Erratum to “Consumer Choice Under Limited Attention When Alternatives Have Different Costs”

Frank Huettner
SKK, GSB Sungkyunkwan University
25-2 Sungkyunkwan-ro, Jongro-gu, Seoul, 03063, Korea
huettner@skku.edu

Tamer Boyacı
ESMT Berlin,
Schlossplatz 1, 10178 Berlin, Germany
tamer.boyaci@esmt.org

Yalçın Akçay
Melbourne Business School, University of Melbourne
200 Leicester Street, Carlton VIC 3053, Australia
y.akcay@mbs.edu

There is an error in one of the results of our paper (Huettner et al. 2019, Oper. Res., 67(3), 671-699). In this erratum, we point the error out and provide a correction based on Walker-Jones (2023, J. Econ. Theory 212, 105688). Our key characterizations, insights and numerical examples do not depend on this error and hence remain valid. The main implication is on the stopping condition used in the algorithm. We propose a fix based on the new sufficient condition and if needed standard convex optimization techniques.

Key words: discrete choice, rational inattention, information acquisition, non-uniform information costs, market inference.

---

1. Background

It has been brought to our attention by David Walker-Jones, of University of Surrey, that there is an error in Theorem 2 of our paper Huettner et al. (2019).

Recall the following model preliminaries and notation. The consumer is learning about the state of the world $\boldsymbol{\omega} \in \boldsymbol{\Omega} = \Omega_1 \times \ldots \times \Omega_k \times \ldots \times \Omega_n$, where each component ω_k can be learned at a different learning cost, λ_k, where $\lambda_1 \leqslant \ldots \leqslant \lambda_k \leqslant \ldots \leqslant \lambda_n$. If the consumer chooses alternative $i \in A$ in state $\boldsymbol{\omega}$, then this gives the utility $u(i, \boldsymbol{\omega})$.

In Huettner et al. (2019) we start with a model that allows the consumer to generate any information strategy that is consistent with the prior belief $g \in \Delta(\boldsymbol{\Omega})$. To this end, the cost of learning is reflected by the expected reduction in Shannon entropy, thereby differentiating the different learning cost for different states. We first show that this model can be simplified to a model that is stated in choice probabilities. Specifically, we arrive at the following optimization problem:

$$\max_{(p(i\,|\,\boldsymbol{\omega}))_{i\in A,\boldsymbol{\omega}\in\boldsymbol{\Omega}}} \sum_{i\in A}\sum_{\boldsymbol{\omega}\in\boldsymbol{\Omega}} u(i,\boldsymbol{\omega})p(i\,|\,\boldsymbol{\omega})g(\boldsymbol{\omega}) - \sum_{k=1}^{n}\lambda_k I_p(\Omega_k, A\,|\,\boldsymbol{\Omega}_{1\cdot\cdot k-1})$$
$$\text{subject to } p(i\,|\,\boldsymbol{\omega}) \geq 0 \quad \text{for all } i\in A, \boldsymbol{\omega}\in\boldsymbol{\Omega},$$
$$\sum_{i\in A} p(i\,|\,\boldsymbol{\omega}) = 1 \quad \text{for all } \boldsymbol{\omega}\in\boldsymbol{\Omega},$$

where $p(i\,|\,\boldsymbol{\omega})$ denotes the conditional probability of chosing alternative i given that the state is $\boldsymbol{\omega}$, and $I_p(\Omega_k, A\,|\,\boldsymbol{\Omega}_{1\cdot\cdot k-1})$ is the conditional mutual information between Ω_k and choice of A, given the states $\boldsymbol{\Omega}_{1\cdot\cdot k-1}$ that are easier to learn.

Theorem 1 in Huettner et al. (2019) identifies the following *necessary conditions for optimality*, obtained from the first order conditions:

$$p(i\,|\,\boldsymbol{\omega}) = \frac{e^{\frac{u(i,\boldsymbol{\omega})}{\lambda_n}} p(i)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p(i\,|\,\boldsymbol{\omega}_{1\cdot\cdot k})^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}}}{\sum_{j\in A} e^{\frac{u(j,\boldsymbol{\omega})}{\lambda_n}} p(i)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p(j\,|\,\boldsymbol{\omega}_{1\cdot\cdot k})^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}}} \quad \text{for all } i\in A, \boldsymbol{\omega}\in\boldsymbol{\Omega}. \tag{1}$$

Note that those equations are void if $p(i\,|\,\boldsymbol{\omega}) = 0$. In contrast, if $p(i\,|\,\boldsymbol{\omega}) > 0$ for all $\boldsymbol{\omega}\in\boldsymbol{\Omega}$ and all $i\in A$, then (1) is also sufficient for the optimality of a solution.

Insertion of (1) into the consumer's optimization yields a simplifcation of the problem. To this end, denote a collection of partial conditional choice probabilities by

$$\mathbf{p} := (p(i\,|\,\boldsymbol{\omega}_{1\cdot\cdot n-1}))_{i\in A, \boldsymbol{\omega}_{1\cdot\cdot n-1}\in\boldsymbol{\Omega}_{1\cdot\cdot n-1}}.$$

Lemma 2 in Huettner et al. (2019) derives this equivalent, alternative formulation of the problem:

$$\max_{\mathbf{p}} \ \lambda_n \sum_{\boldsymbol{\omega}} g(\boldsymbol{\omega}) \log \sum_{i\in A} e^{\frac{u(i,\boldsymbol{\omega})}{\lambda_n}} p(i)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p(i\,|\,\boldsymbol{\omega}_{1\cdot\cdot k})^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}} \tag{2}$$
$$\text{subject to } p(i\,|\,\boldsymbol{\omega}_{1\cdot\cdot n-1}) \geq 0 \quad \text{for all } i\in A, \text{ and all } \boldsymbol{\omega}_{1\cdot\cdot n-1}\in\boldsymbol{\Omega}_{1\cdot\cdot n-1}, \tag{3}$$
$$\sum_{i\in A} p(i\,|\,\boldsymbol{\omega}) = 1 \quad \text{for all } \boldsymbol{\omega}_{1\cdot\cdot n-1}\in\boldsymbol{\Omega}_{1\cdot\cdot n-1}. \tag{4}$$

2. The Error and Correction

In Huettner et al. (2019), we further attempt to solve the alternative problem (2)-(4) and claim in Theorem 2 that the following constitutes a sufficient condition for a collection of partial conditional choice probabilities $\boldsymbol{p}$ to be optimal:

$$\sum_{\boldsymbol{\omega}_{1\cdot\cdot n-1}} g(\boldsymbol{\omega}_{1\cdot\cdot n-1}) \sum_{\omega_n} \frac{g(\omega_n\,|\,\boldsymbol{\omega}_{1\cdot\cdot n-1}) e^{\frac{u(i,\boldsymbol{\omega})}{\lambda_n}}}{\sum_{j\in A} e^{\frac{u(j,\boldsymbol{\omega})}{\lambda_n}} p(j)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p(j\,|\,\boldsymbol{\omega}_{1\cdot\cdot k})^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}}} \leq 1 \quad \text{for all } i\in A. \tag{5}$$

As it turns out, these conditions are necessary, but *not* sufficient for a collection of partial conditional choice probabilities $\mathbf{p}$ to be an optimal solution. We establish this through an example constructed from a scenario that was kindly shared with us by David Walker-Jones. In this example, a set of choice probabilities actually satisfy (5), but they do not constitute an optimum.

Example. Consider the scenario with three options $A = \{1, 2, 3\}$ and the state space $\boldsymbol{\Omega} = \Omega_1 \times \Omega_2$ with $\Omega_1 = \{9.9, 0\}$ indicating the value of alternative 1 and $\Omega_2 = \{10, 0\}$ the value of alternative 2. Let alternative 3 be the mirror image of alternative 2, so its value is implied by alternative 2. Accordingly, there are 4 states and suppose that they are a-priori equally likely.

State	$(9.9, 0)$	$(9.9, 10)$	$(0, 0)$	$(0, 10)$
Prior belief	1/4	1/4	1/4	1/4
Utility Alternative 1	9.9	9.9	0	0
Utility Alternative 2	0	10	0	10
Utility Alternative 3	10	0	10	0

Moreover, let the cost of information be

$$\lambda_1 := 1 < 5 =: \lambda_2.$$

Clearly, the consumer will learn the realization of Ω_1 with high precision, and if alternative 1 has value 9.9, it will be chosen with high probability. Indeed, the optimal solution is as follows:[1]

Alternative	1	2	3
Optimal partial conditional choice probability			
given that the value of alternative 1 is . . .			
9.9	93.04%	3.48%	3.48%
0	0.067%	49.97%	49.97%
Optimal unconditional choice probability	46.6%	26.7%	26.7%

However, the choice probabilities given by $p(1) = 0$, i.e., never select alternative 1, and choice probabilities for alternatives 2 and 3 given as

$$p(2 \mid \omega_1) = p(3 \mid \omega_1) = \frac{1}{2} \quad \text{for } \omega_1 \in \{9.9, 0\}, \tag{6}$$

implying $p(2) = p(3) = \frac{1}{2}$, satisfy the conditions in (5). Specifically, for $i = 1, 2, 3$, we have:

$$\begin{aligned}
&\frac{1}{4}\frac{e^{9.9}}{0.5e^0 + 0.5e^{10}} + \frac{1}{4}\frac{e^{9.9}}{0.5e^{10} + 0.5e^0} + \frac{1}{4}\frac{e^0}{0.5e^0 + 0.5e^{10}} + \frac{1}{4}\frac{e^0}{0.5e^0 + 0.5e^{10}} < 1,\\
&\frac{1}{4}\frac{e^0}{0.5e^0 + 0.5e^{10}} + \frac{1}{4}\frac{e^{10}}{0.5e^{10} + 0.5e^0} + \frac{1}{4}\frac{e^0}{0.5e^0 + 0.5e^{10}} + \frac{1}{4}\frac{e^{10}}{0.5e^{10} + 0.5e^0} = 1,\\
&\frac{1}{4}\frac{e^{10}}{0.5e^0 + 0.5e^{10}} + \frac{1}{4}\frac{e^0}{0.5e^{10} + 0.5e^0} + \frac{1}{4}\frac{e^{10}}{0.5e^0 + 0.5e^{10}} + \frac{1}{4}\frac{e^0}{0.5e^{10} + 0.5e^0} = 1.
\end{aligned}$$

This establishes the fact that $\mathbf{p}$ may satisfy (5) without optimizing the problem (2)-(4).

[1] The conditional choice probabilities $p(i \mid \omega_1 \omega_2)$ can be computed via the necessary conditions in (1).

2.1. Correction

In order to remedy the error, we refer to the sufficient conditions in Walker-Jones (2023), who studies a very similar problem. To this end, we introduce the following notation. For every alternative i, state $\boldsymbol{\omega}$, and a collection of partial conditional choice probabilities $\mathbf{p} = (p(i \mid \boldsymbol{\omega}_{1\cdots n-1}))_{i\in A, \boldsymbol{\omega}_{1\cdots n-1}\in\boldsymbol{\Omega}_{1\cdots n-1}}$, define

$$
\begin{aligned}
\gamma_i^{(0)}(\mathbf{p},\boldsymbol{\omega}) &= \frac{e^{\frac{u(i,\boldsymbol{\omega})}{\lambda}}}{\sum_{j\in A} e^{\frac{u(j,\boldsymbol{\omega})}{\lambda}} p(j)^{\frac{\lambda_1}{\lambda_n}} \prod_{k=1}^{n-1} p(j \mid \boldsymbol{\omega}_{1\cdots k})^{\frac{\lambda_{k+1}-\lambda_k}{\lambda_n}}} \\
\gamma_i^{(1)}(\mathbf{p},\boldsymbol{\omega}_{1\cdots n-1}) &= \left(\sum_{\omega_n\in\Omega_n} g(\omega_n \mid \boldsymbol{\omega}_{1\cdots n-1})\gamma_i^{(0)}(\mathbf{p},\boldsymbol{\omega}_{1\cdots n-1}\omega_n)\right)^{\frac{\lambda_n}{\lambda_{n-1}}} \\
\gamma_i^{(2)}(\mathbf{p},\boldsymbol{\omega}_{1\cdots n-2}) &= \left(\sum_{\omega_{n-1}\in\Omega_{n-1}} g(\omega_{n-1} \mid \boldsymbol{\omega}_{1\cdots n-2})\gamma_i^{(1)}(\mathbf{p},\boldsymbol{\omega}_{1\cdots n-2}\omega_{n-1})\right)^{\frac{\lambda_{n-1}}{\lambda_{n-2}}} \\
&\vdots \\
\gamma_i^{(n-1)}(\mathbf{p},\omega_1) &= \left(\sum_{\omega_2\in\Omega_2} g(\omega_2 \mid \omega_1)\gamma_i^{(n-2)}(\mathbf{p},\omega_1\omega_2)\right)^{\frac{\lambda_2}{\lambda_1}} \\
\gamma_i^{(n)}(\mathbf{p}) &= \sum_{\omega_1\in\Omega_1} g(\omega_1)\gamma_i^{(n-1)}(\mathbf{p},\omega_1)
\end{aligned}
$$

Here, $\gamma_i^{(0)}(\mathbf{p},\boldsymbol{\omega})$ is the relative "score" of alternative i in state $\boldsymbol{\omega}$ if choice probabilities are given by $\mathbf{p}$. These are then aggregated in a step-wise manner (in conjunction with an exponentiation involving the corresponding information costs) to obtain the expected score $\gamma_i^{(n)}(\mathbf{p})$ of each alternative $i \in A$. The correct sufficient conditions follow from Theorem 3 of Walker-Jones (2023) and are presented in the next proposition.

Proposition 1 *A collection of partial conditional choice probabilities* $\mathbf{p}^*$ *solves the alternative problem (2)-(4), whenever*

$$\gamma_i^{(n)}(\mathbf{p}^*) \leqslant 1 \quad \text{for all } i \in A. \tag{7}$$

The sufficient condition tests whether some alternative has an expected score $\gamma_i^{(n)}$ higher than 1; i.e., whether some unchosen alternative outperforms the chosen alternatives (it follows from the necessary conditions that $\gamma_i^{(n)}(\mathbf{p}^*) = 1$ for all chosen alternatives). The sufficient conditions we presented in (5) performed the same test, but relied on a simpler (hence incorrect) aggregation of the scores.

Returning to our example, where $n = 2$, condition (7) becomes

$$\sum_{\omega_1\in\{9.9,0\}} \frac{1}{2}\left(\sum_{\omega_2\in\{10,0\}} \frac{1}{2}\frac{e^{\frac{u(i,\omega_1\omega_2)}{\lambda}}}{\sum_{j\in\{1,2,3\}} e^{\frac{u(j,\omega_1\omega_2)}{\lambda}} p(j)^{\frac{\lambda_1}{\lambda_2}} p(j \mid \omega_1)^{\frac{\lambda_2-\lambda_1}{\lambda_2}}}\right)^{\frac{\lambda_2}{\lambda_1}} \leqslant 1$$

for all $i \in \{1,2,3\}$. Indeed, the optimal solution gives $\gamma_i^{(2)}(\mathbf{p}^*) = 1$ for each $i \in \{1,2,3\}$, whereas for the incorrect set of probabilities examined earlier, we get $\gamma_1^{(2)}(\mathbf{p}) = 7.68$, which violates (7).

3. Impact on Further Results

We remark that our key derivations and characterizations, in particular Theorem 1 (necessary conditions), Corollary 1 (invariant ratio of posterior beliefs), and Lemma 1 (alternative formulation) remain unaffected. We also used these correct results for the computation of our numeric examples and applications.

Regarding the algorithm, Proposition 2 (the improvement in each step) does not rely on this sufficient condition, and hence remains valid. The main implication is on Step 3 of the general algorithm in §6.2, where we apply the test for optimality when probabilities converge using our (incorrect) sufficient condition, and perturb the solution in a way that exploits the violation of the sufficient condition if the optimality test fails, gain improvement and continue. In other words, even though our algorithm will improve the objective function in each step (as stated in Proposition 2) if $\mathbf{p}^{t+1} \neq \mathbf{p}^t$, we cannot guarantee optimality when $\mathbf{p}^{t+1} \approx \mathbf{p}^t$, and hence Theorem 3 (algorithm optimality) would not hold. Naturally, we can apply the correct sufficient conditions presented (7) in Proposition 1 to conduct the test for optimality when $\mathbf{p}^{t+1} \approx \mathbf{p}^t$. If the solution satisfies (7), the optimality would be guaranteed. If not, we would need to resort to standard convex optimization techniques (see e.g., Boyd Vandenberghe 2004) to continue the search.

References

Boyd S, Vandenberghe L (2004) *Convex Optimization* (Cambridge University Press), URL `http://dx.doi.org/10.1017/cbo9780511804441`.

Huettner F, Boyacı T, Akçay Y (2019) Consumer choice under limited attention when alternatives have different information costs. *Operations Research* 67(3):671–699, URL `http://dx.doi.org/10.1287/opre.2018.1828`.

Walker-Jones D (2023) Rational inattention with multiple attributes. *Journal of Economic Theory* 212:105688, URL `http://dx.doi.org/10.1016/j.jet.2023.105688`.