Online Appendix

E Extended Lag Model

E.1 Consumer problem in the extended model

The key results in this paper focus on a simple case of the "short memory habits" (SMH) model in which the effects of addictive characteristics persist for only one period. Here, we show that the results extend easily to a more general L-lag SMH model.

Defining $L \in \mathbb{N}$ to be the length of habit persistence, our model of interest becomes

$$\max_{\{\boldsymbol{x}_t,\, y_t\}_{t=1}^T} \sum_{t=1}^{T} \beta^{t-1}\big(u(\boldsymbol{z}_t^c, \boldsymbol{z}_t^a, \boldsymbol{z}_{t-1}^a, \ldots, \boldsymbol{z}_{t-L}^a) + y_t\big) \quad \text{subject to} \quad \sum_{t=1}^{T} \boldsymbol{\rho}_t' \boldsymbol{x}_t + \sum_{t=1}^{T} \beta^{t-1} y_t = W, \qquad \boldsymbol{z}_t = \boldsymbol{A}\boldsymbol{x}_t, \tag{27}$$

where $\boldsymbol{\rho}_t$ denotes the vector of present-value prices, $\beta = 1/(1+\delta)$ where $\delta \in [0, \infty)$ is the consumer's rate of time preference, and W is the present value of the consumer's lifetime wealth.

With this extended lag dependency, we redefine the augmented vectors and matrices via:

$$\tilde{\boldsymbol{z}}_t := \begin{pmatrix} \boldsymbol{z}_t^c \\ \boldsymbol{z}_t^a \\ \boldsymbol{z}_{t-1}^a \\ \vdots \\ \boldsymbol{z}_{t-L}^a \end{pmatrix} \quad \tilde{\boldsymbol{x}}_t := \begin{pmatrix} \boldsymbol{x}_t \\ \boldsymbol{x}_{t-1} \\ \vdots \\ \boldsymbol{x}_{t-L} \end{pmatrix} \quad \tilde{\boldsymbol{A}} := \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0}_{J\times K} & \cdots & \boldsymbol{0}_{J\times K} \\ \boldsymbol{0}_{J_2\times K} & \boldsymbol{A}^a & \cdots & \boldsymbol{0}_{J_2\times K} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0}_{J_2\times K} & \boldsymbol{0}_{J_2\times K} & \cdots & \boldsymbol{A}^a \end{pmatrix}, \tag{28}$$

so that $\tilde{\boldsymbol{z}}_t$ is now a $J + LJ_2$ column vector, $\tilde{\boldsymbol{x}}_t$ is a $(L+1)K$ column vector, and $\tilde{\boldsymbol{A}}$ is a $(J + LJ_2) \times (L+1)K$ block matrix. Using this augmented notation, the general L-lag model can be written as

$$\max_{\{\boldsymbol{x}_t,\, y_t\}_{t=1}^T} \sum_{t=1}^{T} \beta^{t-1}\big(u(\tilde{\boldsymbol{z}}_t) + y_t\big) \quad \text{subject to} \quad \sum_{t=1}^{T} \boldsymbol{\rho}_t' \boldsymbol{x}_t + \sum_{t=1}^{T} \beta^{t-1} y_t = W, \qquad \tilde{\boldsymbol{z}}_t = \tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t. \tag{29}$$

Notice that by setting $L = 1$, we recover the basic model analysed in Section 2. 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.

E.2 Consistency in the extended model

The Lagrangian for the constrained optimisation problem associated with the extended lag model is

$$\mathcal{L}(\{\boldsymbol{x}_t\}, \lambda) = \sum_{t=1}^{T} \beta^{t-1} u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t) - \lambda\Big\{\sum_{t=1}^{T} \boldsymbol{\rho}_t' \boldsymbol{x}_t - W\Big\}, \tag{30}$$

where quasi-linearity justifies suppressing the outside good and W is now interpreted as lifetime wealth net of outside-good consumption; see Appendix A for details.

The associated first-order necessary conditions follow as before using the chain rule, noting we now have the following changes in dimensionality:

$$\underbrace{\frac{\partial u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t)}{\partial \boldsymbol{x}_t}}_{(K\times 1)} = \underbrace{\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_t}}_{(K\times(L+1)K)} \underbrace{\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{x}}_t}}_{((L+1)K\times 1)} = \underbrace{\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_t}}_{(K\times(L+1)K)} \underbrace{\frac{\partial \tilde{\boldsymbol{z}}_t}{\partial \tilde{\boldsymbol{x}}_t}}_{((L+1)K\times(J+LJ_2))} \underbrace{\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{z}}_t}}_{((J+LJ_2)\times 1)} \tag{31}$$

where, using our notation defined in (28) we have,

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

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

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

where $\big|$ denotes the horizontal concatenation of the $K \times K$ identity matrix and the $K \times LK$ matrix of zeros, and $\partial u(\tilde{\boldsymbol{z}})$ denotes the superderivative of u at $\tilde{\boldsymbol{z}}$. Repeating the chain rule exercise in (31), except this time differentiating with respect to the l-period lag of market goods, $l \in \{1, \ldots, L\}$, we have,

$$\underbrace{\frac{\partial u(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t)}{\partial \boldsymbol{x}_{t-l}}}_{(K\times 1)} = \underbrace{\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_{t-l}}}_{(K\times(L+1)K)} \underbrace{\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{x}}_t}}_{((L+1)K\times 1)} = \underbrace{\frac{\partial \tilde{\boldsymbol{x}}_t}{\partial \boldsymbol{x}_{t-l}}}_{(K\times(L+1)K)} \underbrace{\frac{\partial \tilde{\boldsymbol{z}}_t}{\partial \tilde{\boldsymbol{x}}_t}}_{((L+1)K\times(J+LJ_2))} \underbrace{\frac{\partial u(\tilde{\boldsymbol{z}}_t)}{\partial \tilde{\boldsymbol{z}}_t}}_{((J+LJ_2)\times 1)} \tag{32}$$

where the only new term is,

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

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

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

and for all $l \in \{1, \ldots, L\}$,

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

The first-order necessary conditions associated with the Lagrangian in (30) now follow immediately as,

$$\begin{aligned}
\partial_{\boldsymbol{x}_t}\mathcal{L} = 0 \quad &\Rightarrow \quad \lambda \boldsymbol{\rho}_t = \beta^{t-1}\left[\boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times LK}\right] \tilde{\boldsymbol{A}}' \partial u(\tilde{\boldsymbol{z}}_t) + \sum_{l=1}^{L} \beta^{t-1+l}\left[\boldsymbol{0}_{K\times lK} \,\middle|\, \boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times(L-l)K}\right] \tilde{\boldsymbol{A}}' \partial u(\tilde{\boldsymbol{z}}_{t+l}) \\
&\Rightarrow \quad \boldsymbol{\rho}_t = \frac{\beta^{t-1}}{\lambda}\left(\left[\boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times LK}\right] \tilde{\boldsymbol{A}}' \partial u(\tilde{\boldsymbol{z}}_t) + \sum_{l=1}^{L} \beta^{l}\left[\boldsymbol{0}_{K\times lK} \,\middle|\, \boldsymbol{I}_{K\times K} \,\middle|\, \boldsymbol{0}_{K\times(L-l)K}\right] \tilde{\boldsymbol{A}}' \partial u(\tilde{\boldsymbol{z}}_{t+l})\right).
\end{aligned} \tag{33}$$

But, just as in the simple case of $L = 1$, these first-order conditions can be substantially simplified. Multiplying the conformable block matrices as in Section 2 the first-order conditions in (33) reduce to:

$$\boldsymbol{\rho}_t = \frac{\beta^{t-1}}{\lambda}\left(\boldsymbol{A}'\begin{bmatrix}\partial_{\boldsymbol{z}_t^c} u(\tilde{\boldsymbol{z}}_t)\\ \partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_t)\end{bmatrix} + \sum_{l=1}^{L}\beta^l(\boldsymbol{A}^a)'\left[\partial_{\boldsymbol{z}_t^a} u(\tilde{\boldsymbol{z}}_{t+l})\right]\right). \tag{34}$$

This gives rise to our formal definition of *consistency* in the extended SMH model as follows:

Definition E.1. The data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t\in\{1,\ldots,T\}}$ are *consistent* with the L-lag habits-over-characteristics model for given technology $\boldsymbol{A}$ if they solve the agent's lifetime utility maximisation problem defined in Equation (29), for some locally non-satiated, superdifferentiable, and concave utility function $u(\cdot)$ and positive constants λ and β.

The following lemma provides a set of necessary and sufficient conditions for this extended lag form of consistency to hold.

Lemma E.1. The data $\{\boldsymbol{\rho}_t; \boldsymbol{x}_t\}_{t\in\{1,\ldots,T\}}$ are *consistent* with the L-lag habits-over-characteristics model for given technology $\boldsymbol{A}$ if there exists a locally non-satiated, superdifferentiable, and concave utility function $u(.)$ and positive constants λ and β such that for all $t \in \{1, \ldots, T - L\}$,

$$\boldsymbol{\rho}_t \geq \boldsymbol{A}'\boldsymbol{\pi}_t^0 + \sum_{l=1}^{L}(\boldsymbol{A}^a)'\boldsymbol{\pi}_{t+l}^l, \tag{$\star_L$}$$

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

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

$$\boldsymbol{\pi}_t^l = \frac{\beta^{t-1}}{\lambda}\left[\partial_{\boldsymbol{z}_{t-l}^a} u(\tilde{\boldsymbol{z}}_t)\right], \tag{SP_l}$$

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

As usual, λ denotes the marginal utility of lifetime wealth. We can interpret $\boldsymbol{\pi}_t^l$ as the discounted WTP for the consumption of habit-forming characteristics l periods ago.

Clearly, Definition E.1 and Lemma E.1 nest the simple one-lag habits-over-characteristics model when $L = 1$. Indeed, the latter gives the natural (dynamic) extension of the hedonic pricing equation when habits persist for exactly L periods. It tells us that the discounted prices $\boldsymbol{\rho}_t$ of goods today depend on current discounted shadow prices of the characteristics *as well as* the discounted shadow price of habit-forming characteristics tomorrow, $\boldsymbol{\pi}_{t+1}^1$, the next day, $\boldsymbol{\pi}_{t+2}^2$, and up to L periods in the future, $\boldsymbol{\pi}_{t+3}^3, \ldots, \boldsymbol{\pi}_{t+L}^L$. This is because today's consumption of goods (and the habit-forming characteristics contained therein) affects the agent's marginal utility L periods in the future by building up a habit. A characterisation of this notion of consistency in the extended SMH model follows naturally from Theorem 2.1.

E.3 Afriat conditions in the extended model

Theorem E.1. The following statements are equivalent:

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

(B_L) There exist T J-vector shadow discounted prices $\left\{\boldsymbol{\pi}_t^0\right\}_{t\in\{1,\ldots,T\}}$, T LJ_2-vector shadow discounted prices $\left\{\boldsymbol{\pi}_t^1,\ldots,\boldsymbol{\pi}_t^L\right\}_{t\in\{1,\ldots,T\}}$ and positive constants β and λ such that,

$$0 \leq \sum_{\forall s,t\in\sigma} \tilde{\boldsymbol{\pi}}_s' \left(\tilde{\boldsymbol{z}}_t - \tilde{\boldsymbol{z}}_s\right) \qquad \forall\, \sigma \subseteq \{1,\ldots,T\} \tag{$B1_L$}$$

$$\rho_t^k \geq \boldsymbol{a}_k' \boldsymbol{\pi}_t^0 + \sum_{l=1}^{L} \boldsymbol{a}_k^{a\prime} \boldsymbol{\pi}_{t+l}^l \qquad \forall\, k,t \in \{1,\ldots,T-L\} \tag{$B2_L$}$$

$$\rho_t^k = \boldsymbol{a}_k' \boldsymbol{\pi}_t^0 + \sum_{l=1}^{L} \boldsymbol{a}_k^{a\prime} \boldsymbol{\pi}_{t+l}^l \qquad \text{if } x_t^k > 0,\ \forall\, k,t \in \{1,\ldots,T-L\} \tag{$B3_L$}$$

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

Proof. Identical to Theorem 2.1 with extended lag notation. $\square$

F Testing model consistency via linear programming

Theorem 2.1 is the characteristics model analogue of Theorem 1 in the habits-over-goods model by Crawford (2010). As discussed in Section 2.2, theoretical consistency of the data reduces to a linear programming problem when one commits to a grid search over the discount factor, β. However, in its current form, condition (B) has some disadvantages in its practical implementation. Indeed, since (B1) must hold for all $\sigma \subseteq \{1, \ldots, T\}$, this condition alone requires testing $2^T - 1$ inequalities. This becomes very computationally expensive with the length of panel. To address this issue, we have derived the following equivalent statement to be used when implementing the test for model consistency.

Theorem F.1. The following statements are equivalent:

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

(L) There exist T numbers $\{V_t\}_{t=1,\ldots,T}$, 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 positive constant β such that,

$$V_s - V_t - \frac{1}{\beta^{t-1}} \left[\boldsymbol{\pi}_t^{0\prime}, \boldsymbol{\pi}_t^{1\prime}\right] (\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t) \leq 0 \quad \forall\, s, t \in \{1, \ldots, T\} \tag{L1}$$

$$\left[\boldsymbol{A}' \,\middle|\, (\boldsymbol{A}^a)'\right] \begin{bmatrix} \boldsymbol{\pi}_t^0 \\ \boldsymbol{\pi}_{t+1}^1 \end{bmatrix} \leq \boldsymbol{\rho}_t \qquad \forall\, t \in \{1, \ldots, T-1\} \tag{L2}$$

$$\left[\boldsymbol{B}_t' \,\middle|\, (\boldsymbol{B}_t^a)'\right] \begin{bmatrix} \boldsymbol{\pi}_t^0 \\ \boldsymbol{\pi}_{t+1}^1 \end{bmatrix} = \boldsymbol{\rho}_t^+ \qquad \forall\, t \in \{1, \ldots, T-1\} \tag{L3}$$

where $\boldsymbol{\rho}_t^+$ is a K_t^+ vector equal to the sub-vector of period t prices for which demands are positive, and $\boldsymbol{B}_t$ and $\boldsymbol{B}_t^a$ are the corresponding $J \times K_t^+$ and $J_2 \times K_t^+$ sub-matrices matrices of $\boldsymbol{A}$ and $\boldsymbol{A}^a$, respectively (as introduced in Section 2.4).

Notice that the original (B1) has been converted to the equivalent constraint (L1), which requires testing only $(T-1)^2$ inequalities. This is a strict improvement on the $2^T - 1$ inequalities involved in testing constraint (B1).

Proof.

Set $\lambda = 1$ in Theorem 2.1 without loss of generality (see Section 2.2). Then, condition (A) is identical to that in Theorem 2.1. Accounting for notational differences, conditions (L2) and (L3) are also identical to conditions (B2) and (B3) in Theorem 2.1, respectively. Hence, the proof reduces to showing that condition (L1) is equivalent to condition (B1) in Theorem 2.1.

$(B1) \Rightarrow (L1)$: Assume (B1) holds. This means that for the specific choice of $\sigma = \{s, t\}$, where $s, t \in \{1, \ldots, T\}$ it holds,

$$0 \leq \tilde{\boldsymbol{\pi}}_s'(\tilde{\boldsymbol{z}}_t - \tilde{\boldsymbol{z}}_s) + \tilde{\boldsymbol{\pi}}_t'(\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t) = \tilde{\boldsymbol{\pi}}_s'\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t - \tilde{\boldsymbol{\pi}}_s'\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_s + \tilde{\boldsymbol{\pi}}_t'(\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t). \tag{35}$$

Now, define Afriat numbers,

$$V_t := \tilde{\boldsymbol{\pi}}_s'\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}_t, \quad \forall\, t \in \{1, \ldots T\}.$$

Then, (35) is equivalent to,

$$0 \leq V_t - V_s + \tilde{\boldsymbol{\pi}}_t'(\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t), \quad \forall\, s,t \in \{1,\ldots T\}.$$

Substituting in the definition for $\tilde{\boldsymbol{\pi}}_t$ from Theorem 2.1 this gives us,

$$V_s - V_t - \frac{1}{\beta^{t-1}}\left[\boldsymbol{\pi}_t^{0\prime},\, \boldsymbol{\pi}_t^{1\prime}\right](\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t) \leq 0 \quad \forall\, s,t \in \{1,\ldots,T\},$$

which is constraint (L1).

$(L1) \Rightarrow (B1)$: Assume (L1) holds. Using the augmented notation, this means that

$$0 \leq V_t - V_s + \tilde{\boldsymbol{\pi}}_t'(\tilde{\boldsymbol{z}}_s - \tilde{\boldsymbol{z}}_t), \quad \forall\, s,t \in \{1,\ldots T\}.$$

Summing both sides of the inequality over any choice of $\sigma \subseteq \{1,\ldots T\}$ sees the Afriat numbers cancelling to leave us with,

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

which is constraint (B1). $\square$