Online Appendix

to

Learning from Experience, Simply

by

Song Lin

Juanjuan Zhang

and

John R. Hauser

May 29, 2014

Song Lin is a PhD Candidate, MIT Sloan School of Management, Massachusetts Institute of Technology, E62-582, 77 Massachusetts Avenue, Cambridge, MA 02139, (617) 253-7273, songlin@mit.edu.

Juanjuan Zhang is an Associate Professor of Marketing, MIT Sloan School of Management, Massachusetts Institute of Technology, E62-537, 77 Massachusetts Avenue, Cambridge, MA 02139, (617) 452-2790, jjzhang@mit.edu.

John R. Hauser is the Kirin Professor of Marketing, MIT Sloan School of Management, Massachusetts Institute of Technology, E62-538, 77 Massachusetts Avenue, Cambridge, MA 02139, (617) 253-2929, hauser@mit.edu.

Online Appendix

List of Contents

Online Appendix A. Proof of Proposition 1 (Indexability)

A.1. Proof of Proposition 1 When Consumers Are Risk Neutral

Without loss of generality, we set observable shocks x_t to zero and use ϵ_t to represent all utility shocks. The focus is on the sub-problem where the consumer chooses between an uncertain brand j and a certain reward λ. To simplify notation, we drop the brand identifier j. The Bellman equation for this problem is

$$\text{(A1)} \quad V(s_t, \epsilon_t, \lambda) = \max\{\lambda + \delta \mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)|\epsilon_t],\ \epsilon_t + \mathbb{E}[q_t|s_t] + \delta \mathbb{E}[V(s_{t+1}, \epsilon_{t+1}, \lambda)|s_t, \epsilon_t]\},$$

where s_t summarizes the consumer's belief about brand quality at purchase occasion t. The definition of indexability is that, for any state (s_t, ϵ_t), if it is optimal to choose the fixed reward λ, then it must be also optimal to choose the fixed reward λ' for any $\lambda' > \lambda$. This is equivalent to the following condition:

$$\text{(A2)} \qquad \frac{\partial}{\partial \lambda}\left(\lambda + \delta \mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)|\epsilon_t] - \epsilon_t - \mathbb{E}\left[q_t|s_t\right] - \delta \mathbb{E}\left[V(s_{t+1}, \epsilon_{t+1}, \lambda)|s_t, \epsilon_t\right]\right) \geq 0$$

$$\Leftrightarrow 1 + \delta \frac{\partial}{\partial \lambda} \mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)|\epsilon_t] - \delta \frac{\partial}{\partial \lambda} \mathbb{E}[V(s_{t+1}, \epsilon_{t+1}, \lambda)|s_t, \epsilon_t] \geq 0.$$

Intuitively, condition (A2) requires that, as λ increases, the expected future value of choosing the uncertain brand should not grow too fast compared to that of choosing the fixed reward λ. It turns out that the assumptions of the canonical forward-looking experiential learning problem as specified in §2 are *sufficient* for condition (A2) to hold. We prove this result below.

We first define the expected value function $EV(s_t, \lambda)$ by integrating out ϵ_t:

$$\text{(A3)} \qquad EV(s_t, \lambda) \triangleq \mathbb{E}_{\epsilon_t}[V(s_t, \epsilon_t, \lambda)].$$

Given the assumptions that ϵ_{t+1} and ϵ_t are *i.i.d.* and that ϵ_t is independent of s_t, we have

$$\mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)|\epsilon_t] = \mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)] = \mathbb{E}[V(s_t, \epsilon_t, \lambda)] = EV(s_t, \lambda),$$

and

$$\mathbb{E}[V(s_{t+1},\epsilon_{t+1},\lambda)|s_t,\epsilon_t] = \mathbb{E}[V(s_{t+1},\epsilon_{t+1},\lambda)|s_t] = \mathbb{E}[V(s_{t+1},\epsilon_t,\lambda)|s_t] = \mathbb{E}[EV(s_{t+1},\lambda)|s_t].$$

Therefore, Equation (A1) implies the following fixed point:

$$\text{(A4)} \quad EV(s_t,\lambda) = \int \max\{\lambda + \delta EV(s_t,\lambda),\ \epsilon_t + \mathbb{E}[q_t|s_t] + \delta\mathbb{E}[EV(s_{t+1},\lambda)|s_t]\}\, dH(\epsilon_t).$$

Denote 0 as the option of the certain reward λ, and 1 as the uncertain brand. We define the following quantities:

$$\text{(A5)} \qquad v_0(s_t,\lambda) \triangleq \lambda + \delta EV(s_t,\lambda) \ \text{ and } \ v_1(s_t,\lambda) \triangleq \mathbb{E}\left[q_t|s_t\right] + \delta\mathbb{E}[EV(s_{t+1},\lambda)|s_t].$$

Observe that the conditional probability of choosing 1 is given by

$$\text{(A6)} \qquad \begin{aligned} P(1|s_t,\lambda) &= \int \mathbb{1}\left[v_0(s_t,\lambda) \le \epsilon_t + v_1(s_t,\lambda)\right]\, dH(\epsilon_t) \\ &= \int \frac{\partial}{\partial v_1(s_t,\lambda)} \max\left[v_0(s_t,\lambda), \epsilon_t + v_1(s_t,\lambda)\right]\, dH(\epsilon_t) \\ &= \frac{\partial EV(s_t,\lambda)}{\partial v_1(s_t,\lambda)}. \end{aligned}$$

The last equality is obtained by interchanging integration and differentiation and evoking Equations (A4) and (A5). Similarly we have

$$\text{(A7)} \qquad P(0|s_t,\lambda) = \frac{\partial EV(s_t,\lambda)}{\partial v_0(s_t,\lambda)}.$$

Differentiating both sides of Equation (A4) with respect to λ and using the Chain Rule, we obtain

$$\text{(A8)} \qquad \begin{aligned} EV_\lambda(s_t,\lambda) &\triangleq \frac{\partial EV(s_t,\lambda)}{\partial \lambda} = \frac{\partial EV(s_t,\lambda)}{\partial v_0(s_t,\lambda)}\frac{\partial v_0(s_t,\lambda)}{\partial \lambda} + \frac{\partial EV(s_t,\lambda)}{\partial v_1(s_t,\lambda)}\frac{\partial v_1(s_t,\lambda)}{\partial \lambda} \\ &= P(0|s_t,\lambda)(1 + \delta EV_\lambda(s_t,\lambda)) + P(1|s_t,\lambda)(\delta\mathbb{E}[EV_\lambda(s_{t+1},\lambda)|s_t]), \end{aligned}$$

where the last equality is obtained from Equations (A5), (A6), and (A7). Next, we prove the following lemma.

Lemma 1. For all s, λ, we have

(A9) $$0 \leq EV_{\lambda}(s,\lambda) \leq \frac{1}{1-\delta}.$$

Proof. Fix any s, ϵ, and λ. Let π^* denote the optimal policy that achieves $V(s,\epsilon,\lambda)$. First, if a positive constant c is added only to the fixed reward λ in every period but the uncertain brand remains unchanged, then following π^* yields an expected total utility at least as large as $V(s,\epsilon,\lambda)$. Therefore, $V(s,\epsilon,\lambda+c) \geq V(s,\epsilon,\lambda)$. Second, if a positive constant c is added to both the fixed reward and the uncertain brand in every period, then π^* is still optimal and yields an expected total utility of $V(s,\epsilon,\lambda) + c/(1-\delta)$. By construction, adding a positive constant to both options yields expected utility at least as high as adding the constant only to the fixed reward: $V(s,\epsilon,\lambda) + c/(1-\delta) \geq V(s,\epsilon,\lambda+c)$. Integrating out ϵ we have

$$EV(s,\lambda) \leq EV(s,\lambda+c) \leq EV(s,\lambda) + \frac{c}{1-\delta}.$$

It follows that

$$0 \leq \frac{EV(s,\lambda+c) - EV(s,\lambda)}{c} \leq \frac{1}{1-\delta}.$$

Taking the limit on both sides as $c \to 0$ establishes the lemma.

Lemma 1 implies that $0 \leq EV_{\lambda}(s_t,\lambda) \leq 1 + \delta EV_{\lambda}(s_t,\lambda)$. This result, together with Equation (A8) and the fact that $P(0|s_t,\lambda) + P(1|s_t,\lambda) = 1$, in turn implies that:

(A10) $$1 + \delta EV_{\lambda}(s_t,\lambda) \geq \delta \mathbb{E}[EV_{\lambda}(s_{t+1},\lambda)|s_t],$$

which establishes the indexability condition (A2).

A.2. Proof of Proposition 1 When Consumers Exhibit General Risk Preferences

In this section, we extend the proof of §A.1 to show that the canonical forward-looking experiential learning problem is indexable when consumers exhibit general risk preferences.

The Bellman equation in the case of general risk preferences is

(A1') $$V(s_t, \epsilon_t, \lambda) = \max\{u(\lambda) + \delta \mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)|\epsilon_t],$$
$$\mathbb{E}[u(\epsilon_t + q_t)|s_t] + \delta \mathbb{E}[V(s_{t+1}, \epsilon_{t+1}, \lambda)|s_t, \epsilon_t]\}.$$

For the ease of comparison, we denote the above equation as (A1'), meaning that it corresponds to Equation (A1) of §A.1. The same notational rule applies throughout §A.2. The indexability condition becomes

(A2') $$\frac{\partial}{\partial \lambda}(u(\lambda) + \delta \mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)|\epsilon_t] - \mathbb{E}[u(\epsilon_t + q_t)|s_t] - \delta \mathbb{E}[V(s_{t+1}, \epsilon_{t+1}, \lambda)|s_t, \epsilon_t]) \geq 0$$

$$\Leftrightarrow u'(\lambda) + \delta \frac{\partial}{\partial \lambda} \mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)|\epsilon_t] - \delta \frac{\partial}{\partial \lambda} \mathbb{E}[V(s_{t+1}, \epsilon_{t+1}, \lambda)|s_t, \epsilon_t] \geq 0.$$

We again define the expected value function $EV(s_t, \lambda)$ by integrating out ϵ_t:

(A3') $$EV(s_t, \lambda) \triangleq \mathbb{E}_{\epsilon_t}[V(s_t, \epsilon_t, \lambda)].$$

Given the assumptions that ϵ_{t+1} and ϵ_t are *i.i.d.* and that ϵ_t is independent of s_t, we have:

$$\mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)|\epsilon_t] = \mathbb{E}[V(s_t, \epsilon_{t+1}, \lambda)] = \mathbb{E}[V(s_t, \epsilon_t, \lambda)] = EV(s_t, \lambda),$$

and

$$\mathbb{E}[V(s_{t+1}, \epsilon_{t+1}, \lambda)|s_t, \epsilon_t] = \mathbb{E}[V(s_{t+1}, \epsilon_{t+1}, \lambda)|s_t] = \mathbb{E}[V(s_{t+1}, \epsilon_t, \lambda)|s_t] = \mathbb{E}[EV(s_{t+1}, \lambda)|s_t].$$

Therefore, the Bellman equation of the sub-problem implies the following fixed point:

(A4') $$EV(s_t, \lambda) = \int \max\{u(\lambda) + \delta EV(s_t, \lambda),$$
$$\mathbb{E}[u(\epsilon_t + q_t)|s_t] + \delta \mathbb{E}[EV(s_{t+1}, \lambda)|s_t]\} \, dH(\epsilon_t).$$

Denote 0 as the option of the certain reward λ, and 1 as the uncertain brand. We define the following quantities:

(A5') $$v_0(s_t, \lambda) \triangleq u(\lambda) + \delta EV(s_t, \lambda) \quad \text{and} \quad v_1(s_t, \lambda) \triangleq \delta \mathbb{E}[EV(s_{t+1}, \lambda)|s_t].$$

The conditional probability of choosing 1 is given by

(A6') $$P(1|s_t,\lambda) = \int \mathbb{1}\{v_0(s_t,\lambda) \le \mathbb{E}[u(\epsilon_t+q_t)|s_t] + v_1(s_t,\lambda)\}\, dH(\epsilon_t)$$

$$= \int \frac{\partial}{\partial v_1(s_t,\lambda)} \max\{v_0(s_t,\lambda),\ \mathbb{E}\big[u\big(\epsilon_t+q_t\big)|s_t\big] + v_1(s_t,\lambda)\big\}\, dH(\epsilon_t)$$

$$= \frac{\partial EV(s_t,\lambda)}{\partial v_1(s_t,\lambda)}.$$

The last equality is obtained by interchanging integration and differentiation and evoking Equations (A4') and (A5'). Similarly we have

(A7') $$P(0|s_t,\lambda) = \frac{\partial EV(s_t,\lambda)}{\partial v_0(s_t,\lambda)}.$$

Differentiating both sides of Equation (A4') with respect to λ and using the Chain Rule, we obtain

(A8') $$EV_\lambda(s_t,\lambda) \triangleq \frac{\partial EV(s_t,\lambda)}{\partial \lambda} = \frac{\partial EV(s_t,\lambda)}{\partial v_0(s_t,\lambda)}\frac{\partial v_0(s_t,\lambda)}{\partial \lambda} + \frac{\partial EV(s_t,\lambda)}{\partial v_1(s_t,\lambda)}\frac{\partial v_1(s_t,\lambda)}{\partial \lambda}$$

$$= P(0|s_t,\lambda)\big(u'(\lambda) + \delta EV_\lambda(s_t,\lambda)\big) + P(1|s_t,\lambda)(\delta \mathbb{E}[EV_\lambda(s_{t+1},\lambda)|s_t]),$$

where the last equality is obtained from Equations (A5'), (A6'), and (A7'). Next, we prove the following lemma.

Lemma 1'. For all s,λ, we have

(A9') $$0 \le EV_\lambda(s,\lambda) \le \frac{u'(\lambda)}{1-\delta}.$$

Proof. Fix any s,ϵ, and λ. Let π^* denote the optimal policy that solves $V(s,\epsilon,\lambda)$. Suppose a positive constant c is added only to the fixed reward λ in every period but the uncertain brand remains unchanged. First, the consumer is weakly better off after this change. Even if the consumer maintains π^* – and the consumer can do weakly better – the consumer's expected utility remains unchanged in periods when the uncertain brand is chosen but increases in periods when the fixed reward is chosen. Therefore, $V(s,\epsilon,\lambda+c) \geq V(s,\epsilon,\lambda)$.

Second, let π' denote the optimal policy that solves $V(s,\epsilon,\lambda+c)$. Suppose the consumer adopts π' in state (s,ϵ,λ). After the increase in the fixed reward, the consumer's expected utility from adopting π' remains unchanged in periods when π' indicates choosing the uncertain brand but increases in periods when π' indicates choosing the fixed reward. The improvement in the consumer's expected discounted utility is thus weakly less than if π' indicated choosing the fixed reward in each period, in which case the improvement in the consumer's expected discounted utility would equal $\frac{u(\lambda+c)-u(\lambda)}{1-\delta}$. Recall that π^* is the optimal policy that solves $V(s,\epsilon,\lambda)$. By definition, the consumer is weakly better off choosing π^* than π' in state (s,ϵ,λ). Therefore:

$$V(s,\epsilon,\lambda+c) - V(s,\epsilon,\lambda) \leq \frac{u(\lambda+c)-u(\lambda)}{1-\delta}.$$

Integrating out ϵ we have:

$$EV(s,\lambda) \leq EV(s,\lambda+c) \leq EV(s,\lambda) + \frac{u(\lambda+c)-u(\lambda)}{1-\delta}.$$

It follows that:

$$0 \leq \frac{EV(s,\lambda+c)-EV(s,\lambda)}{c} \leq \frac{u(\lambda+c)-u(\lambda)}{c(1-\delta)}.$$

Taking the limit on both sides as $c \to 0$ establishes the lemma.

Lemma 1' implies that $0 \leq EV_\lambda(s_t,\lambda) \leq u'(\lambda) + \delta EV_\lambda(s_t,\lambda)$. This result, together with Equation (A8') and the fact that $P(0|s_t,\lambda) + P(1|s_t,\lambda) = 1$, in turn implies that:

(A10') $$u'(\lambda) + \delta EV_\lambda(s_t,\lambda) \geq \delta \mathbb{E}[EV_\lambda(s_{t+1},\lambda)|s_t],$$

which establishes the indexability condition (A2').

Online Appendix B. Proof of Proposition 2 (Invariance)

We first prove two useful lemmas. The focus is again on the sub-problem of a single brand and thus we drop the brand identifier j.

Lemma 2. Fix a prior quality belief $s_0 = (\bar{\mu}_0, \bar{\sigma}_0)$, and a sequence of quality draws $\{q_t: t \geq 0\}$. Consider a modified version of the original sub-problem where the utility shocks become $\epsilon_t^m = \epsilon_t + c$ for all t, and the fixed reward becomes $\lambda^m = \lambda + c$. Denote EV^m and W^m as the expected value and index value for the modified problem. Then for any belief state s, we have:

(B1)
$$EV^m(s, \lambda + c) = EV(s, \lambda) + \frac{c}{1-\delta},$$

(B2)
$$W^m(s, \epsilon + c;\ \mu^\epsilon + c, \sigma^\epsilon) = W(s, \epsilon;\ \mu^\epsilon, \sigma^\epsilon) + c.$$

Proof. To prove the first part of the lemma, it suffices to show that the proposed identity in Equation (B1) satisfies the fixed-point relationship implied by the modified problem:

(B3)
$$EV^m(s_t, \lambda^m) = \int \max\{\lambda^m + \delta EV^m(s_t, \lambda^m),\ \epsilon_t^m + \mathbb{E}[q_t|s_t] + \delta \mathbb{E}[EV^m(s_{t+1}, \lambda^m)|s_t]\}\, dH^m(\epsilon_t^m).$$

Suppose Equation (B1) holds. Following the definitions given in Equation (A5) we have

(B4)
$$v_0^m(s_t, \lambda^m) \triangleq \lambda + c + \delta EV^m(s_t, \lambda + c) = \lambda + \delta EV(s_t, \lambda) + \frac{c}{1-\delta} = v_0(s_t, \lambda) + \frac{c}{1-\delta},$$

and

(B5) $$v_1^m(s_t, \lambda^m) \triangleq \mathbb{E}\left[q_t|s_t\right] + \delta\mathbb{E}[EV^m(s_{t+1}, \lambda + c)|s_t]$$

$$= \mathbb{E}[q_t|s_t] + \delta\mathbb{E}[EV(s_{t+1}, \lambda)|s_t] + \frac{c\delta}{1-\delta} = v_1(s_t, \lambda) + \frac{c\delta}{1-\delta}.$$

Define $\Delta(s_t, \lambda) \triangleq v_0(s_t, \lambda) - v_1(s_t, \lambda)$. Then $\Delta^m(s_t, \lambda + c) = \Delta(s_t, \lambda) + c$ for the modified problem. The assumption that the distribution of ϵ has scale and location parameters implies:

(B6) $$H^m(\Delta^m) = \Pr(\epsilon^m \leq \Delta^m) = \Pr(\epsilon \leq \Delta) = H(\Delta), \text{ and}$$

$$\int_{\epsilon^m \geq \Delta^m} \epsilon^m d\,H^m(\epsilon^m) = \int_{\epsilon \geq \Delta} (\epsilon + c) d\,H(\epsilon).$$

The right-hand side of Equation (B3) becomes:

$$\int_{\epsilon^m \geq \Delta^m} \epsilon^m d\,H^m(\epsilon^m) + v_1^m(s_t, \lambda^m)\left(1 - H^m(\Delta^m)\right) + v_0^m(s_t, \lambda^m) H^m(\Delta^m)$$

$$= \int_{\epsilon \geq \Delta} (\epsilon + c) d\,H(\epsilon) + \left(v_1(s_t, \lambda) + \frac{c\delta}{1-\delta}\right)\left(1 - H(\Delta)\right) + \left(v_0(s_t, \lambda) + \frac{c}{1-\delta}\right) H(\Delta)$$

$$= \int_{\epsilon \geq \Delta} \epsilon d\,H(\epsilon) + v_1(s_t, \lambda)\left(1 - H(\Delta)\right) + v_0(s_t, \lambda)\,H(\Delta) + \frac{c}{1-\delta}$$

$$= EV(s, \lambda) + \frac{c}{1-\delta}$$

$$= EV^m(s, \lambda + c)\,,$$

which is the left-hand side of Equation (B3). The first equality follows from Equations (B4) to (B6). The third equality uses the fact that $EV(s, \lambda)$ is the fixed point of Equation (A4). Therefore, $EV^m(s, \lambda^m)$ also satisfies the fixed-point relationship implied by the modified problem.

For the second part of the lemma, we use the definition of Whittle's index, which is the value of λ^m such that

(B7) $$v_0^m(s_t, \lambda^m) = \epsilon^m + v_1^m(s_t, \lambda^m).$$

It suffices to show that the proposed identity in Equation (B2) solves the above equality. Note that the right-hand side of Equation (B7) is:

$$\epsilon^m + v_1^m(s_t, \lambda^m) = \epsilon + c + v_1(s_t, \lambda) + \frac{c\delta}{1-\delta}$$

$$= \epsilon + v_1(s_t, \lambda) + \frac{c\delta}{1-\delta} = v_0(s_t, \lambda) + \frac{c}{1-\delta},$$

which equals the left-hand side of Equation (B7) following Equation (B4). The first equality follows from Equation (B5). The third equality follows from the definition of Whittle's index for the original problem (setting $W(s, \epsilon) = \lambda$). Then by the definition of Whittle's index, we have $W^m(s, \epsilon + c;\ \mu^\epsilon + c, \sigma^\epsilon) = W(s, \epsilon;\ \mu^\epsilon, \sigma^\epsilon) + c$.

Lemma 3. Fix the original sub-problem. Consider a modified problem where the quality sample becomes $\{q_t^m = bq_t + c: t \geq 0\}$, the utility shocks becomes $\epsilon_t^m = b\epsilon_t$ for all t, the prior belief becomes $s_0^m = (\bar{\mu}_0^m, \bar{\sigma}_0^m) = (b\bar{\mu}_0 + c, b\bar{\sigma}_0)$, and the fixed reward becomes $\lambda^m = b\lambda + c$. Then for all t, $s_t^m = (\bar{\mu}_t^m, \bar{\sigma}_t^m) = (b\bar{\mu}_t + c, b\bar{\sigma}_t)$. Denote EV^m and W^m as the expected value and index value for the modified problem. Then for any belief state s, we have:

(B8) $$EV^m(s^m, b\lambda + c) = bEV(s, \lambda) + \frac{c}{1-\delta},$$

(B9) $$W^m(s^m, b\epsilon;\ b\mu^\epsilon, b\sigma^\epsilon) = bW(s, \epsilon;\ \mu^\epsilon, \sigma^\epsilon) + c.$$

Proof. The strategy of the proof is similar to that of Lemma 2. Note that the Bayesian updating implies that for all t the precision w_t^m of the modified problem remains the same as that of the original problem:

$$w_t^m = \frac{\bar{\sigma}_t^{m2}}{\bar{\sigma}_t^{m2} + \sigma^{m2}} = \frac{b^2\bar{\sigma}_t^{\ 2}}{b^2\bar{\sigma}_t^{\ 2} + b^2\sigma^2} = w_t.$$

It follows that the updated posterior mean and variance in the next period become

$$\bar{\mu}_{t+1}^m = w_t^m q_t^m + (1 - w_t^m)\bar{\mu}_t^m = w_t\left(bq_t + c\right) + (1 - w_t)\left(b\bar{\mu}_t + c\right) = b\bar{\mu}_{t+1} + c,$$

$$\bar{\sigma}_{t+1}^m = \bar{\sigma}_t^m \sqrt{1 - w_t^m} = b\bar{\sigma}_t \sqrt{1 - w_t} = b\bar{\sigma}_{t+1}.$$

Therefore, the belief state in the next period preserves the relationship:

$$s_{t+1}^m = \left(\bar{\mu}_{t+1}^m, \bar{\sigma}_{t+1}^m\right) = \left(b\bar{\mu}_{t+1} + c, b\bar{\sigma}_{t+1}\right).$$

For the first part of the lemma we show that the identity in Equation (B8) satisfies the fixed-point relationship implied by the modified problem

(B10)
$$\begin{aligned} &EV^m(s_t^m, \lambda^m) \\ &= \int \max\{\lambda^m + \delta EV^m(s_t^m, \lambda^m),\ \epsilon_t^m + \mathbb{E}[q_t^m | s_t^m] \\ &+ \delta \mathbb{E}[EV^m(s_{t+1}^m, \lambda^m) | s_t^m]\}\, dH^m(\epsilon_t^m). \end{aligned}$$

Suppose Equation (B8) holds, then

(B11)
$$\begin{aligned} v_0^m(s_t^m, \lambda^m) &\triangleq b\lambda + c + \delta EV^m(s_t^m, b\lambda + c) \\ &= b\lambda + \delta bEV(s_t, \lambda) + \frac{c}{1 - \delta} = bv_0(s_t, \lambda) + \frac{c}{1 - \delta}. \end{aligned}$$

Similarly,

(B12) $$v_1^m(s_t^m,\lambda^m) = \mathbb{E}\left[q_t^m|s_t^m\right] + \delta\mathbb{E}[EV^m(s_{t+1}^m,b\lambda+c)|s_t^m]$$

$$= \iint \left(bq_t + c + \delta\left(bEV(s_{t+1},\lambda) + \frac{c}{1-\delta}\right)\right) dF^m(q_t^m|\mu^m,\sigma^m)dB_t^m(\mu^m|s_t^m)$$

$$= \iint \left(bq_t + c + \delta\left(bEV(s_{t+1},\lambda) + \frac{c}{1-\delta}\right)\right) dF(q_t|\mu,\sigma)dB_t(\mu|s_t)$$

$$= b\iint \left(q_t + \delta EV(s_{t+1},\lambda)\right) dF(q_t|\mu,\sigma)dB_t(\mu|s_t) + \frac{c}{1-\delta}$$

$$= bv_1(s_t,\lambda) + \frac{c}{1-\delta}.$$

The first equality uses the fact that $EV^m(s_{t+1}^m, b\lambda + c) = bEV(s_{t+1},\lambda) + c/(1-\delta)$. The second equality follows from normality and conjugate prior assumptions for the distribution of qualities F and beliefs B_t. Then $\Delta^m(s_t^m, b\lambda + c) \triangleq v_0^m(s_t^m, b\lambda + c) - v_1^m(s_t^m, b\lambda + c) = b\Delta(s_t,\lambda)$ for the modified problem. The assumption that the distribution of ϵ has scale and location parameters implies

(B13) $$H^m(\Delta^m) = H(\Delta), \qquad \text{and} \qquad \int_{\epsilon^m \geq \Delta^m} \epsilon^m d\, H^m(\epsilon^m) = b\int_{\epsilon \geq \Delta} (\epsilon + c) d\, H(\epsilon).$$

The right-hand side of Equation (B10) becomes

$$\int_{\epsilon^m \geq \Delta^m} \epsilon^m d\, H^m(\epsilon^m) + v_1^m(s_t^m, b\lambda + c)\left(1 - H^m(\Delta^m)\right) + v_0^m(s_t^m, b\lambda + c)H^m(\Delta^m)$$

$$= b\int_{\epsilon \geq \Delta} \epsilon\, d\, H(\epsilon) + \left(bv_1(s_t,\lambda) + \frac{c}{1-\delta}\right)\left(1 - H(\Delta)\right) + \left(bv_0(s_t,\lambda) + \frac{c}{1-\delta}\right)H(\Delta)$$

$$= b\int_{\epsilon \geq \Delta} \epsilon d\, H(\epsilon) + bv_1(s_t,\lambda)\left(1 - H(\Delta)\right) + bv_0(s_t,\lambda)H(\Delta) + \frac{c}{1-\delta}$$

$$= bEV(s_t,\lambda) + \frac{c}{1-\delta}$$

$$= EV^m(s_t^m, b\lambda + c),$$

which is the left-hand side of Equation (B10). The first equality follows from Equations (B11) to (B13). The third equality uses the fact that $EV(s,\lambda)$ is the fixed point of Equation (A4). Therefore, $EV^m(s_t^m,\lambda^m)$ also satisfies the fixed-point relationship implied by the modified problem.

For the second part of the lemma, we again use the definition of Whittle's index, which is the value of λ^m such that

$$v_0^m(s_t^m,\lambda^m) = \epsilon^m + v_1^m(s_t^m,\lambda^m). \tag{B14}$$

It suffices to show the proposed relation in Equation (B9) solves the above equality. Note that the right-hand side of Equation (B14) is

$$\epsilon^m + v_1^m(s_t^m,\lambda^m) = b\epsilon + bv_1(s_t,\lambda) + \frac{c}{1-\delta} = b\big(\epsilon + v_1(s_t,\lambda)\big) + \frac{c}{1-\delta}$$
$$= bv_0(s_t,\lambda) + \frac{c}{1-\delta},$$

which equals the left-hand side of Equation (B14) following Equation (B11). The first equality follows from Equation (B12). The third equality follows from the definition of Whittle's index for the original problem (setting $W(s,\epsilon)=\lambda$). Then by the definition of Whittle's index, we have $W^m(s^m,b\epsilon;\ b\mu^\epsilon,b\sigma^\epsilon) = bW(s,\epsilon;\ \mu^\epsilon,\sigma^\epsilon)+c$.

To complete the proof of the proposition, note that by Lemma 3 we have:

$$W(b\bar{\mu}+c,b\bar{\sigma},b\epsilon;\ b\sigma,b\mu^\epsilon,b\sigma^\epsilon) = bW(\bar{\mu},\bar{\sigma},\epsilon;\ \sigma,\mu^\epsilon,\sigma^\epsilon)+c.$$

Setting $b=1/\sigma$ and $c=-\bar{\mu}/\sigma$ and evoking Lemma 2 yields

$$W(\bar{\mu}, \bar{\sigma}, \epsilon;\ \sigma, \mu^{\epsilon}, \sigma^{\epsilon}) = \bar{\mu} + \ \sigma W\left(0, \frac{\bar{\sigma}}{\sigma}, \frac{\epsilon}{\sigma};\ 1, \frac{\mu^{\epsilon}}{\sigma}, \frac{\sigma^{\epsilon}}{\sigma}\right)$$

$$= \bar{\mu} + \ \sigma\left[W\left(0, \frac{\bar{\sigma}}{\sigma}, \frac{\epsilon - \mu^{\epsilon}}{\sigma};\ 1, 0, \frac{\sigma^{\epsilon}}{\sigma}\right) + \frac{\mu^{\epsilon}}{\sigma}\right]$$

$$= \bar{\mu} + \mu^{\epsilon} + \ \sigma W\left(0, \frac{\bar{\sigma}}{\sigma}, \frac{\epsilon - \mu^{\epsilon}}{\sigma};\ 1, 0, \frac{\sigma^{\epsilon}}{\sigma}\right),$$

which completes the proof of the proposition.

Online Appendix C. Proof of Proposition 3 (Comparative Statics)

We again focus on the sub-problem of a single brand and thus drop the brand identifier *j*.

Proof of Proposition 3(1). The first part that Whittle's index increases with posterior mean $\bar{\mu}$ is evident from Proposition 2. For the second part, fix some belief state s and consider any $\epsilon' > \epsilon$. Let W' and W be the corresponding Whittle's indices. Recall that $\Delta(s,\lambda) \triangleq v_0(s,\lambda) - v_1(s,\lambda)$. Then by the definition of an index, $\Delta(s,W') = \epsilon' > \epsilon = \Delta(s,W)$. Note that:

$$\Delta_\lambda(s,\lambda) = \frac{\partial\Delta(s,\lambda)}{\partial\lambda} = \frac{\partial v_0(s,\lambda)}{\partial\lambda} - \frac{\partial v_1(s,\lambda)}{\partial\lambda} \geq 0\,, \tag{C1}$$

where the inequality is implied by (A10). It then follows that $W' > W$.

Proof of Proposition 3(2). We will prove the first part. The second part holds following a similar argument. Fix some $\sigma_2^2 > \sigma_1^2$. Let $\{Y_t: t \geq 0\}$ be a sequence of random variables conditionally *i.i.d.* from the distribution $N(\mu, \sigma_1^2)$. Let $\{\omega_t: t \geq 0\}$ be a sequence of random variables conditionally *i.i.d.* from the distribution $N(\mu, \sigma_2^2 - \sigma_1^2)$. The two sequences are independent. Construct a sequence of random variables such that $Z_t = Y_t + \omega_t$ for all t. Then $\{Z_t: t \geq 0\}$ are conditionally *i.i.d.* from the distribution $N(\mu, \sigma_2^2)$. Fix some policy π that solves the problem under $\{Z_t: t \geq 0\}$. Denote $\pi(s_t, \epsilon_t) = 0$ if the fixed reward λ is chosen, and $\pi(s_t, \epsilon_t) = 1$ if the uncertain brand is chosen. Then the value function in state (s, ϵ) when π is applied becomes

$V_{\pi}(s, \epsilon, \lambda; \sigma_2)$

$$= \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \{\mathbb{1}[\pi(s_t, \epsilon_t) = 1](Z_t + \epsilon_t) + \mathbb{1}[\pi(s_t, \epsilon_t) = 0]\lambda\} \mid (s_0, \epsilon_0) = (s, \epsilon)\right]$$

$$= \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \{\mathbb{1}[\pi(s_t, \epsilon_t) = 1](Y_t + \omega_t + \epsilon_t) + \mathbb{1}[\pi(s_t, \epsilon_t) = 0]\lambda\} \mid (s_0, \epsilon_0) = (s, \epsilon)\right]$$

$$= \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \{\mathbb{1}[\pi(s_t, \epsilon_t) = 1](Y_t + \epsilon_t) + \mathbb{1}[\pi(s_t, \epsilon_t) = 0]\lambda\} \mid (s_0, \epsilon_0) = (s, \epsilon)\right]$$

$$+ \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \, \mathbb{1}[\pi(s_t, \epsilon_t) = 1]\omega_t \mid (s_0, \epsilon_0) = (s, \epsilon)\right]$$

$= V_{\pi}(s, \epsilon, \lambda;\ \sigma_1),$

where the last equality uses the fact that the second term is equal to zero. Note that $V_{\pi}(s, \epsilon, \lambda;\ \sigma_1) \ \leq V(s, \epsilon, \lambda;\ \sigma_1)$ because the latter is the optimal value function. Therefore $V_{\pi}(s, \epsilon, \lambda;\ \sigma_2) \ \leq V(s, \epsilon, \lambda; \sigma_1)$ for all π. Taking maximum on the left-hand side gives $V(s, \epsilon, \lambda;\ \sigma_2) \ \leq V(s, \epsilon, \lambda;\ \sigma_1)$. Integrating out ϵ further yields $EV(s, \lambda;\ \sigma_2) \ \leq EV(s, \lambda;\ \sigma_1)$. Then we have $EV_{\sigma}(s, \lambda;\ \sigma) \triangleq \partial EV(s, \lambda;\ \sigma)/\partial\sigma \ \leq 0$ for all s, λ, σ. Differentiating both sides of Equation (A4) with respect to σ and using Chain Rule gives

$$EV_{\sigma}(s_t, \lambda; \sigma) = \frac{\partial EV(s_t, \lambda; \sigma)}{\partial v_0(s_t, \lambda; \sigma)} \frac{\partial v_0(s_t, \lambda; \sigma)}{\partial \sigma} + \frac{\partial EV(s_t, \lambda; \sigma)}{\partial v_1(s_t, \lambda; \sigma)} \frac{\partial v_1(s_t, \lambda; \sigma)}{\partial \sigma}$$

$$= P(0|s_t, \lambda; \sigma)\big(\delta EV_{\sigma}(s_t, \lambda; , \sigma)\big) + \ P(1|s_t, \lambda; , \sigma)(\delta \mathbb{E}[EV_{\sigma}(s_{t+1}, \lambda; , \sigma)|s_t]).$$

The last equality implies that

$$EV_{\sigma}(s_t, \lambda; \sigma) = \frac{\delta P(1|s_t, \lambda; , \sigma)}{1 - \delta P(0|s_t, \lambda; \sigma)} \ \mathbb{E}[EV_{\sigma}(s_{t+1}, \lambda; \sigma)|s_t].$$

It then follows that

$$\frac{\partial\Delta(s_t,\lambda;\sigma)}{\partial\sigma} = \delta(EV_\sigma(s_t,\lambda;\sigma) - \mathbb{E}[EV_\sigma(s_{t+1},\lambda;,\sigma)|s_t])$$

$$= \delta\left(\frac{\delta P(1|s_t,\lambda;\sigma)}{1-\delta P(0|s_t,\lambda;\sigma)} - 1\right)\mathbb{E}[EV_\sigma(s_{t+1},\lambda;\sigma)|s_t]$$

$$= \delta\left(\frac{\delta - 1}{1-\delta P(0|s_t,\lambda;\sigma)}\right)\mathbb{E}[EV_\sigma(s_{t+1},\lambda;\sigma)|s_t] \geq 0,$$

where the last inequality uses the fact that $\mathbb{E}[EV_\sigma(s_{t+1},\lambda;\sigma)|s_t] \leq 0$. Let W_2 and W_1 be the Whittle's indices corresponding to σ_2 and σ_1. This inequality implies $\Delta(s,W_1;\sigma_2) \geq \Delta(s,W_1;\sigma_1)$. Since by the definition of an index, $\Delta(s,W_2;\sigma_2) = \Delta(s,W_1;\sigma_1) = \epsilon$, we have $\Delta(s,W_1;\sigma_2) \geq \Delta(s,W_2;\sigma_2)$. It then follows that $W_2 \leq W_1$ by Equation (C1).

Proof of Proposition 3(3). Consider any $\bar{\sigma}' > \bar{\sigma}$. By the invariance property we have

$$W(\bar{\mu},\bar{\sigma},\epsilon;\ \sigma,\mu^\epsilon,\sigma^\epsilon) = \frac{\bar{\sigma}}{\bar{\sigma}'}\ W\left(\bar{\mu},\bar{\sigma}',\epsilon;\ \frac{\bar{\sigma}'}{\bar{\sigma}}\sigma,\mu^\epsilon,\frac{\bar{\sigma}'}{\bar{\sigma}}\sigma^\epsilon\right)$$

$$\leq \frac{\bar{\sigma}}{\bar{\sigma}'}\ W(\bar{\mu},\bar{\sigma}',\epsilon;\ \sigma,\mu^\epsilon,\ \sigma^\epsilon)$$

$$< W(\bar{\mu},\bar{\sigma}',\epsilon;\ \sigma,\mu^\epsilon,\ \sigma^\epsilon)\,,$$

where the first inequality follows from Proposition 3(2).

Online Appendix D. Computation of the Index Function

We can use the invariance property to simplify the computation of Whittle's index. The computation is based on the fixed point problem in Equation (A4) and the definition of Whittle's index. Product identifier j is dropped to simplify notation. Note that the EV function evaluated at $\bar{\mu}_t = 0$ is $EV(0, \bar{\sigma}_t, \lambda) =$

$$\int \max\{\lambda + \delta EV(0, \bar{\sigma}_t, \lambda),\ \epsilon_t + \delta \mathbb{E}[EV(\bar{\mu}_{t+1}, \bar{\sigma}_{t+1}, \lambda)|0, \bar{\sigma}_t]\}\, dH(\epsilon_t)$$

$$= \int \max\left\{\lambda + \delta EV(0, \bar{\sigma}_t, \lambda),\ \epsilon_t + \delta \mathbb{E}_{q_t}[EV(w_t q_t + (1 - w_t)\bar{\mu}_t, \bar{\sigma}_{t+1}, \lambda)|0, \bar{\sigma}_t]\right\} dH(\epsilon_t)$$

$$= \int \max\left\{\lambda + \delta EV(0, \bar{\sigma}_t, \lambda),\ \epsilon_t + \delta \mathbb{E}_{q_t}\left[EV(0, \bar{\sigma}_{t+1}, \lambda - w_t q_t) + \frac{w_t q_t}{1 - \delta}|0, \bar{\sigma}_t\right]\right\} dH(\epsilon_t)$$

$$= \int \max\left\{\lambda + \delta EV(0, \bar{\sigma}_t, \lambda),\ \epsilon_t + \delta \mathbb{E}_{q_t}[EV(0, \bar{\sigma}_{t+1}, \lambda - w_{t+1} q_t)|0, \bar{\sigma}_t]\right\} dH(\epsilon_t),$$

where $w_t = \bar{\sigma}_t^2/(\bar{\sigma}_t^2 + \sigma^2)$ is the precision. The first equality is implied by Bayesian updating formulae for the normal distribution. The second equality uses Equation (B8) from Lemma 3 and the fact that the expectation of q_t and $\bar{\mu}_t$ conditional on $\bar{\mu}_t = 0$ are both zero. The last equality again uses the zero expectation of q_t conditional on $\bar{\mu}_t = 0$.

We now treat λ as a state variable. Let $\lambda_{t+1} = \lambda_t - w_t q_t$. Note that the distribution of q_t conditional on the belief $(\bar{\mu}_t, \bar{\sigma}_t)$ is normal with mean $\bar{\mu}_t$ and standard deviation $\sqrt{\sigma^2 + \bar{\sigma}_t^2}$. Therefore $\lambda_{t+1}|\lambda_t \sim N(\lambda_t, w_t\sqrt{\sigma^2 + \bar{\sigma}_t^2})$. The fixed point problem now only involves the EV function fixed at $\bar{\mu}_t = 0$, and evolves on the state space $(\bar{\sigma}_t, \lambda_t)$:

$$EV(0, \bar{\sigma}_t, \lambda_t) = \int \max\left\{\lambda_t + \delta EV(0, \bar{\sigma}_t, \lambda_t),\ \epsilon_t + \delta \mathbb{E}_{\lambda_{t+1}}[EV(0, \bar{\sigma}_{t+1}, \lambda_{t+1})|0, \bar{\sigma}_t, \lambda_t]\right\} dH(\epsilon_t).$$

Standard dynamic programming algorithms can be used to solve the above fixed point. Given the solution of $EV(0, \bar{\sigma}_t, \lambda_t)$, we can find the value of λ such that the two terms inside the max operator in the above equation are equal for various values of random shocks ϵ under $\bar{\mu} = 0$. This value is then the corresponding Whittle's index : $W(0, \bar{\sigma}, \epsilon;\ \sigma, \mu^{\epsilon}, \sigma^{\epsilon})$. The index evaluated at any value of posterior mean $\bar{\mu}$ is then computed by linear summation as implied by the invariance property. We present further implementation details in Online Appendix G.2.

Online Appendix E. Computation of the Value Function

Given a decision rule Π, we compute the value function (expected total utilities) by forward simulating utilities for a sufficiently long horizon. Starting at a given state $(\vec{s}_0, \vec{\epsilon}_0) = (\vec{s}, \vec{\epsilon})$, we sample a large number D of Markov chains for each brand $\left\{\left(s_{jk}^{(d)},\ \epsilon_{jk}^{(d)}\right): k = 1, 2, \dots K\right\}_{d=1}^{D}$, where K is greater than the truncated horizon T. Bayesian updating of the normal distribution leads to the following state transition probabilities:

$$\bar{\mu}_{j,k+1}^{(d)} \mid \bar{\mu}_{jk}^{(d)} \sim N\left(\bar{\mu}_{jk}^{(d)}, w_{jk}^{(d)}\sqrt{\sigma_j^2 + \bar{\sigma}_{jk}^{2\,(d)}}\right),\ \ \bar{\sigma}_{j,k+1}^{(d)} \mid \bar{\sigma}_{jk}^{(d)} = \bar{\sigma}_{jk}^{(d)}\sqrt{1 - w_{jk}^{(d)}}\text{, and}$$

$$\epsilon_{jk}^{(d)} \sim H(\epsilon;\ \mu^{\epsilon}, \sigma^{\epsilon}),\text{ where } w_{jk}^{(d)} = \bar{\sigma}_{jk}^{2\,(d)} / (\sigma_j^2 + \bar{\sigma}_{jk}^{2\,(d)}).$$

These sequences of belief states are then fixed in advance and reused for each decision rule. Under a decision rule Π, the empirical estimate of its expected total utility for a truncated horizon T is given by

$$\mathbb{V}_{\Pi}(\vec{s}, \vec{\epsilon}) = \frac{1}{D}\sum_{d=1}^{D}\left\{\sum_{t=0}^{T-1}\delta^t \sum_{j=1}^{J} \mathbb{1}\left[\Pi\left(\vec{s}_t^{(d)}, \vec{\epsilon}_t^{(d)}\right) = j\right]\left(\bar{\mu}_{j,n_{jt}}^{(d)} + \epsilon_{jt}^{(d)}\right)\middle|\, (\vec{s}_0, \vec{\epsilon}_0) = (\vec{s}, \vec{\epsilon})\right\},$$

where n_{jt} is the cumulative number of trials for brand j up to period t. Note that the realized state values are chosen from the pre-drawn sample paths, with n_{jt} indicating which state in the sample path is chosen.

Online Appendix F. Maximum Simulated Likelihood Estimation

We estimate each model's parameters with maximum simulated likelihood estimation. To simplify notation let $\vec{\alpha}$ denote the vector of parameters to be estimated. Let $d_{it} \in A$ denote household i's decision at period t and let $\vec{d}_i^t = \{d_{ik}\}_{k=1}^t$ denote i's decision sequence up to period t. The likelihood of observing the choice sequences as a function of $\vec{\alpha}$ is:

$$L(\vec{\alpha}) = \prod_{i=1}^{I} L_i(\vec{\alpha}) = \prod_{i=1}^{I} \Pr\left(\vec{d}_i^{T_i}; \vec{\alpha}\right).$$

Learning strategies depend upon the evolution of the unobserved belief states, which complicates the inference process. If we were to write the likelihood function as a function of each consumer's unobserved belief states and shocks over periods, we would need to sample from an extremely complicated joint density of belief states and shocks. Instead, we augment the data and sample directly from the more-fundamental unobservables – the quality experiences q_{ijt} that are drawn conditionally *i.i.d.* from normal distributions with mean μ_j and standard deviation σ_j. Given a set of quality experiences and a set of prior beliefs, we obtain the unobserved belief states, $s_{ijt} = (\bar{\mu}_{ijt}, \bar{\sigma}_{ijt})$, by conjugate updating formulae:

$$\bar{\mu}_{ijt} = \frac{\bar{\sigma}_{j0}^2}{\bar{\sigma}_{j0}^2 + \sigma_j^2/n_{ijt}} \bar{q}_{ijt} + \frac{\sigma_j^2/n_{ijt}}{\bar{\sigma}_{j0}^2 + \sigma_j^2/n_{ijt}} \bar{\mu}_{j0}, \quad \text{and} \quad \bar{\sigma}_{ijt}^2 = \frac{\bar{\sigma}_{j0}^2 \sigma_j^2/n_{ijt}}{\bar{\sigma}_{j0}^2 + \sigma_j^2/n_{ijt}},$$

where n_{ijt} is the cumulative number of purchases of brand j by consumer i through period t. We use $\bar{q}_{ijt} = \frac{1}{n_{ijt}} \sum_{k=1}^{n_{ijt}} q_{ijk}$ to denote the average quality experience observed by the consumer through period t. Note that the posterior variance $\bar{\sigma}_{ijt}^2$ decreases over time towards zero as more information is incorporated and the speed of convergence depends on the prior and the true variance. When the prior $\bar{\sigma}_{j0}^2$ is much larger than the true variance σ_j^2, then after just one update, the posterior quickly converges to σ_j^2.[1]

We introduce vector notation to simplify exposition. Let $\vec{\mu}$ be the vector (over j) of mean qualities, let $\vec{\sigma}$ be the vector (over j) of the standard deviations of quality draws, and let $\vec{\sigma}^{\epsilon}$ be the vector (over j) of the standard deviations of unobservable shocks. Let the sequence of quality draws through period t be $\vec{q}_i^{\,t} = \{\vec{q}_{ik}\}_{k=1}^{t}$. Let $\vec{x}_t$ and $\vec{\epsilon}_{it}$ be the vectors (over j) of prices and unobservable utility shocks. Finally, let $f_q(\cdot)$ and $p_\epsilon(\cdot)$ be the probability density functions for the quality draws and the unobservable shocks. The likelihood for household i is given by

$$\text{(F1)}\quad L_i(\vec{\alpha}) = \int\int \prod_{t=1}^{T_i} \mathbb{1}\left\{\Pi\left[s_{it}\left(\vec{d}_i^{\,t-1}, \vec{q}_i^{\,t-1}\right), \vec{x}_t, \vec{\epsilon}_{it}\right] = d_{it}\middle|\vec{q}_i^{\,T_i}; \vec{\alpha}\right\} f_q\left(\vec{q}_i^{\,T_i}; \vec{\mu}, \vec{\sigma}\right) p_\epsilon(\vec{\epsilon}_{it}) d\vec{q}_i^{\,T_i}\, d\vec{\epsilon}_{it}\,.$$

To compute the likelihood we integrate over quality draws and unobservable shocks. To integrate numerically we sample R sequences of quality draws (each sequence has T_i draws for consumer i) from a multivariate normal distribution with parameters $\vec{\mu}$ and $\vec{\sigma}$. We assume that the unobservable shocks follow zero-mean Gumbel distribution with homogenous variance σ^{ϵ} for all brands. This assumption allows us to use the well-known logit formula to substantially simplify the computation of choice probabilities for all models. Based on Proposition 2, we specify the index function as a linear function of the unobserved shocks ϵ_{ijt}, while preserving monotonicity:

$$\widehat{\Pi}_W = argmax_j\left\{\bar{\mu}_{ijt} + \mu_{ij}^{x,\epsilon} + \epsilon_{ijt} + \sigma_{ij}\breve{W}_j\left(0, \frac{\bar{\sigma}_{ijt}}{\sigma_{ij}}, \frac{\beta x_{ijt} - \mu_{ij}^{x,\epsilon}}{\sigma_{ij}}, 1, 0, \frac{\sigma_{ij}^{x,\epsilon}}{\sigma_{ij}}, \delta\right)\right\}.$$

We provide further implementation details of the estimation procedures in Online Appendix G.

[1] This introduces difficulty in estimating the variance of prior quality beliefs. For example, we encounter difficulty when estimating the mixture model. The likelihood function is flat over the regions where the values of σ_j^2 are large. As a solution, we terminate the iterations of estimation when the likelihood improves by less than 0.02%. We expect the same difficulty to apply to the approximately optimal solution.

Online Appendix G. User's Guide to Implementation

This user's guide documents the implementation details of the optimization solutions and the estimation procedures. §§G.1 and G.2 provide the details of solving the single-agent problem via the approximately optimal solution and the index solution using discrete approximation. §§G.3 and G.4 summarize the details of estimating the approximately optimal solution and the index solution.

G.1. Computing the Approximately Optimal Solution

G.1.1. Overview

Recall that the goal is to solve the following Bellman equation of the overall problem:

$$V(\vec{s}_t, \vec{x}_t, \vec{\epsilon}_t) = \max_{j \in A} \{\vec{\beta}' \vec{x}_{jt} + \epsilon_{jt} + \mathbb{E}[q_{jt} + \delta V(\vec{s}_{t+1}, \vec{x}_{t+1}, \vec{\epsilon}_{t+1}) | \vec{s}_t, j]\}.$$

Under the assumption that unobservable shocks ϵ_{jt} are *i.i.d.*, we can integrate out this component and transform the problem to (Rust 1994):

$$\text{(G1)} \qquad EV(\vec{s}_t, \vec{x}_t) = \int_{\vec{\epsilon}_t} \max_{j \in A} \{\vec{\beta}' \vec{x}_{jt} + \epsilon_{jt} + \mathbb{E}[q_{jt} + \delta EV(\vec{s}_{t+1}, \vec{x}_{t+1}) | \vec{s}_t, j]\} dH(\vec{\epsilon}_t).$$

Further assumption on the distribution of $\vec{\epsilon}_t$ can simplify the above integration. If ϵ_{jt} follows *i.i.d.* Gumbel distribution, then we have a closed-form expression (Rust 1994):

$$\text{(G2)} \qquad EV(\vec{s}_t, \vec{x}_t) = \mu^\epsilon + \sigma^\epsilon \gamma + \sigma^\epsilon \log \left\{ \sum_{j=1}^{J} \exp\left(\frac{\vec{\beta}' \vec{x}_{jt} + \mathbb{E}[q_{jt} + \delta EV(\vec{s}_{t+1}, \vec{x}_{t+1}) | \vec{s}_t, j]}{\sigma^\epsilon} \right) \right\},$$

where γ is the Euler constant. Notice that for this simplification to hold, we need to assume the distribution of the unobservable shocks is the same for all brands: $(\mu_j^\epsilon, \sigma_j^\epsilon) = (\mu^\epsilon, \sigma^\epsilon)$ for all *j*.

The observable shocks $\vec{x}_t$ remain in the state space. One can also integrate out $\vec{x}_t$ given the independence assumption, so that Equation (G1) becomes

(G3) $$EV(\vec{s}_t) = \int_{(\vec{x}_t,\vec{\epsilon}_t)} \max_{j\in A}\{\vec{\beta}'\vec{x}_{jt} + \epsilon_{jt} + \mathbb{E}[q_{jt} + \delta EV(\vec{s}_{t+1})|\vec{s}_t, j]\}dH(\vec{x}_t, \vec{\epsilon}_t).$$

We solve Equation (G3) with simulation by integrating over the joint distribution of $(\vec{\beta}'\vec{x}_{jt} + \epsilon_{jt})$.[2] The modified Bellman equation still cannot be solved exactly because the state space of each brand $s_j = (\bar{\mu}_j, \bar{\sigma}_j)$ is continuous. There are many algorithms to approximate the solution (see a survey by Rust 1996 for methods to solve continuous-state Markov decision processes). We will use a "discrete approximation" approach that first discretizes the state space, then solves the discrete-state dynamic programming problem, and finally finds the value function of the continuous-state problem by aggregating the discrete-state solution using interpolation. While discrete approximation may be slower than "smooth approximation" (e.g., the Keane-Wolpin 1994 algorithm adopted by Erdem and Keane 1996; see Ching et al. 2013b for details of its implementation), it can fully preserve the contraction property of the Bellman equation, and is guaranteed to converge to the true solution as the discretization becomes finer (e.g., Chow and Tsitsiklis 1991). The absolute computation time depends on various factors such as the algorithm, computer memory, software package, coding, etc. In this paper we are interested in comparing the relative computation time of the two solution concepts, the approximately optimal solution and the index solution, using the same discrete approximation approach to solve the fixed-point problem in both solutions.

G.1.2. State Space and Transition Probabilities

We discretize the state space into a finite collection of state points $\mathbb{D}(\bar{\mu}_j;\ M) =$

[2] Alternatively, one can assume that the distribution of the sum $(\vec{\beta}'\vec{x}_{jt} + \epsilon_{jt})$ is Gumbel and is homogenous, which leads to a value function similar to Equation (G3). However, this assumption is not appealing in this setting because (1) we need ϵ_{jt} alone to be Gumbel to obtain a simple logit expression of choice probabilities, and (2) the uncertainty in utility shocks may vary across brands.

$\left\{\hat{\bar{\mu}}_j^{(m)}\right\}_{m=1}^{M}$ and $\mathbb{D}(\bar{\sigma}_j; N) = \left\{\hat{\bar{\sigma}}_j^{(n)}\right\}_{n=1}^{N}$ for each brand. Given the discretized state space, we convert the transition probabilities from the continuous problem to the discrete problem. Recall that the transitions in the continuous problem are given as follows. Suppose the consumer's current belief about brand *j* at period *t* is $(\bar{\mu}_{jt}, \bar{\sigma}_{jt})$. If the consumer chooses brand *j*, then the consumer's beliefs about brand *j* in the next period become

(G4) $$\bar{\mu}_{j,t+1} | \bar{\mu}_{jt} \sim f\left(\bar{\mu}_{jt}, w_{jt}\sqrt{\sigma_j^2 + \bar{\sigma}_{jt}^2}\right) \quad \text{and} \quad \bar{\sigma}_{j,t+1} | \bar{\sigma}_{jt} = \bar{\sigma}_{jt}\sqrt{1 - w_{jt}}\,,$$

where $f(\cdot)$ is the normal density, and $w_{jt} = \bar{\sigma}_{jt}^2 / (\sigma_j^2 + \bar{\sigma}_{jt}^2)$. If the consumer does not choose brand *j*, then the consumer's beliefs about *j* remain the same in the next period $(\bar{\mu}_{j,t+1}, \bar{\sigma}_{j,t+1}) = (\bar{\mu}_{jt}, \bar{\sigma}_{jt})$.

The transition of the variance of belief $\bar{\sigma}_{jt}$ is deterministic. Given the variance of prior belief $\bar{\sigma}_{j0}$, we know the exact values that the future $\bar{\sigma}_{jt}$ would fall in. This pins down the discretization of $\bar{\sigma}_{jt}$. We can set $\hat{\bar{\sigma}}_j^{(1)} = \bar{\sigma}_{j0}$ and $\hat{\bar{\sigma}}_j^{(2)} = \bar{\sigma}_{j1}$, and so on: $\hat{\bar{\sigma}}_j^{(N)} = \bar{\sigma}_{j,N-1}$. Note that for large values of N, $\hat{\bar{\sigma}}_j^{(N)}$ will be close to zero.

The transition of the posterior mean of belief $\bar{\mu}_{jt}$ is probabilistic on the entire unbounded continuous space. We choose a bound $[-B, B]$ and discretize it uniformly into M points. The size B is chosen to be large enough so that the states outside the bounds are rarely visited. We set B to be 5 deviation from the mean of the conditional distribution. We then define the transition probabilities on the discretized state space from one state point $(\hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)})$ to another $(\hat{\bar{\mu}}_j^{(m')}, \hat{\bar{\sigma}}_j^{(n')})$ as $p(\hat{\bar{\mu}}_j^{(m')}, \hat{\bar{\sigma}}_j^{(n')} \,|\, \hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)})$. If the consumer does not choose brand *j*, then $(\hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)})$ transits to $(\hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)})$ with probability one. If the consumer chooses brand *j*, then the transition is as follows. First, the variance of posterior belief $\hat{\hat{\sigma}}_j^{(n)}$ transits to $\hat{\hat{\sigma}}_j^{(n')}$ deterministically with $n' = n + 1$ for $n < N$. Note that the difference between $\hat{\hat{\sigma}}_j^{(n)}$ and $\hat{\hat{\sigma}}_j^{(n+1)}$ converges to zero as *n* becomes larger. Therefore, we will set $N' = N$ such that $\hat{\hat{\sigma}}_j^{(N)}$ transits to itself. Second, the mean of posterior belief $\hat{\bar{\mu}}_j^{(m)} \in \mathbb{D}(\bar{\mu}_j; M)$ transits to any state point $\bar{\mu}_j^{(m')} \in \mathbb{D}(\bar{\mu}_j; M)$ with normalized probability:

(G5) $$p(\hat{\bar{\mu}}_j^{(m')} | \hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)}) = \frac{f(\hat{\bar{\mu}}_j^{(m')} | \hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)})}{\sum_{m'=1}^{M} f(\hat{\bar{\mu}}_j^{(m')} | \hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)})},$$

where $f(\hat{\bar{\mu}}_j^{(m')} | \hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)})$ is the density conditional on $(\hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)})$ defined in the continuous-state problem (see Equation G4).

G.1.3. Algorithm

After obtaining the transition probabilities for the discrete problem we can then solve the discrete-state dynamic programming problem using any standard dynamic programming algorithm. Here we use value iteration, which is easy to implement albeit not particularly fast. One can adopt the multi-grid approach (Chow and Tsitsiklis 1991) or the random-grid approach (Rust 1996) to speed up the algorithm. The value iteration algorithm proceeds as follows:

Step 1: Initialize the $(M \times N)^J$ matrix $\widehat{EV}_0\left(\vec{\hat{\bar{\mu}}}, \vec{\hat{\bar{\sigma}}}\right)$.

Step 2: Iterate the modified Bellman equation until $\left\|\widehat{EV}_{k+1} - \widehat{EV}_k\right\| < Tolerance$:

$$\widehat{EV}_{k+1}\left(\vec{\hat{\bar{\mu}}}, \vec{\hat{\bar{\sigma}}}\right) = \int_{(\vec{x},\vec{\epsilon})} \max_{j \in A} \left\{ \vec{\beta}'\vec{x}_j + \epsilon_j + \hat{\bar{\mu}}_j^{(m)} \right.$$
$$\left. + \delta \sum_{m'=1}^{M} \widehat{EV}_k\left(\hat{\bar{\mu}}_j^{(m')}, \hat{\bar{\sigma}}_j^{(n')}\right) p(\hat{\bar{\mu}}_j^{(m')} | \hat{\bar{\mu}}_j^{(m)}, \hat{\bar{\sigma}}_j^{(n)}) \right\} dH(\vec{x}, \vec{\epsilon}),$$

In Step 2, numerical integration is needed. We use direct simulation to integrate the function over the joint distribution of utility shocks. Once we have found the solution $\widehat{EV}$ to the discrete problem, we can aggregate it to produce the solution EV defined on the entire continuous state space using interpolation. In this paper, we use linear interpolation.

G.2. Computing the Index Solution

G.2.1. Overview

Recall that in Appendix D we have shown that the index function reduces to solving the following modified Bellman equation (after integrating out both $\vec{x}_j$ and ϵ_j):

$$EV(0,\bar{\sigma}_{jt},\lambda_{jt}) = \int_{(\vec{x}_{jt},\epsilon_{jt})} \max\Big\{\lambda_{jt}+\delta EV(0,\bar{\sigma}_{jt},\lambda_{jt}),\ \vec{\beta}'\vec{x}_{jt}+\epsilon_{jt} + \delta\mathbb{E}_{\lambda_{j,t+1}}\big[EV(0,\bar{\sigma}_{j,t+1},\lambda_{j,t+1})|0,\bar{\sigma}_{jt},\lambda_{jt}\big]\Big\}\, dH(\vec{x}_{jt},\epsilon_{jt}).$$

where the *EV* function is fixed at $\bar{\mu}_{jt}=0$, and λ_{jt} is treated as a state variable with the following transition probabilities:

(G6)
$$\lambda_{j,t+1}|\lambda_{jt} \sim Normal\left(\lambda_{jt},\ w_{jt}\sqrt{\sigma_j^2+\bar{\sigma}_{jt}^2}\right),\ \text{ and }\ \bar{\sigma}_{j,t+1}|\,\bar{\sigma}_{jt} = \bar{\sigma}_{jt}\sqrt{1-w_{jt}}\,,$$
where $w_{jt} = \bar{\sigma}_{jt}^2/(\sigma_j^2+\bar{\sigma}_{jt}^2)$.

We will remove 0 from the *EV* function for notational convenience, with a slight abuse of notation of the *EV* function:

(G7)
$$EV(\bar{\sigma}_{jt},\lambda_{jt}) = \int_{(\vec{x}_{jt},\epsilon_{jt})} \max\Big\{\lambda_{jt}+\delta EV(\bar{\sigma}_{jt},\lambda_{jt}),\ \vec{\beta}'\vec{x}_{jt}+\epsilon_{jt} + \delta\mathbb{E}_{\lambda_{j,t+1}}\big[EV(\bar{\sigma}_{j,t+1},\lambda_{j,t+1})|\bar{\sigma}_{jt},\lambda_{jt}\big]\Big\}\, dH(\vec{x}_{jt},\epsilon_{jt}).$$

The integration over the distribution of utility shocks ($\vec{\beta}'\vec{x}_{jt}+\epsilon_{jt}$) in general has no closed-form expression. We can, however, compute its value given any distributional assumptions using simulation or Gaussian quadrature. For example, we can assume that $\vec{x}_{jt}$ is normal and ϵ_{jt} is Gumbel, or the sum $\vec{\beta}'\vec{x}_{jt} + \epsilon_{jt}$ is normal or Gumbel. The computation is relatively easy given this is a one-dimensional problem (i.e., involving one brand). For the synthetic-data analysis, we assume that $\vec{x}_{jt}$ is 0 and ϵ_{jt} is Gumbel. For the field-data analysis on diaper purchases, we assume that $\vec{x}_{jt}$ is normal and ϵ_{jt} is Gumbel. These assumptions are made simply to ease the computation of the approximately optimal solution, to which the index solution is compared, rendering ours a conservative test of the relative simplicity of the index strategy.

The procedure of computing the index function involves two stages. The first stage is to find the solution to the *EV* function from Equation (G7). We will use discrete approximation, the same method used for the approximately optimal solution. The second stage is to find the value of λ such that the two terms inside the max operator in Equation (G7) are equal. This λ value then equals Whittle's index evaluated at $\bar{\mu}_{jt} = 0$ and for a particular variance $\bar{\sigma}_{jt}$ and utility shocks $\vec{\beta}'\vec{x}_{jt} + \epsilon_{jt}$, that is, $W\left(0, \bar{\sigma}_{jt}, \vec{\beta}'\vec{x}_{jt} + \epsilon_{jt}; \sigma_j, \mu_j^{x,\epsilon}, \sigma_j^{x,\epsilon}\right)$.

G.2.1. Algorithm

The discretization of the state space is the same as that for the approximately optimal solution described in §G.1.2, except that we now treat λ as the state variable rather than $\bar{\mu}_{jt}$.

Step 1: Initialize $\widehat{EV}_0\left(\hat{\bar{\sigma}}_j^{(n)}, \hat{\lambda}_j^{(m)}\right)$, $\forall\, m, n$.

Step 2: Iterate the modified Bellman equation until $\left\|\widehat{EV}_{k+1} - \widehat{EV}_k\right\|$ < *Tolerance*:

$$\widehat{EV}_{k+1}\left(\hat{\bar{\sigma}}_j^{(n)}, \hat{\lambda}_j^{(m)}\right)$$

$$= \int_{(\vec{x}_j,\epsilon_j)} \max\left\{\hat{\lambda}_j^{(m)} + \delta\widehat{EV}_k\left(\hat{\bar{\sigma}}_j^{(n)}, \hat{\lambda}_j^{(m)}\right), \vec{\beta}'\vec{x}_j + \epsilon_j \right.$$

$$\left. + \delta \sum_{m=1}^{M} \widehat{EV}_k\left(\hat{\bar{\sigma}}_j^{(n')}, \hat{\lambda}_j^{(m')}\right) p(\hat{\lambda}_j^{(m')} \mid \hat{\bar{\sigma}}_j^{(n)}, \hat{\lambda}_j^{(m)}) \right\} dH(\vec{x}_j, \epsilon_j).$$

Step 3: Obtain $\widehat{EV}^*\left(\hat{\bar{\sigma}}_j^{(n)}, \hat{\lambda}_j^{(m)}\right)$ after Step 2. For every n and value of $\vec{\beta}'\vec{x}_j + \epsilon_j$, find the root λ^* such that the difference $\Delta(\lambda) = 0$, where

$$\Delta(\lambda) = \hat{\lambda}_j^{(m)} + \delta\widehat{EV}^*\left(\hat{\bar{\sigma}}_j^{(n)}, \hat{\lambda}_j^{(m)}\right) - \vec{\beta}'\vec{x}_j - \epsilon_j$$

$$- \delta \sum_{m=1}^{M} \widehat{EV}_k\left(\hat{\bar{\sigma}}_j^{(n')}, \hat{\lambda}_j^{(m')}\right) p(\hat{\lambda}_j^{(m')} \mid \hat{\bar{\sigma}}_j^{(n)}, \hat{\lambda}_j^{(m)}) .$$

Note that the difference $\Delta(\lambda)$ increases with λ. One easy way is to locate the $(\hat{\lambda}^{(a)}, \hat{\lambda}^{(b)}) \in \left\{\hat{\lambda}_j^{(m)}\right\}_{m=1}^{M}$ such that $\Delta\left(\hat{\lambda}_j^{(a)}\right) > 0$ and $\Delta\left(\hat{\lambda}_j^{(b)}\right) < 0$ and then use linear interpolation to find λ_j^* such that $\Delta\left(\lambda_j^*\right) = 0$. λ_j^* is then Whittle's index for brand j, $\widehat{W}_j$, evaluated at $\hat{\bar{\sigma}}_j^{(n)}$ and $\vec{\beta}'\vec{x}_j + \epsilon_j$.

Once we have the $\widehat{EV}$ function and Whittle's index $\widehat{W}_j$ defined on the discrete state space, we can use interpolation to find values of EV and W_j over the entire continuous state space.

G.3. Estimating the Approximately Optimal Solution

We estimate the model using maximum simulated likelihood. For the approximately optimal solution, the likelihood function in Equation (F1) can be expressed as:

$$L_i(\vec{\alpha}) =$$

$$\int_{\vec{q}_i^{T_i}} \left(\prod_{t=1}^{T_i} \frac{\exp\left(\beta x_{d_{it},t} + \bar{\mu}_{id_{it},t}\left(\vec{d}_i^{t-1}, \vec{q}_i^{t-1}\right) + \delta \mathbb{E}\left[EV\left(\bar{\mu}_{id_{it},t+1}, \bar{\sigma}_{id_{it},t+1}\right) \middle| \bar{\mu}_{id_{it},t}, \bar{\sigma}_{id_{it},t}, d_{it}\right]\right)}{\sum_{j=1}^{J} \exp\left(\beta x_{jt} + \bar{\mu}_{ijt}\left(\vec{d}_i^{t-1}, \vec{q}_i^{t-1}\right) + \delta \mathbb{E}\left[EV\left(\bar{\mu}_{ij,t+1}, \bar{\sigma}_{ij,t+1}\right) \middle| \bar{\mu}_{ijt}, \bar{\sigma}_{ij,t}, j\right]\right)} \right)$$

$f_q\left(\vec{q}_i^{\;T_i}; \vec{\mu}, \vec{\sigma}\right) d\vec{q}_i^{\;T_i}$,

where the last equation uses the assumption that ϵ_{ijt} follows the *i.i.d.* Gumbel distribution. The outer integration is computed by simulating R sequences of quality signals $\left\{\vec{q}_i^{\;T_i(r)}\right\}_{r=1}^{R}$ from a multivariate normal distribution of true quality $f_q(\cdot)$ parameterized by $(\vec{\mu}, \vec{\sigma})$.

To find the maximum simulated likelihood estimates $\vec{\alpha}^* = argmax \prod_{i=1}^{I} L_i(\vec{\alpha})$, we adopt the nested fixed point algorithm (Rust 1994):

- In the inner loop, for each guess of parameters $\vec{\alpha}$, solve for the EV function based on Equation (G3) using the procedure in §G.1 and then evaluate the likelihood function.
- In the outer loop, find the parameters $\vec{\alpha}^*$ that maximize the likelihood value.

The inner loop is computationally intense. For each parameter guess, we obtain an EV function, which is then used to initialize the EV function for the next parameter guess. This allows for faster convergence.

Another important issue is the choice of the size of discretization (i.e., the values of M and N). The larger these numbers, the greater accuracy we obtain. But computation memory and time will increase exponentially as well. If we choose $M = N = 10$ then the size of the state space is $(M \times N)^4 = 10^8$. If $M = N = 100$ the size is 10^{16}. In the estimation, we set $M = N = 5$ which yields a state space of size $5^8 = 390{,}635$.

G.4. Estimating the Index Strategy

The estimation procedure is similar to the one for the approximately optimal solution. Recall that the index rule is defined as

(G8) $$\Pi_W = argmax_j \left\{ \bar{\mu}_{ijt} + \mu_j^{x,\epsilon} + \sigma_j \breve{W}_j \left(0, \frac{\bar{\sigma}_{ijt}}{\sigma_j}, \frac{\beta x_{jt} + \epsilon_{ijt} - \mu_j^{x,\epsilon}}{\sigma_j}, 1, 0, \frac{\sigma_j^{x,\epsilon}}{\sigma_j}, \delta \right) \right\}.$$

We rewrite the index as a linear function of ϵ_{ijt} :

(G9) $$\widehat{\Pi}_W = argmax_j \left\{ \bar{\mu}_{ijt} + \beta x_{jt} + \epsilon_{ijt} + \sigma_j \breve{W}_j \left(0, \frac{\bar{\sigma}_{ijt}}{\sigma_j}, 0, 1, 0, \frac{\sigma_j^{x,\epsilon}}{\sigma_j}, \delta \right) \right\}.$$

This transformed function preserves all the properties of Whittle's index and simplifies the computation of choice probabilities.

Assuming ϵ_{ijt} follows *i.i.d.* Gumbel, the likelihood in Equation (F1) for household i is

$$L_i(\vec{\alpha})$$

$$= \int_{\vec{q}_i^{T_i}} \left(\prod_{t=1}^{T_i} \frac{\exp\left(\beta x_{d_{it},t} + \bar{\mu}_{id_{it},t}\left(\vec{d}_i^{t-1}, \vec{q}_i^{t-1}\right) + \sigma_j \breve{W}_j \left(0, \frac{\bar{\sigma}_{id_{it}t}}{\sigma_{d_{it}}}, 0, 1, 0, \frac{\sigma_{d_{it}}^{x,\epsilon}}{\sigma_{d_{it}}}, \delta \right) \right)}{\sum_{j=1}^{J} \exp\left(\beta x_{jt} + \bar{\mu}_{ijt}\left(\vec{d}_i^{t-1}, \vec{q}_i^{t-1}\right) + \sigma_j \breve{W}_j \left(0, \frac{\bar{\sigma}_{ijt}}{\sigma_j}, 0, 1, 0, \frac{\sigma_j^{x,\epsilon}}{\sigma_j}, \delta \right) \right)} \right) f_q\left(\vec{q}_i^{T_i}; \vec{\mu}, \vec{\sigma} \right) d\vec{q}_i^{T_i}.$$

We again use the nested fixed point algorithm to find the maximum simulated likelihood estimates:

- In the inner loop, for each guess of parameters $\vec{\alpha}$, first compute Whittle's index $\breve{W}_j$ using procedures described in §G.2 with utility shocks normalized to zero. Then evaluate the likelihood given the index function.
- In the outer loop, find the parameters $\vec{\alpha}^*$ that maximize the likelihood value.

We need to determine the number of grid points for the posterior mean and the posterior variance (i.e., the values of M and N). For an "apples-to-apples" comparison between the approximately optimal solution and the index strategy, we set $M = 5$ and $N = 5$ for both models. The state space is $(5 \cdot 5)^4 = 390{,}625$ for the approximately optimal solution and $5 \cdot 5 = 25$ for the index strategy, a ratio of $15{,}625$. (We solve for $J = 4$ indices.)

Because the size of the state space does not grow exponentially with the number of brands under the index strategy, it is feasible to choose finer grids for the index strategy than for the approximately optimal solution. Thus, for greater accuracy we set $M = 200$ and set N to the maximum number of repeat purchases of households in the sample, which is 75. The size of the state space for the index strategy is $200 \cdot 75 = 15{,}000$. If we were to attempt this finer grid for the approximately optimal solution, the state space would be $(200 \cdot 75)^4 = 50{,}625{,}000{,}000{,}000{,}000$. This is about 130 billion times the approximately optimal solution's state space under the original grid of $M = N = 5$. It is unlikely that computations would be feasible with such a large state space, even with a more efficient search of the grid.

Online Appendix H. Predicted Switching Matrices

Table H presents the predicted switching matrices among diaper brands, using the parameter estimates from the (a) no-learning model, (b) myopic learning model, and (c) approximately optimal solution model, respectively. The predicted switching matrix based on the parameter estimates from the index strategy model is presented in Table 2b.

Table H. Switching among Diaper Brands

(a) Predicted Switching Matrix – No-Learning Model

	Percent of times that Row Brand is purchased at Occasion t and Column Brand is purchased at Occasion $t+1$[†]			
	Pampers	Huggies	Luvs	Other Brands
Within the first 13 purchases				
Pampers	17.5%	4.4%	4.7%	0.0%
Huggies	8.1%	19.8%	3.6%	0.0%
Luvs	4.4%	2.9%	9.2%	0.0%
Other Brands	n/a[‡]	n/a	n/a	n/a
After the first 13 purchases				
Pampers	5.3%	1.7%	1.6%	0.0%
Huggies	2.5%	7.5%	1.0%	0.0%
Luvs	1.2%	0.7%	3.7%	0.0%
Other Brands	n/a	n/a	n/a	n/a

[†] Switching percentages are weighted by market share so that the percentages in the same table add up to 100%.
[‡] Switching probability not applicable because the model predicts no purchase of Other Brands.

Table H. Switching among Diaper Brands (continued)

(b) Predicted Switching Matrix – Myopic Learning Model

	Percent of times that Row Brand is purchased at Occasion t and Column Brand is purchased at Occasion $t+1$[†]			
	Pampers	Huggies	Luvs	Other Brands
Within the first 13 purchases				
Pampers	23.3%	2.6%	1.6%	0.1%
Huggies	1.8%	23.9%	1.1%	0.1%
Luvs	1.7%	1.2%	15.7%	0.1%
Other Brands	0.3%	0.1%	0.1%	1.0%
After the first 13 purchases				
Pampers	6.2%	0.6%	0.5%	0.0%
Huggies	0.3%	12.4%	0.1%	0.0%
Luvs	0.5%	0.2%	4.3%	0.0%
Other Brands	0.0%	0.0%	0.0%	0.1%

(c) Predicted Switching Matrix – Approximately Optimal Solution Model

	Percent of times that Row Brand is purchased at Occasion t and Column Brand is purchased at Occasion $t+1$[†]			
	Pampers	Huggies	Luvs	Other Brands
Within the first 13 purchases				
Pampers	20.2%	3.0%	2.0%	0.2%
Huggies	2.0%	25.9%	1.4%	0.1%
Luvs	1.5%	1.6%	15.3%	0.1%
Other Brands	0.3%	0.1%	0.1%	1.0%
After the first 13 purchases				
Pampers	4.6%	0.4%	0.4%	0.0%
Huggies	0.3%	15.0%	0.1%	0.0%
Luvs	0.5%	0.2%	3.4%	0.0%
Other Brands	0.0%	0.0%	0.0%	0.1%