymbol{A}$ if and only if there exist a locally non-satiated, concave, superdifferentiable utility function $u(\cdot)$ and a discount factor $\beta \in (0,1]$ such that for all $t \in \{1,\ldots,T-1\}$,

$$\boldsymbol{\rho}_t \geq \boldsymbol{A}'\boldsymbol{\pi}_t^0 + (\boldsymbol{A}^a)'\boldsymbol{\pi}_{t+1}^1, \tag{$\star$}$$

with equality for all goods k such that $x_t^k > 0$, where the discounted shadow prices are

$$\boldsymbol{\pi}_t^0 = \beta^{t-1}\begin{bmatrix}\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t) \\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)\end{bmatrix}, \tag{3}$$

$$\boldsymbol{\pi}_t^1 = \beta^{t-1}\partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t), \tag{4}$$

and $\tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t$.

Proof: See Appendix A. □

Lemma 2.1 links observed discounted market prices to shadow prices that measure the consumer's discounted marginal valuations of characteristics (Gorman, 1956). Current goods prices therefore reflect both contemporaneous utility from characteristics and the intertemporal effects induced by habit formation. In the running cereal example above, the price of a high-salt or high-sugar cereal must reflect not only the consumer's current taste for nutty-ness and sensory intensity, but also how today's intensity alters tomorrow's marginal utility of that same sensory characteristic. The first-order condition ($\star$) formalises this intuition: goods prices equal the sum of contemporaneous shadow values and the discounted continuation value generated by habit-forming attributes.

Formally, the shadow price $\boldsymbol{\pi}_t^0$ captures WTP for contemporaneous characteristics, while $\boldsymbol{\pi}_t^1$ captures the marginal utility impact of past consumption of habit-forming characteristics. The key empirical implication is a price wedge: current goods prices internalise future utility effects whenever habits are present. When lagged consumption lowers future marginal utility (i.e., $\partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_{t+1}) < 0$), goods prices satisfy

$$\rho_t^k = \boldsymbol{a}_k'\boldsymbol{\pi}_t^0 + \boldsymbol{a}_k^{a\prime}\boldsymbol{\pi}_{t+1}^1 < \boldsymbol{a}_k'\boldsymbol{\pi}_t^0,$$

so ignoring intertemporal dependence understates true WTP for goods with negatively reinforcing characteristics, such as sensory fatigue. If lagged consumption instead raises future marginal utility, the inequality reverses. Because many hedonic and IO applications maintain intertemporal separability, this wedge implies systematic mismeasurement of WTP—and hence welfare—whenever habit-forming characteristics are present.

A further implication is that under a linear characteristics technology, observed prices for goods consumed in strictly positive amounts must lie in the column space of the technology matrix. By complementary slackness, the first-order conditions bind on the support of consumption, so the intertemporal budget constraint holds equivalently when expressed in goods space or in characteristics space with shadow prices:

$$\sum_{t=1}^T \boldsymbol{\rho}_t'\boldsymbol{x}_t = \sum_{t=1}^T \boldsymbol{z}_t'\boldsymbol{\pi}_t^0 + \boldsymbol{z}_t^{a\prime}\boldsymbol{\pi}_{t+1}^1.$$

In sum, the model's empirical content is governed by two restrictions: a *structural* requirement that observed goods prices admit characteristic shadow prices consistent with the maintained goods-to-characteristics technology, and a *behavioural* requirement that these shadow prices be consistent with concave, dynamically stable preferences.

2.2 Afriat conditions for habits-over-characteristics

I now state the central theoretical result of the paper. It shows that the model has two distinct sources of empirical content: a structural requirement that observed goods prices admit characteristic shadow prices given $\boldsymbol{A}$, and a behavioural requirement that these shadow prices be consistent with concave, dynamically stable preferences.

Theorem 2.1. The following statements are equivalent:

(A) The data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t\in\{1,\dots,T\}}$ are consistent with the one-lag habits model for given technology $\boldsymbol{A}$.

(B) There exist T J-vector discounted shadow prices $\{\boldsymbol{\pi}_t^0\}_{t\in\{1,\dots,T\}}$, T J_2-vector discounted shadow prices $\{\boldsymbol{\pi}_t^1\}_{t\in\{1,\dots,T\}}$ and a discount factor $\beta \in (0,1]$ such that,

$$0 \leq \sum_{\forall s,t\in\sigma} \tilde{\boldsymbol{\pi}}_s' \left(\tilde{\boldsymbol{z}}_t - \tilde{\boldsymbol{z}}_s\right) \qquad \forall\, \sigma \subseteq \{1,\dots,T\} \tag{B1}$$

$$\rho_t^k \geq \boldsymbol{a}_k'\boldsymbol{\pi}_t^0 + \boldsymbol{a}_k^{a\prime}\boldsymbol{\pi}_{t+1}^1 \qquad \forall\, k,t \in \{1,\dots,T-1\} \tag{B2}$$

$$\rho_t^k = \boldsymbol{a}_k'\boldsymbol{\pi}_t^0 + \boldsymbol{a}_k^{a\prime}\boldsymbol{\pi}_{t+1}^1 \qquad \text{if } x_t^k > 0,\ \forall\, k,t \in \{1,\dots,T-1\} \tag{B3}$$

where $\boldsymbol{a}_k$ is the J-vector corresponding to the k-th column of $\boldsymbol{A}$, $\boldsymbol{a}_k^a$ is the J_2-vector corresponding to the last J_2 rows of the k-th column of $\boldsymbol{A}$, and $\tilde{\boldsymbol{\pi}}_t := \frac{1}{\beta^{t-1}}\left[\boldsymbol{\pi}_t^{0\prime}, \boldsymbol{\pi}_t^{1\prime}\right]'$.

Proof: See Appendix A. □

Theorem 2.1 delivers a complete RP characterisation of the one-lag habits-over-characteristics model: the data are rationalisable if and only if there exist shadow prices and a discount factor satisfying conditions (B1)–(B3). When such objects exist, one can construct a concave, locally non-satiated utility function over characteristics that rationalises observed choices; when they do not, no such representation is possible. The theorem therefore delivers a sharp, nonparametric test of dynamic consistency in characteristics space.

It is useful to interpret the economic content of the three conditions. Condition (B1) imposes cyclical monotonicity on the shadow prices. Economically, it rules out "cycles" in revealed marginal valuations over augmented characteristic bundles: there should be no sequence of observed trades in characteristics space that would allow a costless improvement by returning to the starting point. Formally, cyclical monotonicity is equivalent to concavity of the instantaneous utility function (Rockafellar, 1970). This condition is precisely the behavioural discipline of the model.

Conditions (B2) and (B3) impose the structural pricing restrictions implied by the habits-over-characteristics model. They require that observed goods prices be representable as linear combinations of contemporaneous and forward-looking shadow prices. Economically, current goods prices must internalise both current marginal utility from characteristics and the continuation value induced by habit formation. In the running cereal example, the price of a high salt or sugar cereal must reflect not only current taste for nutty-ness and sensory intensity, but also how today's sensory intensity alters tomorrow's marginal utility. These equalities therefore encode the intertemporal wedge introduced by past consumption directly into the price system.

Finally, note that the mechanism tested here differs from inventory-driven persistence (Hendel and Nevo, 2006): stockpiling generates serial correlation in purchases through intertemporal substitution and storage under time-separable preferences. Contrastingly, persistence in my framework operates through state dependence in utility over characteristics, so current consumption of habit-forming attributes carries a continuation value by shifting future marginal valuations.

Testing consistency reduces to an empirical search for shadow prices and a discount factor satisfying (B1)–(B3). The system is nonlinear jointly in shadow prices and β, but becomes linear conditional on β. For any fixed discount factor, feasibility can therefore be assessed via a linear programme. Repeating this feasibility check over a grid of candidate discount factors yields a computationally straightforward implementation strategy.

A practical complication arises from condition (B1). In its raw form, cyclical monotonicity requires the inequalities to hold for all subsets $\sigma \subseteq \{1, \dots, T\}$, implying $2^T - 1$ constraints. This exponential growth in the number of inequalities pertains specifically to checking condition (B1). Online Appendix G derives an equivalent linear-programming formulation that replaces this exponential family with a quadratic number of constraints in T, making implementation feasible in panel settings.

Interpreting the Afriat test requires care. A positive result establishes existence: there exists *some* concave utility function and discount factor consistent with the data. The representation, however, is not unique. Distinct utility functions—beyond simple monotone transformations—may rationalise the same dataset. Moreover, rationalisability depends on the level of temporal and product aggregation; for example, time aggregation may smooth consumption in a way that mimics habit persistence.

A negative result is likewise not diagnostic of the precise source of failure. Rejection may reflect habit persistence extending beyond one lag, non-concavities in preferences, misspecification of the characteristics technology, or an incorrect partition of $\boldsymbol{A}$ into contemporaneous and habit-forming components (for instance, treating "taste" preferences like nutty-ness as static when the data in fact suggest that exposure today shifts future marginal valuations). The Afriat test should therefore be interpreted as a sharp but reduced-form diagnostic of dynamic rationalisability, rather than as a structural identification device.

2.3 A necessary rank condition for consistency

Although Theorem 2.1 provides a complete characterisation, its direct implementation can become computationally burdensome as the dimension of goods, characteristics, or time increases. This subsection therefore derives a low-dimensional diagnostic implied by consistency: a necessary rank condition that can be checked period by period. Failure of the condition immediately falsifies the model, while satisfaction is necessary but not sufficient.

The restriction is *structural* in the sense that it concerns only whether the maintained goods-to-characteristics technology can in principle support *any* shadow-price representation of observed prices. If the condition fails, no candidate preferences—static or dynamic—can rationalise the data because the implied shadow-price system does not exist.

Let $\boldsymbol{\rho}_t^+$ denote the $K_t^+ \leq K$-dimensional vector of discounted prices of goods consumed in strictly positive quantities in period t. Let $\boldsymbol{B}_t$ denote the $J \times K_t^+$ submatrix of $\boldsymbol{A}$ collecting the contemporaneous characteristics of those goods, and let $\boldsymbol{B}_t^a$ denote the $J_2 \times K_t^+$ submatrix collecting the corresponding habit-forming characteristics. By Theorem 2.1, consistency implies that for every $t \leq T - 1$,

$$\boldsymbol{\rho}_t^+ = \boldsymbol{B}_t' \boldsymbol{\pi}_t^0 + (\boldsymbol{B}_t^a)' \boldsymbol{\pi}_{t+1}^1 = \left[\boldsymbol{B}_t' \,\middle|\, (\boldsymbol{B}_t^a)' \right] \begin{bmatrix} \boldsymbol{\pi}_t^0 \\ \boldsymbol{\pi}_{t+1}^1 \end{bmatrix} =: \tilde{\boldsymbol{B}}_t \begin{bmatrix} \boldsymbol{\pi}_t^0 \\ \boldsymbol{\pi}_{t+1}^1 \end{bmatrix}, \tag{5}$$

where $\tilde{\boldsymbol{B}}_t$ is the $K_t^+ \times (J + J_2)$ augmented technology matrix. Because $\boldsymbol{B}_t^a$ is formed by rows of $\boldsymbol{B}_t$, the augmented matrix $\tilde{\boldsymbol{B}}_t$ has the same column space as $\boldsymbol{B}_t'$, so habit formation tilts shadow prices within the same J-dimensional price manifold rather than expanding it.

Consistency therefore requires the observed price vector $\boldsymbol{\rho}_t^+$ to lie in the column space of $\tilde{\boldsymbol{B}}_t$, yielding the following necessary condition.

Proposition 2.1. If the data $\{\boldsymbol{\rho}_t, \boldsymbol{x}_t\}_{t=1}^T$ are consistent with the one-lag habits-over-characteristics model for technology $\boldsymbol{A}$, then for every $t \leq T-1$,

$$\operatorname{rank}(\tilde{\boldsymbol{B}}_t \mid \boldsymbol{\rho}_t^+) = \operatorname{rank}(\tilde{\boldsymbol{B}}_t) \leq \min\{K_t^+,\, J\}, \tag{NC}$$

where $\mid$ denotes horizontal concatenation.

Violation of (NC) at any single date is sufficient to reject the model, providing a sharp, low-dimensional falsification criterion that can be evaluated period by period. The restriction coincides with the necessary spanning condition under intertemporal separability in Blow et al. (2008): habit formation changes the *level* of shadow prices but does not expand the price manifold implied by the mapping from goods to characteristics. As such, (NC) is a structural constraint driven by the geometry of $\boldsymbol{A}$ rather than by the curvature or stability of preferences. Because (NC) is only necessary, however, it does not guarantee consistency: even when (NC) holds, the inequalities for goods consumed at zero quantities may still be infeasible. Nevertheless, (NC) substantially reduces the feasible set and provides a fast diagnostic for empirical implementation.

Beyond its falsification role, the rank restriction also has identification content. For any t, (5) requires the observed price vector $\boldsymbol{\rho}_t^+$ to lie in the column space of $\tilde{\boldsymbol{B}}_t$. Because $\boldsymbol{B}_t^a$ is formed by rows of $\boldsymbol{B}_t$, the augmented matrix $\tilde{\boldsymbol{B}}_t$ has the same column space as $\boldsymbol{B}_t'$, so the feasible set of price vectors is at most J-dimensional. When habit formation is present ($J_2 > 0$), the mapping from $(\boldsymbol{\pi}_t^0, \boldsymbol{\pi}_{t+1}^1)$ to $\boldsymbol{\rho}_t^+$ is therefore not one-to-one: reallocating shadow value between contemporaneous and lag components along the habit-forming directions leaves $\boldsymbol{\rho}_t^+$ unchanged. Even when (NC) holds, the decomposition into contemporaneous and habit components is thus generically set-identified. In the running cereal example, prices may identify the overall shadow value of "nutty-ness" and "sensory intensity," but not separately how much of the value of sensory intensity reflects current taste versus its continuation value through habits.

2.4 Missing prices

I now relax the assumption that prices are observed for all market goods. In many applications, prices are recorded only for goods that are actually purchased, giving rise to a missing price problem.[4] Missing prices complicate RP analysis because imputing unobserved prices requires auxiliary assumptions. An alternative is to treat missing prices as unknowns and ask whether there exist values that render the data rationalisable. This existence approach should be understood as a partial-identification device: without further restrictions, one can always rationalise non-purchases by assigning prohibitively high unobserved prices, so the goal is to characterise when the observed prices alone already force a violation (or allow rationalisation) under the maintained technology. I now formalise this approach.

[4]Throughout, I assume that the technology matrix $\boldsymbol{A}$ is known and time-invariant. This reflects settings where characteristics can be directly observed or constructed (e.g., nutritional content, design features, emissions ratings), even when market prices are only recorded for purchased items; in practice, missing prices are far more common than missing characteristics.

Let $\boldsymbol{\rho}_t^+$ denote the $K_t^+ \leq K$ sub-vector of period-t discounted prices for goods with strictly positive demand, and let $\boldsymbol{B}_t$ and $\boldsymbol{B}_t^a$ denote the corresponding $J \times K_t^+$ and $J_2 \times K_t^+$ sub-matrices of $\boldsymbol{A}$ and $\boldsymbol{A}^a$. Let $\boldsymbol{\rho}_t^0$, $\boldsymbol{B}_t^0$, and $\boldsymbol{B}_t^{a,0}$ denote the complementary sub-vectors and sub-matrices associated with zero demand. The full discounted price vector is $\boldsymbol{\rho}_t = (\boldsymbol{\rho}_t^+, \boldsymbol{\rho}_t^0)$. I can then state an Afriat-type characterisation for the habits-over-characteristics model with missing prices.

Theorem 2.2. The following statements are equivalent:

(A$^+$) There exist prices $\{\boldsymbol{\rho}_t^0\}_{t=1}^T$ such that the data $\{(\boldsymbol{\rho}_t, \boldsymbol{x}_t)\}_{t=1}^T$ satisfy the one-lag habits-over-characteristics model for given technology $\boldsymbol{A}$.

(B$^+$) There exist shadow discounted prices $\{\boldsymbol{\pi}_t^0\}_{t=1}^T$, $\{\boldsymbol{\pi}_t^1\}_{t=1}^T$ and a discount factor $\beta \in (0, 1]$ such that,

$$0 \leq \sum_{\forall s,t \in \sigma} \tilde{\boldsymbol{\pi}}_s' \left(\tilde{\boldsymbol{z}}_t - \tilde{\boldsymbol{z}}_s\right) \qquad \forall \sigma \subseteq \{1, \ldots, T\} \tag{B1$^+$}$$

$$\boldsymbol{\rho}_t^+ = \boldsymbol{B}_t' \boldsymbol{\pi}_t^0 + (\boldsymbol{B}_t^a)' \boldsymbol{\pi}_{t+1}^1 \quad \forall t \in \{1, \ldots, T-1\} \tag{B2$^+$}$$

where $\tilde{\boldsymbol{\pi}}_t := \frac{1}{\beta^{t-1}} \left[\boldsymbol{\pi}_t^{0\prime}, \boldsymbol{\pi}_t^{1\prime}\right]'$.

Proof: See Appendix A. □

Relative to Theorem 2.1, the missing-price test conditions only on purchased-good prices and is therefore weaker.

3 Corollaries and nesting results

Before turning to the empirical application, I record two immediate corollaries that situate the habits-over-characteristics model within the RP literature. First, when characteristics coincide with market goods, the framework collapses to the habits-over-goods model of Crawford (2010). Second, under intertemporal separability and exponential discounting, it reduces to the characteristics-based model of Blow et al. (2008). Both results follow directly from Theorem 2.1 once the model is specialised to the relevant limiting cases. Formal definitions and derivations for these special cases are provided in Appendix A.

3.1 Habits-over-goods as a special case

When characteristics coincide with market goods—that is, when $J = K$ and the technology matrix satisfies $\boldsymbol{A} = \boldsymbol{I}_J$—the habits-over-characteristics framework reduces to a standard habits-over-goods model. In this case, the distinction between goods and characteristics disappears, and the intertemporal first-order conditions involve current and lagged consumption of habit-forming goods directly.

Definition 3.1. The data $\{\boldsymbol{\rho}_t^c, \boldsymbol{\rho}_t^a; \boldsymbol{x}_t^c, \boldsymbol{x}_t^a\}_{t=1}^T$ are *consistent* with the one-lag habits-over-goods model if there exists a locally non-satiated, superdifferentiable, and concave utility function $u(\cdot)$ and positive constants λ and β such that for all $t \in \{1, \ldots, T-1\}$:

$$\boldsymbol{\rho}_t^c \geq \frac{\beta^{t-1}}{\lambda} \partial_{\boldsymbol{x}_t^c} u(\bar{\boldsymbol{x}}_t), \tag{6}$$

$$\boldsymbol{\rho}_t^a \geq \frac{\beta^{t-1}}{\lambda} \partial_{\boldsymbol{x}_t^a} u(\bar{\boldsymbol{x}}_t) + \frac{\beta^t}{\lambda} \partial_{\boldsymbol{x}_t^a} u(\bar{\boldsymbol{x}}_{t+1}), \tag{7}$$

with equality whenever $x_t^k > 0$, and where $\boldsymbol{\rho}_t^c$ and $\boldsymbol{\rho}_t^a$ denote present-value prices of non-habit-forming and habit-forming goods, respectively.

This corresponds to Definition 1 in Crawford (2010), where it is assumed that all goods are consumed in strictly positive quantities. Given this definition, I obtain the following result, equivalent to that in Crawford (2010).[5]

Corollary 3.1. The following statements are equivalent:

(A') The data $\{\boldsymbol{\rho}_t^c, \boldsymbol{\rho}_t^a; \boldsymbol{x}_t^c, \boldsymbol{x}_t^a\}_{t=1}^T$ are consistent with the one-lag habits-over-goods model.

(B') There exist $2T$ shadow price vectors $\{\boldsymbol{\rho}_t^{a,r}\}_{t=1,\ldots,T}^{r=0,1}$ and positive constants λ, β such that:

$$0 \leq \sum_{s,t \in \sigma} \tilde{\boldsymbol{\rho}}_s' (\bar{\boldsymbol{x}}_t - \bar{\boldsymbol{x}}_s) \qquad \forall \sigma \subseteq \{1, \ldots, T\}, \tag{B1'}$$

$$\boldsymbol{\rho}_t^a \geq \boldsymbol{\rho}_t^{a,0} + \boldsymbol{\rho}_{t+1}^{a,1} \qquad \forall t \in \{1, \ldots, T-1\}, \tag{B2'}$$

$$\rho_t^{a,k} = \boldsymbol{e}_k \boldsymbol{\rho}_t^{a,0} + \boldsymbol{e}_k \boldsymbol{\rho}_{t+1}^{a,1} \qquad \text{if } x_t^k > 0, \quad \forall k, t \in \{1, \ldots, T-1\}, \tag{B3'}$$

where $\bar{\boldsymbol{x}}_t := (\boldsymbol{x}_t^{c\prime}, \boldsymbol{x}_t^{a\prime}, \boldsymbol{x}_{t-1}^{a\prime})'$, $\boldsymbol{e}_k$ is the k-th standard basis vector, and:

$$\tilde{\boldsymbol{\rho}}_t := \frac{\lambda}{\beta^{t-1}} \left[\boldsymbol{\rho}_t^{c\prime}, \boldsymbol{\rho}_t^{a,0\prime}, \boldsymbol{\rho}_t^{a,1\prime} \right]'.$$

Proof: See Appendix A. □

3.2 Intertemporally separable preferences over characteristics

If habit formation is absent so that $\boldsymbol{z}_t = \boldsymbol{z}_t^c$ for all t, the model reduces to a characteristics-based framework with intertemporally separable preferences. Unlike Blow et al. (2008), who remain agnostic about intertemporal allocation, my formulation embeds this static characteristics model within a lifecycle problem with exponential discounting.[6] The presence of a single intertemporal budget constraint implies a single shadow value of lifetime wealth, so observed choices must satisfy both intratemporal utility maximisation over characteristics and intertemporal optimality. Consistency with this lifecycle formulation is defined as follows.

Definition 3.2. The data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t=1}^T$ are *consistent* with intertemporally separable preferences over characteristics and a life-cycle model for given technology $\boldsymbol{A}$ if there exists a locally non-satiated, superdiffer-

[5]Equivalence requires setting $\lambda = 1$, which can be done without loss of generality; see Section 2.2.

[6]By a lifecycle problem I mean that the consumer chooses the entire consumption path to maximise lifetime utility subject to a single present-value budget constraint under exponential discounting.

entiable, and concave utility function $u(\cdot)$ and positive constant λ such that, for all $t \in \{1, \ldots, T-1\}$:

$$\boldsymbol{\rho}_t \geq \boldsymbol{A}'\boldsymbol{\pi}_t, \tag{8}$$

with equality for all k such that $x_t^k > 0$, where $\boldsymbol{z}_t = \boldsymbol{A}\boldsymbol{x}_t$ and $\boldsymbol{\rho}_t$ denotes the vector of present-value prices.

This recovers the hedonic pricing equation of Gorman (1956) and the first-order condition of Blow et al. (2008) when the static characteristics model is embedded in a lifecycle framework. From this, I obtain the following characterisation.

Corollary 3.2. The following are equivalent:

(A") The data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t=1}^T$ are consistent with intertemporally separable preferences over characteristics and a life-cycle model.

(B") There exist T shadow price vectors $\{\boldsymbol{\pi}_t\}_{t=1}^T$ and a positive constant λ such that:

$$0 \leq \sum_{s,t \in \sigma} \boldsymbol{\pi}_s'(\boldsymbol{z}_t - \boldsymbol{z}_s) \qquad \forall\, \sigma \subseteq \{1, \ldots, T\}, \tag{B1"}$$

$$\rho_t^k \geq \boldsymbol{a}_k'\boldsymbol{\pi}_t \qquad \forall\, k, t \in \{1, \ldots, T-1\}, \tag{B2"}$$

$$\rho_t^k = \boldsymbol{a}_k'\boldsymbol{\pi}_t \qquad \text{if } x_t^k > 0, \quad \forall\, k, t \in \{1, \ldots, T-1\}, \tag{B3"}$$

where $\boldsymbol{a}_k$ is the kth column of $\boldsymbol{A}$.

Proof: See Appendix A. □

Together, these corollaries clarify the paper's value added. The framework unifies existing RP results in goods space and in characteristics space within a single dynamic hedonic model, and it shows exactly how habit formation over attributes tightens empirical content relative to intertemporal separability: it imposes additional intertemporal price restrictions and delivers diagnostics—via feasibility of (B1)–(B3) and the rank condition (NC)—that separate failures of the hedonic price representation from failures of dynamically coherent preferences.

4 Habits and the rationalisability of cereal purchases

This section asks: *when do habits matter for explaining dynamic consumption patterns in scanner data?* Using household-level cereal purchases, I show that most differences in raw rationalisability across models are driven by the *structural* price restrictions imposed by the hedonic representation, while allowing for habits systematically improves *behavioural* coherence conditional on those constraints. I illustrate the framework using household scanner data on ready-to-eat cereal, where transaction-level purchases permit dynamic RP restrictions to be evaluated using observed intertemporal variation rather than cross-sectional substitution alone.

The empirical exercise proceeds sequentially. First, I ask a structural question: do observed within-period price vectors admit a hedonic representation? Equivalently, do they satisfy the equalities implied by the goods-to-characteristics technology (i.e., ($B2^+$) in Theorem 2.2), so that characteristic shadow prices are well defined? These equalities act as a gatekeeper: if they fail, the behavioural test is not economically meaningful because the relevant shadow prices do not exist. Second, conditional on structural feasibility, I ask a behavioural question:

do observed purchase sequences satisfy the dynamic RP inequalities (i.e., (B1$^+$)) when evaluated at the implied shadow prices? This second stage isolates whether allowing for habits improves intertemporal coherence relative to static preferences.

The sequential decomposition yields two findings. First, most variation in raw pass rates when moving from goods space to characteristics space reflects the hedonic price-system restrictions rather than differences in behavioural fit; empirically, however, restoring hedonic consistency typically requires only modest price adjustments. Second, conditional on satisfying those structural restrictions, allowing for habits systematically improves behavioural coherence.

Market goods are UPC-level products and characteristics are their nutritional content and a small set of descriptive indicators. The key empirical challenge is that prices are observed only for purchased items. Accordingly, I implement the missing-price test of Section 2.4.[7] In this setting, behavioural discipline is inherently limited by sparse price support and modest intertemporal budget variation, so the diagnostic provides a lower bound on the model's behavioural content. In environments with richer price and quantity support, the same framework may generate sharper behavioural restrictions.

4.1 Data

I use household-level scanner data on cold cereal purchases from the IRI Academic Datasets' *BehaviorScan* panel. The panel covers two U.S. markets (Pittsfield, MA, and Eau Claire, WI) over 2010–2011, with purchases recorded at checkout via household ID cards and Universal Product Codes (UPCs). I restrict attention to static households that participate in all 12 months of a calendar year, so recruitment and attrition occur only at year-end.

I compute household-specific time periods to accommodate heterogeneity in purchase frequencies. For household i, let S_i denote the span from first to last purchase and G_i the longest interpurchase gap (including endpoints). I set $T_i = \lfloor S_i/G_i \rfloor$ and partition S_i into T_i equal-length bins, which guarantees at least one purchase in each period by construction. Households with $T_i < 3$ are excluded to ensure at least two observed transitions in the one-lag model. This aggregation is conservative for a one-lag specification: it ensures that the lagged bundle is observed rather than imputed, hence dynamic restrictions are evaluated on realised purchase transitions. After additionally dropping households with purchases lacking characteristics information (described below), the analysis sample contains $N = 2{,}282$ households.[8] Table 1 summarises the resulting panel structure and the scale and variety of the implied choice problems.

Households purchase on average 7.9 units per period (median 6.0). Combined with a median period length of 98 days, this pattern is consistent with ongoing consumption rather than extreme purchase spikes. While promotions may induce some stockpiling, explaining the observed interpurchase gaps purely through inventory accumulation would require implausibly large stock build-ups relative to total quantities purchased. Most periods also contain multiple UPCs, indicating sustained consumption rather than isolated sale-driven bunching. I therefore interpret each household-period as a meaningful dynamic choice observation: it accommodates idiosyncratic shopping frequencies while preserving the temporal structure required for testing dynamic RP conditions.[9]

[7] Imputing prices for unpurchased goods is feasible (e.g., using regional price indices), but it introduces auxiliary assumptions that can blur whether failures reflect preferences or the imputation procedure.

[8] Appendix C documents the full sequence of sample construction and reports balance tests comparing households excluded due to missing characteristics data with the final analysis sample. Differences are modest in magnitude, suggesting limited scope for selection on observables.

[9] Similar coarse aggregation is common in empirical work on dynamic demand when purchase occasions are intermittent (e.g., Crawford (2010) uses quarterly panels in an application to tobacco). I obtain qualitatively similar comparisons across specifications under a common monthly aggregation (i.e., $T_i = 24$ for all i), which increases time resolution but at the cost of more zero-purchase periods and a smaller effective sample.

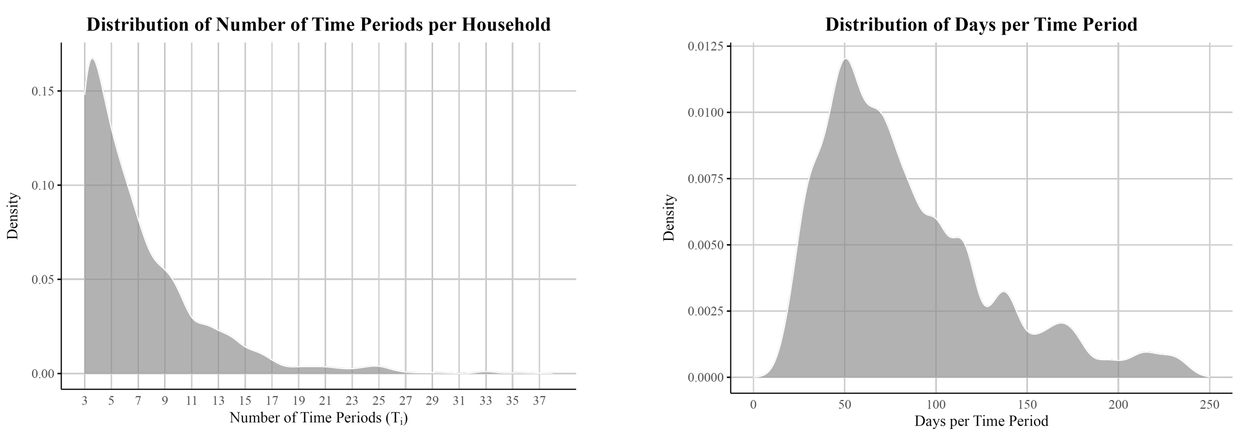
Table 1: Sample and household-level panel structure

Statistic	Mean	Median
Households (N)	2,282	
Months covered	2010–2011 (24 months)	
Distinct products (K)	801	
Characteristics (J)	23	
Time periods per household (T_i)	7.0	6.0
Period length (days)	106	98
Units purchased per household-period	7.9	6.0
Total units purchased (24 months)	55.1	42.0
Expenditure per household-period ($)	22.1	19.5
Total expenditure (24 months, $)	154.2	116.5
Distinct products purchased (24 months)	32.3	25.0

Notes: Statistics are computed across $N = 2{,}282$ households. T_i denotes the number of constructed time periods for household i, defined as $T_i = \lfloor S_i/G_i \rfloor$, where S_i is the span between first and last purchase and G_i the longest interpurchase gap. Units and expenditures are aggregated within household-periods and totals are over 24 months.

Figure 1 shows substantial heterogeneity in shopping frequency (median $T_i = 6$; median period length 98 days). Within each period I aggregate UPC-level quantities and nominal expenditures and compute unit values as expenditure-to-quantity ratios.[10] Because unpurchased goods have no observed unit values, the test treats missing prices as unknowns and searches for completions consistent with rationalisability. Reported pass rates are therefore upper bounds: a household that fails cannot be rescued by any price imputation, while a household that passes does so under at least one completion.

Figure 1: Distribution of constructed time periods and period lengths



(a) Distribution of T_i across households

(b) Distribution of period lengths (days)

Notes: Left panel shows the distribution of T_i across households ($N = 2{,}282$). Right panel shows the implied period length in days. $T_i = \lfloor S_i/G_i \rfloor$, where S_i is the span between first and last purchase and G_i the longest interpurchase gap.

The model is defined over characteristics rather than market goods. For each UPC I construct a vector of

[10] Expenditures are recorded in nominal dollars and converted to present-value terms for the lifecycle formulation using a monthly interest-rate series (30-Year Fixed Rate Mortgage Average, FRED (Freddie Mac, 2025)); results are unchanged under nominal values given the short sample window.

nutritional and descriptive attributes by merging the IRI product file to the (now-defunct) NuVal shelf-labelling database and, where NuVal is missing, supplementing with data from the FatSecret Platform API, following the approach in Barahona et al. (2023). I standardise all nutrients to a 100g basis so that characteristics are linear in quantities. Because IRI assigns placeholder codes to some private-label items, I first map System-88 pseudo-UPCs to their corresponding real UPCs prior to merging. The merged characteristics cover over 97% of purchase-weighted observations; I drop unmatched purchase records and then drop households with any remaining unmatched purchases, so every household in the analysis sample has a complete characteristics mapping for all of its observed purchases.

I construct $J = 23$ characteristics. Eight are continuous Nutrition Facts Panel measures (calories, carbohydrates, total fat, saturated fat, fibre, protein, sodium, and sugar). The remaining 15 are binary indicators capturing salient ingredients and descriptors (10 indicators) and the five most prevalent brands (Kellogg's, General Mills, Post, Quaker, and Kashi). Table 2 reports the binary definitions. In the baseline specification I treat sugar and sodium as habit-forming characteristics ($J_2 = 2$) and take the remaining characteristics as non-habit-forming, motivated by evidence that sugar and salt may activate reward pathways in ways analogous to addictive substances (Avena et al., 2008; Cocores and Gold, 2009). Section 4.2.1 reports robustness to alternative partitions. Additional descriptive evidence on purchase intensity, brand concentration, and the distributions of prices and characteristics is reported in Appendix C.

Table 2: Binary product characteristics used in the hedonic representation

Binary characteristic	Construction
whole grain	contains 'whole grain' or 'wholegrain' in ingredient list
organic	contains the term 'organic' at least once in ingredient list
oat-based	contains 'oat' as part of the first listed ingredient
granola	contains 'granola' in product description
health halos	contains 'source of', 'natural', 'low calorie', or 'low fat' in product description
fruity	contains 'fruit' in the product description
nutty	contains 'nut' in product description
chocolatey	contains 'chocolate' in product description
honey flavour	contains 'honey' in product description
gluten-free	contains 'gluten free' in product description
Kellogg's	cereal brand is Kellogg's
General Mills	cereal brand is General Mills
Post	cereal brand is Post
Quaker	cereal brand is Quaker
Kashi	cereal brand is Kashi

Notes: Characteristics are defined at the UPC level using ingredient lists and product descriptions from the merged IRI–NuVal–FatSecret dataset. Indicators equal one if the stated textual condition is satisfied. Brand indicators correspond to the five most prevalent national brands in the sample. Continuous nutritional characteristics (calories, carbohydrates, total fat, saturated fat, fibre, protein, sodium, and sugar, standardised per 100g) are defined separately in Appendix C.

4.2 Results

4.2.1 Rationalisability scores

Two patterns emerge in the raw rationalisability outcomes: allowing for habits increases pass rates within characteristics space, and goods-based representations exhibit substantially higher pass rates. As the sequential logic of the paper makes clear, however, these raw differences conflate structural and behavioural components.

Section 4.2.2 decomposes these margins and quantifies the severity of violations.

The test is implemented at the household level, allowing full heterogeneity in pass/fail outcomes and in the shadow-price structure of rationalisable households. Because habits enter with a one-period lag, the test uses $t = 2, \ldots, T-1$: period $t = 1$ lacks a lag, while period $t = T$ involves the forward-looking term $\boldsymbol{\pi}^1_{t+1}$. Unless otherwise specified, I impose a lifecycle model with a single intertemporal budget constraint, so the marginal utility of income is constant across periods.

Under the baseline habits-over-characteristics specification, 1,248 of 2,282 households (54.69%) satisfy the test. Varying the set of habit-forming characteristics has essentially no effect on classification: restricting habits to sugar alone (54.60%), sodium alone (54.65%), or allowing all 23 characteristics to be habit-forming (54.69%) changes the pass rate by at most 0.09 percentage points and reclassifies no more than two households. Table 3 reports the full set of specifications.

Table 3: Household-level pass rates: habits-over-characteristics specifications

Model	Description	Lifecycle model	Pass rate
Habits-over-characteristics	$J_2 = 2$ (sugar and sodium)	Yes	54.69%
Habits-over-sugar	$J_2 = 1$ (sugar)	Yes	54.60%
Habits-over-sodium	$J_2 = 1$ (sodium)	Yes	54.65%
Habits-over-all-characteristics	$J_2 = J = 23$	Yes	54.69%

Notes: A household passes if there exists a completion of missing prices and characteristic shadow prices satisfying the dynamic RP conditions (B1$^+$) and structural equalities (B2$^+$) in Theorem 2.2. Pass rates are computed over $N = 2{,}282$ households. $J = 23$ denotes the total number of characteristics and J_2 the number assumed to be habit-forming. All specifications impose a lifecycle model with a single intertemporal budget constraint.

Raw pass rates differ sharply across representations (Table 4). Moving from characteristics to goods sharply increases rationalisability from 54.7% to 99.6% under dynamic preferences. This large difference is driven by the much greater dimensional flexibility of goods-based representations, which impose no cross-good price restrictions. Moreover, many of the apparent "failures" in characteristics space correspond to economically small deviations from the hedonic price restrictions. Section 4.2.2 makes this precise by separating structural and behavioural sources of empirical discipline and quantifying the magnitude of violations.

By contrast, removing habits within characteristics space reduces the pass rate from 54.7% to 52.4%. While this difference in levels is only modest, the paired comparisons below show that the reclassification is entirely directional, with households failing under static preferences but passing once dynamics are introduced. Moreover, I show later that the reduction in the severity of violations is substantially larger than the change in the binary pass rate suggests, indicating that intertemporal coherence improves even among households that remain formally non-rationalisable.

The reclassification pattern is strongly directional. Removing habits from the baseline characteristics model reclassifies 53 households, all of whom fail under static preferences but pass once intertemporal dependence is introduced (p-value $< 10^{-15}$). By contrast, changing the allocation of habit-forming characteristics reclassifies at most three households and yields no statistically significant differences. Table 5 reports the full set of paired comparisons.

Comparisons with goods-based models reveal even larger directional differences. In these cases, most switching households fail the characteristics-based test but pass the corresponding goods-based alternative. As shown in the next subsection, this pattern reflects the much greater dimensional flexibility of the goods representation, which imposes no cross-good price restrictions, rather than tighter behavioural alignment.

Taken together, these results indicate that while the precise *allocation* of habits across characteristics plays

Table 4: Household-level pass rates across characteristics and goods representations

Model	Description	Lifecycle model	Pass rate
Habits-over-characteristics	$J_2 = 2$ (sugar and sodium)	Yes	54.69%
Characteristics (no habits)	$J_2 = 0$	Yes	52.37%
Habits-over-all-goods	$K = J = J_2$, $\boldsymbol{A} = \text{identity}(K)$	Yes	99.56%
Goods (no habits)	$K = J = J_1$, $\boldsymbol{A} = \text{identity}(K)$	Yes	92.59%
Goods (no habits) (GARP)	$K = J = J_1$, $\boldsymbol{A} = \text{identity}(K)$	No	99.69%

Notes: A household passes if there exists a completion of missing prices and characteristic shadow prices satisfying the dynamic RP conditions ($B1^+$) and structural equalities ($B2^+$) in Theorem 2.2. Pass rates are computed over $N = 2{,}282$ households. In characteristics models, $J = 23$ and J_2 denotes the number of habit-forming characteristics. In goods models, K denotes the number of goods and $\boldsymbol{A} = \text{identity}(K)$ implies no cross-good price restrictions. "Lifecycle model" indicates whether a single intertemporal budget constraint is imposed.

Table 5: Paired model comparisons (exact McNemar tests)

Comparison (baseline vs alternative)	Pass_0	Pass_1	Δ (pp)	Switchers	p-value
Baseline: Habits-over-characteristics	54.69				
Habits-over-all-characteristics	54.69	54.69	0.00	0	1.000
Habits-over-sugar	54.69	54.60	0.09	2	0.500
Habits-over-sodium	54.69	54.65	0.04	1	1.000
Characteristics (no habits)	54.69	52.37	2.32	53	$< 10^{-15}$
Habits-over-all-goods	54.69	99.56	-44.87	1024	$< 10^{-15}$
Goods (no habits)	54.69	92.59	-37.91	1039	$< 10^{-15}$
Goods (no habits) (GARP)	54.69	99.69	-45.00	1031	$< 10^{-15}$

Notes: Pass_0 is the baseline pass rate (habits-over-characteristics with $J_2 = 2$), Pass_1 is the alternative model pass rate, and Δ (pp) reports Pass_0 minus Pass_1 in percentage points. "Switchers" is $n(1 \to 0) + n(0 \to 1)$, i.e., the number of households whose pass/fail classification differs across the paired models. p-values are from exact McNemar (binomial) tests based on the switchers.

little role in classification, introducing intertemporal dependence in characteristics space systematically improves rationalisability relative to a static model. In the language of the paper's title, habits matter *conditional on the hedonic representation*: they resolve a directional asymmetry whereby households fail under static characteristics but pass once dynamics are allowed.

4.2.2 Structural and behavioural sources of empirical discipline

Raw pass rates alone are insufficient to interpret fit: they conflate empirical success with permissiveness and provide only a binary measure of failure. As emphasised by Selten (1991), a model may rationalise many datasets either because it captures economically meaningful structure or because it imposes few substantive restrictions on observable outcomes. Moreover, a binary pass/fail outcome does not reveal how *severe* a violation is when the model fails. I therefore separate rationalisability into structural and behavioural components and quantify violations on each margin using continuous discrepancy measures. This approach follows a broader methodological insight that representation theorems naturally induce continuous measures of rationality violations, since their axioms hold if and only if a rationalising object exists (e.g., Andrews (2026)).

In the hedonic setting, this distinction admits a natural decomposition. Rationalisability in the habits-over-characteristics model requires joint satisfaction of two conceptually distinct restrictions. *Structural equalities* ($B2^+$) link observed prices to product characteristics through the hedonic technology and act as overidentifying restrictions on the price system. This is the gatekeeper stage of the empirical analysis. *Behavioural inequalities* ($B1^+$) constrain intertemporal choice through shadow prices and the discount factor (Theorem 2.2). Separating these margins clarifies where empirical discipline originates in characteristics-based valuation and how it differs from more flexible goods-based representations.

The structural equalities require that, in each household-period, the observed price vector $\boldsymbol{\rho}_t^+$ lies in the column space of the augmented characteristics matrix $\tilde{\boldsymbol{B}}_t := \left[\boldsymbol{B}_t' \mid (\boldsymbol{B}_t^a)'\right]$. This restriction is demanding. Even after conditioning on a rich set of nutritional and descriptive characteristics, scanner prices reflect retailer pricing strategies, temporary promotions, and mark-ups driven by market power or shelf placement that do not correspond to attributes households consume. Such components cannot generally be represented as linear combinations of consumption-relevant characteristics. As a result, the rank condition $\text{rank}(\tilde{\boldsymbol{B}}_t \mid \boldsymbol{\rho}_t^+) = \text{rank}(\tilde{\boldsymbol{B}}_t)$ frequently fails in the data.

To quantify the severity of these violations, I compute for each household-period the Euclidean distance

$$d_t = \left\|\boldsymbol{\rho}_t^+ - \tilde{\boldsymbol{B}}_t \tilde{\boldsymbol{B}}_t^+ \boldsymbol{\rho}_t^+\right\|,$$

where $\tilde{\boldsymbol{B}}_t^+$ denotes the Moore–Penrose pseudoinverse and $\tilde{\boldsymbol{B}}_t \tilde{\boldsymbol{B}}_t^+$ is the orthogonal projector onto the equality manifold implied by the hedonic representation. Equivalently, observed prices admit the orthogonal decomposition

$$\boldsymbol{\rho}_t^+ \;=\; \tilde{\boldsymbol{B}}_t \hat{\boldsymbol{\pi}}_t \;+\; \boldsymbol{r}_t,$$

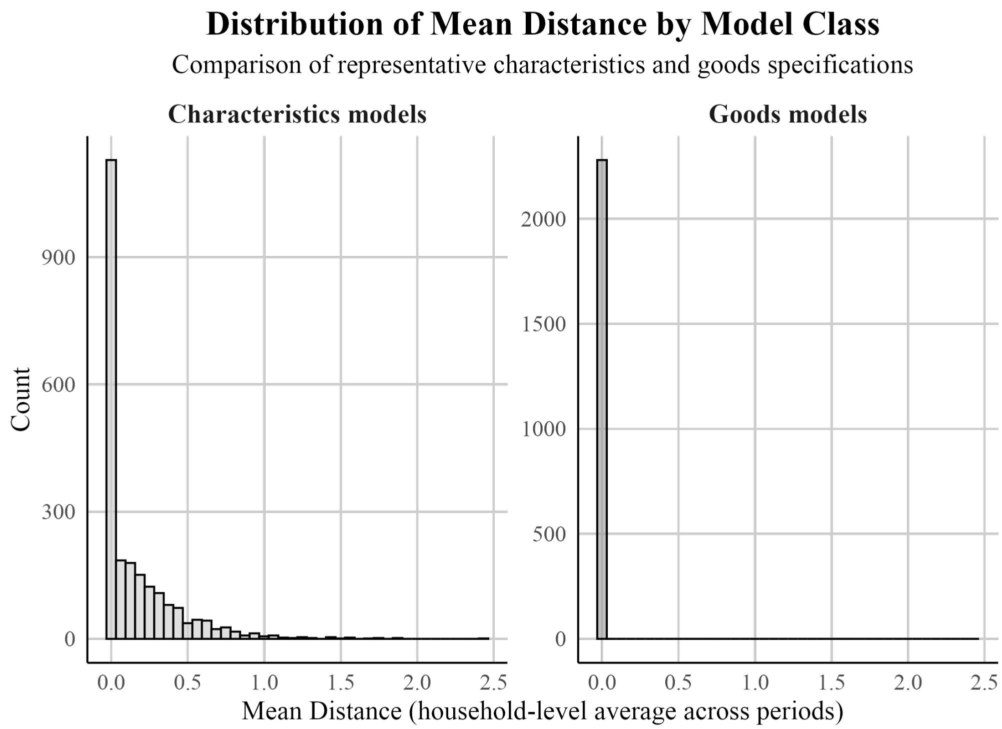
where $\hat{\boldsymbol{\pi}}_t$ minimises $\|\boldsymbol{\rho}_t^+ - \tilde{\boldsymbol{B}}_t \boldsymbol{\pi}\|$ and $\boldsymbol{r}_t$ is orthogonal to the column space of $\tilde{\boldsymbol{B}}_t$. The distance $d_t = \|\boldsymbol{r}_t\|$, measured in dollars, therefore captures the minimal joint price adjustment required for a hedonic price representation to exist. For example, $d_t = 1.0$ indicates that prices would need a total adjustment of 1 dollar (in Euclidean norm across the goods purchased that period) to satisfy the hedonic equalities.

One response to violations of the hedonic equalities is to augment the technology with latent characteristics, as in Blow et al. (2008), thereby expanding the price manifold until the equalities hold. I do not pursue this route here. My objective is not to restore rationalisability by construction, but to quantify how demanding a given hedonic representation is in the data and to separate structural misspecification from behavioural inconsistency.

Figure 2 plots the distribution of mean distances across households for both characteristics- and goods-based specifications. The characteristics models exhibit a wide, right-skewed distribution centred above zero, reflecting the tight geometric constraints imposed by the hedonic representation. Even so, the implied violations are typically small: most household-periods require price adjustments below one dollar to restore hedonic consistency, compared to an average expenditure of \$22.1 per period. By contrast, goods-based models impose no cross-good price restrictions and therefore satisfy the structural equalities mechanically, yielding $d_t = 0$.

An important implication of this geometry is that all characteristics-based specifications impose identical structural restrictions. Habit-forming characteristics enter $\tilde{\boldsymbol{B}}_t$ only as duplicated rows of $\boldsymbol{B}_t$ and therefore neither increase its rank nor enlarge the feasible price set. Structural restrictiveness is thus governed entirely by the hedonic representation itself, not by the allocation of habits across characteristics. Moreover, these structural restrictions operate at the level of prices and characteristics and are conceptually distinct from inventory dynamics: even if households smooth consumption through inventories, prices must still admit a hedonic representation for shadow prices to be well defined.

Figure 2: Mean distance-to-manifold



Notes: Histograms of the household-level mean of the structural distance measure across periods. Distances are measured in dollars and represent the minimal joint price adjustment required for the hedonic price equalities to hold. Left panel reports the characteristics specification; right panel the goods specification.

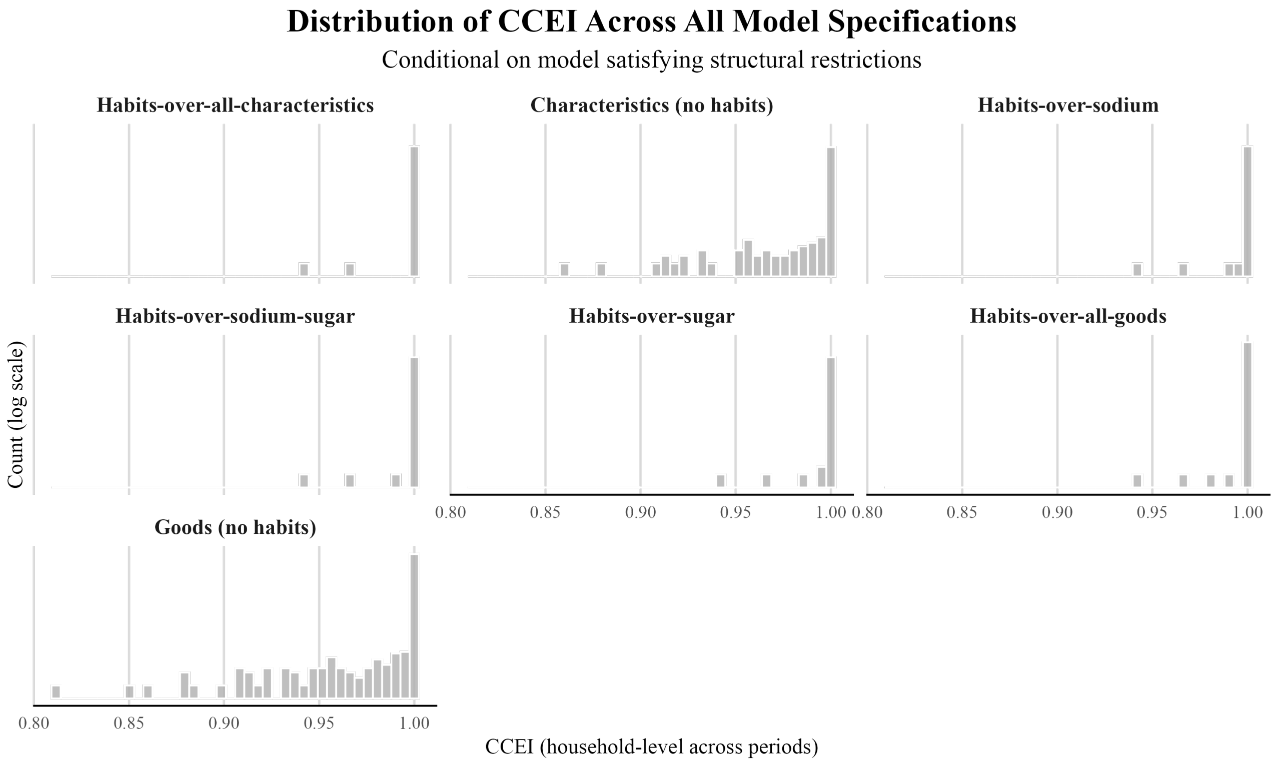
To measure behavioural restrictiveness, I adapt the Critical Cost Efficiency Index (CCEI) following Afriat (1973) and Varian (1990). The CCEI measures the smallest proportional relaxation of revealed affordability required for the behavioural inequalities in (B1$^+$) to admit a solution.[11] Economically, a CCEI of η indicates that shrinking all budgets by at most $(1-\eta) \times 100\%$ suffices to eliminate every intertemporal RP violation. Values close to one therefore indicate that behaviour is nearly dynamically rational, while lower values signal more severe violations. The index is defined only for households whose prices lie exactly on the equality manifold, because characteristic shadow prices—and hence the behavioural inequalities themselves—are well defined only when the hedonic equalities hold. In this sense, the CCEI plays an analogous role to a goodness-of-fit statistic for intertemporal choice: it quantifies how much slack must be introduced before the model can rationalise observed behaviour, holding the price system fixed.

Figure 3 shows that behavioural violations are generally modest, but systematically smaller when habits are allowed. All characteristics-based specifications exhibit substantial mass near unity, indicating that once the hedonic structure is satisfied, only limited perturbations are needed to rationalise behaviour. However, the static characteristics model without habits displays a thicker lower tail, reflecting greater intertemporal inconsistency even when structural feasibility holds. A parallel pattern arises in the goods domain: the habits-over-goods specification yields CCEIs tightly concentrated near one, while the static goods model exhibits a wider

[11] The CCEI is interpreted here strictly as a measure of RP slack, following its original cost-efficiency interpretation in Afriat (1973). As emphasized by Echenique (2022), it should *not* be interpreted as a welfare loss or a measure of foregone surplus.

distribution with more mass below 0.95. These patterns indicate that habit formation strengthens behavioural discipline by reducing intertemporal reversals. This matters for interpretation: habits change the mapping from observed prices to marginal valuations by introducing the dynamic wedge characterised in Section 2.

Figure 3: Critical Cost Efficiency Index (CCEI)



Notes: Histograms of household-level CCEI values across model specifications. The CCEI is the smallest proportional relaxation of the behavioural RP inequalities required for feasibility; values closer to one indicate smaller violations. Computed only for households satisfying the structural price equalities. Vertical axis is on a log scale.

While this decomposition clarifies *where* empirical discipline originates within the model, it does not by itself establish how informative these restrictions are relative to plausible alternatives in the empirical environment. I therefore complement it with a simulation benchmark in the spirit of Fudenberg et al. (2023), which evaluates how unusually well the model fits the observed data relative to nearby feasible perturbations of the scanner environment. The benchmark preserves each household's zero pattern and total expenditure while allowing prices and quantities to vary locally, thereby inducing a comparison distribution over structural and behavioural discrepancy measures. Full details are provided in Appendix E.

On the structural margin, the benchmark confirms that the hedonic equalities impose substantial empirical discipline. Observed scanner prices are, on average, closer to the equality manifold than locally perturbed price systems, yet are rarely extreme outliers relative to the induced comparison distribution. This indicates that the hedonic restrictions are demanding and not trivially satisfied. By contrast, goods-based models impose no cross-good price restrictions and therefore cannot fail on the structural margin, accounting for their much higher raw pass rates.

On the behavioural margin, CCEI values are extremely close to one in both the observed and locally perturbed data, leaving little scope for behaviour to appear unusually efficient in a quantile sense. This reflects the empirical environment—sparse active choice sets and limited intertemporal budget variation—rather than a lack of behavioural content in the model.

Taken together, the price geometry, behavioural slack measures, and simulation benchmark clarify the interpretation of raw pass rates. Structural restrictions embedded in the hedonic representation largely determine differences across goods and characteristics models, while conditional on those restrictions, allowing for habit formation systematically improves intertemporal coherence relative to static preferences. Goods-based models achieve high pass rates primarily because they impose little structural content, not because they deliver a tighter account of behaviour.

4.2.3 Discount factors

The RP conditions are evaluated conditional on the discount factor β. Appendix D reports the fraction of rationalisable households consistent with each value on a fine grid $\beta \in [0.95, 1]$. Acceptance probabilities are uniformly high across the grid under both habits-over-characteristics and habits-over-goods specifications, with no economically meaningful monotonic pattern.

These results indicate that, conditional on the hedonic price restrictions, the behavioural inequalities impose only weak discipline on intertemporal discounting in this environment. The identification sets for β are wide, cautioning against interpreting discount factors recovered from nonparametric dynamic RP tests as tightly identified structural parameters in scanner data.

4.2.4 Predictors of rationalisability

Pass/fail outcomes are primarily associated with the scale and complexity of the observed choice problem rather than with demographics per se. Table 6 reports average marginal effects from probit regressions of the pass indicator on household characteristics. Column (1) includes demographics only; Column (2) adds measures of purchasing intensity and variety.

In the demographic-only specification, households with children are 10.6 percentage points less likely to pass. Once purchasing controls are added, this effect attenuates and becomes statistically insignificant. By contrast, scale measures remain economically large and significant: an interquartile increase in two-year cereal expenditure is associated with roughly a 25 percentage point lower probability of passing, while a comparable increase in product variety reduces the probability by about 8 percentage points.

Taken together, these results suggest that household composition matters primarily through the scale and complexity of the observed choice problem. Larger and more diverse purchasing patterns generate a greater number of structural and behavioural constraints, increasing the likelihood that at least one is violated. For instance, when a household purchases products with sharply different nutritional profiles across periods, the implied cross-attribute shadow-price system must rationalise a wider range of price–quantity trade-offs than in households that repeatedly buy a narrow set of similar cereals. Demographics per se have limited explanatory power once this effective dimensionality is accounted for. The McFadden pseudo-R^2 of 0.176 indicates moderate overall fit.

Table 6: Determinants of rationalisability: probit average marginal effects

	Dependent variable: Pr(Pass)	
	(1) demographics only	(2) w/ purchasing controls
young child	−0.106	−0.081
	(0.084)	(0.079)
children	−0.106**	−0.012
	(0.033)	(0.030)
age group HH head	0.004	−0.006
	(0.012)	(0.011)
education HH head	−0.00004	0.001
	(0.008)	(0.007)
Pittsfield	0.005	−0.001
	(0.021)	(0.022)
high income HH†	−0.036	0.008
	(0.025)	(0.023)
total units purchased		0.00002
		(0.001)
total expenditure		−0.002***
		(0.0003)
total unique products		−0.003**
		(0.001)
Pseudo-R^2 (McFadden)	0.176	0.176
Observations	2,187	2,187

*p<0.1; **p<0.05; ***p<0.01.

†High-income indicator equals one for pre-tax income ≥ 75k.

Notes: Entries report average marginal effects from probit models of an indicator for passing the habits-over-characteristics test. Column (1) includes demographic controls only; column (2) additionally includes purchasing intensity measures. Standard errors in parentheses are based on the delta method. 95 observations omitted due to missing demographic data.

5 Conclusion

This paper provides a nonparametric foundation for dynamic hedonic valuation. I characterise exactly when observed prices and choices admit a coherent life-cycle interpretation in which utility depends on current characteristics and lagged consumption of a habit-forming subset. The main result is an Afriat-type theorem in characteristics space that delivers necessary and sufficient conditions for rationalisability, together with a missing-price extension suited to scanner environments.

The framework sharpens the empirical content of hedonic valuation along three margins. First, it clarifies how habits alter the economic interpretation of hedonic prices by introducing a dynamic continuation-value wedge. Second, it separates two distinct sources of discipline: structural equalities that restrict prices through the maintained characteristics technology and behavioural inequalities that restrict intertemporal choice conditional on that technology. Third, it provides quantitative diagnostics that distinguish the incidence of model rejection from its economic severity.

When do habits matter for hedonic valuation? The answer is sequential. Habits may influence behaviour in any dynamic environment, but they matter for hedonic valuation only once a coherent characteristics representation exists. A dynamic hedonic interpretation is admissible only if the maintained characteristics technology spans observed prices so that characteristic shadow prices are well defined. Conditional on this structural admissibility, intertemporal non-separability becomes economically testable. Habits matter for valuation precisely when they change behavioural coherence and the welfare interpretation of hedonic prices, rather than serving as a residual explanation for structural misspecification in the price system.

A central implication of the framework is that dimensionality reduction and dynamics play fundamentally different roles. Moving from goods to characteristics yields parsimony but imposes geometric discipline on prices, as observed price vectors must lie on a low-dimensional hedonic manifold. Habits do not relax that discipline. Instead, they reshape the interpretation of prices within the feasible hedonic system by introducing a dynamic wedge: the present-value price of a good reflects both contemporaneous marginal utility from its characteristics and the continuation value induced by past consumption. Static hedonic valuations can therefore confound contemporaneous marginal valuations with intertemporal effects, mismeasuring WTP for habit-forming attributes and potentially mis-ranking policies that target them.

The empirical application illustrates this logic. In scanner data on cereal purchases, goods-based benchmarks pass almost universally, while characteristics-based models fail frequently because observed prices often violate the hedonic spanning restriction. Yet the implied violations are typically economically modest: restoring hedonic consistency generally requires only small price adjustments relative to observed expenditures. Conditional on structural admissibility, behaviour is close to dynamically rational, and allowing for habit formation delivers a systematic improvement in behavioural fit relative to static characteristics models. When the structural price restriction holds, this improvement reflects a genuine dynamic reinterpretation of hedonic prices, with consequences for welfare and WTP.

From an applied perspective, the framework functions as a diagnostic that guides model refinement. If the price-spanning restriction fails, the inconsistency originates in the hedonic representation itself, and no dynamic extension can restore consistency; the failure points to the characteristics technology or to pricing components unrelated to consumption-relevant attributes. If the price-spanning restriction holds, the behavioural test isolates whether time separability is the binding assumption and whether habits alter welfare conclusions relative to static valuation.

The approach also has clear limitations. The test is silent about the source of rejection: failure may reflect misspecification of the characteristics technology, omitted attributes, aggregation over time or products, longer memory in habits, or departures from concavity rather than the absence of rational behaviour. Rationalisability is not identification; multiple dynamic preference representations can rationalise the same data. The framework also abstracts from stochastic choice and unobserved heterogeneity, treating the data as deterministic realisations of a representative decision problem. Empirically, scanner environments impose additional constraints: prices are often observed only for purchased goods, intertemporal budget variation is limited, and time aggregation smooths short-run substitution patterns. Richer data with denser price support, more frequent observation, and greater budget variation would sharpen both the structural and behavioural content of the test. These features make the framework a sharp diagnostic of internal coherence rather than a structural estimator of primitives.

Taken together, the results provide a disciplined benchmark for dynamic hedonic modelling. By separating structural from behavioural restrictions, the framework clarifies when habits genuinely matter for valuation and when apparent failures instead reflect misspecification of the price-characteristics mapping. It thus provides a theory-grounded and empirically implementable benchmark for distinguishing behavioural departures from structural misspecification in dynamic hedonic environments.

References

Afriat, S. N. (1967). The Construction of Utility Functions from Expenditure Data. *International Economic Review*, 8(1):67–77.

Afriat, S. N. (1973). On a system of inequalities in demand analysis: An extension of the classical method. *International Economic Review*, 14(2):460–472.

Allcott, H., Gentzkow, M., and Song, L. (2022). Digital addiction. *American Economic Review*, 112(7):2424–2463.

Andrews, I. (2026). Revealed rationality: Label-free regularization from representation theorems. First version: February 21, 2026. Working paper, Department of Economics, MIT and NBER.

Avena, N. M., Rada, P., and Hoebel, B. G. (2008). Evidence for sugar addiction: Behavioral and neurochemical effects of intermittent, excessive sugar intake. *Neuroscience & Biobehavioral Reviews*, 32(1):20–39.

Bajari, P. and Benkard, C. L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy*, 113(6):1239–1276.

Bajari, P. and Kahn, M. E. (2005). Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities. *Journal of Business & Economic Statistics*, 23(1):20–33.

Barahona, N., Otero, C., and Otero, S. (2023). Equilibrium effects of food labeling policies. *Econometrica*, 91(3):839–868.

Becker, G. S., Grossman, M., and Murphy, K. M. (1994). An Empirical Analysis of Cigarette Addiction. *The American Economic Review*, 84(3):396–418.

Becker, G. S. and Murphy, K. M. (1988). A Theory of Rational Addiction. *Journal of Political Economy*, 96(4):675–700.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841–890.

Blow, L., Browning, M., and Crawford, I. (2008). Revealed preference analysis of characteristics models. *The Review of Economic Studies*, 75(2):371–389.

Boyer, M. (1978). A Habit Forming Optimal Growth Model. *International Economic Review*, 19(3):585–609.

Boyer, M. (1983). Rational demand and expenditures patterns under habit formation. *Journal of Economic Theory*, 31(1):27–53.

Browning, M. (1989). A Nonparametric Test of the Life-Cycle Rational Expections Hypothesis. *International Economic Review*, 30(4):979–992.

Cocores, J. A. and Gold, M. S. (2009). The Salted Food Addiction Hypothesis may explain overeating and the obesity epidemic. *Medical Hypotheses*, 73(6):892–899.

Crawford, I. (2010). Habits Revealed. *The Review of Economic Studies*, 77(4):1382–1402.

Demuynck, T. and Verriest, E. (2013). I'll never forget my first cigarette: A revealed preference analysis of the "habits as durables" model. *International Economic Review*, 54(2):717–738.

Diewert, W. E. (1973). Afriat and Revealed Preference Theory. *The Review of Economic Studies*, 40(3):419.

Dubois, P., Griffith, R., and O'Connell, M. (2020). How Well Targeted Are Soda Taxes? *American Economic Review*, 110(11):3661–3704.

Echenique, F. (2022). On the meaning of the critical cost efficiency index.

Felippa, C. A. (2004). *Introduction to finite element methods*. University of Colorado.

Freddie Mac (2025). 30-year fixed rate mortgage average in the united states (mortgage30us). FRED, Federal Reserve Bank of St. Louis. Retrieved August 16, 2025.

Fudenberg, D., Gao, W., and Liang, A. (2023). How Flexible is that Functional Form? Quantifying the Restrictiveness of Theories. *SSRN Electronic Journal*.

Gibbons, S. and Machin, S. (2003). Valuing English primary schools. *Journal of Urban Economics*, 53(2):197–219.

Gorman, W. M. (1956). A Possible Procedure for Analysing Quality Differentials in the Egg Market. *The Review of Economic Studies*, 47(5):843–856.

Greenstone, M. and Gallagher, J. (2008). Does hazardous waste matter? evidence from the housing market and the superfund program. *The Quarterly Journal of Economics*, 123(3):951–1003.

Grossman, M., Chaloupka, F. J., and Sirtalan, I. (1998). An Empirical Analysis of Alcohol Addiction: Results from the Monitoring the Future Panels. *Economic Inquiry*, 36(1):39–48.

Gruber, J. and Kőszegi, B. (2004). Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes. *Journal of Public Economics*, 88(9-10):1959–1987.

Haeck, C., Lawson, N., and Poirier, K. (2022). Estimating consumer preferences for different beverages using the BLP approach. *Economics & Human Biology*, 46:101138.

Heckman, J. and Scheinkman, J. (1987). The Importance of Bundling in a Gorman-Lancaster Model of Earnings. *The Review of Economic Studies*, 54(2):243–255.

Hendel, I. and Nevo, A. (2006). Sales and consumer inventory. *The RAND Journal of Economics*, 37(3):543–561.

Houthakker, H. S. (1950). Revealed Preference and the Utility Function. *Economica*, 17(66):159–174.

Koob, G. F. and Moal, M. L. (1997). Drug Abuse: Hedonic Homeostatic Dysregulation. *Science*, 278(5335):52–58.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *The Journal of Political Economy*, 74(2):132–157.

Le Fur, E. and Outreville, J. F. (2022). The Willingness to Pay for Cider Products: Results of a Survey on Habits and Consumption Behavior. *Wine Business Journal*, 5(1).

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*, 69(2):307–342.

Richards, T. J. and Patterson, P. M. (2006). Native American Obesity: An Economic Model of the "Thrifty Gene" Theory. *American Journal of Agricultural Economics*, 88(3):542–560.

Rockafellar, R. T. (1970). *Convex Analysis*, volume 28 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1):34–55.

Samuelson, P. A. (1948). Consumption Theory in Terms of Revealed Preference. *Economica*, 15(60):243–253.

Selten, R. (1991). Properties of a measure of predictive success. *Mathematical Social Sciences*, 21(2):153–167.

Smith, V. K. and Desvousges, W. H. (1986). The Value of Avoiding a Lulu: Hazardous Waste Disposal Sites. *The Review of Economics and Statistics*, 68(2):293–299.

Varian, H. R. (1982). The Nonparametric Approach to Demand Analysis. *Econometrica*, 50(4):945–973.

Varian, H. R. (1990). Goodness-of-fit in optimizing models. *Journal of Econometrics*, 46(1-2):125–140.

Appendix

A Proofs

Proof of Lemma 2.1.

To define *consistency* more formally, I solve the consumer's lifetime maximisation problem in (1). Substituting the technology constraint $\tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t$ into period utility yields

$$\max_{\{\boldsymbol{x}_t,\, y_t\}_{t=1}^T} \sum_{t=1}^{T} \beta^{t-1}\big(u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t) + y_t\big) \quad \text{subject to} \quad \sum_{t=1}^{T} \boldsymbol{\rho}_t'\boldsymbol{x}_t + \sum_{t=1}^{T} \beta^{t-1} y_t = W,$$

where $\boldsymbol{x}_0$ is treated as fixed. Using the lifetime budget constraint to substitute out the outside good, the consumer's problem can be written as the unconstrained maximisation

$$\max_{\{\boldsymbol{x}_t\}_{t=1}^T} \sum_{t=1}^{T} \beta^{t-1} u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t) + W - \sum_{t=1}^{T} \boldsymbol{\rho}_t'\boldsymbol{x}_t.$$

Since W enters only as an additive constant, it does not affect the maximising choice of $\{\boldsymbol{x}_t\}_{t=1}^T$. I therefore equivalently represent the problem in constrained form by reintroducing the lifetime budget constraint and an associated Lagrange multiplier. The consumer's problem thus reduces to

$$\max_{\{\boldsymbol{x}_t\}_{t=1}^T} \sum_{t=1}^{T} \beta^{t-1} u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t) \quad \text{subject to} \quad \sum_{t=1}^{T} \boldsymbol{\rho}_t'\boldsymbol{x}_t = W,$$

where W is interpreted as lifetime wealth net of outside-good consumption.

The associated Lagrangian is

$$\mathcal{L}(\{\boldsymbol{x}_t\}) = \sum_{t=1}^{T} \beta^{t-1} u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t) - \left\{ \sum_{t=1}^{T} \boldsymbol{\rho}_t'\boldsymbol{x}_t - W \right\}, \tag{9}$$

where I normalise $\lambda = 1$ without loss of generality.

To define the first-order necessary conditions for an interior solution to this constrained optimisation problem, I require several vector derivatives. Applying the chain rule for both scalar and vector functions (Felippa, 2004) and using the "denominator layout" as my notational choice I have,

$$\underbrace{\frac{\partial u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t)}{\partial \boldsymbol{x}_t}}_{(K\times 1)} = \underbrace{\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_t}}_{(K\times 2K)} \underbrace{\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{x}}_t}}_{(2K\times 1)} = \underbrace{\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_t}}_{(K\times 2K)} \underbrace{\frac{\partial \tilde{\boldsymbol{z}}_t}{\partial \tilde{\boldsymbol{x}}_t}}_{(2K\times (J+J_2))} \underbrace{\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{z}}_t}}_{((J+J_2)\times 1)} \tag{10}$$

where, recalling my notation defined in (2) I have,

$$\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_t} = \left[\boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times K}\right],$$

$$\frac{\partial \tilde{\boldsymbol{z}}_t}{\partial \tilde{\boldsymbol{x}}_t} = \frac{\partial(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t)}{\partial \tilde{\boldsymbol{x}}_t} = \tilde{\boldsymbol{A}}',$$

$$\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{z}}_t} = \partial u(\tilde{\boldsymbol{z}}_t) := \left[\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t)', \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)', \partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t)'\right]',$$

where $\big|$ denotes the horizontal concatenation of the $K\times K$ identity matrix and the $K\times K$ matrix of zeros, and $\partial u(\tilde{\boldsymbol{z}})$ denotes the superderivative of u at $\tilde{\boldsymbol{z}}$. Repeating the chain rule exercise in (10), except this time differentiating with respect to the one period lag of market goods, I have,

$$\underbrace{\frac{\partial u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t)}{\partial \boldsymbol{x}_{t-1}}}_{(K\times 1)} = \underbrace{\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_{t-1}}}_{(K\times 2K)} \underbrace{\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{x}}_t}}_{(2K\times 1)} = \underbrace{\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_{t-1}}}_{(K\times 2K)} \underbrace{\frac{\partial \tilde{\boldsymbol{z}}_t}{\partial \tilde{\boldsymbol{x}}_t}}_{(2K\times(J+J_2))} \underbrace{\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{z}}_t}}_{((J+J_2)\times 1)} \tag{11}$$

where the only new term is,

$$\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_{t-1}} = \left[\boldsymbol{0}_{K\times K} \,\middle|\, \boldsymbol{I}_{K\times K}\right].$$

It follows from these intermediate calculations of the vector derivatives that,[12]

$$\frac{\partial u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t)}{\partial \boldsymbol{x}_t} = \left[\boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times K}\right] \tilde{\boldsymbol{A}}' \partial u(\tilde{\boldsymbol{z}}_t)$$

and,

$$\frac{\partial u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_{t+1})}{\partial \boldsymbol{x}_t} = \left[\boldsymbol{0}_{K\times K} \,\middle|\, \boldsymbol{I}_{K\times K}\right] \tilde{\boldsymbol{A}}' \partial u(\tilde{\boldsymbol{z}}_{t+1}).$$

The relevant first-order necessary conditions associated with the Lagrangian in (9) now follow immediately as,

$$\partial_{\boldsymbol{x}_t}\mathcal{L} = 0 \quad \Rightarrow \quad \boldsymbol{\rho}_t = \beta^{t-1}\left(\left[\boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times K}\right] \tilde{\boldsymbol{A}}' \partial u(\tilde{\boldsymbol{z}}_t) + \beta \left[\boldsymbol{0}_{K\times K} \,\middle|\, \boldsymbol{I}_{K\times K}\right] \tilde{\boldsymbol{A}}' \partial u(\tilde{\boldsymbol{z}}_{t+1})\right). \tag{12}$$

But recall from (2) that $\tilde{\boldsymbol{A}}$ is a $2K\times(J+J_2)$ block matrix. Hence, these first-order conditions can be substantially simplified. Indeed, since I have conformable partitions of the block matrices,

$$\begin{aligned}
\left[\boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times K}\right] \tilde{\boldsymbol{A}}' &= \left[\boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times K}\right] \begin{bmatrix} \boldsymbol{A}' & \boldsymbol{0}_{K\times J_2} \\ \boldsymbol{0}_{K\times J} & (\boldsymbol{A}^a)' \end{bmatrix} \\
&= \left[\boldsymbol{I}_{K\times K}\boldsymbol{A}' + \boldsymbol{0}_{K\times K}\boldsymbol{0}_{K\times J} \,\middle|\, \boldsymbol{I}_{K\times K}\boldsymbol{0}_{K\times J_2} + \boldsymbol{0}_{K\times K}(\boldsymbol{A}^a)'\right] \\
&= \left[\boldsymbol{A}' \,\middle|\, \boldsymbol{0}_{K\times J_2}\right].
\end{aligned}$$

[12] Since $u(\cdot)$ is assumed to be concave but not necessarily differentiable, each $\partial u(\tilde{\boldsymbol{z}}_t)$ denotes a supergradient. As such, expressions like those below should formally be interpreted as *set inclusions* rather than strict equalities—e.g., $0 \in \partial_{\boldsymbol{x}_t}\mathcal{L}$. When u is differentiable at $\tilde{\boldsymbol{z}}_t$, the supergradient is a singleton and the equality holds exactly.

Analogously,

$$\begin{aligned}\beta\left[\mathbf{0}_{K\times K}\,\middle|\,\boldsymbol{I}_{K\times K}\right]\tilde{\boldsymbol{A}}' &= \beta\left[\mathbf{0}_{K\times K}\,\middle|\,\boldsymbol{I}_{K\times K}\right]\begin{bmatrix}\boldsymbol{A}' & \mathbf{0}_{K\times J_2}\\ \mathbf{0}_{K\times J} & (\boldsymbol{A}^a)'\end{bmatrix}\\ &= \beta\left[\mathbf{0}_{K\times K}\boldsymbol{A}' + \boldsymbol{I}_{K\times K}\mathbf{0}_{K\times J}\,\middle|\,\mathbf{0}_{K\times K}\mathbf{0}_{K\times J_2} + \boldsymbol{I}_{K\times K}(\boldsymbol{A}^a)'\right]\\ &= \beta\left[\mathbf{0}_{K\times J}\,\middle|\,(\boldsymbol{A}^a)'\right].\end{aligned}$$

Inserting these simplifications, the first-order conditions in (12) reduce to:

$$\boldsymbol{\rho}_t = \beta^{t-1}\left(\left[\boldsymbol{A}'\,\middle|\,\mathbf{0}_{K\times J_2}\right]\partial u(\tilde{\boldsymbol{z}}_t) + \beta\left[\mathbf{0}_{K\times J}\,\middle|\,(\boldsymbol{A}^a)'\right]\partial u(\tilde{\boldsymbol{z}}_{t+1})\right). \tag{13}$$

But again, since the supergradient $\partial u(\tilde{\boldsymbol{z}}_t)$ can be partitioned as a $J + J_2$ block vector,

$$\partial u(\tilde{\boldsymbol{z}}_t) := \begin{bmatrix}\begin{bmatrix}\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t)\\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)\end{bmatrix}\\ \\ \left[\partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t)\right]\end{bmatrix},$$

the first-order conditions in (13) further simplify to,

$$\begin{aligned}\boldsymbol{\rho}_t &= \beta^{t-1}\left(\left[\boldsymbol{A}'\,\middle|\,\mathbf{0}_{K\times J_2}\right]\begin{bmatrix}\begin{bmatrix}\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t)\\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)\end{bmatrix}\\ \left[\partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t)\right]\end{bmatrix} + \beta\left[\mathbf{0}_{K\times J}\,\middle|\,(\boldsymbol{A}^a)'\right]\begin{bmatrix}\begin{bmatrix}\partial_{\boldsymbol{z}_{t+1}^c} u(\tilde{\boldsymbol{z}}_{t+1})\\ \partial_{\boldsymbol{z}_{t+1}^a} u(\tilde{\boldsymbol{z}}_{t+1})\end{bmatrix}\\ \left[\partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_{t+1})\right]\end{bmatrix}\right)\\ &= \beta^{t-1}\left(\boldsymbol{A}'\begin{bmatrix}\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t)\\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)\end{bmatrix} + \beta(\boldsymbol{A}^a)'\left[\partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_{t+1})\right]\right).\end{aligned} \tag{14}$$

Since I assume u to be concave, the associated KKT conditions are sufficient for a global maximum. Hence, allowing for corner solutions, I replace the stationarity equality in (14) with an inequality to obtain the full set of price and data pairs $\{\boldsymbol{\rho}_t, \boldsymbol{x}_t\}_{t=1}^T$ consistent with an interior or corner solution to the consumer's maximisation problem. This gives rise to my formal definition of *consistency* in Definition 2.1 and Lemma 2.1. □

Proof of Theorem 2.1.

$(A) \Rightarrow (B)$: Assume (A) holds. Then, by Lemma 2.1 of consistency I have that for all $t \in \{1, \ldots, T-1\}$,

$$\boldsymbol{\rho}_t \geq \boldsymbol{A}'\boldsymbol{\pi}_t^0 + (\boldsymbol{A}^a)'\boldsymbol{\pi}_{t+1}^1,$$

with equality for all k such that $x_t^k > 0$. Element wise, this states that for all k and for all $t \in \{1, \ldots, T-1\}$,

$$\rho_t^k \geq \boldsymbol{a}_k'\boldsymbol{\pi}_t^0 + \boldsymbol{a}_k^{a\prime}\boldsymbol{\pi}_{t+1}^1,$$

and where $x_t^k > 0$,

$$\rho_t^k = \boldsymbol{a}_k'\boldsymbol{\pi}_t^0 + \boldsymbol{a}_k^{a\prime}\boldsymbol{\pi}_{t+1}^1.$$

This gives us restrictions (B2) and (B3), respectively.

It remains to show (B1) holds. Using the augmented notation $\tilde{\boldsymbol{z}}_t := ((\boldsymbol{z}_t^c)', (\boldsymbol{z}_t^a)', (\boldsymbol{z}_{t-1}^a)')'$ and $\tilde{\boldsymbol{\pi}}_t := \frac{1}{\beta^{t-1}}\left[\boldsymbol{\pi}_t^{0\prime}, \boldsymbol{\pi}_t^{1\prime}\right]'$ it follows from Definitions (3) and (4) for the shadow prices that for any element of the superdifferential of u at $\tilde{\boldsymbol{z}}_t$,

$$\begin{aligned}
\partial u(\tilde{\boldsymbol{z}}_t)' &= \left[\left[\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t)', \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)'\right], \partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t)'\right] \\
&= \frac{1}{\beta^{t-1}}\left[\beta^{t-1}\left[\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t)', \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)'\right], \beta^{t-1}\partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t)'\right] \\
&= \frac{1}{\beta^{t-1}}\left[\boldsymbol{\pi}_t^{0\prime}, \boldsymbol{\pi}_t^{1\prime}\right] \\
&= \tilde{\boldsymbol{\pi}}_t'.
\end{aligned}$$

The concavity and superdifferentiability of the instantaneous utility function $u(\tilde{\boldsymbol{z}}_t)$ means,

$$u(\tilde{\boldsymbol{z}}_s) - u(\tilde{\boldsymbol{z}}_t) \leq \partial u(\tilde{\boldsymbol{z}}_t)'(\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t) \quad \forall\, s, t \in \{1, \ldots, T\}.$$

Combining this concavity result with that implied by the optimising behaviour above therefore implies,

$$u(\tilde{\boldsymbol{z}}_s) - u(\tilde{\boldsymbol{z}}_t) \leq \tilde{\boldsymbol{\pi}}_t'(\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t) \quad \forall\, s, t \in \{1, \ldots, T\}.$$

Then for any subset σ of observations from $\tau = \{1, \ldots, T\}$, summing across all observations within σ gives,

$$0 \leq \sum_{\forall s,t \in \sigma} \tilde{\boldsymbol{\pi}}_t'(\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t) \quad \forall \sigma \subseteq \{1, \ldots, T\}$$

which is restriction (B1).

$(B) \Rightarrow (A)$: Restriction (B1) imposes that the data $(\tilde{\boldsymbol{\pi}}_t, \tilde{\boldsymbol{z}}_t), t = 1, \ldots, T$ are cyclically monotone as defined by Browning (1989). Given this property, define the function $u : \mathbb{R}^{J+J_2} \to \mathbb{R}$ via,

$$u(\tilde{\boldsymbol{z}}) := \min_{\sigma \subseteq \{1,\ldots,T\}} \left\{\tilde{\boldsymbol{\pi}}_T'(\tilde{\boldsymbol{z}} - \tilde{\boldsymbol{z}}_T) + \cdots + \tilde{\boldsymbol{\pi}}_1'(\tilde{\boldsymbol{z}}_2 - \tilde{\boldsymbol{z}}_1)\right\},$$

where the minimum is taken over all $2^T - 1$ possible (finite, non-empty) subsets σ of $\tau = \{1, \ldots, T\}$, hence is well-defined. Since $u(\cdot)$ is the minimum of a finite set of proper affine functions (one for each choice of $\sigma \subseteq \tau$), $u(\tilde{\boldsymbol{z}})$ is continuous (hence superdifferentiable), non-satiated, and proper concave for all $\tilde{\boldsymbol{z}}$. Since u is a proper concave function, I can apply Rockafellar (1970) Theorem 24.8 to conclude that the superdifferential of u is cyclically monotone. But this means that the $\tilde{\boldsymbol{\pi}}_r$ vectors in the construction of $u(\tilde{\boldsymbol{z}})$ are a $\mathbb{R}^{J+J_2}$ to $\mathbb{R}^{J+J_2}$ mapping belonging to the set of supergradients of u at $\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{z}}_r$. It follows that $u(\tilde{\boldsymbol{z}})$ is a locally non-satiated, superdifferentiable, and concave utility function such that,

$$\tilde{\boldsymbol{\pi}}_t = \partial u(\tilde{\boldsymbol{z}}_t) \quad \Rightarrow \quad \frac{1}{\beta^{t-1}}\left[\boldsymbol{\pi}_t^{0\prime}, \boldsymbol{\pi}_t^{1\prime}\right] = \left[\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t)', \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)', \partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t)'\right].$$

for all $t \in \tau$. Rearranging yields the following relationships between shadow prices and supergradients of the utility function:

$$\boldsymbol{\pi}_t^0 = \beta^{t-1} \begin{bmatrix} \partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t) \\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t) \end{bmatrix} \tag{15}$$

$$\boldsymbol{\pi}_t^1 = \beta^{t-1} \begin{bmatrix} \partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t) \end{bmatrix}. \tag{16}$$

Since (16) holds for all $t \in \tau$, this condition can be forwarded one period to obtain,

$$\boldsymbol{\pi}_{t+1}^1 = \beta^t \begin{bmatrix} \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_{t+1}) \end{bmatrix} \tag{17}$$

for all $t \in \{1, \dots, T-1\}$. Substituting expressions (15) and (17) into condition (B2) re-written in matrix form, $\boldsymbol{\rho}_t \geq \boldsymbol{A}'\boldsymbol{\pi}_t^0 + (\boldsymbol{A}^a)'\boldsymbol{\pi}_{t+1}^1$, I obtain,

$$\boldsymbol{\rho}_t \geq \boldsymbol{A}'\beta^{t-1} \begin{bmatrix} \partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t) \\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t) \end{bmatrix} + (\boldsymbol{A}^a)'\beta^t \begin{bmatrix} \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_{t+1}) \end{bmatrix}. \tag{18}$$

In the event that a good k is consumed in strictly positive amounts, then substituting expressions (15) and (17) into condition (B3) yields the same expression as in (18), but with equality. By Lemma 2.1, this means that the data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t \in \{1,\dots,T\}}$ are consistent with the one-lag habits model for given technology $\boldsymbol{A}$. □

Proof of Theorem 2.2.

$(A^+) \Rightarrow (B^+)$: Given $\{\boldsymbol{\rho}_t^0\}_{t \in \{1,\dots,T\}}$ I have a full set of prices, $\{\boldsymbol{\rho}_t\}_{t \in \{1,\dots,T\}}$, such that the data satisfies the model. Hence, by Theorem 2.1, condition (B) holds. Hence, (B^+) also holds.

$(B^+) \Rightarrow (A^+)$: I have shadow discounted prices $\{\boldsymbol{\pi}_t^0\}_{t \in \{1,\dots,T\}}$ and $\{\boldsymbol{\pi}_t^1\}_{t \in \{1,\dots,T\}}$ such that (B1$^+$) and (B2$^+$) hold. Use these shadow prices to construct the unobserved prices via,

$$\boldsymbol{\rho}_t^0 = (\boldsymbol{B}_t^0)'\boldsymbol{\pi}_t^0 + (\boldsymbol{B}_t^{a,0})'\boldsymbol{\pi}_{t+1}^1.$$

Using these constructed prices, it follows by comparison with Theorem 2.1 that (A$^+$) holds. □

Habits over goods as a special case of the characteristics model (Definition 3.1).

The (one-lag) habits-over-characteristics model nests the habits-over-goods model of Crawford (2010) as a special case. To see this, consider a trivial characteristics model where $J = K$ and the technology matrix $\boldsymbol{A}$ is the $J \times J$ identity matrix. Then:

$$\boldsymbol{z}_t = \begin{bmatrix} \boldsymbol{z}_t^c \\ \boldsymbol{z}_t^a \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_t^c \\ \boldsymbol{x}_t^a \end{bmatrix} = \boldsymbol{x}_t,$$

where $\boldsymbol{x}_t^c$ and $\boldsymbol{x}_t^a$ denote the J_1 non-habit-forming and J_2 habit-forming goods, respectively. Under this assumption, the matrices $(\boldsymbol{A}^c)'$ and $(\boldsymbol{A}^a)'$ take on block-structured forms. The matrix $(\boldsymbol{A}^c)'$ is a $K \times J_1$ matrix whose first J_1 rows form a $J_1 \times J_1$ identity matrix, with all remaining entries equal to zero. Conversely, $(\boldsymbol{A}^a)'$ is a $K \times J_2$ matrix whose bottom J_2 rows comprise a $J_2 \times J_2$ identity matrix, while the entries in the first $K - J_2$ rows are zero.

Substituting the identity structure of $\boldsymbol{A}$ into Equation $(\star)$ and defining the augmented bundle $\bar{\boldsymbol{x}}_t :=$

$\left(\boldsymbol{x}_t^{c\prime}, \boldsymbol{x}_t^{a\prime}, \boldsymbol{x}_{t-1}^{a\prime}\right)' = \tilde{\boldsymbol{z}}_t$, the first-order condition becomes,

$$\boldsymbol{\rho}_t = \begin{bmatrix} \boldsymbol{\rho}_t^c \\ \boldsymbol{\rho}_t^a \end{bmatrix} \geq \begin{bmatrix} \boldsymbol{\pi}_t^{c,0} \\ \mathbf{0}_{K-J_1} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{K-J_2} \\ \boldsymbol{\pi}_t^{a,0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{K-J_2} \\ \boldsymbol{\pi}_{t+1}^{1} \end{bmatrix} = \beta^{t-1} \begin{bmatrix} \partial_{\boldsymbol{x}_t^c} u(\bar{\boldsymbol{x}}_t) \\ \mathbf{0}_{K-J_1} \end{bmatrix} + \beta^{t-1} \begin{bmatrix} \mathbf{0}_{K-J_2} \\ \partial_{\boldsymbol{x}_t^a} u(\bar{\boldsymbol{x}}_t) \end{bmatrix} + \beta^{t} \begin{bmatrix} \mathbf{0}_{K-J_2} \\ \partial_{\boldsymbol{x}_t^a} u(\bar{\boldsymbol{x}}_{t+1}) \end{bmatrix},$$

where $\boldsymbol{\pi}_t^{c,0}$, $\boldsymbol{\pi}_t^{a,0}$ and $\boldsymbol{\rho}_t^c$, $\boldsymbol{\rho}_t^a$ are the subvectors of $\boldsymbol{\pi}_t^0$ and $\boldsymbol{\rho}_t$ corresponding to non-habit-forming and habit-forming goods, respectively, and the final equality follows using Definitions 3 and 4. For all k such that $x_t^k > 0$, these inequalities hold with equality. Thus, under the restriction $J = K$ and $\boldsymbol{A} = \boldsymbol{I}_J$, Definition 2.1 simplifies as in Definition 3.1.

Proof of Corollary 3.1.

Follows directly from Theorem 2.1 under the restriction $J = K$ and $\boldsymbol{A} = \boldsymbol{I}_J$. Equivalently, one may derive it from Definition 3.1 by replicating the proof of Theorem 2.1. □

Intertemporally separable preferences over characteristics (Definition 3.2).

To formalise this nesting result, assume $\boldsymbol{z}_t = \boldsymbol{z}_t^c$ and embed the model in a lifecycle framework with a single intertemporal budget constraint and maintain the normalisation that the marginal utility of lifetime wealth equals one. Then the shadow price in Equation (3) reduces to:

$$\boldsymbol{\pi}_t = \partial_{\boldsymbol{z}_t} u(\boldsymbol{z}_t) \tag{19}$$

noting that β no longer appears due to intertemporal separability.[13] Under these assumptions, Definition 2.1 reduces to that in Definition 3.2.

Proof of Corollary 3.2.

Corollary 3.2 is immediate from Theorem 2.1 upon setting $J_1 = J$ and $\beta = 1$.

I now show that this Corollary is equivalent to Theorem 1 of Blow et al. (2008) when the static characteristics model is embedded in a lifecycle framework. Equivalence between (A") and the P" condition in their theorem follows directly from Definition 3.2. Equivalence of conditions (B2") and (B3") with (A2) and (A3) in Blow et al. (2008) is straightforward after aligning notation.

The key step is to show that condition (B1") is equivalent to condition (A1) in Blow et al. (2008), which states that there exist T scalars V_t and T vectors $\boldsymbol{\pi}_t$ such that:

$$V_s \leq V_t + \lambda_t \boldsymbol{\pi}_t'(\boldsymbol{A}\boldsymbol{x}_s - \boldsymbol{A}\boldsymbol{x}_t) \quad \forall\, s, t. \tag{A1}$$

In Blow et al. (2008), λ_t denotes the (time-t) marginal utility of wealth. Under my single lifetime budget constraint I have $\lambda_t \equiv \lambda$, and I maintain the normalisation $\lambda = 1$ throughout.

To show that (B1") implies (A1), take any $\sigma = \{s, t\}$ so that:

$$0 \leq \boldsymbol{\pi}_s'(\boldsymbol{z}_t - \boldsymbol{z}_s) + \boldsymbol{\pi}_t'(\boldsymbol{z}_s - \boldsymbol{z}_t) = \boldsymbol{\pi}_s'\boldsymbol{A}\boldsymbol{x}_t - \boldsymbol{\pi}_s'\boldsymbol{A}\boldsymbol{x}_s + \boldsymbol{\pi}_t'(\boldsymbol{A}\boldsymbol{x}_s - \boldsymbol{A}\boldsymbol{x}_t). \tag{20}$$

Now define:

$$V_t := \boldsymbol{\pi}_s'\boldsymbol{A}\boldsymbol{x}_t \quad \forall\, t.$$

[13] This is equivalent to setting $\beta = 1$, which is without loss of generality in this setting.

Then (20) becomes:

$$0 \leq V_t - V_s + \boldsymbol{\pi}_t'(\boldsymbol{A}\boldsymbol{x}_s - \boldsymbol{A}\boldsymbol{x}_t),$$

which is equivalent to (A1) when the static characteristics model is embedded in a lifecycle framework.

Conversely, suppose (A1) holds. Under $\lambda = 1$, rearranging yields:

$$0 \leq (V_t - V_s) + \boldsymbol{\pi}_t'(\boldsymbol{A}\boldsymbol{x}_s - \boldsymbol{A}\boldsymbol{x}_t).$$

Summing over any $\sigma \subseteq \{1, \ldots, T\}$ yields (B1”).

I conclude that Corollary 3.2 provides an equivalent characterisation to Blow et al. (2008) when the static characteristics model is embedded in a lifecycle framework. This highlights the additional empirical restrictions imposed by intertemporal optimisation: a consumer consistent with intratemporal utility maximisation over characteristics must also allocate expenditure across periods to maximise lifetime utility. $\square$

B Non-linear characteristics model

B.1 Consumer problem in the non-linear model

This paper focuses on the linear characteristics model, $\boldsymbol{z}_t = \boldsymbol{A}\boldsymbol{x}_t$. However, most of my analysis generalises to a non-linear characteristics setting where I assume $\boldsymbol{z}_t = \boldsymbol{F}(\boldsymbol{x}_t)$ for some concave, increasing function $\boldsymbol{F} : \mathbb{R}^K_{\geq 0} \to \mathbb{R}^J_{\geq 0}$. Here, I provide an analogue notion of consistency and the relevant Afriat Theorem for such non-linear technologies. For convenience, I take $\boldsymbol{F}$ to be differentiable, with an associated (denominator layout) $K \times J$ matrix derivative at $\boldsymbol{x}_t$ given by,

$$\boldsymbol{\nabla F}(\boldsymbol{x}_t) := \left(\frac{\partial \boldsymbol{z}^c_t}{\partial \boldsymbol{x}_t} \quad \Big| \quad \frac{\partial \boldsymbol{z}^a_t}{\partial \boldsymbol{x}_t} \right) = \begin{pmatrix} \frac{\partial z^c_{1,t}}{\partial x_{1,t}} & \cdots & \frac{\partial z^c_{J_1,t}}{\partial x_{1,t}} & \frac{\partial z^a_{1,t}}{\partial x_{1,t}} & \cdots & \frac{\partial z^a_{J_2,t}}{\partial x_{1,t}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z^c_{1,t}}{\partial x_{K,t}} & \cdots & \frac{\partial z^c_{J_1,t}}{\partial x_{K,t}} & \frac{\partial z^a_{1,t}}{\partial x_{K,t}} & \cdots & \frac{\partial z^a_{J_2,t}}{\partial x_{K,t}} \end{pmatrix}.$$

If instead $\boldsymbol{F}$ is only superdifferentiable, one must replace all gradients with the superdifferential.

Use the augmented notation for $\tilde{\boldsymbol{x}}_t$ and $\tilde{\boldsymbol{z}}_t$ from Section 2. In addition, define the augmented technology function $\tilde{\boldsymbol{F}} : \mathbb{R}^{2K}_{\geq 0} \to \mathbb{R}^{J+J2}_{\geq 0}$ via,

$$\tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t) := \begin{pmatrix} \boldsymbol{F}(\boldsymbol{x}_t) \\ \boldsymbol{F}^a(\boldsymbol{x}_{t-1}) \end{pmatrix}, \tag{21}$$

where $\boldsymbol{F}^a : \mathbb{R}^{2K}_{\geq 0} \to \mathbb{R}^{J2}_{\geq 0}$ is the function defined by taking the last J_2 rows of $\boldsymbol{F}$. Note that the derivative of this extended vector-valued function is a $2K \times (J + J_2)$ matrix given by,

$$\boldsymbol{\nabla}\tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t) = \begin{pmatrix} \frac{\partial \boldsymbol{z}^c_t}{\partial \boldsymbol{x}_t} & \Big| & \frac{\partial \boldsymbol{z}^a_t}{\partial \boldsymbol{x}_t} & \Big| & \boldsymbol{0}_{K \times J_2} \\ \boldsymbol{0}_{K \times J_1} & \Big| & \boldsymbol{0}_{K \times J_2} & \Big| & \frac{\partial \boldsymbol{z}^a_{t-1}}{\partial \boldsymbol{x}_{t-1}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\nabla F}(\boldsymbol{x}_t) & \Big| & \boldsymbol{0}_{K \times J_2} \\ \boldsymbol{0}_{K \times J} & \Big| & \boldsymbol{\nabla F}^a(\boldsymbol{x}_{t-1}) \end{pmatrix} \tag{22}$$

where $\boldsymbol{\nabla F}^a(\boldsymbol{x}_{t-1})$ is the $K \times J_2$ submatrix found by taking the last J_2 columns of $\boldsymbol{\nabla F}(\boldsymbol{x}_{t-1})$ and I use the fact that $\frac{\partial \boldsymbol{z}^a_{t-1}}{\partial \boldsymbol{x}_t} = \boldsymbol{0}_{K \times J_2}$ and $\left(\frac{\partial \boldsymbol{z}^c_t}{\partial \boldsymbol{x}_{t-1}} \quad \Big| \quad \frac{\partial \boldsymbol{z}^a_t}{\partial \boldsymbol{x}_{t-1}} \right) = \left(\boldsymbol{0}_{K \times J_1} \quad \Big| \quad \boldsymbol{0}_{K \times J_2} \right)$.

Using this augmented notation, my quasilinear model of interest becomes:

$$\max_{\{\boldsymbol{x}_t,\, y_t\}_{t=1}^T} \sum_{t=1}^T \beta^{t-1}\big(u(\tilde{\boldsymbol{z}}_t) + y_t\big) \quad \text{subject to} \quad \sum_{t=1}^T \boldsymbol{\rho}_t' \boldsymbol{x}_t + \sum_{t=1}^T \beta^{t-1} y_t = W, \qquad \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t). \tag{23}$$

By quasi-linearity, the outside good can be suppressed and the analysis can be conducted in terms of $\{\boldsymbol{x}_t\}$ and the present-value expenditure constraint; see Appendix A for details on this suppression step.

B.2 Consistency in the non-linear model

I now formalise the notion of *consistency* when the transformation technology is non-linear. As in Section 2, this amounts to solving the consumer's constrained maximisation problem defined in (23). Indeed, combining the technology constraint $\tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t)$ with quasi-linearity allows the outside good to be suppressed (Appendix A). The consumer's problem can therefore be written as

$$\max_{\{\boldsymbol{x}_t\}_{t=1}^T} \sum_{t=1}^T \beta^{t-1} u(\tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t)) \quad \text{subject to} \quad \sum_{t=1}^T \boldsymbol{\rho}_t' \boldsymbol{x}_t = W,$$

where W is now interpreted as lifetime wealth net of outside-good consumption. The associated Lagrangian is

$$\mathcal{L}(\{\boldsymbol{x}_t\}) = \sum_{t=1}^{T} \beta^{t-1} u(\tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t)) - \{ \sum_{t=1}^{T} \boldsymbol{\rho}_t' \boldsymbol{x}_t - W \}, \tag{24}$$

where I normalise $\lambda = 1$ without loss of generality.

The first-order necessary conditions then follow analogously to Section 2, except that the derivative of the augmented characteristic vector with respect to the augmented market goods vector is now,

$$\frac{\partial \tilde{\boldsymbol{z}}_t}{\partial \tilde{\boldsymbol{x}}_t} = \frac{\partial \tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t)}{\partial \tilde{\boldsymbol{x}}_t} = \boldsymbol{\nabla} \tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t),$$

as defined in (22). Hence, following the same simplification steps as in Section 2, the first-order conditions reduce to,

$$\boldsymbol{\rho}_t = \beta^{t-1} \left(\boldsymbol{\nabla F}(\boldsymbol{x}_t) \begin{bmatrix} \partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t) \\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t) \end{bmatrix} + \beta \boldsymbol{\nabla F}^a(\boldsymbol{x}_t) \left[\partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_{t+1}) \right] \right), \tag{25}$$

where $\boldsymbol{\nabla F}^a(\boldsymbol{x}_t)$ is the $K \times J_2$ submatrix found by taking the last J_2 columns of $\boldsymbol{\nabla F}(\boldsymbol{x}_t)$. This gives rise to my formal definition of *consistency* in the non-linear model as follows.

Definition B.1. The data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t \in \{1,\dots,T\}}$ are said to be *consistent* with the non-linear one-lag habits-over-characteristics model given the increasing, concave technology $\boldsymbol{F}$ if they solve the agent's lifetime utility maximisation problem defined in Equation (23), for some locally non-satiated, superdifferentiable, and concave utility function $u(\cdot)$ and discount factor $\beta \in (0, 1]$.

The following lemma provides a set of necessary and sufficient conditions for this non-linear form of consistency to hold.

Lemma B.1. The data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t \in \{1,\dots,T\}}$ are *consistent* with the non-linear one-lag habits-over-characteristics model given the increasing, concave technology $\boldsymbol{F}$ if there exists a locally non-satiated, superdifferentiable, and concave utility function $u(\cdot)$ and a discount factor $\beta \in (0, 1]$ such that for all $t \in \{1, \dots, T-1\}$,

$$\boldsymbol{\rho}_t \geq \boldsymbol{\nabla F}(\boldsymbol{x}_t) \boldsymbol{\pi}_t^0 + \boldsymbol{\nabla F}^a(\boldsymbol{x}_t) \boldsymbol{\pi}_{t+1}^1, \tag{$\star_N$}$$

with equality for all k such that $x_t^k > 0$, and where discounted shadow prices are defined as:

$$\boldsymbol{\pi}_t^0 = \beta^{t-1} \begin{bmatrix} \partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t) \\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t) \end{bmatrix}, \tag{SP_0}$$

$$\boldsymbol{\pi}_t^1 = \beta^{t-1} \left[\partial_{\boldsymbol{z}_{t-1}^a} u(\tilde{\boldsymbol{z}}_t) \right]. \tag{SP_1}$$

where $\tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{F}}(\tilde{\boldsymbol{x}}_t)$ for all $t \in \{1, \dots, T\}$ and $\boldsymbol{\rho}_t$ denotes the vector of present-value prices.

Clearly, Definition B.1 and Lemma B.1 nest the linear one-lag habits-over-characteristics model when $\boldsymbol{F}(\boldsymbol{x}_t) = \boldsymbol{A}\boldsymbol{x}_t$. The key difference when $\boldsymbol{F}(\boldsymbol{x}_t) \neq \boldsymbol{A}\boldsymbol{x}_t$ is that the marginal product of the market goods in terms of the characteristics is no longer independent of demand. This of course also means that the price of a market good is no longer a linear combination of the shadow discounted prices. Rather, prices become a non-linear function of shadow prices that vary with demand. Using this more general, non-linear notion of consistency, I now derive testable empirical conditions involving only observables.

B.3 Afriat conditions in the non-linear model

I now generalise my main Theorem 2.1 to the non-linear model.

Theorem B.1. The following statements are equivalent:

(A_N) The data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t\in\{1,\ldots,T\}}$ are consistent with the one-lag habits model given the increasing, concave technology $\boldsymbol{F}$.

(B_N) There exist T J-vector shadow discounted prices $\left\{\boldsymbol{\pi}_t^0\right\}_{t\in\{1,\ldots,T\}}$, T J_2-vector shadow discounted prices $\left\{\boldsymbol{\pi}_t^1\right\}_{t\in\{1,\ldots,T\}}$ and a discount factor $\beta \in (0,1]$ such that,

$$0 \leq \sum_{\forall s,t\in\sigma} \tilde{\boldsymbol{\pi}}_s' \left(\tilde{\boldsymbol{z}}_t - \tilde{\boldsymbol{z}}_s\right) \qquad \forall\, \sigma \subseteq \{1,\ldots,T\} \tag{$B1_N$}$$

$$\rho_t^k \geq [\boldsymbol{\nabla F}(\boldsymbol{x}_t)]_k \boldsymbol{\pi}_t^0 + [\boldsymbol{\nabla F}^a(\boldsymbol{x}_t)]_k \boldsymbol{\pi}_{t+1}^1 \qquad \forall\, k,t \in \{1,\ldots,T-1\} \tag{$B2_N$}$$

$$\rho_t^k = [\boldsymbol{\nabla F}(\boldsymbol{x}_t)]_k \boldsymbol{\pi}_t^0 + [\boldsymbol{\nabla F}^a(\boldsymbol{x}_t)]_k \boldsymbol{\pi}_{t+1}^1 \qquad \text{if } x_t^k > 0,\ \forall\, k,t \in \{1,\ldots,T-1\} \tag{$B3_N$}$$

where $[\boldsymbol{\nabla F}(\boldsymbol{x}_t)]_k$ is the J-column vector corresponding to the k-th row of $\boldsymbol{\nabla F}(\boldsymbol{x}_t)$, $[\boldsymbol{\nabla F}^a(\boldsymbol{x}_t)]_k$ is the J_2-column vector corresponding to the k-th row of $\boldsymbol{\nabla F}^a(\boldsymbol{x}_t)$, and $\tilde{\boldsymbol{\pi}}_t := \frac{1}{\beta^{t-1}}\left[\boldsymbol{\pi}_t^{0\prime}, \boldsymbol{\pi}_t^{1\prime}\right]'$.

Proof. Identical to Theorem 2.1 using the updated notion of consistency given in Definition B.1. □

C Additional data description and summary statistics

This appendix reports additional descriptive statistics for the scanner panel, including sample selection details, purchase intensity, brand concentration, price distributions, and characteristics distributions.

C.1 Sample selection

I begin by documenting the construction of the estimation sample from the raw 2010–2011 scanner panel. Table 7 reports the sequential filtering steps and resulting sample sizes. The first four restrictions impose minimum activity and panel-length requirements required for feasibility of the revealed preference framework and are therefore purely mechanical. The final restriction—dropping households whose purchased UPCs lack complete characteristics data—is the only potentially selective cut and is evaluated separately using balance tests.

Table 7: Sample construction and filtering

Step	Households	UPCs (K)	Purely mechanical
raw scanner (2010–11 combined)	4697	1073	Yes
drop HHs with only one observed purchase occasion	4505	4505	Yes
drop HHs with only one constructed time period ($T_i \leq 1$)	3673	1132	Yes
drop HHs with $T_i \leq 2$ (not meaningful for habit formation)	2972	1122	Yes
drop HHs who purchase UPCs with missing characteristics data	2282	801	No

Notes: Rows report sequential sample restrictions applied to the 2010–11 scanner panel. T_i denotes the number of household-specific constructed time periods. The restriction $T_i \leq 2$ ensures at least two transitions for identification of one-lag habit formation. "Purely mechanical" indicates whether the restriction is driven by panel structure rather than missing product characteristics.

As shown in Table 7, most sample attrition occurs due to mechanical feasibility restrictions, while the potentially selective exclusion due to missing characteristics data affects a comparatively smaller subset of households.

To assess whether exclusion due to missing characteristics data induces observable selection, Table 8 compares demographic characteristics between the baseline sample—defined as households satisfying all mechanical feasibility restrictions—and the final analysis sample. Reported p-values correspond to Pearson chi-squared tests of equality in distributions across groups.

Across most dimensions, observable characteristics are similar between excluded and retained households. While some differences arise along geographic location and household composition, these differences are modest in magnitude, and joint tests fail to reject equality across the majority of demographic characteristics. Overall, the balance results suggest limited scope for selection on observables arising from the final exclusion.

Table 8: Demographic balance: excluded versus final sample

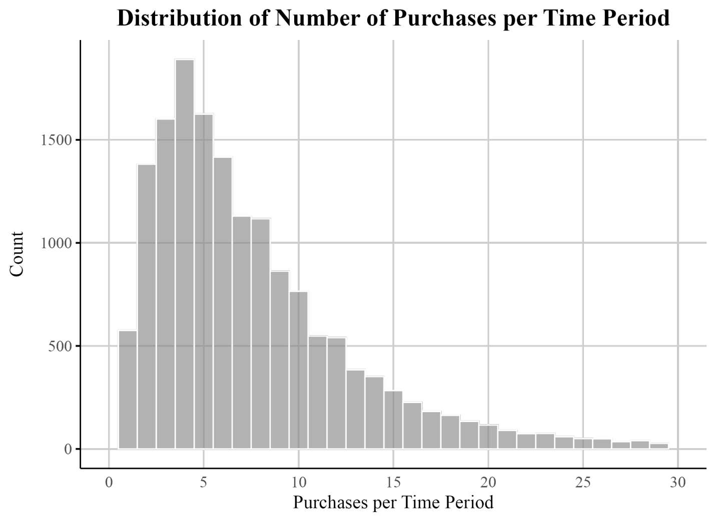
Characteristic	**Overall** N = 2,972	**Excluded** N = 690	**Final Sample** N = 2,282	**p-value**
Family size				0.012
1	550 (19%)	118 (17%)	432 (19%)	
2	1,362 (46%)	293 (42%)	1,069 (47%)	
3+	1,060 (36%)	279 (40%)	781 (34%)	
Marital status				> 0.9
Married	307 (10%)	70 (10%)	237 (10%)	
Not married	2,658 (90%)	617 (90%)	2,041 (90%)	
Region				< 0.001
Eau Claire	1,597 (54%)	278 (40%)	1,319 (58%)	
Pittsfield	1,375 (46%)	412 (60%)	963 (42%)	
Children				0.032
Any children	667 (22%)	176 (26%)	491 (22%)	
No children	2,305 (78%)	514 (74%)	1,791 (78%)	
Occupation				0.9
Blue collar/service	1,149 (39%)	267 (39%)	882 (39%)	
Office/sales (sales/clerical)	454 (15%)	100 (14%)	354 (16%)	
Other/unknown	215 (7.2%)	55 (8.0%)	160 (7.0%)	
Retired	254 (8.5%)	62 (9.0%)	192 (8.4%)	
White collar (prof/manager)	900 (30%)	206 (30%)	694 (30%)	
Education				0.8
≤ High school	1,737 (61%)	400 (60%)	1,337 (61%)	
College+	203 (7.1%)	46 (6.9%)	157 (7.2%)	
Some college/tech	911 (32%)	219 (33%)	692 (32%)	

Notes: Counts are n (%). The “Excluded” column reports households removed due to missing UPC characteristic information; all other panel-length restrictions are satisfied in both groups. p-values are from Pearson chi-squared tests of equality across excluded and retained households.

C.2 Purchase intensity and brand concentration

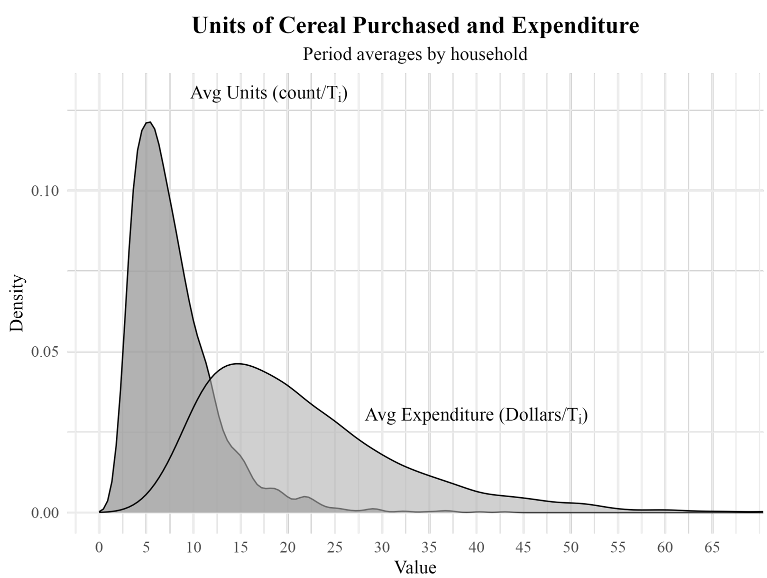
Figure 4 shows the distribution of purchase counts per household-period. Figure 5 reports the distributions of average period consumption and expenditure. Figure 6 reports the distribution of the number of distinct brands purchased per period.

Figure 4: Units purchased per household-period



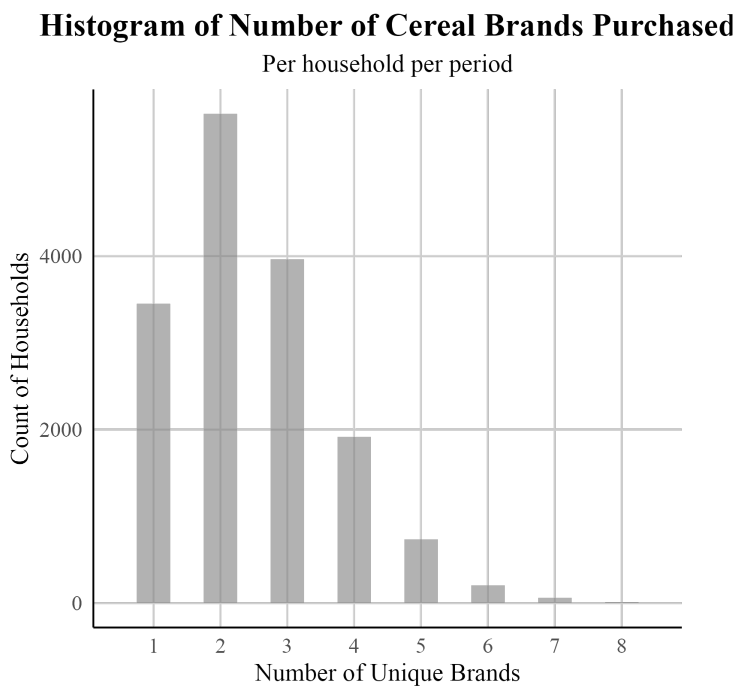
Notes: Histogram of total cereal units purchased by household i in period t, pooled across all households and constructed time periods.

Figure 5: Average purchasing intensity across households



Notes: Distributions of household-level averages of units purchased per period and expenditure per period across the sample.

Figure 6: Distinct brands purchased per household-period

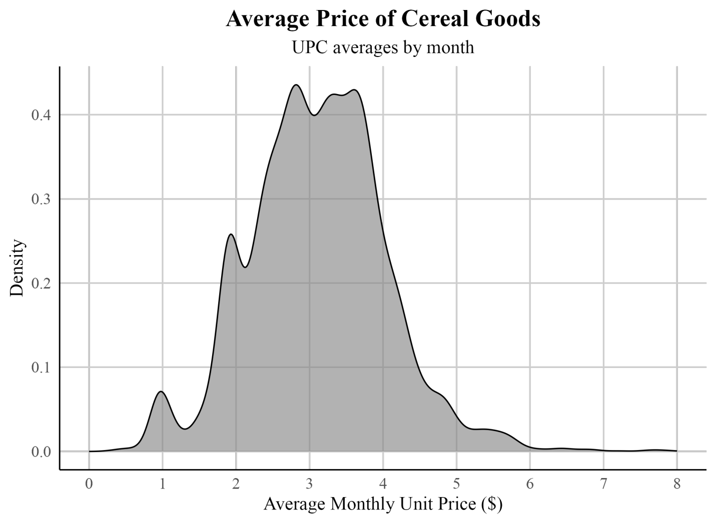


Notes: Histogram of the number of distinct cereal brands purchased by a household within a constructed time period.

C.3 Price distributions

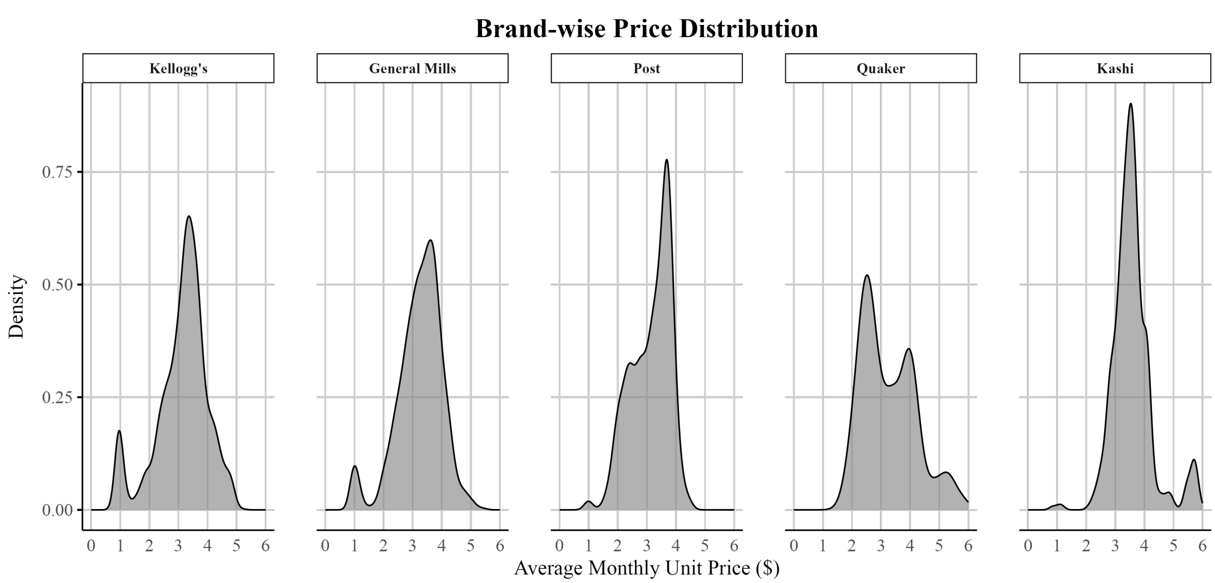
Figure 7 plots the distribution of average unit values across UPCs. Figure 8 shows the corresponding distributions for the five most prevalent brands.

Figure 7: Average unit values by UPC



Notes: Distribution of average unit values (expenditure divided by quantity) computed at the UPC level over the sample period.

Figure 8: Average unit values by brand



Notes: Distribution of UPC-level average unit values for the five most prevalent cereal brands in the sample.

C.4 Characteristics summary statistics

Tables 9 and 10 report descriptive statistics for the 23 characteristics used in the analysis. Nutritional attributes are expressed per 100g serving. Calories cluster tightly across products, while carbohydrates and sugar are generally high and fat content is low. Sodium exhibits substantial cross-product variation. Among binary attributes, whole-grain indicators and health-related descriptors are common, whereas organic and gluten-free labels are rare. Brand indicators reflect the dominance of Kellogg's and General Mills in the sample.

Table 9: Nutritional characteristics of cereal products (per 100g)

Characteristic	units	mean	sd	q1	median	q3
calories	kcal	405	534	364	380	400
carbohydrates	g	87	114	78	82	85
total fat	g	5	11	2	3	5
saturated fat	g	1	1	0	0	1
fiber	g	7	5	3	7	10
protein	g	8	10	6	7	10
sodium	mg	482	959	281	467	600
sugar	g	27	44	18	25	33

Notes: Summary statistics computed across the $K = 801$ UPCs in the final sample. Nutritional values are standardised per 100g serving and rounded to the nearest integer.

Table 10: Binary product characteristics

Characteristic	Share (%)	Characteristic	Share (%)
whole grain	57	honey flavour	15
organic	8	gluten-free	3
oat-based	24	Kellogg's	18
granola	10	General Mills	13
health halos	33	Post	11
fruity	6	Quaker	7
nutty	11	Kashi	4
chocolatey	5		

Notes: Shares denote the percentage of the $K = 801$ UPCs in the final sample exhibiting the indicated characteristic. Values are rounded to the nearest integer.

D Identification of the Discount Factor

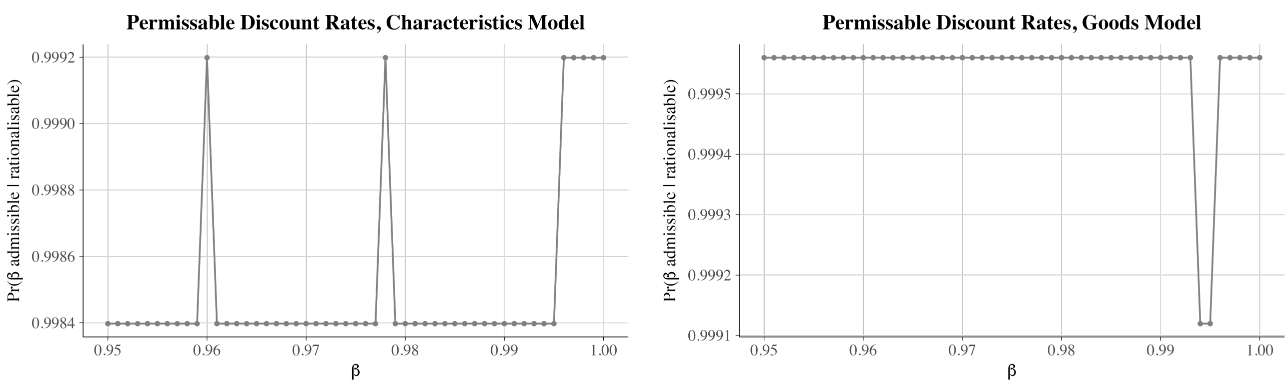
The dynamic RP inequalities are conditional on the discount factor β. To assess informativeness, I evaluate each household over a grid $\beta \in \{0.950, 0.951, \ldots, 1.000\}$ and record the set of values for which the inequalities in Theorem 2.2 are feasible. A household is classified as rationalisable if at least one β is admissible.

Figure 9a reports, for each β, the fraction of rationalisable households consistent with that value under the habits-over-characteristics specification. Figure 9b reports the corresponding results under habits-over-goods.

Acceptance probabilities are uniformly high across the grid, with only isolated spikes and dips reflecting discrete feasibility changes for marginal households. There is no systematic monotonic relationship between β and pass rates. In particular, the goods-based model admits nearly the full grid for almost all rationalisable households.

Overall, the identification sets for β are wide. In scanner environments with limited effective price variation across adjacent periods, the dynamic inequalities provide weak discipline on intertemporal discounting once the structural hedonic restrictions are satisfied.

Figure 9: Admissible discount factors among rationalisable households



(a) Habits-over-characteristics ($J_2 = 2$ (sugar and sodium)) (b) Habits-over-all-goods ($K = J = J_2$, $\boldsymbol{A} = \text{identity}(K)$)

Notes: For each discount factor β in the grid $[0.95, 1]$, the figure reports the share of households that remain feasible at β, conditional on being rationalisable for at least one value of β. Panel (a) corresponds to the habits-over-characteristics specification with $J_2 = 2$ (sugar and sodium). Panel (b) corresponds to the goods specification with $\boldsymbol{A} = \text{identity}(K)$.

E Structural and behavioural restrictiveness: additional results

To assess the empirical restrictiveness of each model, I follow the logic of Fudenberg et al. (2023) by evaluating how unusually well the model fits the observed data relative to nearby feasible alternatives. I specify discrepancy measures for the structural and behavioural restrictions implied by the model and compare their realised values in the observed data to those generated by locally perturbed versions of the scanner environment. This comparison yields quantile-based measures of restrictiveness, which record the position of the observed discrepancy within the distribution induced by the eligible set of perturbations.

E.1 Simulation design

The simulation procedure preserves the observed zero pattern in each period, meaning that only goods actually purchased in period t receive positive simulated quantities or prices. This is not a restriction on the underlying choice set but a requirement of the LP characterisation: both the equality system (B2$^+$) and the behavioural inequalities (B1$^+$) are defined only over goods with strictly positive quantities, and altering the active set would change the dimension and geometry of the equality manifold. Preserving zeros therefore fixes the empirical mapping between goods and characteristics, the dimensionality of $\widetilde{\boldsymbol{B}}_t$, and the effective choice environment that generated the observed behaviour. For each household and model, I generate $M = 10{,}000$ such locally perturbed datasets.

Within each period, simulated prices for active goods are drawn from a uniform distribution centred on the household's observed price range, scaled by multiplicative factors 0.5 and 1.5 to allow local variation. Simulated quantities are drawn as Dirichlet shares to ensure positivity and summation to one across active goods, then rescaled to match the household's observed total expenditure across periods. This yields a locally perturbed but structurally comparable environment in which both structural and behavioural margins vary across simulations.

E.2 Observed versus simulated moments

Table 11 reports means of the corresponding discrepancy measures, averaged across households in the observed data and across households and simulation draws in the simulated data. Two features stand out. First, all five characteristics-based models produce identical distributions for the distance statistic, as expected given that they share the same equality manifold. The observed structural distances are substantially smaller than those generated under local perturbations (mean 0.167 observed vs. 0.289 simulated), indicating that real scanner prices lie closer to the hedonic equality system than typical simulated price vectors. Second, behavioural variation arises only through the CCEI. For the characteristics models, CCEI values are extremely high in both observed and simulated data, with the observed values slightly higher. The goods-based models yield distance equal to zero by construction and CCEI extremely close to one in both observed and simulated data, reflecting their lack of structural and behavioural content.

E.3 Quantile-based measures of restrictiveness

For each discrepancy measure, I evaluate the position of the observed data within the distribution induced by locally perturbed environments. For the structural margin, the relevant quantity is the empirical quantile

$$q^{\text{dist}}_{n,m} = \frac{1}{M} \sum_{j=1}^{M} \mathbf{1}\{d^{\text{sim}}_{n,m,j} \leq d^{\text{obs}}_{n,m}\},$$

Table 11: Observed and simulated restrictiveness statistics

Model	Dist (real)	Dist (sim)	CCEI (real)	CCEI (sim)
Habits-over-all-characteristics	0.167	0.289	1.000	0.951
Characteristics (no habits)	0.167	0.289	0.998	0.955
Habits-over-sodium	0.167	0.289	1.000	0.950
Habits-over-sodium-sugar	0.167	0.289	1.000	0.951
Habits-over-sugar	0.167	0.289	1.000	0.956
Habits-over-all-goods	0.000	0.000	1.000	1.000
Goods (no habits)	0.000	0.000	0.998	0.999

Notes: "Dist" denotes the household-level mean structural distance-to-manifold; "CCEI" denotes the household-level behavioural efficiency index. "Real" columns report averages across households in the observed data. "Sim" columns report averages across households and simulation draws under the local perturbation design.

which records the fraction of simulated datasets whose prices lie at least as close to the hedonic equality manifold as the observed data. Smaller values indicate that the observed prices are unusually well aligned with the structural restrictions relative to nearby feasible price systems.

For the behavioural margin, I analogously compute the quantile

$$q_{n,m}^{\text{CCEI}} = \frac{1}{M}\sum_{j=1}^{M} \mathbf{1}\{\text{CCEI}_{n,m,j}^{\text{sim}} \geq \text{CCEI}_{n,m}^{\text{obs}}\},$$

which measures how frequently locally perturbed datasets exhibit behavioural efficiency at least as high as that observed in the data. Smaller values indicate that the observed behaviour is unusually close to dynamic rationality relative to nearby perturbations.

Table 12: Empirical quantile-based restrictiveness measures

Model	Mean quantile ($dist$)	$\Pr(q_{n,m}^{\text{dist}} < 0.05)$	Mean quantile ($CCEI$)	$\Pr(q_{n,m}^{\text{CCEI}} < 0.05)$
Habits-over-all-characteristics	0.579	0.117	0.998	0.00206
Characteristics (no habits)	0.579	0.117	0.990	0.00206
Habits-over-sodium	0.579	0.117	0.998	0.00206
Habits-over-sodium-sugar	0.579	0.117	0.998	0.00206
Habits-over-sugar	0.579	0.117	0.998	0.00206
Habits-over-all-goods	1.000	0.000	1.000	0.000
Goods (no habits)	1.000	0.000	0.994	0.000

Notes: For each household and model, $q_{n,m}^{\text{dist}}$ is the empirical left-tail quantile of the observed structural distance relative to its simulated distribution; $q_{n,m}^{\text{CCEI}}$ is the empirical right-tail quantile of the observed CCEI relative to its simulated distribution. Smaller quantiles indicate greater model restrictiveness. Entries report means across households and the fraction with quantile below 0.05.

E.4 Interpretation and implications for raw pass rates

These simulation results clarify the patterns in Tables 3 and 4. The raw pass rate of roughly 55% for the habits-over-characteristics model reflects the fact that the hedonic equalities in (B2$^+$) impose a demanding geometric restriction on observed retail prices. Table 12 shows that only around 12% of households lie in the lower tail of the structural discrepancy distribution: for most households, observed prices are not especially close to the hedonic equality manifold relative to nearby feasible perturbations. In the sense of Fudenberg et al. (2023), the structural discrepancy quantiles therefore indicate that the hedonic price restrictions are demanding, but that the observed data are typically not extreme outliers relative to locally perturbed price systems. This is unsurprising given that scanner prices embed promotional activity, retailer mark-ups, and other components orthogonal to consumption-relevant characteristics.

Conditional on satisfying the structural equalities, the behavioural inequalities in ($B1^+$) almost never bind. Observed CCEI values are extremely close to one across all characteristics specifications, indicating that only minimal budget perturbations are required to rationalise behaviour. Relative to locally smooth perturbations of the empirical environment, observed behaviour is therefore not unusually efficient. In this setting—where active choice sets are sparse and intertemporal budgets vary little—the behavioural inequalities in ($B1^+$) add limited additional restrictiveness beyond the structural constraints.

The goods-based models achieve raw pass rates above 90% not because they provide a better description of observed behaviour, but because they impose no structural discipline: their distance statistic is identically zero by construction. The quantile-based restrictiveness measures in Table 12 therefore indicate that differences in raw pass rates primarily reflect differences in structural restrictiveness rather than behavioural content. Within the class of characteristics-based models, alternative allocations of habit-forming attributes generate nearly identical structural and behavioural p-values, implying that differences in raw pass rates primarily reflect the tight structural restrictions imposed by the hedonic representation rather than meaningful differences in behavioural discipline. In the sense of Fudenberg et al. (2023), the goods-based models are empirically permissive, while the characteristics-based models impose tighter structural restrictions through the hedonic price representation.