Online Appendix

Andrei Hagiu and Julian Wright

We provide detailed proofs of material that is not included in the main text of our paper "Data-enabled learning, network effects and competitive advantage".

A Viability of the losing firm

In the framework presented in the main paper, the losing firm remains active despite obtaining no revenue. This reflects our assumption that any fixed costs are sunk and there are no ongoing operational costs. In the case of pure within-user learning (Section 4), this was not an issue, since each firm could each sell to different consumers (or consumer segments) and so could play the role of an incumbent for some consumers and an entrant for others. With pure across-user learning (Section 3), this assumption was made in order to keep the analysis streamlined. As explained in the text, our results continue to hold when we extend our model in ways that allow the losing firm to make positive expected revenues in every period, so that it remains active even when facing some (sufficiently small) operating costs.

In this section, we discuss two such extensions. In both cases we obtain conditions for either firm to win that are very similar to what we have obtained in the main text, and are identical in the limit, thus showing that our baseline results are not knife-edge in nature. Finally, we consider what happens in our baseline model when we assume the losing firm exits because of some small fixed cost, showing that the cutoff for E to win still remains identical in this case.

A.1 Publicly available discovery

Suppose that in every period there is a fixed probability $\lambda > 0$ of a publicly available technological breakthrough that makes any previous advantage from data-enabled learning irrelevant, and that results in the two firms having expected PDV of profits equal to $\widetilde{V}_I > 0$ and $\widetilde{V}_E > 0$ from the perspective of the period in which the discovery is made. For instance, this would follow if there was an exogenous distribution of possible standalone values s'_I and s'_E for the two firms after the discovery, with positive probabilities for either firm to have the higher standalone value. Note, however, that s'_I and s'_E need not be constant over time, i.e. we can allow for new learning to occur after the discovery. The only requirements we impose on $\widetilde{V}_I$ and $\widetilde{V}_E$ is that they are both positive (e.g. due to sufficient uncertainty as to which firm will end up offering higher value) and do not depend on each firm's cumulated learning prior to discovery.

With probability $1 - \lambda$, the game proceeds as before. The method of proof follows exactly the same two-dimensional backwards induction as the proof of Proposition 1. However, the end point is now different. Once both firms are in state $(N_I, N_E) = (\overline{N}_I, \overline{N}_E)$ (assuming the technological breakthrough has not yet happened), there is no further learning, and the value function of firm i's profit in this case is given by

$$\begin{aligned} V^i(\overline{N}_I, \overline{N}_E) &= \max\left\{s_i - s_j + f_i(\overline{N}_i) - f_j(\overline{N}_j), 0\right\} + \delta\left(\lambda\widetilde{V}_i + (1-\lambda)\max\left\{s_i - s_j + f_i(\overline{N}_i) - f_j(\overline{N}_j), 0\right\}\right) \\ &\quad +\delta^2(1-\lambda)\left(\lambda\widetilde{V}_i + (1-\lambda)\max\left\{s_i - s_j + f_i(\overline{N}_i) - f_j(\overline{N}_j), 0\right\}\right) \\ &\quad +\delta^3(1-\lambda)^2\left(\lambda\widetilde{V}_i + (1-\lambda)\max\left\{s_i - s_j + f_i(\overline{N}_i) - f_j(\overline{N}_j), 0\right\}\right) + ... \\ &= \frac{\max\left\{s_i - s_j + f_i(\overline{N}_i) - f_j(\overline{N}_j), 0\right\}}{1-\delta(1-\lambda)} + \frac{\delta\lambda\widetilde{V}_i}{1-\delta(1-\lambda)}. \end{aligned}$$

As we go back each step in the induction process, we now have to take into account the probability that the discovery is realized and the two firms obtain $\widetilde{V}_I$ and $\widetilde{V}_E$, as above. In this way, it is straightforward to show that the modified condition for E to win (prior to the discovery) is that $s_E - s_I$ must exceed

$$\begin{aligned} \widetilde{\Delta}(N_I, N_E) &= (1-\delta(1-\lambda))\left(\sum_{j=0}^{\overline{N}_I - N_I} \delta^j (1-\lambda)^j f_I(N_I + j) - \sum_{j=0}^{\overline{N}_E - N_E} \delta^j (1-\lambda)^j f_E(N_E + j)\right) \\ &\quad + (\delta(1-\lambda))^{\overline{N}_I - N_I + 1} f_I\left(\overline{N}_I\right) - (\delta(1-\lambda))^{\overline{N}_E - N_E + 1} f_E\left(\overline{N}_E\right) \end{aligned}$$

and that the two firms' value functions are

$$V^I(N_I, N_E) = \begin{cases} \frac{\delta\lambda\widetilde{V}_I}{1-\delta(1-\lambda)} + \frac{s_I - s_E}{1-\delta(1-\lambda)} + \frac{\widetilde{\Delta}(N_I,N_E) - \delta(1-\lambda)\widetilde{\Delta}(N_I,N_E+1)}{(1-\delta(1-\lambda))^2} & \text{if} \quad s_E - s_I < \widetilde{\Delta}(N_I, N_E+1) \\ \frac{\delta\lambda\widetilde{V}_I}{1-\delta(1-\lambda)} + \frac{s_I - s_E + \widetilde{\Delta}(N_I,N_E)}{(1-\delta(1-\lambda))^2} & \text{if} \quad \widetilde{\Delta}(N_I, N_E+1) \le s_E - s_I < \widetilde{\Delta}(N_I, N_E) \\ \frac{\delta\lambda\widetilde{V}_I}{1-\delta(1-\lambda)} & \text{if} \quad s_E - s_I \ge \widetilde{\Delta}(N_I, N_E) \end{cases}$$

$$V^E(N_I, N_E) = \begin{cases} \frac{\delta\lambda\widetilde{V}_E}{1-\delta(1-\lambda)} & \text{if} \quad s_E - s_I < \widetilde{\Delta}(N_I, N_E) \\ \frac{\delta\lambda\widetilde{V}_E}{1-\delta(1-\lambda)} + \frac{s_E - s_I - \widetilde{\Delta}(N_I,N_E)}{(1-\delta(1-\lambda))^2} & \text{if} \quad \widetilde{\Delta}(N_I, N_E) \le s_E - s_I < \widetilde{\Delta}(N_I+1, N_E) \\ \frac{\delta\lambda\widetilde{V}_E}{1-\delta(1-\lambda)} + \frac{s_E - s_I}{1-\delta(1-\lambda)} - \frac{\widetilde{\Delta}(N_I,N_E) - \delta(1-\lambda)\widetilde{\Delta}(N_I+1,N_E)}{(1-\delta(1-\lambda))^2} & \text{if} \quad s_E - s_I \ge \widetilde{\Delta}(N_I+1, N_E) \end{cases}.$$

Thus, the cutoff is the same as before, once δ is replaced by $\delta(1-\lambda)$, while each firm i's value functions is simply modified by replacing δ with $\delta(1-\lambda)$ and adding the constant term $\frac{\delta\lambda\widetilde{V}_i}{1-\delta(1-\lambda)}$, for $i = I, E$.

Clearly then, Propositions 3, 7, 8 and 9 all go through unchanged with this setup when applied to the pre-discovery game.

To see that Proposition 4 (social efficiency of the outcome) also continues to hold, note that the outcome in the event of discovery is assumed to be the same regardless of the path to that moment, so it does not matter for efficiency. Which means the socially efficient condition for E to win in periods before the discovery happens is also independent of what happens after discovery.

And it is then easily seen that the socially efficient condition for E to win prior to discovery is

$$\sum_{j=0}^{\overline{N}_E-N_E}(\delta(1-\lambda))^j\left(s_E+f_E(N_E+j)\right)+\frac{(\delta(1-\lambda))^{\overline{N}_E-N_E+1}\left(s_E+f_E\left(\overline{N}_E\right)\right)}{1-\delta}$$
$$\geq \sum_{j=0}^{\overline{N}_I-N_I}(\delta(1-\lambda))^j\left(s_I+f_I(N_I+j)\right)+\frac{(\delta(1-\lambda))^{\overline{N}_I-N_I+1}\left(s_I+f_I\left(\overline{N}_I\right)\right)}{1-\delta},$$

which is exactly equivalent to $s_E - s_I \geq \widetilde{\Delta}(N_I, N_E)$. So the outcome remains socially efficient.

By the same logic, the expressions of consumer surplus in (4) and (5) are modified simply by replacing δ with $\delta(1-\lambda)$ and adding a term equal to the expected consumer surplus after discovery, which does not depend on (N_I, N_E). This implies Proposition (5) also goes through unchanged. And so does Proposition 6.

A.2 Multiple customer segments or multiple products

In practice, each firm may not only use the learnings it obtains from data to improve the core product that we have focused on, but also versions of the same product sold to other consumer segments or other products sold in different consumer markets. In these situations, even the losing firm may continue to make money (and so cover its fixed costs) from serving these other consumer segments or markets.

To model this, suppose each firm $i \in \{I, E\}$ gets an additional profit of $\alpha_i g_i(f_i(N_i))$ every period from other consumer segments or markets it may serve, when it has sold its core product to N_i previous consumers. The parameter $\alpha_i > 0$ measures how important those adjacent segments or markets are compared to the firm's core market, and g_i is some increasing function with $g_i(0) \geq 0$. To keep things as simple as possible, we assume the firms do not learn anything from selling to the adjacent consumer segments or markets. This setting can also capture other ways in which firms can monetize their accumulated data, such as selling a copy of the data or their learnings from it, or selling access to past customers' data (e.g. for targeted advertising).[1]

The results for this extension are summarized in the following proposition, the proof of which is quite similar to our existing analysis, and is available from the authors upon request.

Proposition 15. *Suppose I has previously sold to* $0 \leq N_I \leq \overline{N}_I$ *consumers and E has previously sold to* $0 \leq N_E \leq \overline{N}_E$ *consumers. Provided* α_I *and* α_E *are sufficiently close to zero, there exists a unique MPE in which E wins sales of the core product in all periods if and only if* $s_E - s_I \geq$

[1] We have also solved the case when the firms compete (imperfectly) with each other in their adjacent segments/markets, so that firm i gets additional profit of $\sigma_i + \alpha_i(f_i(N_i) - f_j(N_j))$ from its outside product. The results for this case are available from the authors upon request.

$\widetilde{\Delta}(N_I, N_E)$, *and I wins in all periods otherwise, where*

$$\begin{aligned}\widetilde{\Delta}(N_I, N_E) &= \Delta(N_I, N_E) + \alpha_I (1-\delta) \sum_{j=0}^{\overline{N}_I - N_I - 1} \delta^j g_I(f_I(N_I + j)) + \alpha_I \delta^{\overline{N}_I - N_I} g_I(f_I(\overline{N}_I)) + \alpha_E g_E(f_E(N_E)) \\ &\quad - \left(\alpha_E (1-\delta) \sum_{j=0}^{\overline{N}_E - N_E - 1} \delta^j g_E(f_E(N_E + j)) + \alpha_E \delta^{\overline{N}_E - N_E} g_E(f_E(\overline{N}_E)) + \alpha_I g_I(f_I(N_I)) \right).\end{aligned}$$

The outcome is socially optimal. Moreover, the firms' value functions are determined as in Corollary 1, except $\Delta(N_I, N_E)$ *is replaced everywhere with* $\widetilde{\Delta}(N_I, N_E)$, *and* $\frac{\alpha_i g_i(f_i(N_i))}{1-\delta}$ *is added to firm i's value function regardless of the range of* $s_E - s_I$.

Proposition 15 highlights that both firms obtain positive value in all cases, reflecting their profits from the adjacent consumer segments or markets. One complication relative to our baseline setting is that here, when α_i is large enough, it is possible that firm i is more likely to win when it is further away from the threshold than when it is closer. This arises for functional forms of g_i and f_i such that firm i stands to gain more from winning in terms of increasing revenue from the adjacent segments/markets when it is further away from the threshold compared to when it is closer. However, provided α_i is not too large, any such effect will be dominated by the property that winning improves the core product, and makes it easier for the winning firm to win in the future. In Proposition 15, the cutoff $\widetilde{\Delta}(N_I, N_E)$ is proportional to the difference between the two firms in the present discounted value (PDV) of gross surplus generated from learning in both the core market and in the adjacent segments/markets, comparing the paths where each firm wins the core market in every period from the current period onwards. This explains why the outcome remains socially optimal. Moreover, it is easily seen that taking $\alpha_i \to 0$ for $i = \{I, E\}$ brings us back to our baseline results.

A.3 Exit of the losing firm

Finally, we consider what happens in our baseline model of across-user learning if there is a small fixed operating cost $K > 0$ for each firm. Bertrand-type competition implies one of the firms will stop being active (either it would not enter in the first place, or if already present, would exit). If I can win this period by pricing at p^I and E exits, the PDV of I's profit becomes $p^I - c - K + \sum_{j=1}^{\infty} \delta^j \left(s_I + f_I\left(\min\left\{N_I + j, \overline{N}_I\right\}\right) - c - K\right)$ given that in all future periods I will be a monopolist and can extract consumers' full willingness to pay but must pay the costs $c + K$ each period. Otherwise, if I cannot win this period, it will exit and obtain nothing. Thus, I is willing to price down to $p^I = c + K - \sum_{j=1}^{\infty} \delta^j \left(s_I + f_I\left(\min\left\{N_I + j, \overline{N}_I\right\}\right) - c - K\right)$ to win in the current period. Likewise, E is willing to price down to $p^E = c + K - \sum_{j=1}^{\infty} \delta^j \left(s_E + f_E\left(\min\left\{N_E + j, \overline{N}_E\right\}\right) - c - K\right)$ to win in the current period. Given E wins iff $s_E + f_E(N_E) - p^E \geq s_I + f_I(N_I) - p^I$, plugging in the lowest price each firm is willing to price down to in order to win, it is straightforward to see that the resulting condition for E to win is identical to that in Proposition 1. And this is exactly the condition for it to be efficient for E to win, which is not surprising here given that the winning firm extracts consumers' full willingness to pay. This means all our comparative statics results continue to hold in this alternative setting. However, the fact that in every period the winning firm leaves consumers with no surplus seems unreasonable to us (the extensions above, in which the losing firm makes some profit every period and remains active, are more realistic), which is why we focused on a setting in which the losing firm remains active and still disciplines the winning firm.

B Finite number of periods with pure across-user learning

In this section, we derive the cutoff for the game with pure across-user learning and a finite number of periods. Denote the two firms' profits when there are T periods remaining and the current state is (N_I, N_E) by $\Pi^I(N_I, N_E, T)$ and $\Pi^E(N_I, N_E, T)$.

For $T = 1$ we have

$$\begin{aligned}\Pi^I(N_I, N_E, 1) &= \max\{s_I + f_I(N_I) - s_E - f_E(N_E), 0\} \\ \Pi^E(N_I, N_E, 1) &= \max\{s_E + f_E(N_E) - s_I - f_I(N_I), 0\},\end{aligned}$$

so

$$\Delta(N_I, N_E, 1) = f_I(N_I) - f_E(N_E).$$

Suppose now that for some $T > 0$ and any state (N_I, N_E) we have

$$\Delta(N_I, N_E, T) = (1-\delta)\left(\sum_{j=0}^{T-1} \frac{\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T + T\delta^{T+1}}\left(f_I(N_I+j) - f_E(N_E+j)\right)\right)$$

$$\Pi^I(N_I, N_E, T) = \begin{cases} \sum_{j=0}^{T-1}\delta^j\left(s_I + f_I(N_I+j) - s_E - f_E(N_E)\right) & \text{if} \quad s_E - s_I < \Delta(N_I, N_E+1, T-1) \\ \left(1 + 2\delta + ... + T\delta^{T-1}\right)\left(s_I - s_E + \Delta(N_I, N_E, T)\right) & \text{if} \quad \begin{array}{c}\Delta(N_I, N_E+1, T-1) \le s_E - s_I \\ < \Delta(N_I, N_E, T)\end{array} \\ 0 & \text{if} \quad s_E - s_I \ge \Delta(N_I, N_E, T)\end{cases}$$

$$\Pi^E(N_I, N_E, T) = \begin{cases} 0 & \text{if} \quad s_E - s_I < \Delta(N_I, N_E, T) \\ \left(1 + 2\delta + ... + T\delta^{T-1}\right)\left(s_E - s_I - \Delta(N_I, N_E, T)\right) & \text{if} \quad \begin{array}{c}\Delta(N_I, N_E, T) \le s_E - s_I \\ < \Delta(N_I+1, N_E, T-1)\end{array} \\ \sum_{j=0}^{T-1}\delta^j\left(s_E + f_E(N_E+j) - s_I - f_I(N_I)\right) & \text{if} \quad s_E - s_I \ge \Delta(N_I+1, N_E, T-1)\end{cases}.$$

At the end of this section we will confirm that $\Delta(N_I, N_E+1, T-1) < \Delta(N_I, N_E, T) < \Delta(N_I+1, N_E, T-1)$ for any (N_I, N_E) and $T \ge 1$.

Consider the game with $T+1$ periods starting from state (N_I, N_E). I wins iff

$$\begin{array}{rl} & s_I + f_I(N_I) + \delta\left(\Pi^I(N_I+1, N_E, T) - \Pi^I(N_I, N_E+1, T)\right) \\ > & s_E + f_E(N_E) + \delta\left(\Pi^E(N_I, N_E+1, T) - \Pi^E(N_I+1, N_E, T)\right), \end{array}$$

so

$$\begin{aligned} \Pi^I(N_I, N_E, T+1) &= \delta\Pi^I(N_I, N_E+1, T) + \max\left\{ \begin{array}{c} s_I - s_E + f_I(N_I) - f_E(N_E) \\ +\delta\left(\begin{array}{c} \Pi^I(N_I+1, N_E, T) + \Pi^E(N_I+1, N_E, T) \\ -\Pi^I(N_I, N_E+1, T) - \Pi^E(N_I, N_E+1, T) \end{array}\right), 0 \end{array}\right\} \\ \Pi^E(N_I, N_E, T+1) &= \delta\Pi^E(N_I+1, N_E, T) + \max\left\{ \begin{array}{c} s_E - s_I + f_E(N_E) - f_I(N_I) \\ +\delta\left(\begin{array}{c} \Pi^I(N_I, N_E+1, T) + \Pi^E(N_I, N_E+1, T) \\ -\Pi^I(N_I+1, N_E, 1) - \Pi^E(N_I+1, N_E, T) \end{array}\right), 0 \end{array}\right\}. \end{aligned}$$

Using the induction result for the game with T periods to write the expressions of $\Pi^I(N_I+1, N_E, T)$ and $\Pi^E(N_I, N_E+1, T)$, straightforward calculations yield

$$\Pi^I(N_I, N_E, T+1) = \begin{cases} \sum_{j=0}^{T} \delta^j \left(s_I + f_I(N_I+j) - s_E - f_E(N_E)\right) & \text{if} \quad s_E - s_I < \Delta(N_I, N_E+1, T) \\ \left(1 + 2\delta + \ldots + (T+1)\delta^T\right)\left(\begin{array}{c} s_I - s_E \\ +\Delta(N_I, N_E, T+1) \end{array}\right) & \text{if} \quad \begin{array}{c} \Delta(N_I, N_E+1, T) \le s_E - s_I \\ < \Delta(N_I, N_E, T+1) \end{array} \\ 0 & \text{if} \quad s_E - s_I \ge \Delta(N_I, N_E, T+1) \end{cases}$$

$$\Pi^E(N_I, N_E, T+1) = \begin{cases} 0 & \text{if} \quad s_E - s_I < \Delta(N_I, N_E, T+1) \\ \left(1 + 2\delta + \ldots + (T+1)\delta^T\right)\left(\begin{array}{c} s_E - s_I \\ -\Delta(N_I, N_E, T+1) \end{array}\right) & \text{if} \quad \begin{array}{c} \Delta(N_I, N_E, T+1) \le s_E - s_I \\ < \Delta(N_I+1, N_E, T) \end{array} \\ \sum_{j=0}^{T} \delta^j \left(s_E + f_E(N_E+j) - s_I - f_I(N_I)\right) & \text{if} \quad s_E - s_I \ge \Delta(N_I+1, N_E, T) \end{cases}$$

so

$$\Delta(N_I, N_E, T+1) = (1-\delta)\left(\sum_{j=0}^{T} \frac{\delta^j\left(1 - \delta^{T+1-j}\right)}{1 - (T+2)\delta^{T+1} + (T+1)\delta^{T+2}} \left(f_I(N_I+j) - f_E(N_E+j)\right)\right)$$

Thus, by induction, the expressions above hold for any $T \ge 1$.

Finally, we need to confirm that $\Delta(N_I, N_E+1, T-1) < \Delta(N_I, N_E, T) < \Delta(N_I+1, N_E, T-1)$

for any (N_I, N_E) and $T \geq 1$. To do so, write

$$\begin{aligned}\Delta(N_I, N_E, T) - \Delta(N_I, N_E+1, T-1) &= \sum_{j=0}^{T-1} \frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}}\left(f_I(N_I+j)-f_E(N_E+j)\right) \\ &\quad -\sum_{j=0}^{T-2} \frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T}\left(f_I(N_I+j)-f_E(N_E+1+j)\right)\end{aligned}$$

$$\begin{aligned}= &\sum_{j=0}^{T-1} \frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}} f_I(N_I+j) - \sum_{j=0}^{T-1} \frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T} f_I(N_I+j) \\ &+\sum_{j=0}^{T-2} \frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T} f_E(N_E+1+j) - \sum_{j=0}^{T-1} \frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}} f_E(N_E+j)\end{aligned}$$

It is easily verified that this expression is positive for $T = 1$ and $T = 2$. So assume $T \geq 3$. Recall that

$$\sum_{j=0}^{T-1} \frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}} = \sum_{j=0}^{T-1} \frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T} = \sum_{j=0}^{T-2} \frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T} = 1.$$

Furthermore, it is straightforward to verify that there exists $j^* \in [1, ..., T-2]$ such that

$$\frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T} > \frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}}$$

for all $0 \leq j \leq j^*$ and

$$\frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T} \leq \frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}}$$

for all $j^* < j \leq T-1$.

Thus, we have

$$\begin{aligned}&\sum_{j=0}^{T-1} \frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}} f_I(N_I+j) - \sum_{j=0}^{T-1} \frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T} f_I(N_I+j) \\ = &\sum_{j=0}^{j^*}\left(\frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}} - \frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T}\right) f_I(N_I+j) \\ &+\sum_{j=j^*+1}^{T-1}\left(\frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}} - \frac{(1-\delta)\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T}\right) f_I(N_I+j)\end{aligned}$$

$$\begin{aligned}
&> \sum_{j=0}^{j^*}\left(\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}-\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\,\delta^T}\right)f_I\left(N_I+j^*\right)\\
&\quad+\sum_{j=j^*+1}^{T-1}\left(\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}-\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\,\delta^T}\right)f_I\left(N_I+j^*\right)\\
&= 0.
\end{aligned}$$

And

$$\begin{aligned}
&\sum_{j=0}^{T-2}\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\,\delta^T}f_E\left(N_E+1+j\right)-\sum_{j=0}^{T-1}\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}f_E\left(N_E+j\right)\\
&= \sum_{j=1}^{T-1}\frac{(1-\delta)\,\delta^{j-1}\left(1-\delta^{T-j}\right)}{1-T\delta^{T-1}+(T-1)\,\delta^T}f_E\left(N_E+j\right)-\sum_{j=0}^{T-1}\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}f_E\left(N_E+j\right)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^{T-1}\left(\frac{(1-\delta)\,\delta^{j-1}\left(1-\delta^{T-j}\right)}{1-T\delta^{T-1}+(T-1)\,\delta^T}-\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}\right)f_E\left(N_E+j\right)-\frac{(1-\delta)\left(1-\delta^T\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}f_E\left(N_E\right)\\
&> \sum_{j=1}^{T-1}\left(\frac{(1-\delta)\,\delta^{j-1}\left(1-\delta^{T-j}\right)}{1-T\delta^{T-1}+(T-1)\,\delta^T}-\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}\right)f_E\left(N_E\right)-\frac{(1-\delta)\left(1-\delta^T\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}f_E\left(N_E\right)\\
&= 0,
\end{aligned}$$

where the inequality follows from the observation that $\frac{(1-\delta)\delta^{j-1}\left(1-\delta^{T-j}\right)}{1-T\delta^{T-1}+(T-1)\delta^T} > \frac{(1-\delta)\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\delta^T+T\delta^{T+1}}$ for all j.

We can therefore conclude that

$$\begin{aligned}
&\sum_{j=0}^{T-1}\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}f_I\left(N_I+j\right)-\sum_{j=0}^{T-1}\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\,\delta^T}f_I\left(N_I+j\right)\\
&+\sum_{j=0}^{T-2}\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-1-j}\right)}{1-T\delta^{T-1}+(T-1)\,\delta^T}f_E\left(N_E+1+j\right)-\sum_{j=0}^{T-1}\frac{(1-\delta)\,\delta^j\left(1-\delta^{T-j}\right)}{1-(T+1)\,\delta^T+T\delta^{T+1}}f_E\left(N_E+j\right)\\
&> 0,
\end{aligned}$$

so $\Delta\left(N_I,N_E,T\right) > \Delta\left(N_I,N_E+1,T-1\right)$.

The symmetry of I and E then implies that $\Delta\left(N_I,N_E,T\right) < \Delta\left(N_I+1,N_E,T-1\right)$.

Taking the difference,

$$\begin{aligned}
&\Delta(N_I,N_E,T)-\Delta^S(N_I,N_E,T) \qquad \text{(B.1)}\\
&= (1-\delta)\sum_{j=0}^{T-1}\delta^j\left(\frac{1-\delta^{T-j}}{1-(T+1)\delta^T+T\delta^{T+1}}-\frac{1}{1-\delta^T}\right)\left(f_I(N_I+j)-f_E(N_E+j)\right).
\end{aligned}$$

Note that for $T = 1$, which is the case with standard one-period Bertrand competition, there is no distortion, i.e. $\Delta(N_I, N_E, 1) = \Delta^S(N_I, N_E, 1)$. The term in large brackets in (B.1) is positive when $j = 0$, everywhere decreasing in j, and negative when $j = T - 1$. Given that $\sum_{j=0}^{T-1} \delta^j \left(\frac{1-\delta^{T-j}}{1-(T+1)\delta^T + T\delta^{T+1}} - \frac{1}{1-\delta^T} \right) = 0$, the overall distortion depends on a weighted average of the difference in the learning functions at consecutive steps, where the weights sum to zero. Thus, in general, for a finite number of periods greater than one, the distortion can go in either direction depending on the shapes of the learning curves.

C Data acquisition

The cutoff Δ^* in Proposition 7 is determined as follows:

- if

$$\Delta(N_I, N_E + N_A) + \Delta(N_I + N_A, N_E) < \delta\Delta(N_I + N_A, N_E + 1) + (2 - \delta)\Delta(N_I + 1, N_E + N_A),$$

  then

$$\Delta^* = \frac{\Delta(N_I, N_E + N_A) + \Delta(N_I + N_A, N_E) - \delta\Delta(N_I + N_A, N_E + 1)}{2 - \delta}$$

- if

$$\delta\Delta(N_I + N_A, N_E + 1) + (2 - \delta)\Delta(N_I + 1, N_E + N_A) \leq \Delta(N_I, N_E + N_A) + \Delta(N_I + N_A, N_E)$$

$$\leq \delta\Delta(N_I + 1, N_E + N_A) + (2 - \delta)\Delta(N_I + N_A, N_E + 1),$$

  then

$$\Delta^* = \frac{\Delta(N_I, N_E + N_A) + \Delta(N_I + N_A, N_E) - \delta\Delta(N_I + 1, N_E + N_A) - \delta\Delta(N_I + N_A, N_E + 1)}{2(1 - \delta)}$$

- if

$$\Delta(N_I, N_E + N_A) + \Delta(N_I + N_A, N_E) > \delta\Delta(N_I + 1, N_E + N_A) + (2 - \delta)\Delta(N_I + N_A, N_E + 1),$$

  then

$$\Delta^* = \frac{\Delta(N_I, N_E + N_A) + \Delta(N_I + N_A, N_E) - \delta\Delta(N_I + 1, N_E + N_A)}{2 - \delta}$$

Here we prove that the existence of the new data N_A can make I worse off despite the fact I acquires the new data in equilibrium.

We already know that when $s_E - s_I < \Delta(N_I, N_E + N_A)$ or when $s_E - s_I \geq \Delta(N_I + N_A, N_E)$, the losing firm does not subsidize before and after the data acquisition by the winning firm. Therefore, in these cases the winning firm is made better off by the appearance of the data acquisition opportunity: it pays a price of zero to acquire the data and the new data puts it in an even better competitive position.

Now assume we are in the case with $\Delta(N_I, N_E + N_A) \leq s_E - s_I < \Delta(N_I + N_A, N_E)$. Suppose I wins the data, so $\Delta(N_I, N_E + N_A) \leq s_E - s_I < \Delta^*$. I will have to pay $W_E(s_E - s_I)$ for the data and ends up better off as a whole iff

$$V^I(N_I + N_A, N_E) - V^I(N_I, N_E) > W_E(s_E - s_I) = V^E(N_I, N_E + N_A) - V^E(N_I + N_A, N_E)$$

Note that I wins the data iff

$$V^I(N_I + N_A, N_E) - V^I(N_I, N_E + N_A) > V^E(N_I, N_E + N_A) - V^E(N_I + N_A, N_E).$$

So it's possible that I wins the data but is worse off, which happens whenever

$$\begin{aligned} V^I(N_I + N_A, N_E) - V^I(N_I, N_E + N_A) &> V^E(N_I, N_E + N_A) - V^E(N_I + N_A, N_E) \\ &> V^I(N_I + N_A, N_E) - V^I(N_I, N_E). \end{aligned}$$

D Finite number of periods with pure within-user learning

In this section, we prove that the outcome of the game with pure within-user learning and a finite time horizon is socially optimal (unlike the case with across-user learning and a finite time horizon).

Recall that the state (N_I, N_E) here means that the representative consumer has previously purchased N_i times from firm i, for $i = I, E$. If there are t periods remaining and the state is (N_I, N_E), the two firms' profits are denoted $\Pi^I(N_I, N_E, t)$ and $\Pi^E(N_I, N_E, t)$, while the PDV of the net surplus derived by the representative consumer is denoted $u(N_I, N_E, t)$.

Consider first $t = 1$. We have

$$\begin{aligned} \Pi^I(N_I, N_E, 1) &= \max\{s_I + f_I(N_I) - s_E - f_E(N_E), 0\} \\ \Pi^E(N_I, N_E, 1) &= \max\{s_E + f_E(N_E) - s_I - f_I(N_I), 0\}. \end{aligned}$$

So

$$\begin{aligned} \Delta(N_I, N_E, 1) &= f_I(N_I) - f_E(N_E) \\ u(N_I, N_E, 1) &= \min\{s_I - c + f_I(N_I), s_E - c + f_E(N_E)\}. \end{aligned}$$

Thus, the result holds for $t = 1$.

Now suppose for some $t > 0$, we have

$$\begin{aligned}
\Pi^I (N_I, N_E, t) &= \max \left\{ \sum_{j=0}^{t-1} \delta^j \left(s_I + f_I (N_I + j) - s_E - f_E (N_E + j) \right), 0 \right\} \\
\Pi^E (N_I, N_E, t) &= \max \left\{ \sum_{j=0}^{t-1} \delta^j \left(s_E + f_E (N_E + j) - s_I - f_I (N_I + j) \right), 0 \right\}
\end{aligned}$$

$$\begin{aligned}
\Delta (N_I, N_E, t) &= \frac{\sum_{j=0}^{t-1} \delta^j \left(f_I (N_I + j) - f_E (N_E + j) \right)}{\sum_{j=0}^{t-1} \delta^j} \\
u (N_I, N_E, t) &= \min \left\{ \sum_{j=0}^{t-1} \delta^j \left(s_I - c + f_I (N_I + j) \right), \sum_{j=0}^{t-1} \delta^j \left(s_E - c + f_E (N_E + j) \right) \right\}
\end{aligned}$$

for all (N_I, N_E). Note that $\Delta (N_I, N_E, t)$ defined above is the socially optimum cutoff for the game starting in state (N_I, N_E) and with t periods left.

Consider now the game starting with state (N_I, N_E) and having $t+1$ periods left. I wins iff

$$\begin{aligned}
& s_I + f_I (N_I) + \delta \left(\Pi^I (N_I + 1, N_E, t) - \Pi^I (N_I, N_E + 1, t) \right) + \delta u (N_I + 1, N_E, t) \\
> \;& s_E + f_E (N_E) + \delta \left(\Pi^E (N_I, N_E + 1, t) - \Pi^E (N_I + 1, N_E, t) \right) + \delta u (N_I, N_E + 1, t),
\end{aligned}$$

so

$$\begin{aligned}
\Pi^I (N_I, N_E, t+1) &= \delta \Pi^I (N_I, N_E + 1, t) \\
& \quad + \max \left\{ \begin{array}{c} s_I - s_E + f_I (N_I) - f_E (N_E) \\ + \delta \left(\begin{array}{c} \Pi^I (N_I + 1, N_E, t) + \Pi^E (N_I + 1, N_E, t) + u (N_I + 1, N_E, t) \\ - \Pi^I (N_I, N_E + 1, t) - \Pi^E (N_I, N_E + 1, t) - u (N_I, N_E + 1, t) \end{array} \right), 0 \end{array} \right\}
\end{aligned}$$

$$\begin{aligned}
\Pi^E (N_I, N_E, t+1) &= \delta \Pi^E (N_I + 1, N_E, t) \\
& \quad + \max \left\{ \begin{array}{c} s_E - s_I + f_E (N_E) - f_I (N_I) \\ + \delta \left(\begin{array}{c} \Pi^I (N_I, N_E + 1, t) + \Pi^E (N_I, N_E + 1, t) + u (N_I, N_E + 1, t) \\ - \Pi^I (N_I + 1, N_E, t) - \Pi^E (N_I + 1, N_E, t) - u (N_I + 1, N_E, t) \end{array} \right), 0 \end{array} \right\}.
\end{aligned}$$

Using the induction hypothesis for the case with t periods left and states $(N_I + 1, N_E)$ and

$(N_I, N_E + 1)$, straightforward calculations lead to

$$\Pi^I (N_I, N_E, t+1) = \max \left\{ \sum_{j=0}^{t} \delta^j (s_I + f_I (N_I + j) - s_E - f_E (N_E + j)), 0 \right\}$$
$$\Pi^E (N_I, N_E, t+1) = \max \left\{ \sum_{j=0}^{t} \delta^j (s_E + f_E (N_E + j) - s_I - f_I (N_I + j)), 0 \right\}$$

$$\Delta (N_I, N_E, t+1) = \frac{\sum_{j=0}^{t} \delta^j (f_I (N_I + j) - f_E (N_E + j))}{\sum_{j=0}^{t-1} \delta^j}$$
$$u (N_I, N_E, t+1) = \min \left\{ \sum_{j=0}^{t} \delta^j (s_I - c + f_I (N_I + j)), \sum_{j=0}^{t} \delta^j (s_E - c + f_E (N_E + j)) \right\}$$

Thus, the result holds for $t+1$. By induction, we have thus proven the result for any $t \geq 1$.

E Proof of Proposition 14

If both firms have reached their learning thresholds, then the value functions are

$$V^E (1,1,1,1) = \max \left\{ \frac{s_E - s_I + f_E (1,1) - f_I (1,1)}{1-\delta}, 0 \right\}$$
$$V^I (1,1,1,1) = \max \left\{ \frac{s_I - s_E + f_I (1,1) - f_E (1,1)}{1-\delta}, 0 \right\}$$

and consumer net surplus is

$$u (1,1,1,1) = \frac{\min \{s_E + f_E (1,1), s_I + f_I (1,1)\} - c}{1-\delta}.$$

E.1 Pareto beliefs

All consumers start in state $(1,1,0,0)$. Under Pareto beliefs, consumers choose I in the current period iff

$$s_I + f_I (1,1) - p_I + \delta u (1,1,0,0) > s_E - p_E + \delta u (1,1,1,1).$$

This condition says that consumers jointly prefer to choose I over E. We also need to impose that no individual consumer wants to unilaterally deviate and choose E, obtaining $s_E - p_E + \delta u (1,1,0,1)$. This is implied by the condition above provided $u (1,1,0,1) \leq u (1,1,1,1)$, which we will show is true at the end.

Let us determine conditions for choosing I to be an equilibrium. E is willing to price down to $c - \delta \left(V^E (1,1,1,1) - V^E (1,1,0,0)\right)$, while I is willing to price down to $c - \delta \left(V^I (1,1,0,0) - V^I (1,1,1,1)\right)$.

So I wins iff

$$\begin{aligned} & s_I - c + f_I(1,1) + \delta\left(V^I(1,1,0,0) - V^I(1,1,1,1) + u(1,1,0,0)\right) \\ \geq \ & s_E - c + \delta\left(V^E(1,1,1,1) - V^E(1,1,0,0) + u(1,1,1,1)\right). \end{aligned}$$

Assuming I wins, we have

$$\begin{aligned} V^I(1,1,0,0) &= s_I - s_E + f_I(1,1) + \delta\left(V^I(1,1,0,0) - V^E(1,1,1,1) + u(1,1,0,0) - u(1,1,1,1)\right) \\ V^E(1,1,0,0) &= 0 \end{aligned}$$

$$\begin{aligned} u(1,1,0,0) &= s_E - c + \delta\left(V^E(1,1,1,1) + u(1,1,1,1)\right) = s_E - c + \delta\frac{s_E - c + f_E(1,1)}{1-\delta} \\ &= \frac{s_E - c + \delta f_E(1,1)}{1-\delta}. \end{aligned}$$

So

$$V^I(1,1,0,0) = \frac{s_I - s_E + f_I(1,1) - \delta f_E(1,1)}{1-\delta},$$

which means the cutoff is

$$\Delta^I(1,1,0,0) = f_I(1,1) - \delta f_E(1,1).$$

Finally we need to make sure no consumer has an incentive to unilaterally deviate from the proposed equilibrium, which as explained above comes down to showing $u(1,1,0,1) \leq u(1,1,1,1)$.

Suppose a consumer is in state $(1,1,0,1)$, and all other consumers are in state $(1,1,0,0)$. Because the consumer who is in state $(1,1,0,1)$ has a different history from the other consumers, she can be charged a different price by each of the two firms—we denote those prices $\widetilde{p}_I$ and $\widetilde{p}_E$. Since both firms have maximized within-user learning for this consumer and she is too small to matter for total firm profits and across-user learning, neither firm gains anything from subsidizing this consumer. Thus, both $\widetilde{p}_I$ and $\widetilde{p}_E$ are greater or equal to c. This consumer chooses firm I iff (recall this consumer expects firm I wins all other consumers, who are in state $(1,1,0,0)$)

$$s_I + f_I(1,1) - \widetilde{p}_I + \delta u(1,1,0,1) \geq s_E + f_E(0,1) - \widetilde{p}_E + \delta u(1,1,0,1).$$

So we have

$$\begin{aligned} u(1,1,0,1) &= \frac{\min\{s_I + f_I(1,1) - \widetilde{p}_I, s_E + f_E(0,1) - \widetilde{p}_E\}}{1-\delta} \\ &\leq \frac{\min\{s_I + f_I(1,1) - c, s_E + f_E(0,1) - c\}}{1-\delta} \\ &\leq \frac{\min\{s_I + f_I(1,1) - c, s_E + f_E(1,1) - c\}}{1-\delta} = u(1,1,1,1). \end{aligned}$$

Now let us determine the conditions for choosing E to be an equilibrium in state $(1,1,0,0)$

under Pareto beliefs. E wins iff

$$s_E - p_E + \delta u(1,1,1,1) \geq s_I + f_I(1,1) - p_I + \delta u(1,1,0,0).$$

This condition says all consumers jointly prefer to choose E over I. At the end we will show that the equilibrium prices are such that

$$s_E - p_E + \delta u(1,1,1,1) \geq s_I + f_I(1,1) - p_I + \delta u(1,1,1,0),$$

i.e. no consumer wants to unilaterally deviate.

Using the same logic as before to determine how low E and I are willing to price, E wins iff

$$\begin{aligned} & s_E - c + \delta\left(V^E(1,1,1,1) - V^E(1,1,0,0) + u(1,1,1,1)\right) \\ \geq\ & s_I - c + f_I(1,1) + \delta\left(V^I(1,1,0,0) - V^I(1,1,1,1) + u(1,1,0,0)\right). \end{aligned}$$

Assuming E wins, we have

$$\begin{aligned} V^I(1,1,0,0) &= \delta V^I(1,1,1,1) = 0 \\ V^E(1,1,0,0) &= s_E - s_I - f_I(1,1) + \delta\left(V^E(1,1,1,1) + u(1,1,1,1) - u(1,1,0,0)\right) \\ u(1,1,0,0) &= \frac{s_I - c + f_I(1,1)}{1-\delta}. \end{aligned}$$

So

$$V^E(1,1,0,0) = \frac{s_E - s_I + \delta f_E(1,1) - f_I(1,1)}{1-\delta},$$

which means

$$\Delta^E(1,1,0,0) = \Delta^I(1,1,0,0) = f_I(1,1) - \delta f_E(1,1).$$

Finally, we have to verify that no consumer wants to unilaterally deviate from the proposed equilibrium with E winning. At equilibrium prices, we have

$$s_E - p_E + \delta u(1,1,1,1) = s_I + f_I(1,1) - p_I + \delta u(1,1,0,0),$$

so in order to show that

$$s_E - p_E + \delta u(1,1,1,1) \geq s_I + f_I(1,1) - p_I + \delta u(1,1,1,0)$$

it is sufficient to show that $u(1,1,1,0) = u(1,1,0,0)$ in equilibrium. And this is true because by definition, in this equilibrium, E wins all consumers when they are in state $(1,1,0,0)$, so E must also win any consumer who is in state $(1,1,1,0)$ because E is in a stronger position relative to $(1,1,0,0)$. And in both cases, consumer surplus is the maximum surplus offered by the losing firm, i.e. I, so we have

$$u(1,1,1,0) = u(1,1,0,0) = \frac{s_I - c + f_I(1,1)}{1-\delta}.$$

E.2 Favorable beliefs for I

Again, we start with all consumers in state $(1,1,0,0)$. Under favorable beliefs for I, consumers choose I in the current period whenever that is an equilibrium, i.e. iff

$$s_I + f_I(1,1) - p_I + \delta u(1,1,0,0) \geq s_E - p_E + \delta u(1,1,0,1).$$

Let us determine conditions for choosing I to be an equilibrium. E is willing to price down to $c - \delta\left(V^E(1,1,1,1) - V^E(1,1,0,0)\right)$, while I is willing to price down to $c - \delta\left(V^I(1,1,0,0) - V^I(1,1,1,1)\right)$. So I wins iff

$$\begin{aligned} & s_I - c + f_I(1,1) + \delta\left(V^I(1,1,0,0) - V^I(1,1,1,1) + u(1,1,0,0)\right) \\ \geq\ & s_E - c + \delta\left(V^E(1,1,1,1) - V^E(1,1,0,0) + u(1,1,0,1)\right). \end{aligned}$$

Assuming I wins, we have

$$\begin{aligned} V^I(1,1,0,0) &= s_I - s_E + f_I(1,1) + \delta\left(V^I(1,1,0,0) - V^E(1,1,1,1) + u(1,1,0,0) - u(1,1,0,1)\right) \\ V^E(1,1,0,0) &= 0 \\ u(1,1,0,0) &= s_E - c + \delta\left(V^E(1,1,1,1) + u(1,1,0,1)\right) \\ &= s_E - c + \delta\left(\frac{\max\{s_E - s_I + f_E(1,1) - f_I(1,1), 0\}}{1-\delta} + u(1,1,0,1)\right) \end{aligned}$$

and prices are

$$\begin{aligned} p_I &= c + s_I - s_E + f_I(1,1) - \delta\left(\frac{\max\{s_E - s_I + f_E(1,1) - f_I(1,1), 0\}}{1-\delta} + u(1,1,0,1) - u(1,1,0,0)\right) \\ p_E &= c - \delta\frac{\max\{s_E - s_I + f_E(1,1) - f_I(1,1), 0\}}{1-\delta}. \end{aligned}$$

Suppose a consumer is in state $(1,1,0,1)$, and all other consumers are in state $(1,1,0,0)$. If there was no within-user learning for firm E (i.e. $f_E(0,1) = 0$), then the state $(1,1,0,1)$ would be exactly the same as the state $(1,1,0,0)$, so the focal consumer would be just like the others, and would therefore face the same prices (p_I, p_E) determined above. Note that those prices could be below cost. However, as soon as $f_E(0,1) > 0$, the consumer who is in state $(1,1,0,1)$ can be treated differently and therefore charged a different price by each of the two firms—we denote those prices $\widetilde{p}_I$ and $\widetilde{p}_E$. And because the focal consumer in state $(1,1,0,1)$ is too small to matter for total firm profits, the two firms can choose any prices for that consumer. Furthermore, since both firms have maximized within-user learning for the focal consumer, neither firm gains anything from subsidizing that consumer. So we impose the restriction that $\widetilde{p}_I$ and $\widetilde{p}_E$ are both greater or equal to c whenever $f_E(0,1) > 0$.

The focal consumer chooses firm I iff (recall this consumer expects firm I wins all other consumers, who are in state $(1,1,0,0)$)

$$s_I + f_I(1,1) - \widetilde{p}_I + \delta u(1,1,0,1) \geq s_E + f_E(0,1) - \widetilde{p}_E + \delta u(1,1,0,1).$$

So we have

$$u(1,1,0,1) = \frac{\min\{s_I + f_I(1,1) - \widetilde{p}_I, s_E + f_E(0,1) - \widetilde{p}_E\}}{1-\delta}$$

$$u(1,1,0,0) = s_E - c + \delta\left(\frac{\max\{s_E - s_I + f_E(1,1) - f_I(1,1), 0\}}{1-\delta} + u(1,1,0,1)\right)$$

$$V^I(1,1,0,0) = \frac{s_I - s_E + f_I(1,1) - \delta\left(\frac{\max\{s_E - s_I + f_E(1,1) - f_I(1,1), 0\}}{1-\delta} + u(1,1,0,1) - u(1,1,0,0)\right)}{1-\delta}.$$

Suppose first $s_E - s_I < f_I(1,1) - f_E(1,1)$. Then E has no chance of winning any consumer even if it were to win all consumers this period. Furthermore, I has already reached the threshold for both types of learning. Thus, E does not subsidize any consumer, so we must have (given our restriction above)

$$\widetilde{p}_E = c.$$

This implies

$$u(1,1,0,1) = \frac{\min\{s_I + f_I(1,1) - \widetilde{p}_I, s_E + f_E(0,1) - c\}}{1-\delta} \leq \frac{s_E + f_E(0,1) - c}{1-\delta}$$

$$u(1,1,0,0) = s_E - c + \delta u(1,1,0,1)$$

$$\begin{aligned} V^I(1,1,0,0) &= \frac{s_I - s_E + f_I(1,1) - \delta((1-\delta)u(1,1,0,1) - (s_E - c))}{1-\delta} \\ &\geq \frac{s_I - s_E + f_I(1,1) - \delta f_E(0,1)}{1-\delta} > 0. \end{aligned}$$

So the threshold must be higher than $f_I(1,1) - f_E(1,1)$.

Suppose now $s_E - s_I \geq f_I(1,1) - f_E(0,1)$. This means that any consumer that deviates to E will stay with E, regardless of what everyone else does. Thus, I has no chance of ever winning the consumer that is in state $(1,1,0,1)$, so it won't subsidize her, which means

$$\widetilde{p}_I = c.$$

This implies

$$u(1,1,0,1) = u(1,1,1,1) = \frac{s_I - c + f_I(1,1)}{1-\delta},$$

which in turn implies

$$\begin{aligned} u(1,1,0,0) &= s_E - c + \delta\left(\frac{s_E - s_I + f_E(1,1) - f_I(1,1)}{1-\delta} + u(1,1,0,1)\right) \\ &= \frac{s_E - c + \delta f_E(1,1)}{1-\delta} \end{aligned}$$

$$\begin{aligned} V^I(1,1,0,0) &= \frac{s_I - s_E + f_I(1,1) - \delta\left(\frac{s_E - s_I + f_E(1,1) - f_I(1,1)}{1-\delta} + \frac{s_I - c + f_I(1,1)}{1-\delta} - \frac{s_E - c + \delta f_E(1,1)}{1-\delta}\right)}{1-\delta} \\ &= \frac{s_I - s_E + f_I(1,1) - \delta f_E(1,1)}{1-\delta}. \end{aligned}$$

Thus, $V^I(1,1,0,0)$ is decreasing in $s_E - s_I$ in this region. And when $s_E - s_I = f_I(1,1) - f_E(0,1)$, we have

$$V^I(1,1,0,0) = \frac{f_E(0,1) - \delta f_E(1,1)}{1-\delta},$$

which is positive iff $f_E(0,1) \geq \delta f_E(1,1)$. In this case, we have

$$\Delta^I(1,1,0,0) = f_I(1,1) - \delta f_E(1,1).$$

Note that as the importance of across-user learning for E goes to zero (i.e. $f_E(1,1)$ converges to $f_E(0,1)$), $V^I(1,1,0,0)$ converges to $\frac{s_I - s_E + f_I(1,1) - \delta f_E(0,1)}{1-\delta}$, consistent with the result in the pure within-user learning case.

Otherwise, if $f_E(0,1) < \delta f_E(1,1)$, then $V^I(1,1,0,0) < 0$ for all $s_E - s_I \geq f_I(1,1) - f_E(0,1)$, so it must be that

$$f_I(1,1) - f_E(1,1) < \Delta^I(1,1,0,0) < f_I(1,1) - f_E(0,1).$$

Suppose then

$$f_I(1,1) - f_E(1,1) < s_E - s_I < f_I(1,1) - f_E(0,1).$$

Note that now

$$f_I(1,1) - f_E(1,1) < f_I(1,1) - \delta f_E(1,1) < f_I(1,1) - f_E(0,1).$$

And we have

$$u(1,1,0,1) = \frac{\min\{s_I + f_I(1,1) - \widetilde{p}_I, s_E + f_E(0,1) - \widetilde{p}_E\}}{1-\delta}$$

$$u(1,1,0,0) = s_E - c + \delta\left(\frac{s_E - s_I + f_E(1,1) - f_I(1,1)}{1-\delta} + u(1,1,0,1)\right)$$

$$V^I(1,1,0,0) = \frac{s_I - s_E + f_I(1,1) - \delta\left(s_E - s_I + f_E(1,1) - f_I(1,1) + (1-\delta)\,u(1,1,0,1) - (s_E - c)\right)}{1-\delta}.$$

Thus, when $s_E - s_I = f_I(1,1) - \delta f_E(1,1)$, we have

$$V^I(1,1,0,0) = \delta \frac{\delta f_E(1,1) - \min\{s_I + f_I(1,1) - \widetilde{p}_I, s_E + f_E(0,1) - \widetilde{p}_E\} + s_E - c}{1-\delta}.$$

Since we are in the region $s_E - s_I < f_I(1,1) - f_E(0,1)$, we know that E cannot win the lone consumer that is in state $(1,1,0,1)$ since everyone else is in state $(1,1,0,0)$ and goes with I. As a result, since E has already maximized its within-user learning on the consumer that is in state $(1,1,0,1)$, it will never subsidize this consumer, so $\widetilde{p}_E = c$ and

$$\begin{aligned} V^I(1,1,0,0) &\geq \delta \frac{\delta f_E(1,1) - (s_E + f_E(0,1) - c) + s_E - c}{1-\delta} \\ &= \delta \frac{\delta f_E(1,1) - f_E(0,1)}{1-\delta} > 0. \end{aligned}$$

Thus, we have shown $V^I(1,1,0,0) > 0$ when $s_E - s_I = f_I(1,1) - \delta f_E(1,1)$. Since $V^I(1,1,0,0)$ is decreasing in $s_E - s_I$, it must be that

$$\Delta^I(1,1,0,0) > f_I(1,1) - \delta f_E(1,1).$$

Now let us determine the conditions for choosing E to be an equilibrium in state $(1,1,0,0)$ under favorable beliefs for I. Consumers choose E in the current period iff

$$s_E - p_E + \delta u(1,1,0,1) \geq s_I + f_I(1,1) - p_I + \delta u(1,1,0,0).$$

Note that at the end we will need to verify that the equilibrium prices are such that

$$s_E - p_E + \delta u(1,1,1,1) \geq s_I + f_I(1,1) - p_I + \delta u(1,1,1,0),$$

i.e. no consumer wants to unilaterally deviate.

Again, E is willing to price down to $c - \delta\left(V^E(1,1,1,1) - V^E(1,1,0,0)\right)$, while I is willing to price down to $c - \delta\left(V^I(1,1,0,0) - V^I(1,1,1,1)\right)$. So E wins iff

$$\begin{aligned} & s_E - c + \delta\left(V^E(1,1,1,1) - V^E(1,1,0,0) + u(1,1,0,1)\right) \\ \geq\ & s_I - c + f_I(1,1) + \delta\left(V^I(1,1,0,0) - V^I(1,1,1,1) + u(1,1,0,0)\right). \end{aligned}$$

Assuming E wins, we have

$$\begin{aligned} V^I(1,1,0,0) &= \delta V^I(1,1,1,1) = 0 \\ V^E(1,1,0,0) &= s_E - s_I - f_I(1,1) + \delta\left(V^E(1,1,1,1) + u(1,1,0,1) - u(1,1,0,0)\right) \\ u(1,1,0,0) &= \frac{s_I - c + f_I(1,1)}{1-\delta} \end{aligned}$$

and prices are

$$\begin{aligned} p_I &= c \\ p_E &= c + s_E - s_I - f_I(1,1) + \delta(u(1,1,0,1) - u(1,1,0,0)). \end{aligned}$$

Suppose a consumer is in state $(1,1,0,1)$, so all other consumers are in state $(1,1,0,0)$. Again, assume the prices charged to the consumer in state $(1,1,0,1)$ are $\widetilde{p}_I$ and $\widetilde{p}_E$. And like above, we assume neither firm would subsidize the focal consumer when that does not help it win that consumer, which means both $\widetilde{p}_I$ and $\widetilde{p}_E$ are greater or equal to c.

The consumer in state $(1,1,0,1)$ chooses firm I iff (recall firm E wins all other consumers, who are in state $(1,1,0,0)$)

$$s_I + f_I(1,1) - \widetilde{p}_I + \delta u(1,1,1,1) \geq s_E + f_E(0,1) - \widetilde{p}_E + \delta u(1,1,1,1).$$

So we have

$$u(1,1,0,1) = \min\{s_I + f_I(1,1) - \widetilde{p}_I, s_E + f_E(0,1) - \widetilde{p}_E\} + \delta u(1,1,1,1).$$

Since $V^I(1,1,1,1) = 0$, we already know that we must be in the region $s_E - s_I \geq f_I(1,1) - f_E(1,1)$, so

$$\begin{aligned} V^E(1,1,1,1) &= \frac{s_E - s_I + f_E(1,1) - f_I(1,1)}{1-\delta} \\ u(1,1,1,1) &= \frac{s_I - c + f_I(1,1)}{1-\delta}. \end{aligned}$$

This implies

$$u(1,1,0,1) = \min\{s_I + f_I(1,1) - \widetilde{p}_I, s_E + f_E(0,1) - \widetilde{p}_E\} + \delta\frac{s_I - c + f_I(1,1)}{1-\delta}$$

$$V^E(1,1,0,0) = s_E - s_I - f_I(1,1) + \delta\left(\frac{s_E - s_I + f_E(1,1) - f_I(1,1)}{1-\delta} + u(1,1,0,1) - \frac{s_I - c + f_I(1,1)}{1-\delta}\right).$$

Suppose $s_E - s_I \geq f_I(1,1) - f_E(0,1)$. Then, by the same logic as above, E wins the deviating consumer no matter what, so I has no incentive to subsidize that consumer. This means we have $\widetilde{p}_I = c$, so

$$V^E(1,1,0,0) = \frac{s_E - s_I + \delta f_E(1,1) - f_I(1,1)}{1-\delta}.$$

So $V^E(1,1,0,0)$ is increasing in $s_E - s_I$ when $s_E - s_I \geq f_I(1,1) - f_E(0,1)$. And when $s_E - s_I = f_I(1,1) - f_E(0,1)$, we have

$$V^E(1,1,0,0) = \frac{\delta f_E(1,1) - f_E(0,1)}{1-\delta}.$$

Thus, if $f_E(0,1) \geq \delta f_E(1,1)$, then we know for sure that

$$\Delta^E(1,1,0,0) = f_I(1,1) - \delta f_E(1,1) \geq f_I(1,1) - f_E(0,1).$$

Now suppose $f_E(0,1) < \delta f_E(1,1)$, which means we must have

$$f_I(1,1) - f_E(1,1) < \Delta^I(1,1,0,0) < f_I(1,1) - f_E(0,1).$$

Note that in this case we also have

$$f_I(1,1) - f_E(1,1) < f_I(1,1) - \delta f_E(1,1) < f_I(1,1) - f_E(0,1).$$

So assume

$$f_I(1,1) - f_E(1,1) < s_E - s_I < f_I(1,1) - f_E(0,1).$$

We have

$$\begin{aligned} u(1,1,0,1) &= \min\{s_I + f_I(1,1) - \widetilde{p}_I, s_E + f_E(0,1) - \widetilde{p}_E\} + \delta\frac{s_I - c + f_I(1,1)}{1-\delta} \\ &\leq s_I + f_I(1,1) - c + \delta\frac{s_I - c + f_I(1,1)}{1-\delta} \\ &= \frac{s_I - c + f_I(1,1)}{1-\delta}. \end{aligned}$$

So

$$\begin{aligned} V^E(1,1,0,0) &= s_E - s_I - f_I(1,1) + \delta\left(\frac{s_E - s_I + f_E(1,1) - f_I(1,1)}{1-\delta} + u(1,1,0,1) - \frac{s_I - c + f_I(1,1)}{1-\delta}\right) \\ &\leq \frac{s_E - s_I + \delta f_E(1,1) - f_I(1,1)}{1-\delta}. \end{aligned}$$

Thus, when $s_E - s_I = f_I(1,1) - \delta f_E(1,1)$, we have

$$V^E(1,1,0,0) \leq 0,$$

so we must have

$$\Delta^E(1,1,0,0) \geq f_I(1,1) - \delta f_E(1,1).$$

Finally, we need to make sure that no consumer wants to unilaterally deviate from the proposed equilibrium, i.e. we want to show

$$s_E - p_E + \delta u(1,1,1,1) \geq s_I + f_I(1,1) - p_I + \delta u(1,1,1,0)$$

at equilibrium prices. But at these prices we must have

$$s_E - p_E + \delta u(1,1,0,1) = s_I + f_I(1,1) - p_I + \delta u(1,1,0,0),$$

so we need to verify that

$$u(1,1,1,1) - u(1,1,0,1) \geq u(1,1,1,0) - u(1,1,0,0).$$

But by the same logic explained in the case with Pareto beliefs, since E wins when consumers are in state $(1,1,0,0)$, E must also win when any consumer is in state $(1,1,0,1)$, $(1,1,1,0)$ or $(1,1,1,1)$. So we have

$$u(1,1,1,1) = u(1,1,0,1) = u(1,1,1,0) = u(1,1,0,0) = \frac{s_I - c + f_I(1,1)}{1-\delta},$$

which means the inequality holds.

F Additional details for the proof of Proposition 1

Recall that the induction hypothesis is that the result in Proposition 1 and Corollary 1 holds for the states (N_I+1, N_E) and (N_I, N_E+1), i.e. that

$$V^I(N_I{+}1, N_E) = \begin{cases} \frac{s_I - s_E}{1-\delta} + \frac{\Delta(N_I+1,N_E) - \delta\Delta(N_I+1,N_E+1)}{(1-\delta)^2} & \text{if} \quad s_E - s_I < \Delta(N_I+1, N_E+1) \\ \frac{s_I - s_E + \Delta(N_I+1,N_E)}{(1-\delta)^2} & \text{if} \quad \Delta(N_I+1,N_E+1) \leq s_E - s_I \leq \Delta(N_I+1,N_E) \\ 0 & \text{if} \quad s_E - s_I \geq \Delta(N_I+1,N_E) \end{cases}$$

$$V^E(N_I, N_E{+}1) = \begin{cases} 0 & \text{if} \quad s_E - s_I < \Delta(N_I, N_E+1) \\ \frac{s_E - s_I - \Delta(N_I,N_E+1)}{(1-\delta)^2} & \text{if} \quad \Delta(N_I,N_E+1) \leq s_E - s_I \leq \Delta(N_I+1,N_E+1) \\ \frac{s_E - s_I}{1-\delta} - \frac{\Delta(N_I,N_E+1) - \delta\Delta(N_I+1,N_E+1)}{(1-\delta)^2} & \text{if} \quad s_E - s_I > \Delta(N_I+1,N_E+1) \end{cases}.$$

Also recall from the proof in the main appendix that $V^I(N_I, N_E)$ and $V^E(N_I, N_E)$ can be written (after substituting in the expression for $\Omega(N_I, N_E)$)

$$V^I(N_I, N_E) = \delta V^I(N_I, N_E+1) + \max\left\{ \begin{array}{c} s_I - s_E + f_I(N_I) - f_E(N_E) \\ +\delta \left(\begin{array}{c} V^I(N_I+1, N_E) + V^E(N_I+1, N_E) \\ -V^I(N_I, N_E+1) - V^E(N_I, N_E+1) \end{array} \right), 0 \end{array} \right\}$$

$$V^E(N_I, N_E) = \delta V^E(N_I+1, N_E) + \max\left\{ \begin{array}{c} s_E - s_I + f_E(N_E) - f_I(N_I) \\ +\delta \left(\begin{array}{c} V^E(N_I, N_E+1) + V^I(N_I, N_E+1) \\ -V^E(N_I+1, N_E) - V^I(N_I+1, N_E) \end{array} \right), 0 \end{array} \right\}.$$

There are two possibilities. If

$$\max\left\{\begin{array}{c} s_I - s_E + f_I(N_I) - f_E(N_E) \\ +\delta(V^I(N_I+1,N_E) + V^E(N_I+1,N_E) - V^I(N_I,N_E+1) - V^E(N_I,N_E+1)), 0 \end{array}\right\} > 0,$$

then $V^E(N_I,N_E) = \delta V^E(N_I+1,N_E) = 0$, because we must have $V^E(N_I,N_E) \geq V^E(N_I+1,N_E)$. In this case, we have

$$V^I(N_I,N_E) = s_I - s_E + f_I(N_I) - f_E(N_E) + \delta(V^I(N_I+1,N_E) - V^E(N_I,N_E+1)).$$

On the other hand, if

$$\max\left\{\begin{array}{c} s_I - s_E + f_I(N_I) - f_E(N_E) \\ +\delta(V^I(N_I+1,N_E) + V^E(N_I+1,N_E) - V^I(N_I,N_E+1) - V^E(N_I,N_E+1)), 0 \end{array}\right\} = 0,$$

then $V^I(N_I,N_E) = \delta V^I(N_I,N_E+1) = 0$, because we must have $V^I(N_I,N_E) \geq V^I(N_I,N_E+1)$. In this case, we have

$$V^E(N_I,N_E) = s_E - s_I + f_E(N_E) - f_I(N_I) + \delta(V^E(N_I,N_E+1) - V^I(N_I+1,N_E)).$$

Focusing on the first possibility and using the above expressions, we have

$$s_I - s_E + f_I(N_I) - f_E(N_E) + \delta(V^I(N_I+1,N_E) - V^E(N_I,N_E+1))$$

$$= \begin{cases} \begin{array}{c} s_I - s_E + f_I(N_I) - f_E(N_E) \\ +\delta(\frac{s_I - s_E}{1-\delta} + \frac{\Delta(N_I+1,N_E) - \delta\Delta(N_I+1,N_E+1)}{(1-\delta)^2}) \end{array} & \text{if} \quad s_E - s_I < \Delta(N_I, N_E+1) \\ \\ \begin{array}{c} s_I - s_E + f_I(N_I) - f_E(N_E) \\ +\delta(\frac{s_I - s_E}{1-\delta} + \frac{\Delta(N_I+1,N_E) - \delta\Delta(N_I+1,N_E+1)}{(1-\delta)^2}) - \delta(\frac{s_E - s_I - \Delta(N_I,N_E+1)}{(1-\delta)^2}) \end{array} & \text{if} \quad \begin{array}{c} \Delta(N_I,N_E+1) \leq s_E - s_I \\ \leq \Delta(N_I+1,N_E+1) \end{array} \\ \\ \begin{array}{c} s_I - s_E + f_I(N_I) - f_E(N_E) \\ +\delta(\frac{s_I - s_E + \Delta(N_I+1,N_E)}{(1-\delta)^2}) - \delta(\frac{s_E - s_I}{1-\delta} - \frac{\Delta(N_I,N_E+1) - \delta\Delta(N_I+1,N_E+1)}{(1-\delta)^2}) \end{array} & \text{if} \quad \begin{array}{c} \Delta(N_I+1,N_E+1) \leq s_E - s_I \\ \leq \Delta(N_I+1,N_E) \end{array} \\ \\ \begin{array}{c} s_I - s_E + f_I(N_I) - f_E(N_E) \\ -\delta(\frac{s_E - s_I}{1-\delta} - \frac{\Delta(N_I,N_E+1) - \delta\Delta(N_I+1,N_E+1)}{(1-\delta)^2}) \end{array} & \text{if} \quad s_E - s_I \geq \Delta(N_I+1,N_E) \end{cases}$$

Straightforward calculations reveal that the expression in the first line is equal to

$$\frac{s_I - s_E}{1-\delta} + \frac{\Delta(N_I,N_E) - \delta\Delta(N_I,N_E+1)}{(1-\delta)^2},$$

while the expressions in the second and third lines are identical and equal to

$$\frac{s_I - s_E + \Delta(N_I, N_E)}{(1-\delta)^2}.$$

Since $s_I - s_E + f_I(N_I) - f_E(N_E) + \delta(V^I(N_I+1, N_E) - V^E(N_I, N_E+1))$ is continuous in $(s_I - s_E)$ and $\Delta(N_I, N_E) < \Delta(N_I+1, N_E)$, we can conclude that the expression of $V^I(N_I, N_E)$ given in Corollary 1 holds. By symmetry in I and E, the same is true for the expression of $V^E(N_I, N_E)$ given in Corollary 1. Thus, the induction hypothesis holds for (N_I, N_E).