Appendix B: Online Appendix for "Learning From Reviews: The Selection Effect and the Speed of Learning"

B.1 Omitted Proofs from Appendix A

Proof of Proposition 1

We first prove that if θ is uniform and $\bar{\theta} \geq \max\left\{p - \mathbb{E}[\zeta], \lambda_{\overline{K}} - \underline{\zeta} + p\right\}$, then for $i' < i$ we have

$$\mathbb{P}[a \in \cup_{j=i}^{m} T_j \mid a \in \cup_{j=i'}^{m} T_j, q, Q = 1] > \mathbb{P}[a \in \cup_{j=i}^{m} T_j \mid a \in \cup_{j=i'}^{m} T_j, q, Q = 0]. \tag{B-1}$$

We can rewrite (B-1) as

$$\frac{\mathbb{P}\left[\theta + \zeta + 1 - p \geq \lambda_i, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right]}{\mathbb{P}\left[\theta + \zeta + 1 - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right]} > \frac{\mathbb{P}\left[\theta + \zeta - p \geq \lambda_i, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right]}{\mathbb{P}\left[\theta + \zeta - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right]},$$

which we prove next. We have that

$$\begin{aligned}\frac{\mathbb{P}\left[\theta + \zeta + 1 - p \geq \lambda_i, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right]}{\mathbb{P}\left[\theta + \zeta + 1 - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right]} &= \frac{1}{1 + \frac{\mathbb{P}[\lambda_i \geq \theta + \zeta + 1 - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0]}{\mathbb{P}[\theta + \zeta + 1 - p \geq \lambda_i, \theta + q + \mathbb{E}[\zeta] - p \geq 0]}} \\ &\overset{(a)}{\geq} \frac{1}{1 + \frac{\mathbb{P}[\lambda_i \geq \theta + \zeta + 1 - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0]}{\mathbb{P}[\theta + \zeta - p \geq \lambda_i, \theta + q + \mathbb{E}[\zeta] - p \geq 0]}} \\ &\overset{(b)}{\geq} \frac{1}{1 + \frac{\mathbb{P}[\lambda_i \geq \theta + \zeta - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0]}{\mathbb{P}[\theta + \zeta - p \geq \lambda_i, \theta + q + \mathbb{E}[\zeta] - p \geq 0]}}\end{aligned}$$

where (a) follows from

$$\mathbb{P}\left[\theta + \zeta + 1 - p \geq \lambda_i, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right] \geq \mathbb{P}\left[\theta + \zeta - p \geq \lambda_i, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right]$$

and we next prove (b) by establishing

$$\mathbb{P}\left[\lambda_i \geq \theta + \zeta - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right] > \mathbb{P}\left[\lambda_i \geq \theta + \zeta + 1 - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right]. \tag{B-2}$$

We can write

$$\begin{aligned}&\mathbb{P}\left[\lambda_i \geq \theta + \zeta - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right] \\ &= \int_{\underline{\zeta}}^{\bar{\zeta}} f_\zeta(x) \mathbb{P}\left[\theta \in \left[\max\left\{p - q - \mathbb{E}[\zeta], \lambda_{i'} + p - x\right\}, \lambda_i - x + p\right]\right] dx \\ &\overset{(a)}{>} \int_{\underline{\zeta}}^{\bar{\zeta}} f_\zeta(x) \mathbb{P}\left[\theta \in \left[\max\left\{p - q - \mathbb{E}[\zeta], \lambda_{i'} + p - 1 - x\right\}, \lambda_i - x + p - 1\right]\right] dx \\ &= \mathbb{P}\left[\lambda_i \geq \theta + \zeta + 1 - p \geq \lambda_{i'}, \theta + q + \mathbb{E}[\zeta] - p \geq 0\right],\end{aligned} \tag{B-3}$$

where (a) holds because for all $x \in [\underline{\zeta}, \bar{\zeta}]$ we have

$$\begin{aligned}
&\mathbb{P}\left[\theta \in \left[\max\left\{p - q - \mathbb{E}[\zeta], \lambda_{i'} + p - x\right\}, \lambda_i - x + p\right]\right] \\
\overset{(a)}{=}&\frac{\lambda_i - x + p - \max\left\{p - q - \mathbb{E}[\zeta], \lambda_{i'} + p - x\right\}}{\bar{\theta} - \underline{\theta}} \\
=&\frac{\min\left\{\lambda_i - x + q + \mathbb{E}[\zeta], \lambda_i - \lambda_{i'}\right\}}{\bar{\theta} - \underline{\theta}} \\
\overset{(c)}{\geq}&\frac{\min\left\{\lambda_i - 1 - x + q + \mathbb{E}[\zeta], \lambda_i - \lambda_{i'}\right\}}{\bar{\theta} - \underline{\theta}} \\
=&\frac{\lambda_i - 1 - x + p - \max\left\{p - q - \mathbb{E}[\zeta], \lambda_{i'} - 1 + p - x\right\}}{\bar{\theta} - \underline{\theta}} \\
\overset{(b)}{=}&\mathbb{P}\left[\theta \in \left[\max\left\{p - q - \mathbb{E}[\zeta], \lambda_{i'} - 1 + p - x\right\}, \lambda_i - 1 - x + p\right]\right],
\end{aligned}$$

where (a) and (b) follow from $\bar{\theta} \geq p - \mathbb{E}[\zeta] \geq p - \mathbb{E}[\zeta] - q$ and $\bar{\theta} \geq \lambda_{\overline{K}} - \underline{\zeta} + p \geq \lambda_i - \underline{\zeta} + p$. Finally, note that inequality (a) in (B-3) is strict because the above inequality (c) holds strictly for $x \geq \lambda_{i'} + q + \mathbb{E}[\zeta]$. Before proceeding with the rest of the proof, let us show that complete learning happens under $\bar{\theta} \geq \max\left\{p - \mathbb{E}[\zeta], \lambda_{\overline{K}} - \underline{\zeta} + p\right\}$. For complete learning to happen, similar to Assumptions 2 and 1, we must have a $\tilde{\theta} \leq \bar{\theta}$ such that

$$\tilde{\theta} + \mathbb{E}[\zeta] - p > 0 \text{ and } \tilde{\theta} + \underline{\zeta} + 1 - p < \lambda_{\underline{K}}.$$

Such a $\tilde{\theta}$ exists if

$$-\mathbb{E}\left[\zeta\right] \leq \lambda_{\underline{K}} - \underline{\zeta} - 1. \tag{B-4}$$

Therefore if the support of ζ and θ are wide enough such that (B-4) and $\bar{\theta} \geq \max\left\{p - \mathbb{E}[\zeta], \lambda_{\overline{K}} - \underline{\zeta} + p\right\}$ hold, then learning happens (i.e., the analogue of Assumptions 2 and 1 hold) and also the assumption of Proposition 2 holds.

We next prove part (a) by showing that if $\frac{f_\zeta(\cdot)}{1 - F_\zeta(\cdot)}$ is increasing, then for $i' < i$ the probability $\mathbb{P}[a \in \cup_{j=i}^m T_j \mid a \in \cup_{j=i'}^m T_j, q, Q]$ is decreasing in q (for both $Q = 0$ and $Q = 1$). Using Lemma A-2, there exist thresholds $\lambda_i \geq \lambda_{i'}$ such that we can write this probability as

$$\frac{\mathbb{P}\left[\theta + \zeta - p + Q \geq \lambda_i, \theta + \mathbb{E}[\zeta] + q - p \geq 0\right]}{\mathbb{P}\left[\theta + \zeta + q - p + Q \geq \lambda_{i'}, \theta + \mathbb{E}[\zeta] - p \geq 0\right]} = \frac{\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)\left(1 - F_\zeta(\lambda_i - x + p - Q)\right) dx}{\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)\left(1 - F_\zeta(\lambda_{i'} - x + p - Q)\right) dx}. \tag{B-5}$$

After taking derivative of (B-5) with respect to q, we see that its derivative is positive if and only if

$$\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x) \frac{1 - F_\zeta(\lambda_{i'} - x + p - Q)}{1 - F_\zeta(\lambda_{i'} + q + \mathbb{E}[\zeta] - Q)} dx \geq \int_{p-q-\mathbb{E}[\zeta]} f_\theta(x) \frac{1 - F_\zeta(\lambda_i - x + p - Q)}{1 - F_\zeta(\lambda_i + q + \mathbb{E}[\zeta] - Q)} dx. \tag{B-6}$$

Inequality (B-6) holds if

$$\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)\frac{1-F_\zeta(\lambda-x+p-Q)}{1-F_\zeta(\lambda+q+\mathbb{E}[\zeta]-Q)}dx \tag{B-7}$$

is decreasing in λ. The derivative of (B-7) with respect to λ is

$$\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)\frac{1}{\left(1-F_\zeta(\lambda+q+\mathbb{E}[\zeta]-Q)\right)^2}\times$$
$$\left(f_\zeta(\lambda+q+\mathbb{E}[\zeta]-Q)\left(1-F_\zeta(\lambda-x+p-Q)\right)-f_\zeta(\lambda-x+p-Q)\left(1-F_\zeta(\lambda+q+\mathbb{E}[\zeta]-Q)\right)\right)dx, \tag{B-8}$$

which is non-positive if $\frac{f_\zeta(\cdot)}{1-F_\zeta(\cdot)}$ is decreasing, i.e., ζ has decreasing hazard rate.
The proof of part (b) follows from a similar argument by noting that if $\frac{f_\zeta(\cdot)}{1-F_\zeta(\cdot)}$ is increasing (i.e., ζ has increasing hazard rate), then (B-8) becomes non-negative and therefore the expression in (B-7) becomes increasing in λ, which in turn implies that (B-6) becomes increasing in q. ■

Proof of Lemma 1

We prove that when the distribution of θ has an increasing hazard rate, the support of $\pi(\overline{K}; F_{\theta,\zeta}, Q = 1, q, T)$ and $\pi(\overline{K}; F_{\theta,\zeta}, Q = 0, q, T)$ are separated. In particular, we prove that

$$\min_{q\in[0,1]} \pi(\overline{K}; F_{\theta,\zeta}, Q=1, q, T) > \max_{q\in[0,1]} \pi(\overline{K}; F_{\theta,\zeta}, Q=0, q, T),$$

establishing that the strict separation condition holds for $\mathcal{S} = \{\overline{K}\}$. We present the proof in two steps.
Step 1: We have

$$\max_q \pi(\overline{K}; F_{\theta,\zeta}, Q=0, q, T) = \pi(\overline{K}; F_{\theta,\zeta}, Q=0, q=0, T),$$

and

$$\min_q \pi(\overline{K}; F_{\theta,\zeta}, Q=1, q, T) = \pi(\overline{K}; F_{\theta,\zeta}, Q=1, q=1, T).$$

Proof of Step 1: Both inequalities follows from the fact that $\pi(\overline{K}; F_{\theta,\zeta}, Q, q, T)$ for $Q \in \{0,1\}$ is decreasing in q, which we show next. We can write

$$\begin{aligned}\pi(\overline{K}; F_{\theta,\zeta}, Q, q, T) &= \mathbb{P}\left[\theta+\mathbb{E}[\zeta]-p+q\geq 0, \theta+\zeta-p+Q\geq\lambda_{\overline{K}} \mid \theta+\mathbb{E}[\zeta]-p+q\geq 0\right]\\ &= \frac{\mathbb{P}\left[\theta+\mathbb{E}[\zeta]-p+q\geq 0, \theta+\zeta-p+Q\geq\lambda_{\overline{K}}\right]}{\mathbb{P}\left[\theta+\mathbb{E}[\zeta]-p+q\geq 0\right]}\\ &= \frac{\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)\left(1-F_\zeta(\lambda_{\overline{K}}-Q+p-x)\right)dx}{\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)dx},\end{aligned}$$

whose derivative with respect to q becomes

$$\frac{f_\theta(p-q-\mathbb{E}[\zeta])}{\left(\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)dx\right)^2}\times$$
$$\left(\left(1-F_\zeta(\lambda_{\overline{K}}-Q+q+\mathbb{E}[\zeta])\right)\left(\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)dx\right)-\left(\int_{p-q-\mathbb{E}[\zeta]} f_\theta(x)\left(1-F_\zeta(\lambda_{\bar{K}}-Q+p-x)\right)dx\right)\right).$$

This derivative is non-positive because $1-F_\zeta(\lambda_{\overline{K}}-Q+q+\mathbb{E}[\zeta]) \leq 1-F_\zeta(\lambda_{\overline{K}}-Q+p-x)$ for $x \geq p-q-\mathbb{E}[\zeta]$. This completes the proof of Step 1.

Step 2: When θ has increasing hazard rate, we have

$$\pi(\overline{K}; F_{\theta,\zeta}, Q=1, q=1) > \pi(\overline{K}; F_{\theta,\zeta}, Q=0, q=0).$$

Proof of Step 2: We first show that for any $\lambda \geq 0$, $\frac{\mathbb{P}_\theta[\theta\geq\lambda+x]}{\mathbb{P}_\theta[\theta\geq x]}$ is monotonically decreasing in x. Taking derivative with respect to x yields

$$\frac{-f_\theta(x+\lambda)(1-F_\theta(x))+f_\theta(x)(1-F_\theta(x+\lambda))}{(1-F_\theta(x))^2},$$

which is negative (from the increasing hazard rate condition, the numerator is negative). Using this property, we have

$$\begin{aligned}
&\pi(\overline{K}; F_{\theta,\zeta}, Q=1, q=1, T) = \mathbb{P}_{\theta,\zeta}[\theta-p+1+\zeta \geq \lambda_{\overline{K}}, \theta+\mathbb{E}[\zeta]-p+1 \geq 0 \mid \theta+\mathbb{E}[\zeta]-p+1 \geq 0] \\
&= \int_y f_\zeta(y)\mathbb{P}_\theta[\theta-p+1 \geq \lambda_{\overline{K}}-y, \theta+\mathbb{E}[\zeta]-p+1 \geq 0 \mid \theta+\mathbb{E}[\zeta]-p+1 \geq 0]dy \\
&= \int_y f_\zeta(y)\frac{\mathbb{P}_\theta[\theta \geq \max\{p+\lambda_{\overline{K}}-y, -\mathbb{E}[\zeta]+p\}-1]}{\mathbb{P}_\theta[\theta \geq -\mathbb{E}[\zeta]+p-1]}dy \\
&\overset{(a)}{>} \int_y f_\zeta(y)\frac{\mathbb{P}_\theta[\theta \geq \max\{p+\lambda_{\overline{K}}-y, -\mathbb{E}[\zeta]+p\}]}{\mathbb{P}_\theta[\theta \geq -\mathbb{E}[\zeta]+p]}dy \\
&= \int_y f_\zeta(y)\mathbb{P}_\theta[\theta-p \geq \lambda_{\overline{K}}-y, \theta+\mathbb{E}[\zeta]-p \geq 0 \mid \theta+\mathbb{E}[\zeta]-p \geq 0]dy \\
&= \mathbb{P}_{\theta,\zeta}[\theta-p+\zeta \geq \lambda_{\overline{K}}, \theta+\mathbb{E}[\zeta]-p \geq 0 \mid \theta+\mathbb{E}[\zeta]-p \geq 0] = \pi(\overline{K}; F_{\theta,\zeta}, Q=0, q=0, T),
\end{aligned}$$

where (a) follows from the fact that $\frac{\mathbb{P}_\theta[\theta\geq x+\lambda]}{\mathbb{P}_\theta[\theta\geq x]}$ is monotonically decreasing in x for $\lambda \geq 0$ and the strict inequality follows from

$$p-\underline{\zeta}+\lambda_{\overline{K}} \overset{(i)}{>} p-\left(\lambda_{-\underline{K}}-1+p-\bar{\theta}\right)+\lambda_{\overline{K}} \overset{(ii)}{>} p-\lambda_{-\underline{K}}+1-\mathbb{E}[\zeta]+\lambda_{\overline{K}} > -\mathbb{E}[\zeta]+p$$

where (i) and (ii) follow from Assumptions 1 and 2, respectively. This completes the proof of Step 2. The proof of Lemma follows from combining Steps 1 and 2. ■

Proof of Proposition 2

Part 1: We present the proof of part 1 for $Q = 0$ (the proof for $Q = 1$ follows from the same argument). Subtracting the speed of learning for full history (Theorem 2) from the speed of learning for summary statistics (Theorem 4), leads to

$$\sum_{i=1}^{m} \mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0] \log\left(\frac{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0]}{\mathbb{P}[a \in T_i | a \in T, q = 1, Q = 1]}\right)$$
$$- \sum_{i=1}^{m} \mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0] \log\left(\frac{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0]}{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 1]}\right)$$
$$= \sum_{i=1}^{m} \mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0] \log\left(\frac{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 1]}{\mathbb{P}[a \in T_i | a \in T, q = 1, Q = 1]}\right).$$

Using inequality $\log x \leq x - 1$ for $x > 0$, this difference is upper bounded by

$$\sum_{i=1}^{m} \mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0] \left(\frac{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 1]}{\mathbb{P}[a \in T_i | a \in T, q = 1, Q = 1]} - 1\right). \tag{B-9}$$

We next show that (B-9) is negative by induction on m.

For $m = 2$, we let

$$x_2 = \mathbb{P}\left[a \in T_2 \mid a \in T, q = 0, Q = 0\right], \quad y_2 = \mathbb{P}\left[a \in T_2 \mid a \in T, q = 0, Q = 1\right],$$
$$z_2 = \mathbb{P}\left[a \in T_2 \mid a \in T, q = 1, Q = 1\right].$$

First note that part 2 of Theorem 3 proves that the strict separation condition holds. For $m = 2$ we only have two review options and therefore

$$x_2 \overset{(a)}{=} \max_{q \in [0,1]} \mathbb{P}\left[a \in T_2 \mid a \in T, q, Q = 0\right] \overset{(b)}{<} \min_{q \in [0,1]} \mathbb{P}\left[a \in T_2 \mid a \in T, q, Q = 1\right] \overset{(c)}{=} z_2, \tag{B-10}$$

where (a) and (c) follow from negative selection and (b) follows from from the fact that probability of more favorable reviews (which is $a \in T_2$ with $m = 2$) is higher for $Q = 1$ compared to $Q = 0$. Then, (B-9) becomes

$$x_2\left(\frac{y_2}{z_2} - 1\right) + (1 - x_2)\left(\frac{1 - y_2}{1 - z_2} - 1\right) = \frac{(y_2 - z_2)(x_2 - z_2)}{z_2(1 - z_2)} \overset{(a)}{<} 0,$$

where (a) follows from negative selection which implies $z_2 < y_2$ and (B-10) which states $x_2 < z_2$. This completes the proof for $m = 2$.

We next prove that if (B-9) is negative for $m - 1$, then it must be negative for m as well. In particular, we establish that by merging the last two terms of the summation, (B-9) increases and we still have a rating system with $m - 1$ review options that satisfies negative selection. In particular, we have

$$
\begin{aligned}
&\sum_{i=1}^{m}\mathbb{P}[a\in T_i|a\in T,q=0,Q=0]\left(\frac{\mathbb{P}[a\in T_i|a\in T,q=0,Q=1]}{\mathbb{P}[a\in T_i|a\in T,q=1,Q=1]}-1\right)\\
&\overset{(a)}{<}\left(\sum_{i=1}^{m-2}\mathbb{P}[a\in T_i|a\in T,q=0,Q=0]\left(\frac{\mathbb{P}[a\in T_i|a\in T,q=0,Q=1]}{\mathbb{P}[a\in T_i|a\in T,q=1,Q=1]}-1\right)\right)\\
&+\mathbb{P}[a\in T_{m-1}\cup T_m|a\in T,q=0,Q=0]\left(\frac{\mathbb{P}[a\in T_{m-1}\cup T_m|a\in T,q=0,Q=1]}{\mathbb{P}[a\in T_{m-1}\cup T_m|a\in T,q=1,Q=1]}-1\right)\overset{(b)}{<}0.
\end{aligned}
$$

We next prove that inequalities (a) and (b) hold.
To establish inequality (a), we define

$$
\begin{aligned}
x_m&=\mathbb{P}\left[a\in T_m\mid a\in T,q=0,Q=0\right], & x_{m-1}&=\mathbb{P}\left[a\in T_{m-1}\mid a\in T,q=0,Q=0\right],\\
y_m&=\mathbb{P}\left[a\in T_m\mid a\in T,q=0,Q=1\right], & y_{m-1}&=\mathbb{P}\left[a\in T_{m-1}\mid a\in T,q=0,Q=1\right],\\
z_m&=\mathbb{P}\left[a\in T_m\mid a\in T,q=1,Q=1\right], & z_{m-1}&=\mathbb{P}\left[a\in T_{m-1}\mid a\in T,q=1,Q=1\right].
\end{aligned}
$$

With this notation, inequality (a) becomes

$$
x_m\left(\frac{y_m}{z_m}-1\right)+x_{m-1}\left(\frac{y_{m-1}}{z_{m-1}}-1\right)<(x_m+x_{m-1})\left(\frac{y_m+y_{m-1}}{z_m+z_{m-1}}-1\right).
$$

The difference between left-hand and right-hand sides of the above inequality can be written as

$$
\begin{aligned}
&x_m\left(\frac{y_mz_{m-1}-y_{m-1}z_m}{z_m(z_m+z_{m-1})}\right)+x_{m-1}\left(\frac{y_{m-1}z_m-y_mz_{m-1}}{z_{m-1}(z_m+z_{m-1})}\right)\\
&=\frac{y_mz_{m-1}-y_{m-1}z_m}{z_{m-1}z_m(z_m+z_{m-1})}\left(x_mz_{m-1}-x_{m-1}z_m\right)\\
&=\frac{z_{m-1}z_m}{z_m+z_{m-1}}\left(\frac{y_m}{z_m}-\frac{y_{m-1}}{z_{m-1}}\right)\left(\frac{x_m}{z_m}-\frac{x_{m-1}}{z_{m-1}}\right)<0,
\end{aligned}
$$

where the inequality follows from: (i) the term $\left(\frac{y_m}{z_m}-\frac{y_{m-1}}{z_{m-1}}\right)$ is positive. This is because using negative selection we have

$$
\begin{aligned}
\frac{y_m}{y_m+y_{m-1}}&=\mathbb{P}[a\in T_m\mid a\in T_m\cup T_{m-1},q=0,Q=1]\\
&>\mathbb{P}[a\in T_m\mid a\in T_m\cup T_{m-1},q=1,Q=1]=\frac{z_m}{z_m+z_{m-1}},
\end{aligned}
$$

and (ii) the term $\left(\frac{x_m}{z_m}-\frac{x_{m-1}}{z_{m-1}}\right)$ is negative. This is because we have

$$
\begin{aligned}
\frac{x_m}{x_{m-1}+x_m}&\overset{(i)}{=}\max_{q\in[0,1]}\mathbb{P}\left[a\in T_m\mid a\in T_{m-1}\cup T_m,q,Q=0\right]\\
&\overset{(ii)}{<}\min_{q\in[0,1]}\mathbb{P}\left[a\in T_m\mid a\in T_{m-1}\cup T_m,q,Q=1\right]\overset{(iii)}{=}\frac{z_m}{z_{m-1}+z_m},
\end{aligned}
$$

where (i) and (iii), again, follows from negative selection and (ii) follows from negative selection and a similar argument to the proof of part 2 of Theorem 3.
Inequality (b) follows from induction hypothesis by noticing that the rating system with review options $T_1' = T_1, \ldots, T_{m-2}' = T_{m-2}, T_{m-1}' = T_{m-1} \cup T_m$ features negative selection. This is because for $q' < q$ and $i' < i$ we have

$$\begin{aligned}
&\mathbb{P}\left[a \in \cup_{j=i}^{m-1} T_j' \mid a \in \cup_{j=i'}^{m-1} T_j', q, Q\right] = \mathbb{P}\left[a \in \cup_{j=i}^{m} T_j \mid a \in \cup_{j=i'}^{m-1} T_j, q, Q\right] \\
&\leq \mathbb{P}\left[a \in \cup_{j=i}^{m} T_j \mid a \in \cup_{j=i'}^{m} T_j, q', Q\right] = \mathbb{P}\left[a \in \cup_{j=i}^{m-1} T_j' \mid a \in \cup_{j=i'}^{m-1} T_j, q', Q\right].
\end{aligned}$$

This completes the proof of part 1.
Part 2: We present the proof of part 2 of proposition for $Q = 0$ (the proof for $Q = 1$ follows from a similar argument). Subtracting the speed of learning for summary statistics (Theorem 4) from the speed of learning for full history (Theorem 2), leads to

$$\begin{aligned}
&\sum_{i=1}^{m} \mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0] \log\left(\frac{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0]}{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 1]}\right) \\
&- \sum_{i=1}^{m} \mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0] \log\left(\frac{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0]}{\mathbb{P}[a \in T_i | a \in T, q = 1, Q = 1]}\right) \\
&= \sum_{i=1}^{m} \mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0] \log\left(\frac{\mathbb{P}[a \in T_i | a \in T, q = 1, Q = 1]}{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 1]}\right).
\end{aligned}$$

Using the fact that for $x > 0$, $\log x \leq x - 1$, this difference is upper bounded by

$$\sum_{i=1}^{m} \mathbb{P}[a \in T_i | a \in T, q = 0, Q = 0] \left(\frac{\mathbb{P}[a \in T_i | a \in T, q = 1, Q = 1]}{\mathbb{P}[a \in T_i | a \in T, q = 0, Q = 1]} - 1\right). \tag{B-11}$$

We next prove that (B-11) is negative by induction on m.

For $m = 2$, we let

$$\begin{aligned}
&x_2 = \mathbb{P}\left[a \in T_2 \mid a \in T, q = 0, Q = 0\right], \quad y_2 = \mathbb{P}\left[a \in T_2 \mid a \in T, q = 0, Q = 1\right], \\
&z_2 = \mathbb{P}\left[a \in T_2 \mid a \in T, q = 1, Q = 1\right].
\end{aligned}$$

Then, (B-11) becomes

$$x_2 \left(\frac{z_2}{y_2} - 1\right) + (1 - x_2)\left(\frac{1 - z_2}{1 - y_2} - 1\right) = \frac{(z_2 - y_2)(x_2 - y_2)}{y_2(1 - y_2)} < 0,$$

where the inequality follows from negative selection which implies $z_2 < y_2$ and the fact that higher Q leads to higher probability of likes, i.e., $x_2 < y_2$. This completes the proof for $m = 2$.

We next prove that if (B-11) is negative for $m - 1$, then it must be negative for m as well. In particular, we prove that by merging the last two terms of the summation, (B-11) increases and we still have a rating system with $m-1$ review options that satisfies positive selection. In particular, we have

$$\begin{aligned}
&\sum_{i=1}^{m}\mathbb{P}[a\in T_i|a\in T,q=0,Q=0]\left(\frac{\mathbb{P}[a\in T_i|a\in T,q=1,Q=1]}{\mathbb{P}[a\in T_i|a\in T,q=0,Q=1]}-1\right)\\
&\overset{(a)}{<}\left(\sum_{i=1}^{m-2}\mathbb{P}[a\in T_i|a\in T,q=0,Q=0]\left(\frac{\mathbb{P}[a\in T_i|a\in T,q=1,Q=1]}{\mathbb{P}[a\in T_i|a\in T,q=0,Q=1]}-1\right)\right)\\
&+\mathbb{P}[a\in T_{m-1}\cup T_m|a\in T,q=0,Q=0]\left(\frac{\mathbb{P}[a\in T_{m-1}\cup T_m|a\in T,q=1,Q=1]}{\mathbb{P}[a\in T_{m-1}\cup T_m|a\in T,q=0,Q=1]}-1\right)\overset{(b)}{<}0.
\end{aligned}$$

We next show that inequalities (a) and (b) hold. To establish inequality (a), we let

$$\begin{aligned}
x_m&=\mathbb{P}\left[a\in T_m\mid a\in T,q=0,Q=0\right], &x_{m-1}&=\mathbb{P}\left[a\in T_{m-1}\mid a\in T,q=0,Q=0\right],\\
y_m&=\mathbb{P}\left[a\in T_m\mid a\in T,q=0,Q=1\right], &y_{m-1}&=\mathbb{P}\left[a\in T_{m-1}\mid a\in T,q=0,Q=1\right],\\
z_m&=\mathbb{P}\left[a\in T_m\mid a\in T,q=1,Q=1\right], &z_{m-1}&=\mathbb{P}\left[a\in T_{m-1}\mid a\in T,q=1,Q=1\right].
\end{aligned}$$

With this notation, inequality (a) becomes

$$x_m\left(\frac{z_m}{y_m}-1\right)+x_{m-1}\left(\frac{z_{m-1}}{y_{m-1}}-1\right)<(x_m+x_{m-1})\left(\frac{z_m+z_{m-1}}{y_m+y_{m-1}}-1\right).$$

The difference between left-hand side and the right-hand side of the above inequality can be written as

$$\begin{aligned}
&x_m\left(\frac{z_m y_{m-1}-z_{m-1}y_m}{y_m(y_m+y_{m-1})}\right)+x_{m-1}\left(\frac{z_{m-1}y_m-z_m y_{m-1}}{y_{m-1}(y_m+y_{m-1})}\right)\\
&=\frac{z_m y_{m-1}-z_{m-1}y_m}{y_{m-1}y_m(y_m+y_{m-1})}\left(x_m y_{m-1}-x_{m-1}y_m\right)\\
&=\frac{y_{m-1}y_m}{y_m+y_{m-1}}\left(\frac{z_m}{y_m}-\frac{z_{m-1}}{y_{m-1}}\right)\left(\frac{x_m}{y_m}-\frac{x_{m-1}}{y_{m-1}}\right)<0,
\end{aligned}$$

where the inequality follows from positive selection which implies $\frac{z_m}{z_m+z_{m-1}}>\frac{y_m}{y_m+y_{m-1}}$ and $\frac{y_m}{y_m+y_{m-1}}>\frac{x_m}{x_m+x_{m-1}}$ which follows from the fact that the probability of having more favorable review is greater for higher Q. Inequality (b) follows from induction hypothesis by noticing that the rating system with review options $T_1'=T_1,\ldots,T_{m-2}'=T_{m-2},T_{m-1}'=T_{m-1}\cup T_m$ features positive selection. This is because for $q'<q$ and $i'<i$ we have

$$\begin{aligned}
&\mathbb{P}\left[a\in\cup_{j=i}^{m-1}T_j'\mid a\in\cup_{j=i'}^{m-1}T_j',q,Q\right]=\mathbb{P}\left[a\in\cup_{j=i}^{m}T_j\mid a\in\cup_{j=i'}^{m-1}T_j,q,Q\right]\\
&\leq\mathbb{P}\left[a\in\cup_{j=i}^{m}T_j\mid a\in\cup_{j=i'}^{m}T_j,q',Q\right]=\mathbb{P}\left[a\in\cup_{j=i}^{m-1}T_j'\mid a\in\cup_{j=i'}^{m-1}T_j,q',Q\right].
\end{aligned}$$

This completes the proof of part 2. ■

Proof of Proposition 3

Suppose the review options in Ω are $\{-\underline{K}, \dots, \overline{K}, \overline{K}+1\}$. It suffices to consider Ω' to have one less review option, for example combining $\overline{K}$ and $\overline{K}+1$, and with a slight abuse of notation, we denote the review options of Ω' by $\{-\underline{K}, \dots, \overline{K}\}$, where $\overline{K}$ stands for either $\overline{K}$ and $\overline{K}+1$ under Ω. By hypothesis, the combination of these two review options in Ω' does not affect customer thresholds.

Let $\pi(a; F_{\theta,\zeta}, Q, q)$ and $\pi'(a; F_{\theta,\zeta}, Q, q)$ denote the probability of leaving review a in rating systems Ω and Ω', respectively. Then, for any review option $a \in \{-\underline{K}, \dots, \overline{K}-1\}$ the probability of leaving this review with any belief q and any $Q \in \{0,1\}$ in these two rating systems are the same, i.e.,

$$\pi(a; F_{\theta,\zeta}, Q, q) = \pi'(a; F_{\theta,\zeta}, Q, q), \quad a \in \{-\underline{K}, \dots, \overline{K}-1\}, \forall q, Q \in \{0,1\}. \tag{B-12}$$

The only difference between the probability distributions in these two rating systems are for review options $\overline{K}$ and $\overline{K}+1$. In particular, we have

$$\pi'(\overline{K}; F_{\theta,\zeta}, Q, q) = \pi(\overline{K}; F_{\theta,\zeta}, Q, q) + \pi(\overline{K}+1; F_{\theta,\zeta}, Q, q), \quad \forall q, Q \in \{0,1\}. \tag{B-13}$$

We will next proceed with the proof of the full history case and summary statistics case separately.
Full history: We present the proof for $Q = 0$ as the proof of $Q = 1$ is similar. Using Theorem 2, for the less refined rating system Ω' the speed of learning is

$$\sum_{a=-\underline{K}}^{\overline{K}} \pi'(a; F_{\theta,\zeta}, Q=0, q=0) \log\left(\frac{\pi'(a; F_{\theta,\zeta}, Q=0, q=0)}{\pi'(a; F_{\theta,\zeta}, Q=1, q=0)}\right). \tag{B-14}$$

For the more refined rating system Ω the speed of learning is

$$\sum_{a=-\underline{K}}^{\overline{K}+1} \pi(a; F_{\theta,\zeta}, Q=0, q=0) \log\left(\frac{\pi(a; F_{\theta,\zeta}, Q=0, q=0)}{\pi(a; F_{\theta,\zeta}, Q=1, q=0)}\right). \tag{B-15}$$

Using (B-12) and (B-13), the difference between the speed of learning of Ω given in (B-15) and the speed of learning of Ω' given in (B-14) becomes

$$\begin{aligned}
&\pi(\overline{K}; F_{\theta,\zeta}, Q=0, q=0) \log\left(\frac{\pi(\overline{K}; F_{\theta,\zeta}, Q=0, q=0)}{\pi(\overline{K}; F_{\theta,\zeta}, Q=1, q=0)}\right) \\
&+ \pi(\overline{K}+1; F_{\theta,\zeta}, Q=0, q=0) \log\left(\frac{\pi(\overline{K}+1; F_{\theta,\zeta}, Q=0, q=0)}{\pi(\overline{K}+1; F_{\theta,\zeta}, Q=1, q=0)}\right) \\
&- \left(\pi(\overline{K}; F_{\theta,\zeta}, Q=0, q=0) + \pi(\overline{K}+1; F_{\theta,\zeta}, Q=0, q=0)\right) \\
&\log\left(\frac{\pi(\overline{K}; F_{\theta,\zeta}, Q=0, q=0) + \pi(\overline{K}+1; F_{\theta,\zeta}, Q=0, q=0)}{\pi(\overline{K}; F_{\theta,\zeta}, Q=1, q=0) + \pi(\overline{K}+1; F_{\theta,\zeta}, Q=1, q=0)}\right),
\end{aligned} \tag{B-16}$$

which is non-negative by using the following inequality

$$x \log \frac{x}{x'} + y \log \frac{y}{y'} \geq (x+y) \log \frac{x+y}{x'+y'}, \quad \forall x, x', y, y' \in (0,1). \tag{B-17}$$

We next find the sufficient and necessary condition under which (B-16) is strictly positive. In particular, we show that the refinement strictly increases the speed of learning if and only if

$$\mathbb{P}\left[a = \overline{K}+1 \mid q=0, Q=0, a \in \{\overline{K}, \overline{K}+1\}\right] \neq \mathbb{P}\left[a = \overline{K}+1 \mid q=0, Q=1, a \in \{\overline{K}, \overline{K}+1\}\right].$$

Intuitively, this condition states that the probability of the top review, conditional on review being one of the top two reviews, is different when the product has high and low quality. This implies that the additional information that the review is $\overline{K}+1$ is informative about the underlying quality. To see this formally, note that inequality (B-17) is equivalent to

$$D\left(\left(\frac{x}{x+y}, \frac{y}{x+y}\right) || \left(\frac{x'}{x'+y'}, \frac{y'}{x'+y'}\right)\right) \geq 0$$

which holds with equality if and only if the two distributions are identical. This implies

$$\begin{aligned}\mathbb{P}\left[a = \overline{K}+1 \mid q=0, Q=0, a \in \{\overline{K}, \overline{K}+1\}\right] &= \frac{y}{x+y} = \frac{y'}{x'+y'} \\ &= \mathbb{P}\left[a = \overline{K}+1 \mid q=0, Q=1, a \in \{\overline{K}, \overline{K}+1\}\right].\end{aligned}$$

This completes the proof for full history.

Summary statistics: The proof follows from Theorem 4 with minor modifications analogous to those for full history. ■

Proof of Lemma A-2

The proof is by induction. Let $\lambda_{\overline{K}}$ be the smallest utility u for which review decision $\overline{K}$ is preferable to all other review options. Formally,

$$\lambda_{\overline{K}} = \inf\{u \in \mathbb{R} \cup \{-\infty\} \;:\; V(u, \overline{K}) \geq V(u, r) \quad \forall r \in \{-\underline{K}, \dots, \overline{K}-1\}\}. \tag{B-18}$$

First note that if the set of utilities u for which review decision $\bar{K}$ is preferable is empty, then no user leaves this review and we can repeat the same argument starting from review $\overline{K}-1$. Therefore, without loss of generality, we assume the set of such utilities is not empty. We claim that all users with utility u larger $\lambda_{\overline{K}}$ leave review $\overline{K}$. This is because for all $r \in \{-\underline{K}, \dots, \overline{K}-1\}$ and $u' \geq \lambda_{\overline{K}}$ we have

$$V(u', \overline{K}) - V(u', r) \overset{(a)}{\geq} V(\lambda_{\bar{K}}, \overline{K}) - V(\lambda_{\overline{K}}, r) \geq 0,$$

where (a) follows from increasing differences in Assumption A-2. Applying this argument inductively, we can determine the remaining thresholds $\lambda_{\overline{K}-1}, \dots, \lambda_{-\underline{K}}$. For example, let the second-highest threshold be the smallest utility for which review decision $\overline{K}-1$ is preferable to the review options below it. That is,

$$\lambda_{\overline{K}-1} = \inf\{u \in \mathbb{R} \cup \{-\infty\} \ : \ V(u, \overline{K}-1) \geq V(u, r) \ \ \forall r \in \{-\underline{K}, \dots, \overline{K}-2\}\}.$$

Then it follows that for all $u \in [\lambda_{\overline{K}-1}, \lambda_{\overline{K}})$ we have $\arg\max_{r \in \{-\underline{K}, \dots, \overline{K}\}} V(u, r) = \overline{K}-1$. Applying the same argument to the other thresholds establishes the existence of thresholds in $\mathbb{R} \cup \{-\infty, \infty\}$. We next prove that these thresholds are finite. Using part 2 of Assumption A-2 we have

$$\lim_{u \to -\infty} \left(V(u, -\underline{K}) - \max_{r \in \{-\underline{K}+1, \dots, \overline{K}\}} V(u, r) \right) > 0.$$

Therefore, there exists $\underline{u} \in \mathbb{R}$ such that for $u \leq \underline{u}$ we have $V(u, -\underline{K}) > V(u, \overline{K})$. Using the definition of $\lambda_{\overline{K}}$ in (B-18) we have $\lambda_{\overline{K}} > \underline{u}$. This completes the proof. ■

Proof of Theorem A-1

We first characterize the Bayes-Nash equilibrium with consequentialist utility and in particular show that there exists a Bayes-Nash equilibrium where review decisions again follow a threshold rule.

Lemma B-1. *Suppose Assumptions A-1 and A-3 hold. Then there exists a sequence of functions* $\{(\lambda_i^{(t)} \ : \ i \in \mathcal{R})\}_{t=1}^{\infty}$ *such that the functions* $\left(\lambda_i^{(t)} \ : \ i \in \mathcal{R}\right)$ *are mappings from the history at time* t *to real numbers where*

$$\lambda_{-\underline{K}}^{(t)}(\Omega_t) \leq \cdots \leq \lambda_{-1}^{(t)}(\Omega_t) \leq \lambda_1^{(t)}(\Omega_t) \leq \cdots \leq \lambda_{\overline{K}}^{(t)}(\Omega_t),$$

and the following review decisions

$$R_t(\Omega_t, u_t) = \begin{cases} -\underline{K} & \text{for} \quad u < \lambda_{-\underline{K}}^{(t)}(\Omega_t), \\ i & \text{for} \quad u_t \in [\lambda_{i-1}^{(t)}(\Omega_t), \lambda_i^{(t)}(\Omega_t)) \ \text{ and } \ -\underline{K} < i < 0, \\ 0 & \text{for} \quad u_t \in [\lambda_{-1}^{(t)}(\Omega_t), \lambda_1^{(t)}(\Omega_t)), \\ i & \text{for} \quad u_t \in [\lambda_i^{(t)}(\Omega_t), \lambda_{i+1}^{(t)}(\Omega_t)) \ \text{ and } 0 < i < \overline{K}, \\ \overline{K} & \text{for} \quad u \geq \lambda_{\overline{K}}^{(t)}(\Omega_t) \end{cases}$$

and purchase decisions $\mathbf{B} = \{B_t\}_{t=1}^{\infty}$ *given by (A-31) with beliefs formed by Bayes' rule constitute a Bayes-Nash equilibrium.*

Proof of Lemma B-1: We first define the sequence of functions $\{(\lambda_i^{(t)} \ : \ i \in \mathcal{R})\}_{t=1}^{\infty}$ and then show that review decisions based on these functions together with purchase decisions based on beliefs that are updated using Bayes' rule form a Bayes-Nash equilibrium.

First note that the purchase probability of customer $t+1$, denoted by pur_{t+1}, is equal to $\mathbb{P}_\theta[\theta_{t+1} + \mathbb{E}_\zeta[\zeta_{t+1}] - p + q_{t+1} \geq 0]$ which is increasing in q_{t+1}. Therefore, instead of $V(\cdot,\cdot)$ we can equivalently work with a utility function $\tilde{V}(\cdot,\cdot)$ that is a function of u_t and q_{t+1} and has increasing differences property.

Claim 1: In this claim we define the aforementioned threshold functions for customer $t \geq 1$. For a given t and history Ω_t we show that there exist thresholds $\lambda^{(t)}_{-\underline{K}} \leq \cdots \leq \lambda^{(t)}_{\overline{K}}$ such that the review decision $R_t : \mathbb{R} \to \mathcal{R}$ defined as

$$R_t(u_t) = \begin{cases} -\underline{K} & \text{for} \quad u < \lambda^{(t)}_{-\underline{K}}, \\ i & \text{for} \quad u_t \in [\lambda^{(t)}_{i-1}, \lambda^{(t)}_{i}) \ \text{ and } \ -\underline{K} < i < 0, \\ 0 & \text{for} \quad u_t \in [\lambda^{(t)}_{-1}, \lambda^{(t)}_{1}), \\ i & \text{for} \quad u_t \in [\lambda^{(t)}_{i}, \lambda^{(t)}_{i+1}) \ \text{ and } 0 < i < \overline{K}, \\ \overline{K} & \text{for} \quad u \geq \lambda^{(t)}_{\overline{K}}, \end{cases}$$

satisfies

$$R_t(u_t) \in \arg\max_{r \in \mathcal{R}} \tilde{V}(u_t, q_{t+1}(R_t, r)), \quad \forall u_t. \tag{B-19}$$

Here $q_{t+1}(R_t, r)$ denotes the belief of customer $t+1$ knowing that customer t is using function $R_t(\cdot)$ to leave a review and her review decision is r. (B-19) means that given customer $t+1$ believes that customer t is using function $R_t(\cdot)$ to leave a review, it is optimal for customer t to follow the review decision based on the function $R_t(\cdot)$. Note that given any review decision of customer t, which is a mapping denoted by $R_t(u_t) \in \mathcal{R}$, the updated belief of customer $t+1$ is a function of this mapping as well as the review decision of customer t denoted by r.

Proof of Claim 1: We show that a mapping that satisfies (B-19) exists for a rating system with two review options $\overline{K}, -\underline{K}$ as the proof of the case with multiple review options is similar. In particular, we show that a threshold rule $\lambda^{(t)}$ exists such that above the threshold the review decision is $\overline{K}$ and below it the review decision is $-\underline{K}$. For a given λ, suppose $R_t(u_t) = \overline{K}$ if and only $u_t \geq \lambda$. We let R_λ be a mapping that maps $u_t \geq \lambda$ to $\overline{K}$ and $u_t < \lambda$ to $-\underline{K}$. We define the function

$$f(\lambda) = \tilde{V}(\lambda, q_{t+1}(R_\lambda, \overline{K})) - \tilde{V}(\lambda, q_{t+1}(R_\lambda, -\underline{K})). \tag{B-20}$$

Note that the function $f(\lambda)$ is continuous in λ. This is because under Assumption A-3, V is continuous and q_{t+1} is a continuous function of the probabilities of different reviews which is continuous in the threshold λ (because under part 1 of Assumption A-3 the CDF of $F_{\theta,\zeta}$ is continuous). We next show that there exists a threshold $\lambda \in [\bar{\theta} + \underline{\zeta} - p + 1, \bar{\theta} + \bar{\zeta} - p]$ for which $f(\lambda) = 0$. For $\underline{\lambda} = \bar{\theta} + \underline{\zeta} - p + 1$, we have $q_{t+1}(R_{\underline{\lambda}}, \overline{K}) > q_t > q_{t+1}(R_{\underline{\lambda}}, -\underline{K})$, which implies that with threshold rule $R_{\underline{\lambda}}$, the purchase probability given review decision $\overline{K}$ is strictly larger than the purchase probability given review decision $-\underline{K}$. Therefore, using Part 3 of Assumption A-3 implies that $f(\underline{\lambda}) < 0$. Similarly, Part 3 of Assumption A-3 implies that $f(\bar{\lambda}) > 0$ for $\bar{\lambda} = \bar{\theta} + \bar{\zeta} - p$. The intermediate value theorem for continuous functions then establishes the existence of $\lambda^{(t)}$ such that

$f(\lambda^{(t)}) = 0$. Since $q_{t+1}(R_{\lambda^{(t)}}, \overline{K}) \geq q_t \geq q_{t+1}(R_{\lambda^{(t)}}, -\underline{K})$, the increasing differences property of V (Part 2 of Assumption A-3) implies that for any $u \geq \lambda^{(t)}$, we have $R_t(u) = \overline{K}$ and for any $u < \lambda^{(t)}$, we have $R_t(u) = -\underline{K}$. Finally, note that since $\lambda \in [\bar{\theta} + \underline{\zeta} - p + 1, \bar{\theta} + \bar{\zeta} - p]$, both $\overline{K}$ and $-\underline{K}$ reviews have a non-zero probability. This completes the proof of Claim 1.

Claim 2: The review decisions based on the thresholds characterized in Claim 1, together with the optimal purchase decisions, constitute a Bayes-Nash equilibrium.

Proof of Claim 2: This claim follows from the one-stage deviation principle. Consider customer t with material utility u_t who makes a review decision at time t. Using (B-19), and noting that customer $t+1$ will update her belief assuming that customer t's review is based on function R_t, we have

$$\tilde{V}(u_t, q_{t+1}(R_t, R_t(u_t))) \geq \tilde{V}(u_t, q_{t+1}(R_t, r')),$$

for all $r' \in \mathcal{R}$. This completes the proof of claim 2 and the proposition. ■

We now proceed with the proof of theorem, which follows those of Theorems 1-4 in the text, but now taking the strategic interaction between adjacent customers and the time-varying thresholds into account. In particular, the thresholds that determine the reviews at time t depend on the information available to customer t, i.e., Ω_t. We therefore provide a sketch of how the proof proceeds but do not repeat parts of the proof that are identical to those of Theorems 1-4. In the full history case, the public belief is still a martingale and converges to a limiting random variable. Moreover, as shown in Lemma B-1, under Assumptions A-3 both positive and negative reviews are left with positive probability. Therefore, a similar argument to that of Theorem 1 establishes that the public belief converges to the true quality almost surely. A similar argument to the proof of Theorem 2 establishes that the convergence is exponentially fast and characterizes its speed. Specifically, for any t and any belief of customer t denoted by q_t, the mapping $q_{t+1}(\cdot, \cdot)$ used in the proof of Lemma B-1 is a function that only depends on q_t, the review decision strategy of customer t, and the review decision of customer t:

$$q_{t+1}(R_\lambda, \overline{K}) = \frac{\frac{q_t}{1-q_t} \frac{\mathbb{P}[u \geq \lambda | Q=1]}{\mathbb{P}[u \geq \lambda | Q=0]}}{1 + \frac{q_t}{1-q_t} \frac{\mathbb{P}[u \geq \lambda | Q=1]}{\mathbb{P}[u \geq \lambda | Q=0]}}.$$

Because this mapping depends on t only through the belief of customer t, q_t, as $t \to \infty$, the belief of customers and therefore the thresholds converge. An identical argument to the proof of Proposition B-8 establishes that, even though the thresholds are time varying, learning is exponentially fast with the same speed of learning as Theorem 2 where the reviews are left based on the limiting thresholds.

For a rating system with summary statistics, when the strict separation condition holds, the same argument as in Theorem 3 applies and establishes that there is complete learning. The speed of learning follows with a similar argument to Theorem 4. In particular, for any t and any summary statistics $\mathbf{S}_t$, the mapping $q_{t+1}(\cdot, \cdot)$ used in the proof of Lemma B-1 is a function that depends on time $t+1$, the fraction of different reviews $\mathbf{S}_{t+1}$, the review decision strategy of customer t, and the review decision of customer t. We denote this function by $g(\mathbf{S}_{t+1}, t+1, R_\lambda, \overline{K})$. First, note that for a given fraction of reviews $\mathbf{S}_{t+1} = \mathbf{S}$, as $t \to \infty$, the belief of customer $t+1$ and the fractions of different reviews converge to limit point. Consequently, the thresholds converge as well, and a similar argument to the proof of Proposition B-8 establishes that, even though the thresholds are time varying, learning is exponentially fast where the reviews are left based on the limiting thresholds. ■

B.2 Examples and Discussion

B.2.1 Fast and Slow Learning From Reviews

The first example shows that a refinement of the review options in the rating system can significantly increase the speed of learning.

Example B-1. Suppose that θ has a uniform distribution over $[-2, 2]$, $p = 0$, and ζ has the following symmetric bimodal distribution

$$f_\zeta(z) = \begin{cases} \frac{1}{24} & z \leq -b - \frac{1}{4}, \\ \frac{1}{24} + 4\left(z + b + \frac{1}{4}\right) & -b - \frac{1}{4} < z \leq -b, \\ \frac{1}{24} + 4\left(-z - b + \frac{1}{4}\right) & -b < z \leq -b + \frac{1}{4}, \\ \frac{1}{24} & -b + \frac{1}{4} < z \leq 0, \end{cases}$$

over the support $[-6, 6]$. This distribution implies that the likelihood of a realization of ζ is very high around b and $-b$. We next compute the speed of learning of rating system with full history and the following review options:

Case 1: "dislike." "no review," and "like" and the thresholds are 1 and -1.

Case 2: "dislike," "no review," "like," and "love" and the thresholds are 1, -1, and 5.

The joint distribution of (θ, ζ) and the probabilities of observing the "like" and "love" reviews under $Q = 0$ and $Q = 1$ are depicted in Figure B-1. The speed of learning for these two cases as a function of parameter b are, in turn, plotted in Figure B-2. In Figure B-2, when $b = 5$, the speed of learning without the "love" option is approximately 0.001, which means that the half time for learning the truth would be 1000 periods. In contrast, with the "love" option present, the speed of learning changes to 0.2, which means that the half time for learning the truth would be 5. The speed of learning of the first case is slow because for each review, the ratio of the probability of observing that review when $Q = 0$ to the probability of observing it when $Q = 1$ is close to 1. In fact, as b increases, this ratio becomes closer to 1 and the speed of learning declines further (towards zero). The speed of learning in the presence of the "love" option is shown in Figure B-2, and is significantly higher. This is because the probability of observing the "love" review is very close to zero when $Q = 0$, and is positive when $Q = 1$. This leads to a large value of KL divergence and thus to relatively fast learning.

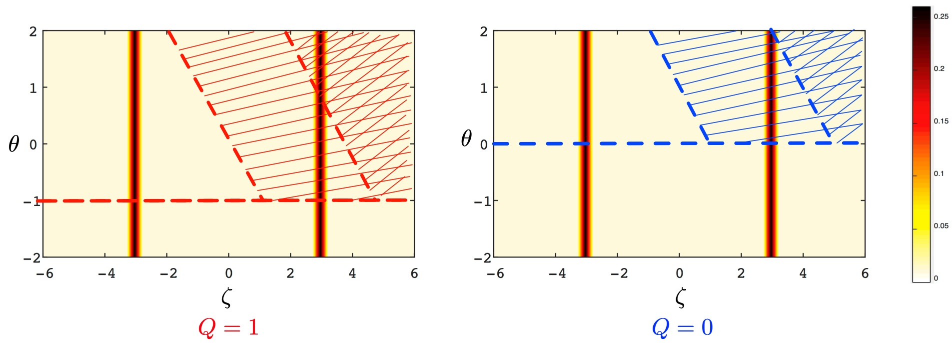

Figure B-1: The joint distribution of (θ, ζ) and the probabilities of the review options "like" and "love" for two cases under $Q = 0$ and $Q = 1$ for $b = 3$ for Example B-1. The double shaded areas correspond to the "love" option, and the shaded areas correspond to the "like" option.

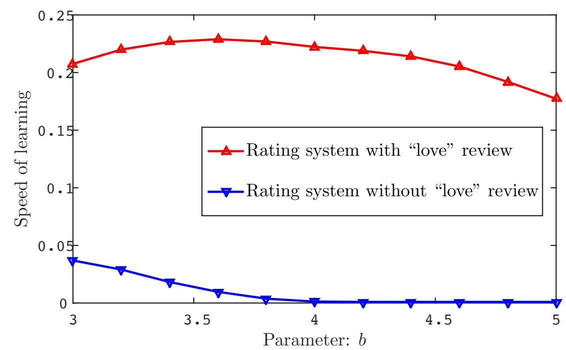

Figure B-2: The speed of learning as a function of parameter b for Example B-1. As b increases, the speed of learning of the rating system without the "love" option becomes very small, while the speed of learning of the rating system with the "love" option remains large.

In the next example, we show that increasing the support of ex post preference parameter, ζ, can significantly decrease the speed of learning.

Example B-2. Suppose θ and ζ have symmetric uniform distributions with diffuse supports such that $p \in \left(-\bar{\theta}+1, \bar{\theta}\right)$ and $\bar{\zeta} \geq \max\{\lambda_{\overline{K}}+1, -\lambda_{-\underline{K}}-p+\bar{\theta}+1\}$. From Theorem 2, the speed of learning with full history can be computed as shown in Figure B-3. We can see that the speed of learning for both $Q = 0$ and $Q = 1$ significantly decrease as the support of ζ widens (i.e., as $\bar{\zeta}$ increases). The characterization of the speed of learning in Theorem 2 provides the intuition for this result. When the true quality is 0 (and similarly, when it is 1) and ζ has a diffuse support, the likelihood of each review under both $(q = 0, Q = 0)$ or $(q = 0, Q = 1)$ is similar, and thus the likelihood ratio of each review is close to 1, leading to a low KL divergence. The differences in the speed of learning can be quite substantial. For example, Figure B-3 shows that when $\bar{\zeta} = 5$, the speed of learning for $Q = 1$ is 0.02, which means that the "half life" of convergence to the true

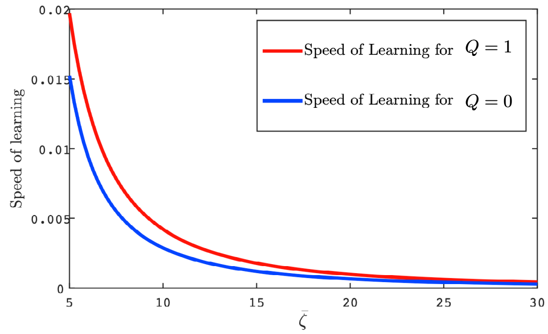

Figure B-3: The speed of learning as a function of the support of ζ for Example B-2 with $p = 0$, $\bar{\theta} = -\underline{\theta} = 2$, and $\lambda_{\overline{K}} = -\lambda_{-\underline{K}} = 1$.

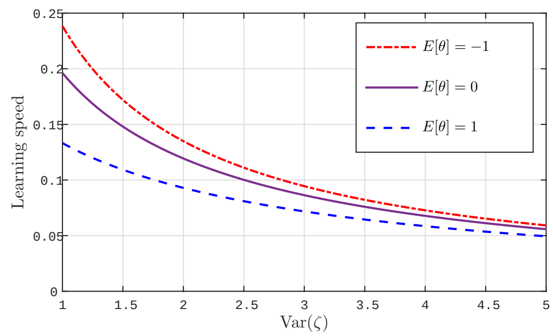

Figure B-4: Speed of learning for the setting in Example B-3 as a function of $\mathrm{var}(\zeta)$ and $\mathbb{E}[\theta]$.

state of nature is 50 periods (meaning that starting from any belief half of the distance to $q = 1$ will be close in 50 periods). In contrast, when $\bar{\zeta}$ is 20, the speed of learning changes to 0.001, and the half life of learning the underlying state is 1000.

Example B-3. We consider a rating system with full history of reviews and review options "like" and "dislike" that shows history of reviews among all purchases . To focus on the effects of distributions, we suppose $p = 0$ and the review is "like" if and only if the material utility is positive. We consider normal independent distribution for θ and ζ, with $\mathbb{E}[\zeta] = 0$ and $\mathrm{var}(\theta) = 1$. Figure B-4 shows the speed of learning as a function of the variance of ζ for different values of $\mathbb{E}[\theta]$. As we can see, increasing $\mathrm{var}(\zeta)$ decreases the the speed of learning. This is because the impact of the product's quality on the material utility becomes smaller. Another observation is that as $\mathbb{E}[\theta]$ increases the speed of learning decreases. This is because for larger values of $\mathbb{E}[\theta]$, the selection effect is more pronounced and hence learning becomes more difficult. More specifically, for large values of $\mathbb{E}[\theta]$, customers who purchase the product are positively biased toward liking the product. Therefore, the impact of the product's quality on their review is smaller.

Example B-4. We consider a rating system with review options "like" and "dislike" that displays the history of reviews among all purchases. Similar to Example B-3, we suppose that $p = 0$ and all

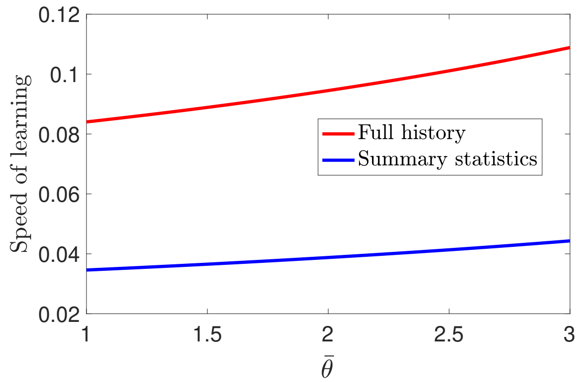

Figure B-5: Speed of learning for the setting in Example B-4 as a function of $\bar{\theta}$ for full history and summary statistics.

users who purchase the product leave a review and this review is "like" if and only if the material utility is positive. We consider uniform independent distribution for θ and ζ, with $\zeta \sim \text{Unif}[-5, 5]$ and $\theta \sim \text{Unif}[-1, \bar{\theta}]$. Figure B-5 shows the speed of learning as a function of $\bar{\theta}$. As we can see, a higher $\bar{\theta}$ increases the speed of learning for both full history and summary statistics. This is because the selection effect becomes less pronounced as $\bar{\theta}$ increases: the preference of user t who has purchased the product is a truncated uniform distribution on $[-q_t, \bar{\theta}]$ where q_t is the belief of user t. As $\bar{\theta}$ increases, whether the users know $q_t \in [0, 1]$ or not becomes less important.

B.2.2 Details of Example 2

The first two claims directly follow from Proposition 1. To see the last claim, note that the rating system with two review options that reports the fraction of "likes" among all customers features positive selection because $\pi(\overline{K}; F_{\theta,\zeta}, Q, q, T) = \mathbb{P}\left[\theta + \zeta + Q - p \geq \lambda_{\overline{K}}, \theta + \mathbb{E}[\zeta] + q - p \geq 0\right]$ is increasing in both Q and q, establishing the claims in Example 2.

B.2.3 Details of Example 3

Let the support of θ be $[-1,1]$ with the following probability density function:

$$f_\theta(x) = \begin{cases} \frac{1}{20} & x \leq \frac{-7}{8} - \frac{1}{10}, \\ \frac{1}{20} + \frac{18/20}{1/10}(x - (-7/8 - 1/10)) & \frac{-7}{8} - \frac{1}{10} \leq x \leq \frac{-7}{8}, \\ \frac{1}{20} - \frac{18/20}{1/10}(x - (-7/8 + 1/10)) & \frac{-7}{8} \leq x \leq \frac{-7}{8} + \frac{1}{10}, \\ \frac{1}{20} & \frac{-7}{8} + \frac{1}{10} \leq x \leq \frac{4}{8} - \frac{1}{10}, \\ \frac{1}{20} + \frac{18/20}{1/10}(x - (4/8 - 1/10)) & \frac{4}{8} - \frac{1}{10} \leq x \leq \frac{4}{8}, \\ \frac{1}{20} - \frac{18/20}{1/10}(x - (4/8 - 1/10)) & \frac{4}{8} \leq x \leq \frac{4}{8} + \frac{1}{10}, \\ \frac{1}{20} & \frac{4}{8} + \frac{1}{10} \leq x \leq \frac{4}{8} + \frac{1}{10}, \end{cases}$$

and let ζ have a symmetric distribution around 0 whose support is $[-2,2]$ with the following probability density function:

$$f_\zeta(x) = \begin{cases} \frac{1}{20} & x \leq \frac{-7}{4} - \frac{1}{10}, \\ \frac{1}{20} + \frac{8/20}{1/10}(x - (-7/4 - 1/10)) & \frac{-7}{4} - \frac{1}{10} \leq x \leq \frac{-7}{4}, \\ \frac{1}{20} - \frac{8/20}{1/10}(x - (-7/4 + 1/10)) & \frac{-7}{4} \leq x \leq \frac{-7}{4} + \frac{1}{10}, \\ \frac{1}{20} & \frac{-7}{4} + \frac{1}{10} \leq x \leq \frac{-1}{4} - \frac{1}{10}, \\ \frac{1}{20} + \frac{8/20}{1/10}(x - (-1/4 - 1/10)) & \frac{-1}{4} - \frac{1}{10} \leq x \leq \frac{-1}{4}, \\ \frac{1}{20} - \frac{8/20}{1/10}(x - (1/4 + 1/10)) & \frac{-1}{4} \leq x \leq \frac{-1}{4} + \frac{1}{10}, \\ \frac{1}{20} & \frac{-1}{4} + \frac{1}{10} \leq x. \end{cases}$$

First consider a rating system that reports the fraction for "likes" among reviews. As depicted in Panel (a) of Figure 1, this rating system features negative selection. We next show that it does not satisfy (weak) separation. To see this, note that the minimum of $\pi(\overline{K}; F_{\theta,\zeta}, Q = 1, q)$ (achieved at $q = 1$) is close to $5/8$ that is smaller than the maximum of $\pi(\overline{K}; F_{\theta,\zeta}, Q = 0, q)$ (achieved at $q = 0$) which is close to $6/8$.

Now consider a rating system that reports the fraction of "likes" among all customers. As we have shown in Example 2 and depicted in Panel (a) of Figure 2, this rating system features positive selection. We next show that it does not satisfy (weak) separation. To see this, note that the minimum of $\pi(\overline{K}; F_{\theta,\zeta}, Q = 1, q)$ (achieved at $q = 0$) is close to $3/8$ that is smaller than the maximum of $\pi(\overline{K}; F_{\theta,\zeta}, Q = 0, q)$ (achieved at $q = 1$) which is close to $4/8$.

Finally, we note that for uniform distribution of valuations with wide enough support, a rating system that reports the fraction of "likes" among reviews features negative selection, a rating system that reports the fraction of "likes" among all customers features positive selection, and we have complete learning for both rating systems (i.e., Assumptions 1 and 2 and the strict separation condition hold).

B.3 Additional Results

This section of Appendix B includes the additional results and extensions discussed in the main text.

B.3.1 Extension of part 2 of Theorem 1

Here, we show that if $\bar{\theta} + \mathbb{E}[\zeta] - p < 0$ and $Q = 0$, then almost surely learning is incomplete. In particular, we prove that

$$\mathbb{P}\left[\lim_{t\to\infty} q_t = 0\right] = 0.$$

To reach a contradiction, suppose $\mathbb{P}\left[\lim_{t\to\infty} q_t = 0\right] > 0$. Under the event $\lim_{t\to\infty} q_t = 0$, there exists s such that for $t \geq s$ we have

$$q_s \leq \frac{1}{2}\left(p - \bar{\theta} - \mathbb{E}[\zeta]\right).$$

We let t_0 be the smallest s with such property. Therefore, for $t \geq t_0$ we obtain

$$q_t + \theta + \mathbb{E}[\zeta] - p \leq q_t + \bar{\theta} + \mathbb{E}[\zeta] - p \leq \frac{1}{2}\left(\bar{\theta} + \mathbb{E}[\zeta] - p\right) \overset{(a)}{<} 0, \tag{B-21}$$

where (a) follows from $\bar{\theta} + \mathbb{E}[\zeta] - p < 0$. The inequality given in (B-21) means that the purchase does not happen for $t \geq t_0$ and the belief remains at q_{t_0}. Finally, note that q_{t_0} is always greater than 0 because, under Assumption 1, if purchase occurs, then the probability of any review is non-zero. Hence, the likelihood ratio at each time is multiplied by a number bounded away from zero and cannot become 0 in finite time. This contradicts $\lim_{t\to\infty} q_t = 0$, proving the claim. ■

B.3.2 Speed of Learning with Observable Valuations

Suppose that customer t (in addition to h_t) observes the ex ante valuation of previous customers, i.e., θ_s for $s = 1, \ldots, t-1$. For any $\theta \in [\underline{\theta}, \bar{\theta}]$, we define

$$\pi_\theta(a; F_{\theta,\zeta}, Q, q) = \mathbb{P}_{\zeta|\theta}\left[a \mid q, Q, \theta\right], \quad \forall q, Q \in \{0,1\}, \forall a \in \mathcal{A},$$

as the the probability of different actions when conditioning on the ex ante valuation. In this setting, the selection effect completely disappears because current customers can condition on past customers' θ's. The next proposition shows that in this case the speed of learning is always greater than in Theorem 2.

Proposition B-1. *Suppose Assumptions 1 and 2 hold, and that in addition to the full history of reviews, customer t observes θ_s for $s = 1, \ldots, t-1$. Complete learning occurs, i.e., $q_t \to Q$ almost surely and is exponentially fast. Moreover, for $Q = 0$, we almost surely have*

$$\lim_{t\to\infty} \frac{1}{t} \log q_t = -\mathbb{E}_\theta\left[D\left(\boldsymbol{\pi}_\theta(F_{\theta,\zeta}, Q=0, q=0) || \boldsymbol{\pi}_\theta(F_{\theta,\zeta}, Q=1, q=0)\right)\right].$$

and for $Q = 1$, we almost surely have

$$\lim_{t\to\infty} \frac{1}{t} \log(1 - q_t) = -\mathbb{E}_\theta \left[D\left(\boldsymbol{\pi}_\theta(F_{\theta,\zeta}, Q = 1, q = 1) || \boldsymbol{\pi}_\theta(F_{\theta,\zeta}, Q = 0, q = 1) \right) \right].$$

In both cases, the speed of learning is greater than in Theorem 2.

In the full history setting of Theorem 2, customers only know the public belief and hence the distribution of ex ante valuations of previous customers who have purchased the product. Proposition B-1 shows that if customers additionally observe ex ante valuations, then the speed of learning would be always greater.[17]

Proof: We prove the proposition for $Q = 0$ as the proof of $Q = 1$ is similar. The learning dynamic in this setting is

$$l_{t+1} = l_t \times \frac{\mathbb{P}_{\zeta|\theta}\left[a_t \mid q = q_t, Q = 1, \theta = \theta_t\right]}{\mathbb{P}_{\zeta|\theta}\left[a_t \mid q = q_t, Q = 0, \theta = \theta_t\right]}, \quad a_t = a \text{ w.p. } \mathbb{P}_{\zeta|\theta}\left[a_t \mid q = q_t, Q = 0, \theta = \theta_t\right].$$

$\{l_t\}_{t=1}^{\infty}$ forms a martingale. Thus almost sure convergence and the speed of learning follow from a similar argument to that of Theorem 1 and 2. In particular, letting $\pi(a; F_{\zeta|\theta}, Q, q, \theta) = \mathbb{P}_{\theta,\zeta}\left[a \mid q, Q, \theta\right]$, the speed of learning becomes $\mathbb{E}_\theta\big[D\big(\boldsymbol{\pi}(F_{\zeta|\theta}, Q = 0, q = 0, \theta) || \boldsymbol{\pi}(F_{\zeta|\theta}, Q = 1, q = 0, \theta)\big)\big]$. We next show that its speed is higher than the one presented in Theorem 2. We have

$$\begin{aligned}
&\mathbb{E}_\theta\left[D\left(\boldsymbol{\pi}(F_{\zeta|\theta}, Q = 0, q = 0, \theta) || \boldsymbol{\pi}(F_{\zeta|\theta}, Q = 1, q = 0, \theta)\right)\right] \\
&\overset{(a)}{\geq} D\left(\mathbb{E}_\theta\left[\boldsymbol{\pi}(F_{\zeta|\theta}, Q = 0, q = 0, \theta)\right] || \mathbb{E}_\theta\left[\boldsymbol{\pi}(F_{\zeta|\theta}, Q = 1, q = 0, \theta)\right]\right) \\
&= D\left(\boldsymbol{\pi}(F_{\theta,\zeta}, Q = 0, q = 0) || \boldsymbol{\pi}(F_{\theta,\zeta}, Q = 1, q = 0)\right),
\end{aligned}$$

where (a) follows from the convexity of KL divergence. ■

We next formally define extra information and then show that providing extra information weakly increases the speed of learning. Consider a rating system I and let $h_t = \{a_1, \ldots, a_{t-1}\}$ denote the history available to customer t. We say that rating system II has extra information compared to rating system I if we have

$$\pi^I(a_t; q, Q) = \mathbb{E}_{z_t}\left[\pi^{II}(a_t; q, Q)\right] \quad \text{for all } t \geq 1, a_t \in \mathcal{A}$$

for some additional information z_t that customers observe in rating system II —the history available to customer t in rating system II is $h_t = \{(a_1, z_1), \ldots, (a_{t-1}, z_{t-1})\}$. Below, we include some examples of extra information:

- $z_t = \theta_t$: this is the setting we described above in which each customer observes the ex ante type of previous customers.

[17] To see this, note that KL divergence is a convex function, and thus the average of the KL divergence for different θ's is greater than the speed of learning in Theorem 2, given by the KL divergence for average θ.

- $z_t = \zeta_t$: this corresponds to a setting in which each customer observes the ex post idiosyncratic preference term of previous customers.
- z_t is any extra information and the customers know the joint distribution $F_{\theta,\zeta,z}$. In rating system II each customer observes the extra information of previous customers while in rating system I each customer does not observe this extra information. An example of such extra information is when there are multiple groups of customers each with a different distribution of preferences (θ, ζ), and z is the group that each customer belongs to. In rating system II, customers know the group of each of previous customers who have left a review. In rating system I, however, customers only observe the previous reviews and not the group of previous customers.

A similar argument to the proof of Proposition B-1 establishes that extra information (weakly) increases the speed of learning, as formally stated in the next proposition.

Proposition B-2. *Suppose rating system II has extra information compared to rating system I, and learning happens for both of them under full history. The speed of learning in rating system II is (weakly) higher than the speed of learning in rating system I.*

B.3.3 When Customers Observe a Subset of Past Actions

We show that the characterization presented for the speed of learning under full history directly generalizes to the case in which the history available to customers only includes a subset of actions. For instance, the history available to customers may only include reviews (and no information about those who have not purchased the product or have not left a review). Similar to the setting in Section 4, we consider a subset of all possible actions $\mathcal{A}$ denoted by T. A rating system is represented by a partition of T, $\{T_1, \ldots, T_m\}$ (i.e., $T = \cup_{i=1}^m T_i$ and $T_i \cap T_j = \emptyset$, $i \neq j \in [m]$) such that for any $i > j$, the reviews in the set T_i are all more positive than the reviews in the set T_j. We also use the notation τ to denote the times t at which users observe an action in T (e.g., if customers observe past reviews, τ corresponds to the number of reviews so far). In other words, the history observed by the customer acting after $\tau - 1$ is $\{\tilde{a}_1, \ldots, \tilde{a}_{\tau-1}\}$, where $\tilde{a}_s = i$ if the s-th action in T belongs to T_i. Customers do not observe any other event and have improper uniform prior over their index in the sequence of all customers. This implies that customers have uniform improper prior over the times at which there have been unobservable actions.

The next proposition generalizes Theorem 2 to this setting.

Proposition B-3. *Suppose that Assumptions 1 and 2 hold, and that customers only observe actions in T and have uniform improper prior over the times at which actions that are not in T have taken place. Then, there is again complete learning. Moreover, for $Q = 0$, we almost surely have*

$$\lim_{\tau \to \infty} \frac{1}{\tau} \log q_\tau = -D\left(\boldsymbol{\pi}(F_{\theta,\zeta}, Q = 0, q = 0, T) || \boldsymbol{\pi}(F_{\theta,\zeta}, Q = 1, q = 0, T)\right),$$

and for $Q = 1$, we almost surely have

$$\lim_{\tau \to \infty} \frac{1}{\tau} \log(1 - q_\tau) = -D\left(\boldsymbol{\pi}(F_{\theta,\zeta}, Q = 1, q = 1, T) || \boldsymbol{\pi}(F_{\theta,\zeta}, Q = 0, q = 1, T)\right).$$

To understand this result, consider a customer at time t who has observed a history with τ actions in T. This customer does not know the number of past users who have joined the platform between the most recent action in T and time t. However, she knows exactly the belief of any user who has taken an action in T because she observes the same history as this user. She can then filter out the selection effect and aggregate information of all past customers who have taken an action in T, and this is sufficient to ensure (exponentially fast) complete learning. The speed of learning is modified in view of the fact that the information of a customer who has not taken an action in T is not being transmitted, but is still given by a similar KL divergence term.

Proposition B-3 provides the speed of learning when only actions in T are observed. In Appendix B.3.6, we provide the speed of learning in terms of calendar time for both full history and summary statistics.

Proof of Proposition B-3: We present the proof for $Q = 0$ (the proof for $Q = 1$ follows from the same argument). Recall that τ denotes the times at which there was an action in T and $\tilde{h}_\tau = \{\tilde{a}_1, \ldots, \tilde{a}_{\tau-1}\}$ denotes the available history after $\tau - 1$ actions in T. Similar to the proof of Theorem 1, we define random variable $\tilde{Z}(\cdot \mid l_\tau) = \frac{\mathbb{P}[\cdot|\tilde{a} \in T, Q=1, l_\tau]}{\mathbb{P}[\cdot|\tilde{a} \in T, Q=0, l_\tau]}$. We first establish that $l_{\tau+1} = l_\tau \tilde{Z}(\cdot \mid l_\tau)$, and then we prove l_τ forms a martingale. Note that between $\tau - 1$-th action in T and τ-th action in T, each customer has observed the same history and therefore has the same belief q_τ. Therefore, each one of them has taken an action $\tilde{a}$ that does not belong to T, independently with probability $\mathbb{P}\left[\tilde{a} \notin T \mid q_\tau, Q\right]$. Also, a new customer does not know the number of customers who joined the platform before her, and by assumption, has a uniform prior over this number. We let $h_{\tau:\tau+1}$ denote the history of actions in between $\tau - 1$-th and τ-th actions in T. We next present a recursive characterization of the likelihood ratio corresponding to belief $\{q_\tau\}_{\tau=0}^\infty$, exploiting the uniform (improper) prior assumption. We have

$$\begin{aligned}
l_{\tau+1} &= \frac{\mathbb{P}[\tilde{h}_\tau, \tilde{a}_\tau | Q = 1]}{\mathbb{P}[\tilde{h}_\tau, \tilde{a}_\tau | Q = 0]} = \frac{\sum_{h_{\tau:\tau+1}} \mathbb{P}[\tilde{h}_\tau, h_{\tau:\tau+1}, \tilde{a}_\tau | Q = 1]}{\sum_{h_{\tau:\tau+1}} \mathbb{P}[\tilde{h}_\tau, h_{\tau:\tau+1}, \tilde{a}_\tau | Q = 0]} \\
&= \frac{\sum_{h_{\tau:\tau+1}} \mathbb{P}[\tilde{h}_\tau \mid Q = 1] \mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q = 1] \mathbb{P}[\tilde{a}_\tau | \tilde{h}_\tau, h_{\tau:\tau+1}, Q = 1]}{\sum_{h_{\tau:\tau+1}} \mathbb{P}[\tilde{h}_\tau \mid Q = 0] \mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q = 0] \mathbb{P}[\tilde{a}_\tau | \tilde{h}_\tau, h_{\tau:\tau+1}, Q = 0]} \\
&= l_\tau \frac{\sum_{h_{\tau:\tau+1}} \mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q = 1] \mathbb{P}[\tilde{a}_\tau | \tilde{h}_\tau, h_{\tau:\tau+1}, Q = 1]}{\sum_{h_{\tau:\tau+1}} \mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q = 0] \mathbb{P}[\tilde{a}_\tau | \tilde{h}_\tau, h_{\tau:\tau+1}, Q = 0]}.
\end{aligned} \tag{B-22}$$

We next consider $\mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q]$. The randomness in $h_{\tau:\tau+1}$ comes from two sources: (i) the number of customers who joined the platform between $\tau - 1$-th and τ-th actions in T and (ii) the actions they took which depends on their θ and ζ. We let C denote the first random variable. By assumption this random variable is uniformly distributed over the integers. We also note that all customers in this time interval have observed the same history (i.e., $\tilde{h}_\tau$), formed the same belief (i.e., q_τ), and took an action independently with probability $\mathbb{P}[a \notin T \mid q_\tau, Q]$. Therefore, we have

$$\mathbb{P}[h_{\tau:\tau+1}|\tilde{h}_\tau, Q] = \sum_{c=0}^{\infty} \mathbb{P}[h_{\tau:\tau+1}|C = c, \tilde{h}_\tau, Q] = \sum_{c=0}^{\infty} \mathbb{P}[a \notin T \mid \tilde{h}_\tau, Q]^c. \tag{B-23}$$

Combining (B-22) and (B-23) yields

$$\begin{aligned} l_{\tau+1} &= l_\tau \frac{\sum_{c=0}^{\infty} \mathbb{P}\left[a \notin T \mid q_\tau, Q=1\right]^c \mathbb{P}\left[\tilde{a}_\tau \mid q_\tau, Q=1\right]}{\sum_{c=0}^{\infty} \mathbb{P}\left[a \notin T \mid q_\tau, Q=0\right]^c \mathbb{P}\left[\tilde{a}_\tau \mid q_\tau, Q=0\right]} = l_\tau \frac{\frac{\mathbb{P}[\tilde{a}_\tau|q_\tau,Q=1]}{1-\mathbb{P}[a\notin T|q_\tau,Q=1]}}{\frac{\mathbb{P}[\tilde{a}_\tau|q_\tau,Q=0]}{1-\mathbb{P}[a\notin T|q_\tau,Q=0]}} \\ &= l_\tau \frac{\frac{\mathbb{P}[\tilde{a}_\tau|q_\tau,Q=1]}{\mathbb{P}[a\in T|q_\tau,Q=1]}}{\frac{\mathbb{P}[\tilde{a}_\tau|q_\tau,Q=0]}{\mathbb{P}[a\in T|q_\tau,Q=0]}} \overset{(a)}{=} l_\tau \frac{\mathbb{P}\left[\tilde{a}_\tau \mid \tilde{a}_\tau \in T, q_\tau, Q=1\right]}{\mathbb{P}\left[\tilde{a}_\tau \mid \tilde{a}_\tau \in T, q_\tau, Q=0\right]} = l_\tau \, \tilde{Z}\left(\tilde{a}_\tau \mid l_\tau\right), \end{aligned}$$

where (a) follows from the definition of conditional probability. Also note that, given $Q = 0$, $\tilde{Z}(\cdot \mid l_\tau)$ has mean 1, which implies that $\{l_\tau\}_{\tau=1}^{\infty}$ is a martingale. Therefore, using $l_\tau \geq 0$ along with the martingale convergence theorem, we conclude that $l_\tau \to l_\infty$ almost surely. The rest of the proof is identical to that of Theorem 1 and Theorem 2. This completes the proof. ■

B.3.4 Extension to Non-Uniform Priors and Formal Definition of Uniform Improper Prior

In this subsection, we show that our characterization of learning and its speed for full history extend to a setting with non-uniform prior over the number of customers and then formalize the uniform improper priors.

We use the same notations as Appendix B.3.3. Following the description of purchase decision in Section 2, the equilibrium purchase decision of a user in a Bayes Nash equilibrium depends on $q_\tau : \Omega_\tau \times \mathbf{R} \to [0,1]$ which is a mapping from the information provided by the rating system to τ-th customer and the review decision strategy profile of all customers to a belief about the true quality. Customers do not observe any other event and have some belief over the times at which there have been unobservable actions. Their prior, however, is that the index of each customer (with respect to the calendar time) in the sequence of customers is geometric with some rate parameter $\lambda > 0$. The memoryless property of geometric distribution implies that for any customer the number of customers between her and the previous action in T is geometric with rate parameter $\lambda > 0$. The following is an extension of Proposition B-3 when the belief of customers about their index in the sequence is geometric. We use the following notation:

$$\tilde{\pi}(i; F_{\theta,\zeta}, Q, q, T) = \mathbb{P}_{\theta,\zeta}\left[a \in T_i \mid q, Q\right] \, \mathbb{E}_{c\sim\text{geo}(\lambda)}\left[\mathbb{P}_{\theta,\zeta}\left[a \notin T \mid q, Q\right]^c\right], \quad \text{for all } Q \in \{0,1\}, q \in [0,1]. \tag{B-24}$$

Proposition B-4. *Suppose that Assumptions 1 and 2 hold, and that customers only observe actions in T and have a geometric prior over their index in the sequence of customers. Then, there is again complete* *learning. Moreover, for $Q = 0$, we almost surely have*

$$\lim_{\tau \to \infty} \frac{1}{\tau} \log q_\tau = -D\left(\tilde{\boldsymbol{\pi}}(F_{\theta,\zeta}, Q = 0, q = 0, T) || \tilde{\boldsymbol{\pi}}(F_{\theta,\zeta}, Q = 1, q = 0, T)\right),$$

and for $Q = 1$, we almost surely have

$$\lim_{\tau \to \infty} \frac{1}{\tau} \log(1 - q_\tau) = -D\left(\tilde{\boldsymbol{\pi}}(F_{\theta,\zeta}, Q = 1, q = 1, T) || \tilde{\boldsymbol{\pi}}(F_{\theta,\zeta}, Q = 0, q = 1, T)\right).$$

Proof: We present the proof for $Q = 0$ (the proof for $Q = 1$ follows from the same argument). We let $\lambda > 0$ denote the geometric prior of each customer over her index in the sequence. Recall that τ is the subsequence of time at which there was an action in T and $\tilde{h}_\tau = \{\tilde{a}_1, \ldots, \tilde{a}_{\tau-1}\}$ denotes the available history after $\tau - 1$ actions in T. Similar to the proof of Theorem 1, we define random variable $\tilde{Z}(\cdot \mid l_\tau) = \frac{\tilde{\pi}(\cdot; F_{\theta,\zeta}, Q=1, q_\tau, T)}{\tilde{\pi}(\cdot; F_{\theta,\zeta}, Q=0, q_\tau, T)}$ where $q_\tau = \frac{l_\tau}{l_\tau + 1}$. We first establish that $l_{\tau+1} = l_\tau \tilde{Z}(\cdot \mid l_\tau)$, and then we show l_τ forms a martingale. Note that between $\tau - 1$-th action in T and τ-th action in T, each customer has observed the same history and therefore has the same belief q_τ. Therefore, each one of them has taken an action $\tilde{a}$ that does not belong to T, independently with probability $\mathbb{P}\left[\tilde{a} \notin T \mid q_\tau, Q\right]$. Also, a new customer does not know the number of customers who joined the platform before her. Since geometric distribution is memory-less, conditioning on any number of customers that could have joined the platform before τ, the number of customers between $\tau - 1$-th and τ-th actions in T has a geometric distribution with rate parameter λ. We let $h_{\tau:\tau+1}$ denote the history of actions in between $\tau - 1$-th and τ-th actions in T. We next present a recursive characterization of the likelihood ratio corresponding to belief $\{q_\tau\}_{\tau=0}^\infty$. We have

$$\begin{aligned}
l_{\tau+1} &= \frac{\mathbb{P}[\tilde{h}_\tau, \tilde{a}_\tau | Q = 1]}{\mathbb{P}[\tilde{h}_\tau, \tilde{a}_\tau | Q = 0]} = \frac{\sum_{h_{\tau:\tau+1}} \mathbb{P}[\tilde{h}_\tau, h_{\tau:\tau+1}, \tilde{a}_\tau | Q = 1]}{\sum_{h_{\tau:\tau+1}} \mathbb{P}[\tilde{h}_\tau, h_{\tau:\tau+1}, \tilde{a}_\tau | Q = 0]} \\
&= \frac{\sum_{h_{\tau:\tau+1}} \mathbb{P}[\tilde{h}_\tau \mid Q = 1] \mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q = 1] \mathbb{P}[\tilde{a}_\tau | \tilde{h}_\tau, h_{\tau:\tau+1}, Q = 1]}{\sum_{h_{\tau:\tau+1}} \mathbb{P}[\tilde{h}_\tau \mid Q = 0] \mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q = 0] \mathbb{P}[\tilde{a}_\tau | \tilde{h}_\tau, h_{\tau:\tau+1}, Q = 0]} \\
&= l_\tau \times \frac{\sum_{h_{\tau:\tau+1}} \mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q = 1] \mathbb{P}[\tilde{a}_\tau | \tilde{h}_\tau, h_{\tau:\tau+1}, Q = 1]}{\sum_{h_{\tau:\tau+1}} \mathbb{P}[h_{\tau:\tau+1} | \tilde{h}_\tau, Q = 0] \mathbb{P}[\tilde{a}_\tau | \tilde{h}_\tau, h_{\tau:\tau+1}, Q = 0]} \\
&= l_\tau \times \frac{\mathbb{P}[\tilde{a}_\tau | q_\tau, Q = 1]}{\mathbb{P}[\tilde{a}_\tau | q_\tau, Q = 0]} \times \frac{\mathbb{E}_{c \sim \text{geo}(\lambda)}\left[\mathbb{P}_{\theta,\zeta}\left[a \notin T \mid q_\tau, Q = 1\right]^c\right]}{\mathbb{E}_{c \sim \text{geo}(\lambda)}\left[\mathbb{P}_{\theta,\zeta}\left[a \notin T \mid q_\tau, Q = 0\right]^c\right]} \\
&= l_\tau \times \frac{\tilde{\pi}(\tilde{a}_\tau; F_{\theta,\zeta}, Q = 1, q_\tau, T)}{\tilde{\pi}(\tilde{a}_\tau; F_{\theta,\zeta}, Q = 0, q_\tau, T)},
\end{aligned}$$

where the last two equalities follow from similar arguments to those of Proposition B-3 and the definition of conditional probability. Also note that, the probability of action $\tilde{a}_\tau$ is the probability of having an exponential number of actions outside of T followed by action $\tilde{a}_\tau$ which, given $Q = 0$, is $\tilde{\pi}(\tilde{a}_\tau; F_{\theta,\zeta}, Q = 0, q_\tau, T)$. Therefore, $\tilde{Z}(\cdot \mid l_\tau)$ has mean 1, showing that $\{l_\tau\}_{\tau=1}^\infty$ forms a martingale. Thus, using $l_\tau \geq 0$ along with the martingale convergence theorem, we conclude that $l_\tau \to l_\infty$ almost surely. The rest of the proof is identical to that of Theorem 1. The speed of learning follows from a similar argument to Theorem 2. ■

The uniform improper prior corresponds to the limit as $\lambda \to 0$. We next formally show this claim by establishing that the learning speed given in Proposition (B-3) is in fact the limit of the learning speed given in Proposition (B-4) as $\lambda \to 0$. For any $q \in [0,1]$ and a, using $\mathbb{P}[C = c] = (1-\lambda)^c \lambda$ for $c = 0, 1, \ldots$, we have

$$\begin{aligned}
\lim_{\lambda \to 0} \frac{\tilde{\pi}(a; F_{\theta,\zeta}, Q=1, q, T)}{\tilde{\pi}(a; F_{\theta,\zeta}, Q=0, q, T)} &= \lim_{\lambda \to 0} \frac{\mathbb{P}[a|q, Q=1]\, \mathbb{E}_{c\sim \text{geo}(\lambda)} \left[\mathbb{P}_{\theta,\zeta}\left[a \notin T \mid q, Q=1\right]^c\right]}{\mathbb{P}[a|q, Q=0]\, \mathbb{E}_{c\sim \text{geo}(\lambda)} \left[\mathbb{P}_{\theta,\zeta}\left[a \notin T \mid q, Q=0\right]^c\right]} \\
&\overset{(a)}{=} \frac{\mathbb{P}[a|q, Q=1]}{\mathbb{P}[a|q, Q=0]} \lim_{\lambda \to 0} \frac{\frac{\lambda}{1-(1-\lambda)\mathbb{P}_{\theta,\zeta}[a \notin T|q, Q=1]}}{\frac{\lambda}{1-(1-\lambda)\mathbb{P}_{\theta,\zeta}[a \notin T|q, Q=0]}} \\
&= \frac{\mathbb{P}[a|q, Q=1]}{\mathbb{P}[a|q, Q=0]} \frac{\frac{1}{\mathbb{P}_{\theta,\zeta}[a \in T|q, Q=1]}}{\frac{1}{\mathbb{P}_{\theta,\zeta}[a \in T|q, Q=0]}} \\
&= \frac{\mathbb{P}[a|a \in T, q, Q=1]}{\mathbb{P}[a|a \in T, q, Q=0]} = \frac{\pi(a; F_{\theta,\zeta}, Q=1, q, T)}{\pi(a; F_{\theta,\zeta}, Q=0, q, T)},
\end{aligned}$$

where (a) follows by using the Moment Generating Function (MGF) of a geometric distribution. We note that in the limit as $\lambda \to 0$, even though the calendar time of a user's arrival does not have a proper distribution, the likelihood ratio (and therefore the belief) is well-defined as we proved in the previous equation.

Finally, we note that by applying the same argument of Proposition B-4 for summary statistics, Lemma A-1 continues to hold when customers have a geometric prior over their index in the sequence of customers. In particular, the characterization of Lemma A-1 and those of Theorems 3 and 4 continue to hold by replacing $\pi(i; F_{\theta,\zeta}, Q, q, T)$ by $\tilde{\pi}(i; F_{\theta,\zeta}, Q, q, T)$ as defined in (B-24).

B.3.5 General Summary Statistics

Here, we consider a general summary statistics and show how our results on complete learning and the speed of learning generalizes. A (general) summary statistics is a function that maps the fraction of each of the reviews to a vector in $\mathbb{R}^m$. That is $f : \Delta^k \to \mathbb{R}^m$, where

$$\Delta^k = \left\{ (x_1, \ldots, x_k) \in [0,1]^k \;:\; \sum_{i=1}^k x_i = 1 \right\},$$

and $k = \underline{K} + \overline{K} + 1$.

Definition 5 (Separation)**.** A general rating system with summary statistics $f : \Delta^k \to \mathbb{R}^m$ satisfies the strict separation condition if there exists $i \in [k]$ and $j \in [m]$ such that the range of the j-th dimension of $f(\mathbf{x}_{-i}, x_i = \pi(i; F_{\theta,\zeta}, Q=0, q))$ and $f(\mathbf{x}_{-i}, x_i = \pi(i; F_{\theta,\zeta}, Q=1, q))$ (as functions of q) are separate. Formally, we have either

$$\min_{q \in [0,1], \mathbf{x}_{-i}} \left[f(\mathbf{x}_{-i}, x_i = \pi(i; F_{\theta,\zeta}, Q=0, q))\right]_j > \max_{q \in [0,1], \mathbf{x}_{-i}} \left[f(\mathbf{x}_{-i}, x_i = \pi(i; F_{\theta,\zeta}, Q=1, q))\right]_j$$

or

$$\min_{q\in[0,1],\mathbf{x}_{-i}} \left[f(\mathbf{x}_{-i}, x_i = \pi(i; F_{\theta,\zeta}, Q = 1, q))\right]_j > \max_{q\in[0,1],\mathbf{x}_{-i}} \left[f(\mathbf{x}_{-i}, x_i = \pi(i; F_{\theta,\zeta}, Q = 0, q))\right]_j$$

Theorem B-1. *Suppose Assumptions 1 and 2 hold. Consider a summary statistics $f : \Delta^k \to \mathbb{R}^m$ and suppose f is continuous and satisfies the strict separation condition. Then complete learning occurs. Moreover, for $Q = 0$ we almost surely have*

$$\lim_{t\to\infty} \frac{1}{t} \log q_t = -\min_{\mathbf{p}\in\mathcal{E}_0} D(\mathbf{p}||\mathbf{q}),$$

where

$$\mathcal{E}_0 = \{(x_1, \dots, x_k) \in \Delta^k \ : \ f(x_1, \dots, x_k) = f(\pi(1; F_{\theta,\zeta}, q = 0, Q = 0), \dots, \pi(k; F_{\theta,\zeta}, q = 0, Q = 0))\},$$

and $\mathbf{q} = (\pi(1; F_{\theta,\zeta}, q = 1, Q = 1), \dots, \pi(k; F_{\theta,\zeta}, q = 1, Q = 1))$. *For $Q = 1$ we almost surely have*

$$\lim_{t\to\infty} \frac{1}{t} \log(1 - q_t) = -\min_{\mathbf{p}\in\mathcal{E}_1} D(\mathbf{p}||\mathbf{q}),$$

where

$$\mathcal{E}_1 = \{(x_1, \dots, x_k) \in \Delta^k \ : \ f(x_1, \dots, x_k) = f(\pi(1; F_{\theta,\zeta}, q = 1, Q = 1), \dots, \pi(k; F_{\theta,\zeta}, q = 1, Q = 1))\},$$

and $\mathbf{q} = (\pi(1; F_{\theta,\zeta}, q = 0, Q = 0), \dots, \pi(k; F_{\theta,\zeta}, q = 0, Q = 0))$.

Proof: The complete learning follows from a similar argument to Theorem 3. In particular, with strict separation, there exists $j \in \{1, \dots, m\}$ such that the j-th dimension of the summary statistics for $Q = 0$ is separated from $Q = 1$ and therefore complete learning happens. We next derive the speed of learning for $Q = 0$, noting that a similar argument provides the speed of learning for $Q = 1$. We use Sanov's Theorem (Cover and Thomas [2012, Chapter 11]) stated below:
Let Q be a probability distribution on finite set $\mathcal{X} = \{1, \dots, k\}$. Also, let $\mathcal{E}$ be a (closed) set of probability distributions on $\mathcal{X}$, then we have

$$\lim_{t\to\infty} \frac{1}{t} \log \left(\sum_{\substack{n_1,\dots,n_k \\ \sum_{i=1}^k n_i = t \\ \left(\frac{n_1}{t},\dots,\frac{n_k}{t}\right)\in\mathcal{E}}} \binom{\tau}{n_1, \dots, n_K} \left(\prod_{i=1}^k Q(i)^{n_i} \right) \right) = -\min_{P\in\mathcal{E}} D(P||Q),$$

Similar to the proof of Theorem 4, letting $\mathcal{H}_t$ indicate the set of all histories for which the summary statistics $(n_1, \dots, n_k)$ is such that $f\left(\frac{n_1}{t}, \dots, \frac{n_k}{t}\right) = f(\pi(1; F_{\theta,\zeta}, 0, 0), \dots, \pi(k; F_{\theta,\zeta}, 0, 0))$. We also define the set of summary statistics $\mathcal{E}_0$ as

$$\mathcal{E}_0 = \Big\{(x_1, \ldots, x_k) \in \Delta^k \ : \ f(x_1, \ldots, x_k) = f(\pi(1; F_{\theta,\zeta}, 0, 0), \ldots, \pi(k; F_{\theta,\zeta}, 0, 0))\Big\}.$$

The continuity of f implies that $\mathcal{E}_0$ is closed and therefore we can use Sanov's theorem for it. For the speed of learning, we can write

$$\begin{aligned}
\lim_{t\to\infty} \frac{1}{t} \log\left(\frac{q_t}{1-q_t}\right) &= \lim_{t\to\infty} \frac{1}{t} \log\left(\frac{\sum_{h_t\in\mathcal{H}_t} \mathbb{P}\left[h_t \mid Q=1\right]}{\sum_{h_t\in\mathcal{H}_t} \mathbb{P}\left[h_t \mid Q=0\right]}\right) \\
&= \lim_{t\to\infty} \frac{1}{t} \log\left(\frac{\sum_{\substack{n_1,\ldots,n_k \\ \sum_{j=1}^{k} n_j = t \\ (n_1/t,\ldots,n_k/t)\in\mathcal{E}_0}} \binom{t}{n_1,\ldots,n_k} \left(\prod_{j=1}^{k} \pi_1(j,1)^{n_j}\right)}{\sum_{\substack{n_1,\ldots,n_k \\ \sum_{j=1}^{K} n_j = t \\ (n_1/t,\ldots,n_k/t)\in\mathcal{E}_0}} \binom{\tau}{n_1,\ldots,n_k} \left(\prod_{j=1}^{k} \pi_0(j,0)^{n_j}\right)}\right) \\
&= -\min_{(a_1,\ldots,a_k)\in\mathcal{E}_0} D\left((a_1,\ldots,a_k) \,||\, (\pi_1(1,1),\ldots,\pi_1(k,1))\right),
\end{aligned}$$

where we used Sanov's Theorem for the set $\mathcal{E}_0$ to obtain the limit. ■

A special case of a general summary statistics is average summary statistics.

Recall, as discussed in the text, that several online platforms report an average score of past reviews rather than depicting detailed fractions of reviews that fall in different categories. In this subsection, we show that this type of average summary statistics leads to slower learning than their detailed counterpart. More formally, consider a rating system with review options $\{1, \ldots, k\}$, leading to review decisions:

$$a_\tau = \begin{cases} 1, & \theta_t + \mathbb{E}[\zeta_t] + q_t - p \geq 0,\ \theta_t + \zeta_t + Q - p < \lambda_1 \\ i, & \theta_t + \mathbb{E}[\zeta_t] + q_t - p \geq 0,\ \lambda_{i-1} \leq \theta_t + \zeta_t + Q - p < \lambda_i, 2 \leq i \leq k-1 \\ k, & \theta_t + \mathbb{E}[\zeta_t] + q_t - p \geq 0,\ \theta_t + \zeta_t + Q - p \geq \lambda_{k-1}. \end{cases}$$

A rating system is said to have an average summary statistics if it reports the number $S_t = \frac{1}{t}\sum_{s=1}^{\tau} a_s$, where t is the number of reviews. In the language of a general summary statistics, the function f is defined as $f(x_1, \ldots, x_k) = \sum_{j=1}^{k} j x_j$.

To distinguish the rating systems we have analyzed so far (which report fractions of users that have left various different reviews) from those with average statistics, we refer to them as "rating systems with vector summary statistics."

Corollary B-1. *Suppose Assumptions 1 and 2 hold, and for all $i = 2, \ldots, k$, $\sum_{j=i}^{k} \pi(j; F_{\theta,\zeta}, 1, q)$ and $\sum_{j=i}^{k} \pi(j; F_{\theta,\zeta}, 0, q)$ are separated, i.e.,*

$$\min_q \sum_{j=i}^{k} \pi(j; F_{\theta,\zeta}, 1, q) > \max_q \sum_{j=i}^{k} \pi(j; F_{\theta,\zeta}, 0, q), \quad i = 2, \ldots, k. \tag{B-25}$$

Then complete learning occurs under average summary statistics, i.e., $q_\tau \to Q$ almost surely, but the speed of learning is slower under average statistics than with vector summary statistics.

Corollary B-1 establishes that the speed of learning for rating systems with average summary statistics is always slower than the comparable rating system with vector summary statistics.
Proof of Corollary B-1: The proof of the complete learning is similar to Theorem 3. Similar to Theorem B-1, for $Q = 0$, we can write the speed of learning as

$$\min_{\substack{a_1,\dots,a_k \\ \sum_{j=1}^k a_j = 1 \\ \sum_{j=1}^k j\, a_j = E_0}} D\left((a_1,\dots,a_k)\,||\,(\pi_1(1,1),\dots,\pi_1(k,1))\right),$$

where $E_0 = \sum_{i=1}^k i\, \pi(i; F_{\theta,\zeta}, 0, 0)$. This speed of learning is (weakly) smaller than

$$D\left((\pi_0(1,0),\dots,\pi_0(k,0))\,||\,(\pi_1(1,1),\dots,\pi_1(k,1))\right),$$

which is the speed of learning for vector summary statistics. ■

We next provide a comparative statics for summary statistics analogue to Proposition 3.

Proposition B-5. *Consider a rating system with summary statistics $f : \Delta^k \to \mathbb{R}^m$ and non-injective function $g : \mathbb{R}^m \to \mathbb{R}^{m'}$ and let $h = g \circ f$ be a coarser summary statistics. Suppose that complete learning happens with both summary statistics. The speed of learning with summary statistics h is smaller than summary statistics f.*

Proof of Proposition B-5: We give the proof for $Q = 0$ (the proof for $Q = 1$ is identical). The speed of learning with summary statistics h is

$$\min_{(a_1,\dots,a_k) \in \mathcal{E}_0^h} D\left((a_1,\dots,a_k)\,||\,(\pi_1(1,1),\dots,\pi_1(k,1))\right)$$

where $\mathcal{E}_0^h = \{(x_1,\dots,x_k) \in \Delta^k \,:\, g(f(x_1,\dots,x_k)) = g(f(\pi(1; F_{\theta,\zeta}, 0, 0),\dots,\pi(k; F_{\theta,\zeta}, 0, 0)))\}$. The speed of learning with summary statistics f is

$$\min_{(a_1,\dots,a_k) \in \mathcal{E}_0^f} D\left((a_1,\dots,a_k)\,||\,(\pi_1(1,1),\dots,\pi_1(k,1))\right)$$

where $\mathcal{E}_0^f = \{(x_1,\dots,x_k) \in \Delta^k \,:\, f(x_1,\dots,x_k) = f(\pi(1; F_{\theta,\zeta}, 0, 0),\dots,\pi(k; F_{\theta,\zeta}, 0, 0))\}$. The proposition follows from the fact that $\mathcal{E}_0^h \subseteq \mathcal{E}_0^f$. ■

B.3.6 The Speed of Learning in terms of the Total Number of Customers

The speed of learning in Theorem 4 and in Proposition B-3 when we considered full history without the no-purchase and no-review decisions is with respect to index τ, the number of customers who left a review. The next corollary shows that once we have the speed of learning with respect to τ, it is straightforward to derive the speed of learning with respect to calendar time.

Corollary B-2. *Consider a rating system in which customers only observe actions in T. Suppose that they have uniform priors over the times at which actions that are not in T have taken place, and that the conditions in Theorems 1 or 3 are satisfied and there is complete learning. Then for $Q = 0$ (and similarly for $Q = 1$) and full history we almost surely have*

$$\lim_{t\to\infty} \frac{1}{t} \log q_t = -\left(\sum_{a\in T} \pi(a; F_{\theta,\zeta}, Q = 0, q = 0)\right) D\left(\boldsymbol{\pi}(F_{\theta,\zeta}, Q = 0, q = 0, T) || \boldsymbol{\pi}(F_{\theta,\zeta}, Q = 1, q = 0, T)\right),$$

and with summary statistics we almost surely have

$$\lim_{t\to\infty} \frac{1}{t} \log q_t = -\left(\sum_{a\in T} \pi(a; F_{\theta,\zeta}, Q = 0, q = 0)\right) D\left(\boldsymbol{\pi}(F_{\theta,\zeta}, Q = 0, q = 0, T) || \boldsymbol{\pi}(F_{\theta,\zeta}, Q = 1, q = 1, T)\right).$$

Unsurprisingly, Corollary B-2 establishes that the speed of learning with respect to the total number of customers has to be scaled down with the probability of taking an action in the set T given $q = Q$, i.e., $\sum_{a\in T} \pi(a; F_{\theta,\zeta}, Q, q = Q)$, and is thus slower. In the special case where all customers purchase and leave a review, we have $\sum_{a\in T} \pi(a; F_{\theta,\zeta}, Q, q = Q) = 1$ and the expressions in Corollary B-2 are identical to those in Theorems 2 or 4.

Proof of Corollary B-2: We present the proof for the full history and $Q = 0$ and the proof for summary statistics follows similarly. For the full history case using Theorem 4, the speed of learning is

$$\lim_{\tau\to\infty} \frac{1}{\tau} \log q_\tau = -D\left(\boldsymbol{\pi}(F_{\theta,\zeta}, Q = 0, q = 0, T) || \boldsymbol{\pi}(F_{\theta,\zeta}, Q = 1, q = 0, T)\right), \tag{B-26}$$

where τ indicates the number of actions in the set T. Now consider the original sequence of customers indexed by t. At time t, we let $\tau(t)$ be the number of reviews in the set T which is given by $\tau(t) = \sum_{i=1}^{t} X_i$, where for all $i = 1, \ldots, t$ we have $X_i \in \{0, 1\}$ with $\mathbb{P}[X_i = 1] = \mathbb{P}[a_i \in T \mid q_{\tau(i)}, Q = 0]$. Since belief q_τ converges almost surely to the true quality $Q = 0$, for any ϵ there exists t_0 such that for all $i \geq t_0$ we have $q_{\tau(i)} \leq \epsilon$ (note that $\tau(i)$ goes to infinity almost surely). Therefore, at any $t \geq t_0$ the random sequence $t_0 + \sum_{i=t_0}^{t} Y_i$ where $Y_i \in \{0, 1\}$ with $\mathbb{P}[Y_i = 1] = \max_{q\in[0,\epsilon]} \mathbb{P}[a \in T \mid q, Q = 0]$ first-order stochastically dominates $\tau(t)$. Moreover, using strong law of large numbers we have

$$\lim_{t\to\infty} \frac{1}{t}\left(t_0 + \sum_{i=t_0}^{t} Y_i\right) = \max_{q\in[0,\epsilon]} \mathbb{P}[a \in T \mid q, Q = 0], \quad \text{almost surely,}$$

establishing that

$$\limsup_{t\to\infty} \frac{1}{t}\tau(t) \leq \max_{q\in[0,\epsilon]} \mathbb{P}[a \in T \mid q, Q = 0], \quad \text{almost surely.} \tag{B-27}$$

Similarly, at any $t \geq t_0$ the random sequence $t_0 + \sum_{i=t_0}^{t} Z_i$ where $Z_i \in \{0, 1\}$ with $\mathbb{P}[Z_i = 1] = \min_{q\in[0,\epsilon]} \mathbb{P}[a \in T \mid q, Q = 0]$ is first-order stochastically dominated by $\tau(t)$. Therefore, using strong law of large numbers we obtain

$$\liminf_{t\to\infty} \frac{1}{t}\tau(t) \geq \min_{q\in[0,\epsilon]} \mathbb{P}[a \in T \mid q, Q = 0], \quad \text{almost surely.} \tag{B-28}$$

Since (B-27) and (B-28) hold for any ϵ, we obtain

$$\lim_{t\to\infty} \frac{1}{t}\tau(t) = \mathbb{P}[a \in T \mid q = 0, Q = 0], \quad \text{almost surely.} \tag{B-29}$$

Finally, from combining (B-26) and (B-29), and noting that every subsequence of a convergent sequence converges to the same limit as the original sequence, we obtain

$$\lim_{t\to\infty} \frac{1}{t} \log q_t = -\mathbb{P}[a \in T \mid q = 0, Q = 0] D\left(\boldsymbol{\pi}(F_{\theta,\zeta}, Q = 0, q = 0, T) || \boldsymbol{\pi}(F_{\theta,\zeta}, Q = 1, q = 0, T)\right),$$

completing the proof. ■

B.3.7 Effects of Targeted Information

In addition to the overall ratings of a product, platforms such as Amazon offer information about the reviews of groups of customers with certain characteristics. For instance, in reviews of a book at the intersection of climate science and economics, Amazon separately shows reviews among customers who are interested in economics as well as reviews among customers who are interested in climate science. We refer to this type of additional information as "targeted information."

To study the effects of targeted information we consider M types of customers where the valuations of customers of type $i \in [M]$ are drawn i.i.d. from distributions $f_\theta^{(i)}$ and $f_\zeta^{(i)}$ for θ and ζ, respectively. We also let ρ_i denote the probability that a new customer belongs to group i for $1 \leq i \leq M$. The platform knows which group of users have left a particular review and can potentially provide this extra information to users.

Our next result shows that additional targeted information always increases the speed of learning.

Proposition B-6. *Consider a rating system with either full history or summary statistics and suppose that there is complete learning. Then the speed of learning is always faster with targeted information.*

The intuition for faster learning with targeted information is related to the selection effect. Recall that the selection effect makes it more difficult for customers to learn from observed reviews. Targeted information increases customers' ability to distinguish the true quality from its opposite, because with additional information they can more successfully filter out the selection effect. Note, however, that for this result targeted information has to be provided in addition to the baseline ratings. If the platform reports only information about the reviews of the subgroup of customers (and suppresses information about reviews from other subgroups), this can slow down learning relative to the baseline without targeted information.

Proof of Proposition B-6: For any $i \in [m]$, we let $\pi(a; F_{\theta,\zeta}^{(i)}, Q, q)$, denote the probability of action a

when the cumulative distribution of θ and ζ is $F^{(i)}_{\theta,\zeta}$. We now present the proof for $Q = 0$ (the proof for $Q = 1$ follows from a similar argument).
Full history: We have $m\times(\overline{K}+\underline{K}+1)$ total possible actions, i.e., $\{1,\ldots,m\}\times\left(\{-\underline{K},\ldots,\overline{K}\}\cup\{N\}\right)$. Therefore, Theorem 2 shows that the speed of learning is

$$\sum_{i=1}^{m}\sum_{a\in\{-\underline{K},\ldots,\overline{K}\}\cup\{N\}}\rho_i\pi(a;F^{(i)}_{\theta,\zeta},Q=0,q=0)\log\left(\frac{\rho_i\pi(a;F^{(i)}_{\theta,\zeta},Q=0,q=0)}{\rho_i\pi(a;F^{(i)}_{\theta,\zeta},Q=1,q=0)}\right)$$
$$=\sum_{i=1}^{m}\rho_i D\left(\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=0,q=0)||\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=1,q=0)\right).$$

We next show that the speed of learning is higher than the case without types. Using convexity of KL divergence and Jensen's inequality, we obtain

$$D\left(\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=0,q=0)||\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=1,q=0)\right)$$
$$\le\sum_{i=1}^{m}\rho_i D\left(\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=0,q=0)||\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=1,q=0)\right),$$

which completes the proof of this part.
Summary statistics: Using Theorem 4, for each type $i\in[m]$ the speed of learning is given by

$$\lim_{t_i\to\infty}\frac{1}{t_i}\log l_{t_i}=-D\left(\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=0,q=0)||\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=1,q=1)\right),\tag{B-30}$$

where t_i denotes the sequence of customers that belong to type i. Each customer that joins the platforms observes the summary statistics of all of m rating systems and forms a belief, which using Bayes' rule can be written as $l_t=\prod_{i=1}^{m}l_{t_i}$. Using (B-30) together with $\frac{t_i}{t}\to\rho_i$ almost surely (using SLLN), we obtain $\lim_{t\to\infty}\frac{1}{t}\log l_t=-\sum_{i=1}^{m}\rho_i D\left(\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=0,q=0)||\boldsymbol{\pi}(F^{(i)}_{\theta,\zeta},Q=1,q=1)\right)$. Again, using Jensen's inequality completes the proof. ■

B.3.8 Platform Decisions: Maximizing Participation

We now outline a simple extension of where the platform chooses the rating system and assume that its objective is to maximize participation.

Let us first describe the joining decision of customers. Suppose that one customer arrives at each time and is denoted by the time of her arrival, $t\in\mathbb{N}$. Customer t faces an entry cost drawn independently from a continuous distribution F_c (representing the cost of time spent on the platform and/or the opportunity cost of using this platform instead of another medium). After observing this cost, the customer decides $j_t\in\{0,1\}$, designating whether she has joined the platform.

The material utility of customer t given in (1) is her "gross" utility (before subtracting the cost of joining the platform c_t, which is already sunk at this point). The "gross" utility from not purchasing the good is normalized to zero.

Before joining the platform, a customer only knows the price p, the cost c_t, and the rating system $\mathbf{\Omega}$ (not the realization of Ω_t). In particular, she does not know the number of people who have joined the platform before her and has an (improper) uniform prior on this number. Therefore, in a Bayes-Nash equilibrium customer t's entry decision is $j_t = 1$ if and only if $\mathbb{E}_{t,\Omega_t,Q,\theta_t,\zeta_t}[b_t\ (\theta_t + \zeta_t + Q - p)] - c_t \geq 0$, where $\mathbb{E}_{t,\Omega_t,Q,\theta_t,\zeta_t}$ is the expectation over all random variables in our model.

We assume that the platform does not know the true quality of the product either and for now its price is set by an outside vendor. Therefore, the only decision of the platform is the rating system, $\mathbf{\Omega}$. Suppose in this subsection that the platform's revenues come from fixed sales commissions, advertisements or other benefits that it derives from the participation of customers. This implies that the platform's objective is to maximize the number of customers joining. Since customers do not know the number of reviews in the system before they join and have uniform priors about the number of past users, they will all join with the same probability, and thus the platform's objective can be written as maximizing this probability:

$$\max_{\mathbf{\Omega}} \mathbb{P}_{\Omega,Q,\theta,\zeta,c}[j = 1].$$

The next proposition shows that in this case the platform maximizes the speed of learning.

Proposition B-7. *Given two rating systems I and II, the platform chooses the one with the greater speed of learning.*

With a rating system that guarantees faster learning, customers know that once they join the platform, they are less likely to make either type I or type II errors in their purchase decisions. This raises their expected utility from joining, and thus increases the joining probability. To see why the probabilities of both types of errors are decreasing in the speed of learning, consider first $Q = 0$. Then the problem for customers is to incorrectly buy the product thinking that it is high quality. If $\theta + \zeta - p \in (-1, 0)$, then with probability $\mathbb{P}_{t,\Omega_t}[q_t \geq p - \theta - \zeta \mid Q = 0]$ they purchase the product and suffer a loss of $\theta + \zeta - p$. This probability of incorrectly purchasing the product becomes smaller when q_t is closer to $Q = 0$. Conversely, suppose that $Q = 1$. Then the problem for customers is to fail to purchase the product when they should be doing so. If $\theta + \zeta - p \in (-1, 0)$, then with probability $\mathbb{P}_{t,\Omega_t}[q_t \leq p - \theta - \zeta \mid Q = 1]$ customers do not purchase the product and forgo a potential utility gain of $1 + \theta + \zeta - p$. The probability of such a mistake becomes smaller when q_t is closer to $Q = 1$. Thus in both cases faster learning ensures smaller probability of mistakes, and this encourages customers to join the platform.

Proof of Proposition B-7: We break the proof into two steps. In the first step we reformulate the joining probability and characterize its dependence on the speed of learning as a key component of it. In the second step we show that increasing speed of learning, increases the joining probability and hence platform's revenue.

Step 1: Maximizing joining probability is equivalent to maximizing

$$
\begin{aligned}
&\mathbb{E}_{\theta,\zeta}\left[\left(\frac{1}{2}-p+\theta+\zeta\right)\mathbf{1}\{\theta+\zeta-p\geq 0\}+\frac{1}{2}(1-p+\theta+\zeta)\mathbf{1}\{\theta+\zeta-p\in(-1,0)\}\right]\\
&+\frac{1}{2}\mathbb{E}_{\theta,\zeta}\left[(-(p-\theta-\zeta)\mathbb{P}_{t,\Omega_t}\left[q_t\geq p-\theta-\zeta\mid Q=0\right])\mathbf{1}\{\theta+\zeta-p\in(-1,0)\}\right]\\
&+\frac{1}{2}\mathbb{E}_{\theta,\zeta}\left[(-(1-(p-\theta-\zeta))\mathbb{P}_{t,\Omega_t}\left[q_t\leq p-\theta-\zeta\mid Q=1\right])\mathbf{1}\{\theta+\zeta-p\in(-1,0)\}\right]. \qquad \text{(B-31)}
\end{aligned}
$$

The platform's problem is to maximize $\mathbb{P}[j_t=1]$. This probability can be written as

$$
\mathbb{P}_c\left[j=1\right]=\mathbb{P}\left[\mathbb{E}_{Q,\theta,\zeta,\Omega,t}\left[b_t=1\left(\theta+\zeta+Q-p\right)\right]-c\geq 0\right]=F_c\left(\mathbb{E}_{Q,\theta,\zeta,\Omega,t}\left[(\theta+\zeta+Q-p)b_t\right]\right).
$$

Since $F_c(\cdot)$ is non-decreasing, the platform's problem is to maximize $\mathbb{E}_{Q,\theta,\zeta,\Omega,t}\left[(\theta+\zeta+Q-p)b_t\right]$. We can further write

$$
\begin{aligned}
&\mathbb{E}_{Q,\theta,\zeta,\Omega,t}\left[(\theta+\zeta+Q-p)\,b_t\right]\\
&=\mathbb{E}_{\theta,\zeta}\left[\left(\frac{1}{2}-p+\theta+\zeta\right)\mathbf{1}\{\theta+\zeta-p\geq 0\}\frac{1}{2}(1-p+\theta+\zeta)\mathbf{1}\{\theta+\zeta-p\in(-1,0)\}\right]\\
&+\frac{1}{2}\mathbb{E}_{\theta,\zeta}\left[(-(p-\theta-\zeta)\mathbb{P}_{t,\Omega}\left[q_t\geq p-\theta-\zeta\mid Q=0\right])\mathbf{1}\{\theta+\zeta-p\in(-1,0)\}\right]\\
&+\frac{1}{2}\mathbb{E}_{\theta,\zeta}\left[(-(1-(p-\theta-\zeta))\mathbb{P}_{t,\Omega}\left[q_t\leq p-\theta-\zeta\mid Q=1\right])\mathbf{1}\{\theta+\zeta-p\in(-1,0)\}\right].
\end{aligned}
$$

This completes the proof of the first step.

Step 2: If system I has a higher speed than system II, then the platform's objective from using system I is higher.

For the second step, note that the terms that depend on the rating system are

$$
\begin{aligned}
&\frac{1}{2}\mathbb{P}_{\theta,\zeta}\left[p-\theta-\zeta\in[0,1]\right]\mathbb{E}_{\theta,\zeta}\left[-\left(p-\theta-\zeta\right)\mathbb{P}_{t,\Omega}\left[q_t\geq p-\theta-\zeta\mid Q=0\right]\mid p-\theta-\zeta\in[0,1]\right]\\
&+\frac{1}{2}\mathbb{P}_{\theta,\zeta}\left[p-\theta-\zeta\in[0,1]\right]\mathbb{E}_{\theta,\zeta}\left[-\left(1-(p-\theta-\zeta)\right)\mathbb{P}_{t,\Omega}\left[q_t\leq p-\theta-\zeta\mid Q=1\right]\mid p-\theta-\zeta\in[0,1]\right].
\end{aligned}
$$

Also, note that if Ω^I has a higher speed of learning than Ω^{II}, then since customers have an improper uniform prior on the number of previous customers (before joining the platform), we obtain $\mathbb{P}_{t,\Omega^I,\Omega^{II}}\left[q_t^I\leq q_t^{II}\mid Q=0\right]=1$, and $\mathbb{P}_{t,\Omega^I,\Omega^{II}}\left[q_t^I\geq q_t^{II}\mid Q=1\right]=1$. This leads to

$$
\begin{aligned}
\mathbb{P}_{t,\Omega^I}\left[q_t^I\geq p-\theta-\zeta\mid Q=0\right]&\leq\mathbb{P}_{t,\Omega^{II}}\left[q_t^{II}\geq p-\theta-\zeta\mid Q=0\right],\\
\mathbb{P}_{t,\Omega^I}\left[q_t^I\leq p-\theta-\zeta\mid Q=1\right]&\leq\mathbb{P}_{t,\Omega^{II}}\left[q_t^{II}\leq p-\theta-\zeta\mid Q=1\right].
\end{aligned}
$$

Since $p-\theta-\zeta\in[0,1]$, we have $-\left(p-\theta-\zeta\right)\leq 0$ and $-\left(1-\left(p-\theta-\zeta\right)\right)\leq 0$, which establishes the joining probability of the rating system Ω^I is higher than the joining probability of the rating system with Ω^{II}. This completes the proof. ■

B.3.9 Platform Decisions: Pricing

We now analyze the platform's pricing decisions, chosen to maximize revenue. We will analyze this setup with both myopic pricing and dynamic pricing and then establish conditions under which learning happens and characterize the speed of learning.

Myopic Pricing. In this case, at time t the firm chooses a price p_t to maximize her current revenue. Hence, the optimal (revenue-maximizing) price at time t can be written as

$$p_t = \arg\max_{p \geq 0} p\,(1 - F_\theta(p - q_t - \mathbb{E}[\zeta])),$$

where q_t is the belief at time t. We next show that even though the evolution of beliefs changes because of the changes in price, there is still complete learning and its speed is the same as before.

Proposition B-8. *Suppose Assumptions 1 and 2 hold for all p in the set $\{p\ :\ \text{there exists } q \in [0,1] \text{ s.t. } p \in \arg\max_x x\,(1 - F_\theta(x - q - \mathbb{E}[\zeta]))\}$. Then under full history and myopic pricing, there is complete learning, that is, $q_t \to Q$ almost surely. Moreover, the speed of learning is the same as in Theorem 2.*

Under summary statistics and myopic learning, if the strict separation condition holds for all p in the set $\{p\ :\ \text{there exists } q \in [0,1] \text{ s.t. } p \in \arg_x x\,(1 - F_\theta(x - q - \mathbb{E}[\zeta]))\}$, then there is complete learning and the speed of learning is the same as in Theorem 4.

For a rating system with full history, the public belief is again a martingale, and an identical argument to that of Theorem 1 establishes complete learning. We also prove that price dynamics do not affect the speed of learning (because once the price sequence is close to its limit, it has a small impact on the distribution of reviews) and hence the speed of learning in Theorem 2 applies. We can also observe that the limiting price in the speed of learning expression is $p^* = \arg\max_p p(1 - F_\theta(p - Q - \mathbb{E}[\zeta]))$, where Q is the true quality. Similar results apply with summary statistics, enabling us to generalize Theorems 3 and 4 to this setting.

The next corollary enables us to characterize whether the convergence of the price sequence to its limit is from above or below depending on the underlying quality.

Corollary B-3. *Suppose that $\frac{f_\theta(p)}{1-F_\theta(p)}$ is non-decreasing in p. Then for $Q = 0$, we have $p_t \geq p^* \geq p_t - q_t$ and for $Q = 1$, we have $p_t \leq p^* \leq p_t + (1 - q_t)$, where $p^* = \arg\max_p p(1 - F_\theta(p - Q - \mathbb{E}[\zeta]))$.*

Therefore, when $Q = 0$ equilibrium price p_t converges to p^* from above, meaning that the firm starts with a higher price in the short run, reducing it as more information about low quality is revealed. Conversely, when $Q = 1$, the equilibrium price p_t converges to p^* from below — as more information about the product being high quality is accumulated, the price increases.

Dynamic Pricing. Suppose now that the platform chooses the price sequence $\{p_t\}_{t=1}^\infty$ to maximize its expected discounted revenue defined as

$$J(q_1) = \max_{\{p_t\}_{t=1}^\infty} \mathbb{E}_{\{\theta_s, \zeta_s\}_{s=1}^\infty} \left[(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} p_t b_t \right],$$

where $\delta \in (0,1)$ is the platform's discount factor and $q_1 = \frac{1}{2}$ is the belief of customers at the time $t = 1$. Note also that p_t can only depend on the history at time t.

The key observation is that in a rating system with full history, the platform and the customers have access to the same information about the history of actions and therefore the price p_t chosen by the platform does not convey any additional information to customers. This is not necessarily true under summary statistics, however, because the platform has access to the full history of reviews and thus knows more than customers about the underlying quality of the product. In this case, p_t may be a signal for quality. For this reason, our results for the case of summary statistics with dynamic pricing strategies will be somewhat weaker.

In a rating system with full history, the public belief q_t is a sufficient statistic for the pricing decision of the platform. Therefore, we can write the Bellman equation for the platform's value (discounted revenue) conditioning just on this public belief:

$$J(q) = \max_p p\left(1 - F_\theta(p - q - \mathbb{E}[\zeta])\right) + \delta \mathbb{E}_{\theta,\zeta}\left[J(q') \mid q\right],$$

where the first term is the current revenue of the platform given current public belief q, and the second term is the expectation of the discounted revenue from next period onwards given next period's public belief q' (conditional on current belief q).

Proposition B-9. *Consider a rating system with full history. Suppose Assumption 1 holds for all $p \in \left[0, \bar{\theta} + 1 + \mathbb{E}[\zeta]\right]$ (where $\bar{\theta}$ is the upper support of the distribution of θ), and Assumption 2 holds for $p = 0$. Then the equilibrium price sequence $\{p_t\}_{t=1}^{\infty}$ is such that in each time period there is a positive probability of purchase, and the price sequence converges to $p^* = \arg\max_p p(1 - F_\theta(p - Q - \mathbb{E}[\zeta]))$. Consequently, there is complete learning and the speed of learning is the same as in Theorem 2.*

Complete learning follows because the likelihood ratio forms a martingale and purchase probability is always non-zero. Zero purchase probability would generate neither any current revenue for the platform nor information for future users. Consequently, after a period in which there is zero probability of purchase, the continuation value of the platform would be the same as at the start of this period, and because of discounting, the objective of the platform would be strictly lower. Given positive purchase probabilities, a similar argument to the one in Theorem 1 establishes complete learning and its speed is determined as before.[18]

Our next result shows that even though under summary statistics, the platform may wish to adopt a more complex pricing strategy in order to transmit information to customers, it can approximately achieve the discounted revenue in Proposition B-9.

Proposition B-10. *Consider a rating system with summary statistics. Suppose Assumption 1 holds for all $p \in \left[0, \bar{\theta} + 1 + \mathbb{E}[\zeta]\right]$, Assumption 2 holds for $p = 0$, and $F_\theta(\cdot)$ is Lipschitz continuous. Then for*

[18] In this proposition, in contrast to Proposition B-8, we need to impose Assumption 1 for all prices $p \in \left[0, \bar{\theta} + 1 + \mathbb{E}[\zeta]\right]$, because the price set by the platform at any point in time need not be one that maximizes current revenue. Assumption 2, on the other hand, is only needed for $p = 0$, because the platform will never charge a price that will lead to zero probability of purchase after some history.

any $\epsilon > 0$, there exists a pricing strategy such that the beliefs of customers are identical to their beliefs in Proposition B-9 and the platform's discounted revenue (after any time t) is at most ϵ below the platform's discounted revenue in the rating system with full history (and the same review options).

Proposition B-10 follows by considering the following strategy: at any time t, the platform chooses a suboptimal price by zeroing out the decimal numbers in the profit-maximizing price with full history. This strategy decreases the current revenue of the platform, but by providing access to the full history to all future customers it enables the platform to replicate the allocation in Proposition B-9. This ensures that the evolution of customer beliefs is the same as in Proposition B-9. While Proposition B-10 shows that a platform with a rating system reporting summary statistics can achieve the same level of discounted revenue as under full history, it leaves open the question of whether it can actually achieve even higher revenues. A full analysis of this issue is beyond the scope of the current paper, in part because the answer depends how much commitment the platform has to a pricing strategy.

Proof of Proposition B-8: We present the proof with full history and summary statistics separately. In the proof of the results of this subsection, without loss of generality, we assume $\mathbb{E}[\zeta] = 0$.

Full history: The proof is similar to the proof of Theorem 1. In particular, since customers observe the full history and the optimal price at any time t, the likelihood ratio sequence $\{l_t\}_{t=1}^{\infty}$ still forms a martingale. Therefore, from the martingale convergence theorem, it converges to some l_∞. This in turn implies the optimal price $\{p_t\}_{t=1}^{\infty}$ also converges to some p_∞ in the set $\{p \,:\, p \in \arg\max_x x\,(1-F_\theta(x-q)) \,:\, q \in [0,1]\}$. We next show that given $Q = 0$, we have $l_\infty = 0$ with probability 1. Suppose $\mathbb{P}[l_\infty = l] > 0$ for some $l > 0$. The corresponding limiting price is the one that maximizes purchase probability given belief $\frac{l}{l+1}$; that is, $p \in \arg\max_x x\,(1 - F_\theta(x - \frac{l}{l+1}))$. With a similar argument to the proof of Theorem 1, when $Q = 0$, we must have $l = 0$, and thus $l_\infty = 0$ with probability one.

The proof of the speed of learning is also similar to the proof of Theorem 2. In particular, for any $\epsilon > 0$, since q_t converges to Q almost surely, there exists t_0 such that for $t \geq t_0$, we have $|p_t - p^*| \leq \epsilon$ with probability one, where p^* is the limiting price, $p^* \in \arg_x x\,(1 - F_\theta(x - Q))$. Using a similar argument to the proof of Theorem 2, we have that almost surely

$$\lim_{t\to\infty} \frac{1}{t} \log q_t \leq \max_{p \,:\, |p-p^*|\leq\epsilon} -D\left(\boldsymbol{\pi}^p(F_{\theta,\zeta}, Q=0, q=0) || \boldsymbol{\pi}^p(F_{\theta,\zeta}, Q=1, q=0)\right), \tag{B-32}$$

where $\boldsymbol{\pi}^p(F_{\theta,\zeta}, Q = 1, q = 0)$ is the probability of different actions when the price is p. Similarly, we have that almost surely

$$\lim_{t\to\infty} \frac{1}{t} \log q_t \geq \min_{p \,:\, |p-p^*|\leq\epsilon} -D\left(\boldsymbol{\pi}^p(F_{\theta,\zeta}, Q=0, q=0) || \boldsymbol{\pi}^p(F_{\theta,\zeta}, Q=1, q=0)\right). \tag{B-33}$$

Since (B-32) and (B-33) hold for any $\epsilon > 0$, we conclude that almost surely

$$\lim_{t\to\infty} \frac{1}{t} \log q_t = -D\left(\boldsymbol{\pi}^{p^*}(F_{\theta,\zeta}, Q=0, q=0) || \boldsymbol{\pi}^{p^*}(F_{\theta,\zeta}, Q=1, q=0)\right).$$

Summary statistics: The proof is similar to the proof of Theorem 3. In particular, under strict separation, we have $q_\tau \to Q$. The proof of the speed of learning follows from a similar argument to the full history case presented above. ■

Proof of Corollary B-3: We first establish the following claim.

Claim: For any q, let p_q be the solution to $p_q = \frac{1-F_\theta(p_q-q)}{f_\theta(p_q-q)}$. Then we have $p \geq p_q$ if and only if $(1-F_\theta(p-q)) - pf_\theta(p-q) \leq (1-F_\theta(p_q-q)) - p_q f_\theta(p_q-q) = 0$. This in turn establishes $p_q = \arg\max_p p\,(1-F_\theta(p-q))$.

Proof of the Claim: For $p > p_q$, we have $(1-F_\theta(p-q))-pf_\theta(p-q) \leq f_\theta(p-q)\left(\frac{1-F_\theta(p_q-q)}{f_\theta(p_q-q)} - p_q\right) = 0$, where we used $p_q \leq p$ and $\frac{1-F_\theta(p_q-q)}{f_\theta(p_q-q)} \geq \frac{1-F_\theta(p-q)}{f_\theta(p-q)}$ which holds since $\frac{1-F_\theta(p)}{f_\theta(p)}$ is non-increasing. Similarly, we can show that if $p_q \geq p$, then $(1-F_\theta(p-q))-pf_\theta(p-q) \geq 0$. Since for any p we have either $p_q \geq p$ or $p_q \leq p$, the claim follows.

We next proceed with the proof of the corollary. Suppose $Q = 0$. We first prove that $p_t \geq p^*$. Using the claim, since p_t maximizes $p(1-F_\theta(p-q_t))$, we have $p_t = \frac{1-F_\theta(p_t-q_t)}{f_\theta(p_t-q_t)} \geq \frac{1-F_\theta(p_t)}{f_\theta(p_t)}$, which in turn implies $(1-F_\theta(p_t)) - p_t f_\theta(p_t) \leq 0$. Again, using the claim, we conclude $p_t \geq p^*$. We next prove $p^* \geq p_t - q_t$. We have $(p^*+q_t) > (p^*+q_t) - q_t = \frac{1-F_\theta((p^*+q_t)-q_t)}{f_\theta((p^*+q_t)-q_t)}$, which in turn proves that $(1-F_\theta((p^*+q_t)-q_t)) - (p^*+q_t)f_\theta((p^*+q_t)-q_t) \leq 0$. Again, using the claim, we can conclude $p^* \geq p_t - q_t$.

Now suppose $Q = 1$. We first prove $p_t \leq p^*$. We have $p_t = \frac{1-F_\theta(p_t-q_t)}{f_\theta(p_t-q_t)} \leq \frac{1-F_\theta(p_t-1)}{f_\theta(p_t-1)}$, which in turn establishes $(1-F_\theta(p_t-1)) - p_t f_\theta(p_t-1) \geq 0$. Again, using the claim, we conclude that $p_t \leq p^*$. Analogously, we next prove $p^* \leq p_t + (1-q_t)$. We have $(p^*-1+q_t) < (p^*-1+q_t) + (1-q_t) = \frac{1-F_\theta((p^*-1+q_t)-q_t)}{f_\theta((p^*-1+q_t)-q_t)}$, which in turn implies $(1-F_\theta((p^*-1+q_t)-q_t)) - (p^*-1+q_t)f_\theta((p^*-1+q_t)-q_t) \geq 0$. Again, using the claim, we conclude that $p^*-1+q_t \leq p_t$, completing the proof of the corollary. ■

Proof of Proposition B-9: We first show that at each point in time the optimal pricing is such that purchase happens with positive probability. Suppose the contrary, i.e., at time t the price is such that no purchase happens with probability one. The belief in time period $t+1$ is then the same as the belief as time t. Therefore, we have $J(q_t) = \delta J(q_{t+1}) = \delta J(q_t)$, which is a contradiction. The rest of the proof follows by using the fact that the likelihood ratio $\{l_t\}_{t=1}^{\infty}$ is still a martingale. Using Assumption 2, at any time the purchase probability is non-zero with a non-negative price. Also, using Assumption 1, probability of "like" with both $Q = 0$ and $Q = 1$ is non-zero. Therefore a similar argument to that of Theorem 1 is sufficient to establish complete learning and then its speed can be characterized using the same steps as in Theorem 2. ■

Proof of Proposition B-10: We first prove that assuming $p(1-F_\theta(p-q))$ is Lipschitz continuous with parameter L, for a given $\epsilon > 0$ there exists $\eta \in \mathbb{R}$ such that for all non-negative $p, p' \leq M$ with $|p-p'| \leq \eta$, we have $|p\,(1-F_\theta(p-q)) - p'\,(1-F_\theta(p'-q))\,| \leq \epsilon, \forall q \in [0,1]$. For $\eta \leq \frac{\epsilon}{1+ML}$, we can write

$$
\begin{aligned}
&|p\,(1-F_\theta(p-q)) - p'\left(1-F_\theta(p'-q)\right)| \\
&= |p\,(1-F_\theta(p-q)) - p'\,(1-F_\theta(p-q)) + p'\,(1-F_\theta(p-q)) - p'\left(1-F_\theta(p'-q)\right)|
\end{aligned}
$$

$$\overset{(a)}{\leq} |p\,(1 - F_\theta(p-q)) - p'\,(1 - F_\theta(p-q))\,| + |p'\,(1 - F_\theta(p-q)) - p'\,\big(1 - F_\theta(p'-q)\big)\,|$$
$$\leq |p - p'|\,(1 - F_\theta(p-q)) + p'|F_\theta(p-q) - F_\theta(p'-q)| \overset{(b)}{\leq} \eta + ML\eta \leq \epsilon$$

where (a) follows from triangle inequality; and (b) follows because $F_\theta(\cdot)$ is Lipschitz and $p' \leq M$. Let $M = \max_{q\in[0,1]} \max_p p(1 - F_\theta(p-q))$ and let k be large enough such that $10^{-k} \leq \eta$.

For any t, let B_t be a sequence of numbers encoding the full history at time t. In particular, if there are K review options, then we use the first $\log_{10} K$ decimal points of B_t to represent a_1, the second $\log_{10} K$ decimal points of B_t to represent a_2 and so on. For any q_t and the corresponding optimal price with full history $p_t^{\mathrm{F}}(q_t)$ with $p' = \frac{\lfloor p_t^{\mathrm{F}}(q_t)10^{-(k-1)}\rfloor}{10^{k-1}} + 10^{-k}B_t$, we obtain

$$|p_t^{\mathrm{F}}(q_t)\left(1 - F_\theta(p_t^{\mathrm{F}}(q_t) - q_t)\right) - p'\left(1 - F_\theta(p' - q_t)\right)| \leq \epsilon.$$

Therefore, at any time t with this choice of price the current revenue of the platform is at most ϵ below the revenue of the platform with full history. Moreover, with the proposed rating system, at any time t the customers know the full history and also recognize that previous customers knew the full history as well. Therefore, all customers will have access to the public belief at all past dates, and hence purchase and review decisions will be as in the case with full history. ■