Defection-Free Collaboration between Competitors in a Learning System

Anonymous Author(s)
Affiliation
Address
email

Abstract

We study collaborative learning systems in which the participants are competitors who will defect from the system if they lose revenue by collaborating. As such, we frame the system as a duopoly of competitive firms who are each training machine learning models and selling their predictions to a market of consumers. We first examine a fully collaborative scheme in which both firms share their models with each other and show that this leads to a market collapse with the revenues of both firms going to zero. We next show that one-sided collaboration in which only the firm with the lower-quality model shares improves the revenue of both firms. Finally, we propose a more equitable, *defection-free* scheme in which both firms share with each other while losing no revenue. We show that for a large range of starting conditions, our algorithm converges to the Nash bargaining solution, and we empirically verify our theory on computer vision datasets.

1 Introduction

When the guarantees of a collaborative learning system are misaligned with the objectives of the learners, it can disincentivize participation and cause the participants to defect. Recent work [4, 2, 21] examines the incentives that clients have to participate in or defect from a collaborative learning system. Such misalignment of incentives can arise in a number of ways. For example, [8] show that some clients might *free-ride*, burdening other participants in the network with all the training work while contributing nothing. [12, 10, 20, 5, 11, 16] show that if there is heterogeneity across clients' data distributions the global model returned by standard collaborative learning protocols might perform poorly for individual clients. To address the misalignment problem, [6] propose an algorithm whose model updates guarantee that client losses degrade sufficiently from step to step to ensure that no client defects (albeit at some cost to the accuracy of the final global model). In this paper, we take an economics-based view of the problem, framing client *utility/revenue* as the determining factor in defection. We frame clients as competitive firms who are selling their models' predictions to consumers and competing for market share. As in the standard collaborative learning protocol, the firms collaboratively train a global model, but if at any point in the process their revenue decreases, they defect from participation.

Motivating Example. Consider two autonomous vehicle companies training self-driving models, each with initial access only to their own training data. Further, suppose their individual training data does not fully reflect the distribution on which the models must perform well at test time. For example, one company might have a lot of urban data and very little rural data and the other company the opposite. Clearly, if these companies combined their models, they could offer safer and better cars to consumers. However, by collaborating they might also lose their competitive advantage in the market, disincentivizing them from participating. Our objective is to design a collaboration scheme such that neither firm loses revenue, thus incentivizing participation.

Submitted to 38th Conference on Neural Information Processing Systems (NeurIPS 2024). Do not distribute.

Our Contributions. We frame the collaborative learning system as a duopoly of competitive firms whose conditions for joining the system are to improve (or at least not lose) revenue, and we show that collaboration is possible under such conditions.

1. We first show surprising outcomes of two possible collaboration schemes. When both firms contribute fully to the collaboration scheme, their model qualities improve maximally but their revenues go to zero. When only the low-quality firm contributes to the collaboration scheme, both firms' model qualities and revenues improve.
2. We next design a defection-free algorithm which allows *both* firms to contribute to the collaborative system without losing revenue at any step.
3. We show that, except in trivial cases, our algorithm converges to the Nash bargaining solution. This is a significant result because we show that even when both firms myopically focus on improving their own revenues, a solution is reached that maximizes the joint surplus of the firms.

1.1 Related Work

Collaborative learning allows multiple clients to collaboratively train a global model without transmitting raw data [13]. In this paper, we characterize the participants in a collaborative learning system as market competitors who will defect from collaboration if they lose revenue by participating. Competitive behavior of firms in markets is a well-established field of study in economics (see [18] for an overview). Particularly relevant to our work is competition in oligopolies [3]. As in [7], we structure our problem as a duopoly of competitive firms. One difference is that they incentivize collaboration with revenue sharing between the firms rather than a guarantee of no-revenue-loss as we do in this paper. Also relevant, [19] parameterize the data sharing problem in terms of competition-type (Bertrand [1] or Cournot [3]) between firms, the number of data points each firm has, and the difficulty of the learning task, and give conditions on these parameters under which collaboration is profitable. As we do, they analyze various data sharing schemes, such as full vs partial collaboration, and propose Nash bargaining [14] as a strategy for partial collaboration. However, we additionally propose a federated optimization algorithm for reaching the Nash bargaining solution, guaranteeing no defections.

2 Collaborative Learning in an Oligopoly

For the rest of the paper, we frame the collaborative learning system as a duopoly (i.e. two firms), but all results can be extended to an oligopoly of more than two firms.

Our setup is the following. Each firm possesses a model whose qualities are initially differentiated by classification accuracy on a target dataset. That is, one firm's model has low accuracy and the other firm's model has high accuracy on the target dataset. The consumers care about performance on the target distribution, which is different from the firms' training distributions. For example, in the autonomous vehicle example above, the target distribution would represent a variety of geographical locations, traffic instances, times of day/night, etc. while the training distributions would not. Additionally we assume that the firms' training distributions are complementary, so the union of their training data is distributed as the target distribution, motivating the benefit of collaboration. Finally, we assume that, prior to collaboration, one firm has better initial model quality than the other (e.g. they have more training resources).

A consumer has one of three options: 1) pay a higher price for the high-quality firm's model, 2) pay a lower price for the low-quality firm's model, or 3) buy neither model. We assume that all consumers would prefer the higher-quality model if the prices of both models were the same – that is, the firms' models are *vertically differentiated*. Consumers would be happiest if both firms collaborated fully since this would give them two maximally good models to choose from, but the initially high-quality firm would have sacrificed revenue in this scenario (we show this formally in Section 3), causing it to defect. Based on this, our motivating question is: can we incentivize firms to join the collaboration scheme, thus benefiting consumers, while giving them no reason to defect due to revenue loss at any stage of the training process? We answer this question affirmatively.

In the following section, we formally describe the duopoly model.

2.1 Duopoly Model

2.1.1 Notation and Assumptions

1. A consumer's type corresponds to how much they value quality of prediction. We assume that consumer-types are uniformly distributed on $\Theta = [0, 1]$, where consumer-type $\theta = 0$ places no value on quality and consumer-type $\theta = 1$ places maximal value on quality.

2. We denote the low-quality firm's loss on its training dataset with model parameters $x \in \mathcal{X}$ as $f(x; l) \in [0, 1]$ and the high-quality firm's loss on its training dataset as $f(x; h) \in [0, 1]$. In the collaborative learning process, both firms want to solve the optimization problem

$$x^* = \arg\min_{x \in \mathcal{X}} f(x), \qquad \text{where } f(x) \overset{\text{def}}{=} \frac{f(x; l) + f(x; h)}{2}. \tag{1}$$

   That is, each firm wants to find the model which has minimal average loss across both firms' training datasets. When the objective (1) is evaluated at the firms' models x_l and x_h, we use the shorthand notation

$$f_l \overset{\text{def}}{=} \frac{f(x_l; l) + f(x_l; h)}{2}, \qquad\qquad f_h \overset{\text{def}}{=} \frac{f(x_h; l) + f(x_h; h)}{2}.$$

   Finally, we define model qualities $q(x) \overset{\text{def}}{=} 1 - f(x)$, $q_l \overset{\text{def}}{=} 1 - f_l$ and $q_h \overset{\text{def}}{=} 1 - f_h$.

3. Consumers pay prices $p_{l/h} \in [0, \infty)$ for the low/high-quality firm's model $x_{l/h}$, where $p_l \leq p_h$.

2.1.2 Equilibrium Quantities

The following definition gives the consumer's utility.

Definition 1. *[Consumer Utility] A type-θ consumer has utility*

$$U_c(\theta) = \begin{cases} \theta q_h - p_h & \text{if buys high-quality firm's model} \\ \theta q_l - p_l & \text{if buys low-quality firm's model} \\ 0 & \text{if buys neither model.} \end{cases} \tag{2}$$

The consumer utilities in Definition 1 induce the following demands for the firms.

Lemma 1 (Consumer Demands)**.** *Given the utilities in Definition 1,*

1. *consumer demand for the low-quality firm is $D_l = \frac{p_h - p_l}{q_h - q_l} - \frac{p_l}{q_l}$, and*
2. *consumer demand for the high-quality firm is $D_h = 1 - \frac{p_h - p_l}{q_h - q_l}$.*

Proof. See Appendix A.1. □

Using the consumer demands in Lemma 1, we can define the utilities of the firms.

Definition 2. *[Firm Utility/Revenue] The low/high firm's utility/revenue from selling its model is*

$$U_{l/h}(q_l, q_h, p_l, p_h) = p_{l/h} D_{l/h}. \tag{3}$$

At equilibrium, the firms will set prices p_l and p_h that maximize (3), yielding price-optimal utilities.

Lemma 2 (Equilibrium Prices and Utilities)**.** *The optimal prices for the low and high firms are*

$$p_l^* = \frac{q_l(q_h - q_l)}{4q_h - q_l}, \qquad p_h^* = \frac{2q_h(q_h - q_l)}{4q_h - q_l},$$

yielding price-optimal utilities

$$U_l(q_l, q_h, p_l^*, p_h^*) = \frac{q_l q_h (q_h - q_l)}{(4q_h - q_l)^2}, \qquad U_h(q_l, q_h, p_l^*, p_h^*) = \frac{4q_h^2 (q_h - q_l)}{(4q_h - q_l)^2}. \tag{4}$$

Proof. See Appendix A.1. □

Going forward, we will use the shorthand $U_{l/h} \overset{\text{def}}{=} U_{l/h}(q_l, q_h, p_l^*, p_h^*)$.

Remark 1. *Since the firms make their pricing decisions simultaneously and compete based on prices, this is the Bertrand model of competition [1]. This is distinct from other forms of oligopolistic competition, such as Cournot competition [3] in which firms compete based on quantity (i.e. the firms independently and simultaneously decide quantities to produce which then determine market price), or Stackelberg competition [17] in which the firms non-independently and sequentially decide quantities to produce.*

The following proposition states how the firms' utilities vary with quality and is key in our analysis going forward.

Proposition 1 (Relationship between utilities and qualities)**.** *For $q_l \leq q_h$,*

1. *U_h is increasing in q_h,*
2. *U_h is decreasing in q_h,*
3. *U_l is increasing in q_h, and*
4. *U_l is increasing in q_l for $q_l \leq \frac{4}{7} q_h$ and decreasing in q_l otherwise.*

Proof. See Appendix A.1 □

In the next section, we examine various collaboration schemes between the firms and observe the impact on their revenues and model qualities.

3 Collaboration Schemes

To motivate our method, we describe two potential collaboration schemes between competitors that have sub-optimal and non-intuitive outcomes.

Sharing Protocol. As in standard federated learning protocols, we do not assume that the firms transmit their raw data to each other. Instead, firm A shares with firm B by evaluating the loss of firm B's model on firm A's training data. Then firm A shares with firm B the loss, or the gradient of the loss, which allows firm B to optimize the objective (1). These exchanges can happen either directly between the firms are through a trusted central coordinator.

3.1 Notation and Assumptions

1. $f(x; {}^{l}/{}_{h})$ is convex and L-smooth in x.
2. We use $q_{l/h,t}$ and $f_{l/h,t}$ to refer to the firms' objectives when the model parameters are $x_{l/h,t}$, i.e. the model parameters at round t of optimization.
3. We define $\rho_t = \frac{q_{l,t}}{q_{h,t}}$, the ratio of the firms' model qualities at round t of optimization.
4. We assume model qualities can only improve or stay the same, not degrade.

3.2 Complete Collaboration

In this arrangement, both firms fully collaborate, sharing their models with each other and therefore obtaining identical-quality models. (Note that this algorithm is just FedAvg [13].) While this collaboration scheme is optimal for the consumer, giving them the choice of two maximally high-quality models, it drives both firms' utilities to zero. With identical-quality models, each firm will continually undercut the other's price by small amounts to capture the entire market share, eventually reaching equilibrium prices $p_l = p_h = 0$.

Lemma 3 (Firm Revenues under Complete Collaboration)**.** *Under Complete Collaboration, the firms' equilibrium utilities are $U_l = U_h = 0$.*

Figure 1 shows that when both firms' qualities increase freely in a Complete Collaboration scheme, their qualities both improve maximally, benefiting the consumer, but their utilities are driven to zero. Therefore, both firms have cause to defect from this collaboration scheme.

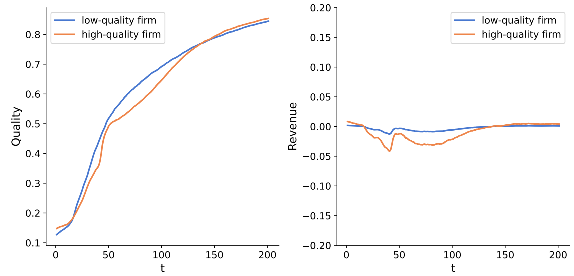


Figure 1: Performance of Complete Collaboration scheme on MNIST. When both firms share with each other, their models converge to the same qualities, driving their revenues to zero.

3.3 One-sided Collaboration

In One-sided Collaboration, one firm shares its model while the other doesn't. There are two possibilities.

Only high-quality firm shares. From Proposition 1, the high-quality firm's revenue increases in q_h but decreases in q_l. Therefore, if the quality of x_h does not increase sufficiently to compensate for the increase in quality of x_l, the high-quality firm will lose revenue, causing it to defect. (In the proof of Proposition 3, we give this increase-threshold precisely.) In our problem setup, the individual firms' training distributions are different than target distribution on which the qualities of their models are evaluated. Therefore, if the low-quality firm benefits from the high-quality firm's model, its performance on the target distribution will outpace the high-quality firm, which is limited to training on its own data. Figure 2a gives an example of this outcome. Due to collaboration, the low-quality firm's model out-performs the high-quality firm's model, causing the high-quality firm's revenue to decrease.

Only low-quality firm shares. From Proposition 1, both firms' utilities increase in q_h. Therefore, both firms will increase their revenue if the low-quality firm shares its model with the high-quality firm. Figure 2b depicts the outcome of this collaboration scheme. Over time, both firms' revenues increase. While this arrangement is defection-free, the low-quality firm is stuck with its own training data, causing it to potentially have lower revenue that it would under a more equitable scheme. To address this, we next propose a defection-free scheme in which *both* firms participate in collaboration without losing revenue.

4 Defection-Free Collaborative Learning

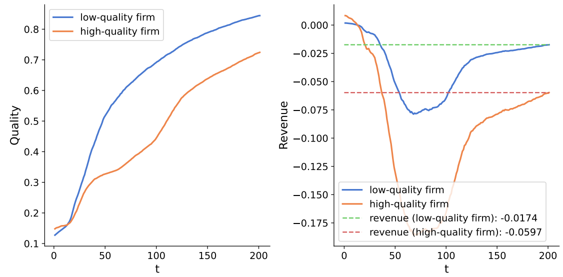
In this section, we introduce our method, Defection-Free Collaborative Learning. Our objectives in designing this algorithm are that

1. for all starting values $(q_{l,0}, q_{h,0})$, neither firm's revenue decreases at any round, and
2. the algorithm converges to the Nash bargaining solution, which we denote (q_l^*, q_h^*). (See Section 4.1.)

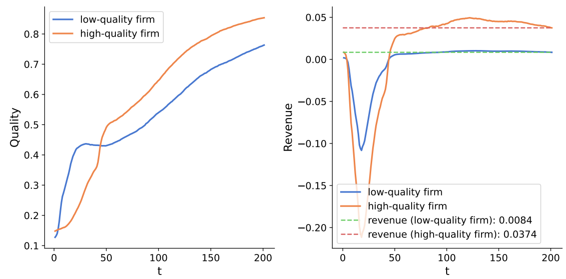
The first objective ensures that the algorithm is defection-free. The second seeks a point of convergence that maximizes the joint surplus of the firms. In Section 4.2, we show that Algorithm 1 achieves 1) entirely and achieves 2) for a large range of starting conditions. Before describing our algorithm, we first motivate the Nash bargaining solution as a suitable convergence goal for our problem setting.

4.1 Nash Bargaining

In cooperative bargaining, agents determine how to share a surplus amongst themselves. If negotiations fail, each agent is guaranteed some fixed surplus, known as the *disagreement point*. A typical application of bargaining involves deciding how to split a firm's profits amongst its employees. The bargaining framework is suitable for our purposes because the firms must agree how to share a



(a) Only high-quality firm shares.



(b) Only low-quality firm shares.

Figure 2: Performance of One-sided Collaboration schemes on MNIST. When only the high-quality firm shares, the high-quality firm's revenue becomes negative. When only the low-quality firm shares, both firms have positive, but less, revenue than with our collaboration scheme (Figure 3).

"surplus of quality" (i.e. set model qualities relative to each other) so that neither firm's revenue decreases at any one round.

An important framework in cooperative bargaining is Nash bargaining [14], a two-person bargaining scheme, which solves for

$$\begin{aligned}(q_l^*, q_h^*) = \arg\max_{(q_l, q_h)} \quad & N(q_l, q_h, q_{l,0}, q_{h,0}) \\ \text{s.t.} \quad & U_l(q_l, q_h) \geq U_l(q_{l,0}, q_{h,0}) \\ & U_h(q_l, q_h) \geq U_h(q_{l,0}, q_{h,0}),\end{aligned}$$

where

$$N(q_l, q_h, q_{l,0}, q_{h,0}) \stackrel{\text{def}}{=} (U_l(q_l, q_h) - U_l(q_{l,0}, q_{h,0}))(U_h(q_l, q_h) - U_h(q_{l,0}, q_{h,0})),$$

and $(q_{l,0}, q_{h,0})$ are the initial model qualities of the firms. The *Nash bargaining solution*, (q_l^*, q_h^*), maximizes the product of the *improvement* in the firms' utilities. Therefore, unlike one-sided collaboration, the Nash objective rewards improvement in the low-quality firm's utility as well as the high-quality firm's utility. In Nash bargaining, the *disagreement point* $(q_{l,0}, q_{h,0})$ determines the surplus for the parties if negotiations fall apart. In our setting, if either firm defects from collaboration, both firms retain their current model qualities. Going forward, we use $N(q_l, q_h)$ as shorthand for $N(q_l, q_h, q_{l,0}, q_{h,0})$. The Nash bargaining solution (q_l^*, q_h^*) has four important properties: 1) it is invariant to affine transformation of the utility functions, 2) it is pareto efficient, 3) it is symmetric, and 4) it is independent of irrelevant alternatives. In fact, the point (q_l, q_h) with these four properties is uniquely the Nash bargaining solution.

The next proposition shows that q_h^* is equivalent to the high-quality firm's maximal quality.

Proposition 2 (Equivalence between maximal quality and the Nash bargaining solution)**.**

$$q_h^* = \max_{x \in \mathcal{X}} q(x).$$

Proof. From Proposition 1, $\frac{\partial U_h}{\partial q_h}$ and $\frac{\partial U_l}{\partial q_h}$ are both non-negative for all $q_l \leq q_h$, and consequently $\frac{\partial N(q_l, q_h)}{\partial q_h} \geq 0$ for all $q_l \leq q_h$. This means that for any q_l, the $N(q_l, q_h)$ can always be improved by increasing q_h. Therefore, q_h^* is necessarily $\max_{x \in \mathcal{X}} q(x)$. □

Algorithm 1 Defection-Free Collaborative Learning

Input: Low-quality model: $x_{l,0}$. High-quality model: $x_{h,0}$.

Note: We assume both firms are trusted parties and will honestly exchange information. For example, to perform the necessary computations, the high-quality firm requires x_l and $\nabla f(x_h; l)$ from the low-quality firm, and the low-quality firm requires x_h, $\nabla f(x_l; h)$, $f(x_h; h)$, and $f(x_l; h)$ from the high-quality firm.

1: **for** $t \in [T]$ **do**

2:     **High-quality Model Update**

3:     Set $\alpha_{h,t} \leq \frac{1}{L}$.

4:     Update: $x_{h,t} = x_{h,t-1} - \alpha_{h,t} \nabla_{x_{h,t-1}} f_{h,t-1}$.

5:     **Low-quality Model Update**

6:     $x_{l,t} = x_{l,t-1}$.

7:     **if** $q_{l,t} < q_l^*$ and $\frac{q_{l,t}}{q_{h,t}} \leq \rho^* = \frac{q_l^*}{q_h^*}$ **then**

8:         Compute: $\hat{q}_{l,t} = B\left(\rho_{t-1}, \frac{q_{h,t}}{q_{h,t-1}}\right) q_{h,t}$, where

$$B(a, b) \stackrel{\text{def}}{=} 4 - \frac{(4-a)^2}{2(1-a)}\left(b - \sqrt{b^2 - \frac{12(1-a)}{(4-a)^2} b}\right).$$

9:         **while** $q_{l,t} \leq \hat{q}_{l,t}$ **do**

10:             Set: $\alpha_{l,t}$.

11:             Update: $x_{l,t} \leftarrow x_{l,t} - \alpha_{l,t} \nabla_{x_{l,t}} f_{l,t}$

12: **Output:** $x_{l,T}, x_{h,T}$

Section 3.3 shows there's a defection-free scheme in which the low-quality firm shares but the high-quality firm doesn't. In Algorithm 1, we give a way for both firms to contribute to collaboration with neither firm losing revenue at any step. Due to the more equitable design of this collaboration scheme, its dynamics mirror those of Nash bargaining which maximizes the joint surplus of the participants.

The difficulty of designing Algorithm 1 is that, in order to reach (q_l^*, q_h^*) without decreasing revenues at any step, neither firm can improve its quality too much in a given step. Given an increase in the high-quality firm's quality $q_{h,t-1} \to q_{h,t}$, the low-quality firm can only improve by some limited amount without decreasing the high-firm's revenue (since U_h is decreasing in q_l by Prop. 1). Because of this capped permissible improvement for the low-quality firm, if the high-quality firm converges to q_h^* too quickly, the low-quality firm will never reach q_l^*.

We describe the key steps of Algorithm 1. We also assume that, prior to the algorithm, both firms have saturated training on their own datasets and will only update their models collaboratively going forward. Since U_l and U_h both increase in q_h, the low-quality firm should always share with the high-quality firm. Step 4 ensures this, where the high-quality firm has access to the low-quality firm's loss on its model $x_{h,t-1}$ when updating. As we show in Section 4.2, in order to converge to the Nash bargaining solution, the low-quality firm should not update if $q_{l,t} \geq q_l^*$ or $\rho_{t-1} > \rho^*$. Step 7 ensures this. Since U_h decreases in q_l, the low-quality firm cannot improve its model beyond a certain threshold before the high-quality firm loses revenue. This threshold $\hat{q}_{l,t}$ is computed in Step 8, and in Steps 9-11, the high-quality firm will only collaborate if the collaborative updates to the low-quality firm's model do not improve its quality beyond $\hat{q}_{l,t}$.

In the next section we prove the two key properties of Defection-Free Collaborative Learning: 1) it guarantees the firms non-decreasing revenue at every step, and 2) it converges to the Nash bargaining solution for all but trivial starting conditions.

4.2 Theory and Analysis

The following proposition shows that Algorithm 1 is defection-free.

Proposition 3 (Non-decreasing revenues). *There exist learning rate schedules $\{\alpha_{l,t}\}_t$ and $\{\alpha_{h,t}\}_t$ such that at no step of Algorithm 1 does either firm's revenue decrease.*

Proof. See Appendix A.2. □

We next examine starting conditions for which Algorithm 1 converges to the Nash bargaining solution. Proposition 4 gives a trivial starting condition for which it does not converge.

Proposition 4 (Impossibility of convergence to the Nash bargaining solution)**.** *If $q_{l,0} > q_l^*$, then Algorithm 1 cannot converge to (q_l^*, q_h^*).*

Proof. Since firms do not degrade their model quality, the low-quality firm cannot converge to q_l^*. □

In the next proposition, we show that for all other starting conditions, Algorithm 1 converges to (q_l^*, q_h^*). Our key insight in the proof of this proposition is that if the high-quality firm converges too quickly to q_h^*, the low-quality firm will not be able to make sufficient progress towards q_l^* without violating the no-revenue-loss condition. Therefore, we must design a learning rate schedule for the high-quality firm $\{\alpha_{h,t}\}_t$ such that convergence to q_h^* is properly paced.

Proposition 5 (Convergence to the Nash bargaining solution)**.** *If $q_{l,0} \leq q_l^*$, then there exist learning rate schedules $\{\alpha_{l,t}\}_{t=1}^T$ and $\{\alpha_{h,t}\}_{t=1}^T$ such that after T rounds Algorithm 1 converges to (q_l^*, q_h^*).*

Proof. See Appendix A.2. □

Proposition 5 shows that even when both firms myopically attend to improving their own revenues, Algorithm 1 converges to the Nash bargaining solution which maximizes joint surplus. The following theorem gives the rate of convergence to the Nash bargaining solution for convex and L-smooth losses.

Theorem 1 (Convergence Rate of Defection-Free Collaborative Learning)**.** *Suppose $q_{l,0} \leq q_l^*$. Then running Algorithm 1 for T rounds ensures*

$$N(q_l^*, q_h^*) - N(q_{l,T}, q_{h,T}) \lesssim \frac{\|x_{h,0} - x_h^*\|^2}{\sum_{t=1}^T \alpha_{h,t}} + |\rho^* - \rho_T|. \tag{5}$$

Proof. See Appendix A.2. □

The first term in the bound (5) shows that the convergence rate to the Nash bargaining solution is determined by the rate at which q_h converges to q_h^*.

The following corollary shows the rate at which the $|\rho^* - \rho_T|$ term in Theorem 1 vanishes with T.

Corollary 1. *Suppose $q_{l,0} \leq q_l^*$. Running Algorithm 1 for $T \gtrsim \frac{L\|x_{h,0} - x_h^*\|^2}{\epsilon}$ rounds ensures that*

$$N(q_l^*, q_h^*) - N(q_{l,T}, q_{h,T}) \lesssim \frac{\|x_{h,0} - x_h^*\|^2}{\sum_{t=1}^T \alpha_{h,t}} + (4 - 5\rho^*) \log\left(\frac{q_h^*}{q_h^* - \epsilon}\right).$$

Proof. See Appendix A.2. □

5 Experiments

All algorithms in our experiments are implemented with PyTorch [15]. Our general experimental setup is the following. We construct three datasets: the low-quality firm's training set $\mathcal{D}_{l,\text{train}}$, the high-quality firm's training set $\mathcal{D}_{h,\text{train}}$, and a common test set for both firms $\mathcal{D}_{\text{target}}$. The datasets are constructed such that $\mathcal{D}_{l,\text{train}} \not\sim \mathcal{D}_{\text{target}}$ and $\mathcal{D}_{h,\text{train}} \not\sim \mathcal{D}_{\text{target}}$, but $\mathcal{D}_{l,\text{train}} \cup \mathcal{D}_{h,\text{train}} \sim \mathcal{D}_{\text{target}}$, i.e. neither firm's training distribution alone matches the target distribution, but their combined training datasets are distributed as the target distribution, incentivizing them to share. We use cross-entropy loss, PyTorch's built-in SGD optimizer, and local compute for all experiments.

MNIST We use a LeNet-5 model [9], set $|\mathcal{D}_{l,\text{train}}| = |\mathcal{D}_{h,\text{train}}| = 1000$, and use the MNIST test set as $\mathcal{D}_{\text{target}}$. We construct $\mathcal{D}_{l,\text{train}}$ so that $\hat{F}(5) = 0.8$ and $\mathcal{D}_{h,\text{train}}$ so that $\hat{F}(5) = 0.2$, where $\hat{F}$ is the empirical CDF over the label space. We train the high-quality firm's model for 10 initial epochs, and for all models and experiments set the learning rate to 0.01.

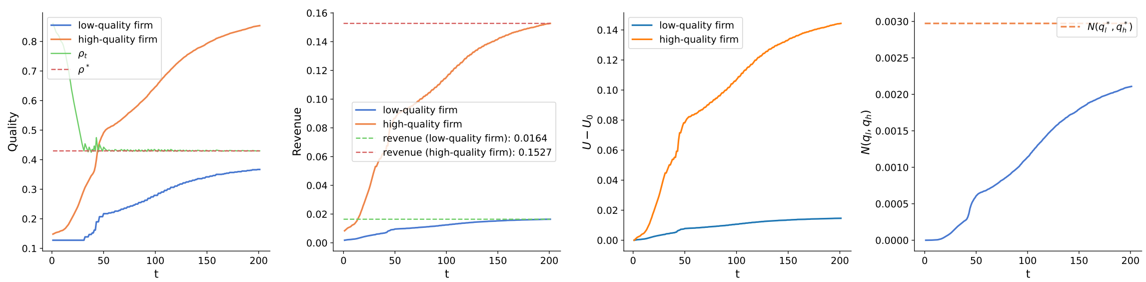


Figure 3: Performance of Defection-Free FL on MNIST. Both firms' qualities increase (figure 1), their revenues increase and approach a higher level than under One-sided Collaboration (figure 2), and the firms' qualities approach the Nash bargaining solution (figure 4).

Defection-Free Collaborative Learning (Fig. 3). Since the low-quality firm shares with the high-quality firm, the high-quality firm improves maximally. The high-quality firm only shares with the low-quality firm to the extent that neither firm's revenue decreases. Under this sharing scheme, we see in the first figure that both firms' qualities increase, and the ratio of their qualities converges to the optimal ratio. The second figure shows that revenues increase (do not decrease), and notably their revenues reach a higher level than under One-sided Collaboration (Section 3.3). Finally, the last figure shows that the Nash bargaining objective approaches its maximal value, showing convergence to the Nash bargaining solution.

6 Conclusion

Contributions. We introduce a defection-free collaborative learning scheme in which participants iteratively optimize their models by sharing training resources, without losing utility at any round and having cause to defect from participation. Framing the collaborative learning system as a duopoly of competitive firms, we show that both firms can improve their model qualities by sharing data with each other without losing revenue at any round. We describe other collaboration schemes for which this is not possible. Notably, even when both firms myopically focus on improving their own revenues, we show that our algorithm converges to the Nash bargaining solution, thus optimizing for joint surplus.

Limitations/Future Work. Future work involves more precise convergence rate analysis (e.g. for a broader class of loss functions besides convex, and a more detailed rate in Theorem 1). We only study a duopoly model, but examining an oligopoly of multiple firms may present different dynamics. Finally, a broader conversation about societal impact on consumers is open for future work.

References

[1] Joseph Bertrand. Theorie mathematique de la richesse sociale. *Journal des Savants*, 68:499–508, 1883.

[2] Avrim Blum, Nika Haghtalab, Richard Lanas Phillips, and Han Shao. One for one, or all for all: Equilibria and optimality of collaboration in federated learning. In *Proceedings of the 38th International Conference on Machine Learning*, 2021.

[3] Augustin Cournot. Recherches sur les principes mathématiques de la théorie des richesses. 1838.

[4] Kate Donahue and Jon Kleinberg. Model-sharing games: Analyzing federated learning under voluntary participation. In *35th AAAI Conference on Artificial Intelligence*, 2021.

[5] Avishek Ghosh, Jichan Chung, Dong Yin, and Kannan Ramchandran. An efficient framework for clustered federated learning. *Advances in Neural Information Processing Systems*, 33:19586–19597, 2020.

[6] Minbiao Han, Kshitij Patel, Kumar, Han Shao, and Lingxiao Wang. On the effect of defections in federated learning and how to prevent them. 2023.

[7] Chao Huang, Shuqi Ke, and Xin Liu. Duopoly business competition in cross-silo federated learning. volume 11, 2023.

[8] Sai Praneeth Karimireddy, Wenshuo Guo, and Michael Jordan. Mechanisms that incentivize data sharing in federated learning. In *Federated Learning Conference at 36th Conference on Neural Information Processing Systems*, 2022.

[9] Yann LeCun, Leon Bottou, Yoshio Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. volume 86, pages 2278–2324, 1998.

[10] T. Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. Fair and robust federated learning through personalization. In *38th International Conference on Machine Learning*, 2021.

[11] Yishay Mansour, Mehryar Mohri, Jae Ro, and Ananda Theertha Suresh. Three approaches for personalization with applications to federated learning. 2021.

[12] Othmane Marfoq, Giovanni Neglia, Laetitia Kameni, and Richard Vidal. Federated multi-task learning under a mixture of distributions. In *Proceedings of the 35th International Conference on Machine Learning*, volume 34, 2021.

[13] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[14] John F. Jr. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950.

[15] Adam Paszke. Pytorch: An imperative style, high-performance deep learning library. In *33rd Conference on Neural Information Processing Systems*, 2019.

[16] Felix Sattler, Klaus-Robert Muller, and Wojciech Samek. Clustered federated learning: Model-agnostic distributed multi-task optimization under privacy constraints. In *IEEE Transactions on Neural Networks and Learning Systems*, volume 32(8), 2021.

[17] Heinrich Freiherr von Stackelberg. Marktform und gleichgewicht. 1934.

[18] Jean Tirole. The theory of industrial organization. 1(0262200716), 1988.

[19] Nikita Tsoy and Nikola Konstantinov. Strategic data sharing between competitors. In *37th Conference on Neural Information Processing Systems*, 2023.

[20] Mariel Werner, Lie He, Michael Jordan, Martin Jaggi, and Sai Praneeth Karimireddy. Provably personalized and robust federated learning. *Transactions on Machine Learning Research*, 2023.

[21] Xiaohu Wu and Han Yu. Mars-fl: Enabling competitors to collaborate in federated learning. pages 1–11, 2022.

A Proofs

A.1 Proofs for Section 2.1

Proof of Lemma 1. Let $\hat{\theta}_l$ be the type of the consumer who is indifferent between buying from the low-quality firm and not buying at all. Then, based on the consumer's utility function (19),

$$\hat{\theta}_l q_l - p_l = 0. \tag{6}$$

Let $\hat{\theta}_h$ be the type of the consumer who is indifferent between buying from the high-quality firm and low-quality firm. Then, from (19),

$$\hat{\theta}_h q_l - p_l = \hat{\theta}_h q_h - p_h. \tag{7}$$

Therefore any consumer with type $\theta \in [\hat{\theta}_l, \hat{\theta}_h)$ will buy from the low-quality firm and any consumer with type $\theta \in [\hat{\theta_h}, 1]$ will buy from the high-quality firm, giving demands $D_l = \hat{\theta}_h - \hat{\theta}_l$ and $D_h = 1 - \hat{\theta_h}$. Solving (6) and (7) for $\hat{\theta}_l$ and $\hat{\theta}_h$ completes the proof. □

Proof of Lemma 7. From Lemma 1, the demand for the low-quality firm is $D_l = \frac{p_h - p_l}{q_h - q_l} - \frac{p_l}{q_l}$, yielding low-quality firm utility

$$U_l = p_l\left(\frac{p_h - p_l}{q_h - q_l} - \frac{p_l}{q_l}\right). \tag{8}$$

To maximize its utility, the low-quality firm sets price

$$\begin{aligned} p_l^* &= \arg\max_{p_l} \frac{\partial U_l}{\partial p_l} \\ &= \arg\max_{p_l} \left(\frac{p_h - 2p_l}{q_h - q_l} - \frac{2p_l}{q_l}\right) \\ &= \frac{q_l p_h}{2q_h}. \end{aligned} \tag{9}$$

Similarly, demand for the high-quality firm is $D_h = 1 - \frac{p_h - p_l}{q_h - q_l}$, yielding high-quality firm utility

$$U_h = p_h\left(1 - \frac{p_h - p_l}{q_h - q_l}\right). \tag{10}$$

To maximize its utility, the high-quality firm sets price

$$\begin{aligned} p_h^* &= \arg\max_{p_h} \frac{\partial U_h}{\partial p_h} \\ &= \arg\max_{p_h} \left(1 - \frac{2p_h - p_l}{q_h - q_l}\right) \\ &= \frac{p_l + (q_h - q_l)}{2}. \end{aligned} \tag{11}$$

Resolving (9) and (11) yields

$$p_l^* = \frac{q_l(q_h - q_l)}{4q_h - q_l} \tag{12}$$

and

$$p_h^* = \frac{2q_h(q_h - q_l)}{4q_h - q_l}. \tag{13}$$

Finally, evaluating (8) and (10) at the optimal prices (12) and (13) yields the price-optimal utilities (20). □

Proof of Proposition 1. The proposition follows from observing the partial derivatives of the firms' utility functions. For $q_l \leq q_h$,

$$\frac{\partial U_h}{\partial q_h} = \frac{4q_h(4q_h^2 - 3q_h q_l + 2q_l^2)}{(4q_h - q_l)^3} \geq 0,$$

$$\frac{\partial U_l}{\partial q_h} = \frac{q_l^2(2q_h + q_l)}{(4q_h - q_l)^3} \geq 0,$$

$$\frac{\partial U_l}{\partial q_l} = \frac{q_h^2(4q_h - 7q_l)}{(4q_h - q_l)^3} \begin{cases} \geq 0 & \text{if } q_l \leq \frac{4}{7}q_h \\ < 0 & \text{if } q_l > \frac{4}{7}q_h \end{cases}$$

and

$$\frac{\partial U_h}{\partial q_l} = -\frac{4q_h^2(2q_h + q_l)}{(4q_h - q_l)^3} \leq 0.$$

Figure 4 provides a graphical representation of this proposition. □

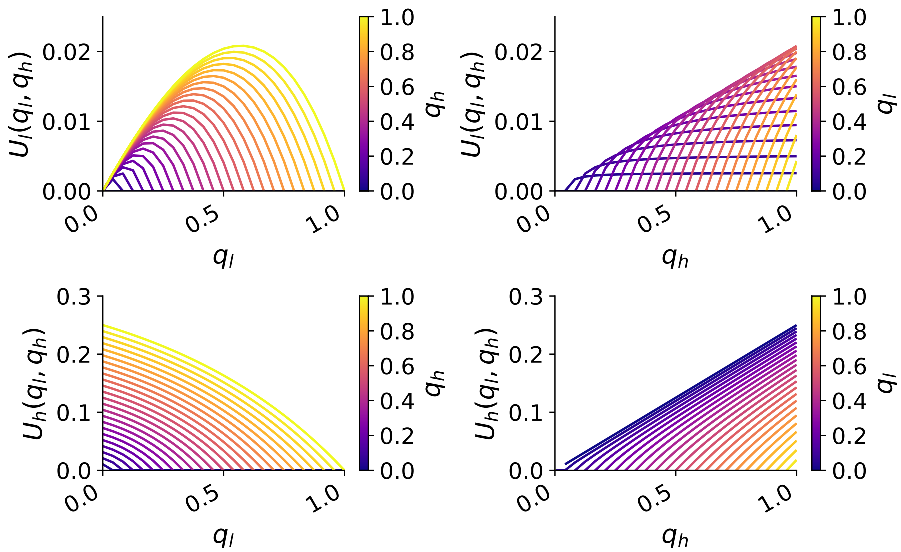


Figure 4: This figure shows how the firms' utilities vary with model quality. U_l and U_h are both increasing in q_h, U_h is decreasing in q_l, and U_l is increasing in q_l for $q_l \leq \frac{4q_h}{7}$ and decreasing in q_l otherwise.

A.2 Proofs for Section 4.2

Proof of Proposition 3. Suppose that at round t, given current qualities $q_{l,t-1}$ and $q_{h,t-1}$, the high-quality firm improves to $q_{h,t}$. Then, in order for neither firm to lose revenue, $q_{l,t}$ must be such that

$$\frac{4q_{h,t}^2(q_{h,t} - q_{l,t})}{(4q_{h,t} - q_{l,t})^2} \geq \frac{4q_{h,t-1}^2(q_{h,t-1} - q_{l,t-1})}{(4q_{h,t-1} - q_{l,t-1})^2} \tag{14}$$

and

$$\frac{q_{l,t}q_{h,t}(q_{h,t} - q_{l,t})}{(4q_{h,t} - q_{l,t})^2} \geq \frac{q_{l,t-1}q_{h,t-1}(q_{h,t-1} - q_{l,t-1})}{(4q_{h,t-1} - q_{l,t-1})^2}. \tag{15}$$

Rearranging terms, (14) can be written as an inequality involving a convex quadratic of $q_{l,t}$:

$$\begin{aligned}
&[4q_{h,t-1}^2(q_{h,t-1} - q_{l,t-1})]q_{l,t}^2 \\
&\quad + [4(4q_{h,t-1} - q_{l,t-1})^2 q_{h,t}^2 - 32q_{h,t-1}^2(q_{h,t-1} - q_{l,t-1})q_{h,t}]q_{l,t} \\
&\quad + [64q_{h,t-1}^2(q_{h,t-1} - q_{l,t-1})q_{h,t}^2 - 4(4q_{h,t-1} - q_{l,t-1})^2 q_{h,t}^3] < 0.
\end{aligned}$$

The right-most root of this quadratic is

$$q_{l,t}^h = 4q_{h,t} - \frac{(4-\rho_{t-1})^2}{2(1-\rho_{t-1})}\left(\frac{q_{h,t}^2}{q_{h,t-1}} - \sqrt{\frac{q_{h,t}^4}{q_{h,t-1}^2} - \frac{12(1-\rho_{t-1})}{(4-\rho_{t-1})^2}\frac{q_{h,t}^3}{q_{h,t-1}}}\right).$$

Similarly, (15) can be written as an inequality involving a convex quadratic of $q_{l,t}$:

$$\begin{aligned}
&[q_{l,t-1}q_{h,t-1}(q_{h,t-1} - q_{l,t-1}) + (4q_{h,t-1} - q_{l,t-1})^2 q_{h,t}]q_{l,t}^2 \\
&\quad + [-8q_{l,t-1}q_{h,t-1}(q_{h,t-1} - q_{l,t-1})q_{h,t} - (4q_{h,t-1} - q_{l,t-1})^2 q_{h,t}^2]q_{l,t} \\
&\quad + [16q_{l,t-1}q_{h,t-1}(q_{h,t-1} - q_{l,t-1})q_{h,t}^2] < 0.
\end{aligned}$$

The right-most root of this quadratic is

$$q_{l,t}^l = \frac{8(1-\rho_{t-1})\rho_{t-1}q_{h,t-1} + (4-\rho_{t-1})^2 q_{h,t} + (4-\rho_{t-1})\sqrt{(4-\rho_{t-1})^2 q_{h,t}^2 - 48\rho_{t-1}(1-\rho_{t-1})q_{h,t-1}q_{h,t}}}{2((1-\rho_{t-1})\rho_{t-1}q_{h,t-1} + (4-\rho_{t-1})^2 q_{h,t})}.$$

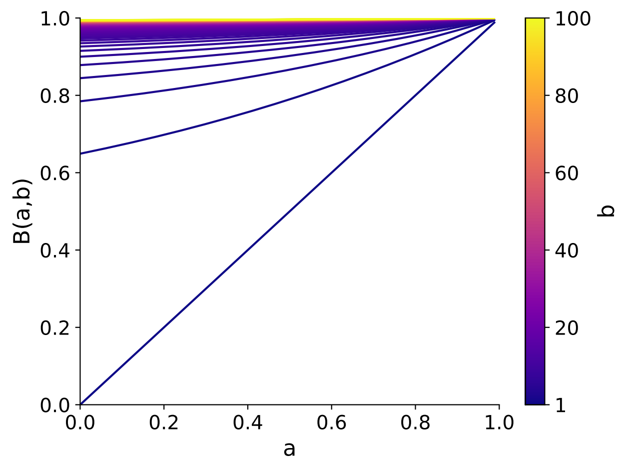


Figure 5: $B(a, b) \geq a$ for all $b \geq 1$.

It can be verified with graphing software that for all feasible parameters, $q^h_{l,t} \leq q^l_{l,t}$. Therefore, the low-quality firm can improve its quality to at most

$$\hat{q}_{l,t} = 4q_{h,t} - \frac{(4-\rho_{t-1})^2}{2(1-\rho_{t-1})}\left(\frac{q^2_{h,t}}{q_{h,t-1}} - \sqrt{\frac{q^4_{h,t}}{q^2_{h,t-1}} - \frac{12(1-\rho_{t-1})}{(4-\rho_{t-1})^2}\frac{q^3_{h,t}}{q_{h,t-1}}}\right),$$

before at least one of the firms loses revenue. Algorithm 1 ensures that $q_{l,t}$ does not exceed $\hat{q}_{l,t}$.

It remains to prove that there exist learning rate sequences $\{\alpha_{l,t}\}_t$ and $\{\alpha_{h,t}\}_t$ that respect the constraint $q_{l,t} \leq \hat{q}_{l,t}$. Since improvement in q_h increases the revenues of both firms (Prop. 1), the high-quality firm can set any learning rate schedule $\{\alpha_{h,t}\}_t$ without violating the no-revenue-loss constraints (14) and 15. For the low-quality firm's learning rate schedule, note that $f_l(x)$, as the average of convex functions $f(x; l)$ and $f(x; h)$, is also convex. Therefore,

$$\begin{aligned} f_{l,t} &\geq f_{l,t-1} + \nabla_{x_{l,t-1}} f^T_{l,t-1}(x_{l,t} - x_{l,t-1}) \\ &= f_{l,t-1} - \alpha_{l,t}\|\nabla_{x_{l,t-1}} f_{l,t-1}\|^2. \end{aligned}$$

Rearranging terms,

$$\begin{aligned} \alpha_{l,t} &\geq \frac{f_{l,t-1} - f_{l,t}}{\|\nabla_{x_{l,t-1}} f_{l,t-1}\|^2} \\ &= \frac{q_{l,t} - q_{l,t-1}}{\|\nabla_{x_{l,t-1}} f_{l,t-1}\|^2}. \end{aligned}$$

Therefore, setting $\alpha_{l,t} = \min\left\{\frac{\hat{q}_{l,t} - q_{l,t-1}}{\|\nabla_{x_{l,t-1}} f_{l,t-1}\|^2}, 1\right\}$ ensures that the low-quality firm's updated quality $q_{l,t}$ does not exceed $\hat{q}_{l,t}$. □

Proof of Proposition 5. We handle the proof in cases.

Case 1: $q_{l,0} \leq q^*_l$ and $\rho_0 \geq \rho^*$.

When $\frac{q_{l,t-1}}{q_{h,t}} \geq \rho^*$, the low-quality firm does not update (line 7 of Alg. 1). Once the high-quality firm improves sufficiently so that $\frac{q_{l,t}}{q_{h,t}} = \rho^*$ (note that such a t exists if $q_{l,0} \leq q^*_l$), then convergence is guaranteed. To see this, we use the following lemma.

Lemma 4. *$B(a, b) \geq a$ for all $b \geq 1$. (See Figure 5 for pictorial proof.)*

Consider step $t+1$ at which $\rho_t = \frac{q_{l,t}}{q_{h,t}} = \rho^*$. Given the high-quality firm's improvement $q_{h,t} \to q_{h,t+1}$, if the low-quality firm improves to $q_{l,t+1} = \hat{q}_{l,t+1}$, by Lemma 4, $\rho_{t+1} \geq \rho_t$. Therefore the

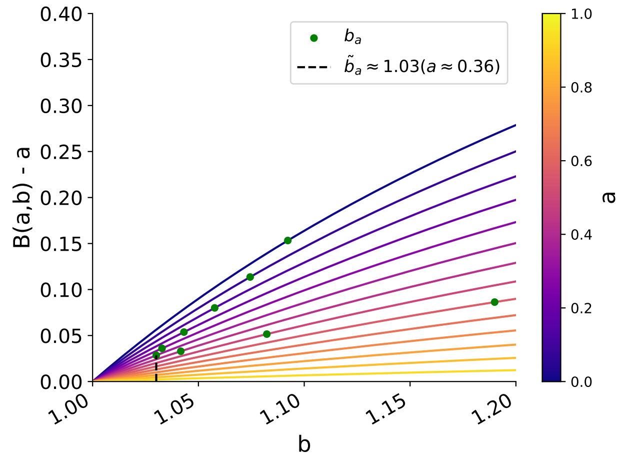


Figure 6: The green dots indicate, for a given $q_{l,t-1}/q_{h,t-1}$ (symbolized by a), the upper bound on $q_{h,t}/q_{h,t-1}$ that ensures convergence to the Nash bargaining solution.

low-quality firm can always improve to some level $q_{l,t+1} \in [q_{l,t}, \hat{q}_{l,t+1}]$ and ensure that $\rho_{t+1} = \rho^*$ with neither firm losing revenue. Maintaining this improvement schedule, once the high-quality firm improves to q_h^* (using any sequence of learning rates $\{\alpha_{h,t}\}_t$), the low-quality firm will be able to reach q_l^* by observing the constraint in lines 9-11 of Alg. 1.

Case 2: $q_{l,0} \leq q_l^*$ and $\rho_0 < \rho^*$.

Our strategy for this case will be to show there exist sequences of learning rates $\{\alpha_{h,t}\}_t$ and $\{\alpha_{l,t}\}_t$ such that $\sum_{t=1}^{T}(\rho_t - \rho_{t-1}) = \rho_T - \rho_0 \geq \rho^* - \rho_0$. We will do this by lower-bounding the quality-ratio gaps $\rho_t - \rho_{t-1} = B(\rho_{t-1}, q_{h,t}/q_{h,t-1}) - \rho_{t-1}$.

For each $\rho \leq 1$, there is a point (possibly infinite)

$$b_\rho \stackrel{\text{def}}{=} \max\{b \geq 1 : (4 - 5\rho)\log_{10} b \leq B(\rho, b) - \rho\}.$$

That is, for a given ρ, b_ρ is the point at which $(4 - 5\rho)\log b$ goes from being a lower to an upper bound on $B(\rho, b) - \rho$. Define $\tilde{b}$ as the smallest such point over all $\rho \leq 1$, so

$$\tilde{b} \stackrel{\text{def}}{=} \min_{\rho \leq 1} b_\rho.$$

Figure 6 plots b_ρ for various values of ρ and shows that $\tilde{b} \approx 1.03 = b_{\rho \approx 0.33}$.

By definition of $\tilde{b}$, $(4 - 5\rho)\log_{10} b \leq B(\rho, b) - \rho$ for any $\rho \leq 1$ and $b \leq \tilde{b}$. Suppose the high-quality firm maintains a learning rate schedule $\{\alpha_{h,t}\}_t$ such that $q_{h,t}/q_{h,t-1} \leq \tilde{b}$ for all t and T is such that $q_h^* - q_{h,T} \leq \epsilon$. Then

$$\begin{aligned}
\sum_{t=1}^{T}(\rho_t - \rho_{t-1}) &= \sum_{t=1}^{T}(B(\rho_{t-1}, q_{h,t}/q_{h,t-1}) - \rho_{t-1}) \\
&\stackrel{(i)}{\geq} \sum_{t=1}^{T}(4 - 5\rho_{t-1})\log_{10}(q_{h,t}/q_{h,t-1}) \\
&\stackrel{(ii)}{\geq} (4 - 5\rho^*)\log_{10}(q_{h,T}/q_{h,0}) \\
&\geq (4 - 5\rho^*)\log_{10}((q_h^* - \epsilon)/q_{h,0}),
\end{aligned}$$

where (i) is due to $q_{h,t}/q_{h,t-1} \leq \tilde{b}$, and (ii) is due to the fact that $\rho_0 \leq \rho^*$ and Lemma 4.

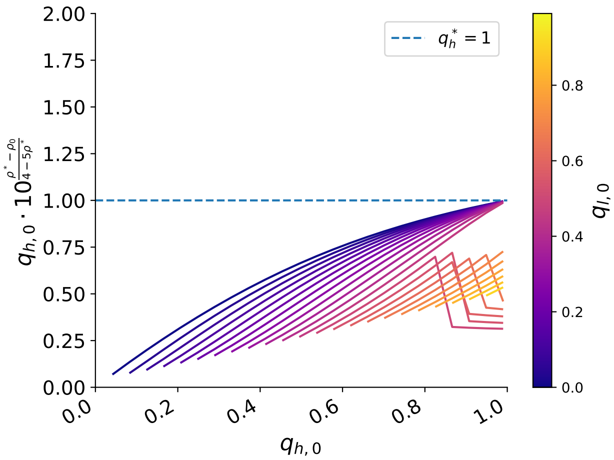


Figure 7: Empirical verification of the inequality: $(4 - 5\rho^*)\log_{10}(q_h^*/q_{h,0}) \geq \rho^* - \rho_0$

Figure 7 shows that $(4 - 5\rho^*)\log_{10}(q_h^*/q_{h,0}) \geq \rho^* - \rho_0$, so

$$(4 - 5\rho^*)\log_{10}\left(\frac{q_h^* - \epsilon}{q_{h,0}}\right) = (4 - 5\rho^*)\left(\log_{10}\left(\frac{q_h^*}{q_{h,0}}\right) - \log_{10}\left(\frac{q_h^*}{q_h^* - \epsilon}\right)\right)$$
$$\geq (\rho^* - \rho_0) - (4 - 5\rho^*)\log_{10}\left(\frac{q_h^*}{q_h^* - \epsilon}\right).$$

Therefore $\rho^* - \rho_T \leq (4 - 5\rho^*)\log_{10}\left(\frac{q_h^*}{q_h^* - \epsilon}\right)$.

It remains to show that there exists a sequence of learning rates $\{\alpha_{h,t}\}_t$ such that $q_{h,t}/q_{h,t-1} \leq \tilde{b}$, and T such that $q_h^* - q_{h,T} \leq \epsilon$. Let $\alpha_{h,t} = \min\left\{\frac{(\tilde{b}-1)q_{h,t-1}}{\|\nabla_{x_{h,t-1}} f_{h,t-1}\|^2}, \frac{1}{L}\right\}$. We analyze what happens when $\alpha_{h,t}$ is each of the values in the *min* expression.

First, suppose $\alpha_{h,t} = \frac{(\tilde{b}-1)q_{h,t-1}}{\|\nabla_{x_{h,t-1}} f_{h,t-1}\|^2}$ for all t. f_h, as the average of L-smooth and convex functions, is also L-smooth and convex, so that

$$\frac{q_{h,t-1} + \frac{\alpha_{h,t}}{2}\|\nabla_{x_{h,t-1}} f_{h,t-1}\|^2}{q_{h,t-1}} \leq \frac{q_{h,t}}{q_{h,t-1}} \leq \frac{q_{h,t-1} + \alpha_{h,t}\|\nabla_{x_{h,t-1}} f_{h,t-1}\|^2}{q_{h,t-1}}.$$

Therefore, the choice of $\alpha_{h,t}$ guarantees that $\frac{\tilde{b}+1}{2} \leq \frac{q_{h,t}}{q_{h,t-1}} \leq \tilde{b}$, giving $\frac{q_{h,T}}{q_{h,0}} \geq \left(\frac{\tilde{b}+1}{2}\right)^T$. From this we see that setting $T \geq \frac{\log(q_h^*/q_{h,0})}{\log((\tilde{b}+1)/2)}$ guarantees convergence to q_h^* in T' steps.

Now suppose $\alpha_{h,t} = \frac{1}{L}$ for all t. Under this condition, standard convergence analysis for gradient descent on convex and L-smooth functions gives

$$f_{h,T} - f_h^* \leq \frac{L\|x_{h,0} - x_h^*\|^2}{2T}.$$

Therefore, $f_{h,T} - f_h^* \leq \epsilon$ after $T = \frac{L\|x_{h,0} - x_h^*\|^2}{2\epsilon}$ rounds.

From the above analysis, we see that after at most $T = \frac{\log(q_h^*/q_{h,0})}{\log((\tilde{b}+1)/2)} + \frac{L\|x_{h,0} - x_h^*\|^2}{2\epsilon}$ rounds, $f_{h,T} - f_h^* = q_h^* - q_{h,T} \leq \epsilon$, completing the proof. □

Proof of Theorem 1. By Taylor's theorem,

$$
\begin{aligned}
N(q_l^*, q_h^*) &\leq N(q_{l,T}, q_{h,T}) + \frac{\partial N(q_l, q_h)}{\partial q_l}(q_l^* - q_{l,T}) + \frac{\partial N(q_l, q_h)}{\partial q_h}(q_h^* - q_{h,T}) \\
&\quad + \left(\max_{q_l, q_h} \frac{\partial^2 N(q_l, q_h)}{\partial q_l^2} \right) \frac{(q_l^* - q_{l,T})^2}{2} + \left(\max_{q_l, q_h} \frac{\partial^2 N(q_l, q_h)}{\partial q_h^2} \right) \frac{(q_h^* - q_{h,T})^2}{2} \\
&\quad + \left(\max_{q_l, q_h} \frac{\partial^2 N(q_l, q_h)}{\partial q_h \partial q_l} \right) (q_l^* - q_{l,T})(q_h^* - q_{h,T}) \\
&\overset{(i)}{\leq} c_1(q_h^* - q_{h,T}) + c_2(\rho^*(q_h^* - q_{h,T}) + q_{h,T}|\rho^* - \rho_T|) \\
&\lesssim (q_h^* - q_{h,T}) + |\rho^* - \rho_T|,
\end{aligned}
$$

where (i) follows from the fact that the gradients of N are bounded by small constants (can be verified with graphing software), qualities $q \in [0, 1]$, and $q_l^* - q_{l,T} = \rho^* q_h^* - \rho_T q_{h,T} \leq \rho^*(q_h^* - q_{h,T}) + q_{h,T}|\rho^* - \rho_T|$.

We now bound $q_h^* - q_{h,T}$. Note that f_h, as the average of L-smooth and convex functions, is also L-smooth and convex. Therefore,

$$
\begin{aligned}
f_{h,t} &\overset{(i)}{\leq} f_{h,t-1} + \left(-\alpha_{h,t} + \frac{L\alpha_{h,t}^2}{2} \right) \|\nabla_{x_{h,t-1}} f_{h,t-1}\|^2 \\
&\overset{(ii)}{\leq} f_{h,t-1} - \frac{\alpha_{h,t}}{2} \|\nabla_{x_{h,t-1}} f_{h,t-1}\|^2 \\
&\overset{(iii)}{\leq} f_h^* + \nabla_{x_{h,t-1}} f_{h,t-1}^T (x_{h,t-1} - x_h^*) - \frac{\alpha_{h,t}}{2} \|\nabla_{x_{h,t-1}} f_{h,t-1}\|^2 \\
&= f_h^* + \frac{2}{\alpha_{h,t}} (\|x_{h,t-1} - x_h^*\|^2 - \|x_{h,t} - x_h^*\|^2),
\end{aligned}
$$

where (i) is due to L-smoothness of f_h, (ii) is due to $\alpha_{h,t} \leq \frac{1}{L}$, and (iii) is due to convexity of f_h. Rearranging terms and summing over t,

$$
\begin{aligned}
\sum_{t=1}^{T} \frac{\alpha_{h,t}}{2} (f_{h,t} - f_h^*) &\leq \sum_{t=1}^{T} \|x_{h,t-1} - x_h^*\|^2 - \|x_{h,t} - x_h^*\|^2 \\
&\leq \|x_{h,0} - x_h^*\|^2. \qquad (16)
\end{aligned}
$$

Since $\{f_{h,t}\}_t$ are decreasing, (16) implies that

$$
f_{h,T} - f_h^* \leq \frac{2\|x_{h,0} - x_h^*\|^2}{\sum_{t=1}^{T} \alpha_{h,t}}.
$$

Noting that $f_{h,T} - f_h^* = q_h^* - q_{h,T}$ completes the proof. □

Proof of Corollary 1. Due to Theorem 1, showing that $|\rho^* - \rho_T| \leq (4 - 5\rho^*) \log \left(\frac{q_h^*}{q_h^* - \epsilon} \right)$ if $T \gtrsim \frac{L\|x_{h,0} - x_h^*\|^2}{\epsilon}$ completes the proof. We handle it in the same cases as in the proof of Proposition 5.

Case 1: $\rho_0 \geq \rho^*$. From lines 9-11 of Algorithm 1, the low-quality firm will not update its model until after round T, where $\rho_T = \rho^*$. With only the high-quality firm updating before this point, the firms' qualities will have reached a ratio ρ^* by T steps if $\frac{q_{l,0}}{q_{h,T}} = \rho^*$. Dividing both sides of this equation by $q_{h,0}$ and rearranging terms, $\frac{q_{h,T}}{q_{h,0}} = \frac{\rho_0}{\rho^*}$. As we showed for this case in the proof of Proposition 5, $\frac{q_{h,t}}{q_{h,t-1}} \leq \tilde{b}$. Therefore,

$$
\frac{q_{h,T}}{q_{h,0}} = \frac{\rho_0}{\rho^*} \leq \tilde{b}^T,
$$

which gives $T \geq \frac{\log(\rho_0/\rho^*)}{\log(\tilde{b})}$. That is, after $\frac{\log(\rho_0/\rho^*)}{\log(\tilde{b})}$ steps, $\rho_T = \rho^*$. As discussed in the proof of Proposition 5, the firms can maintain a quality ration of ρ^* for all future rounds, making $|\rho^* - \rho_T| = 0$.

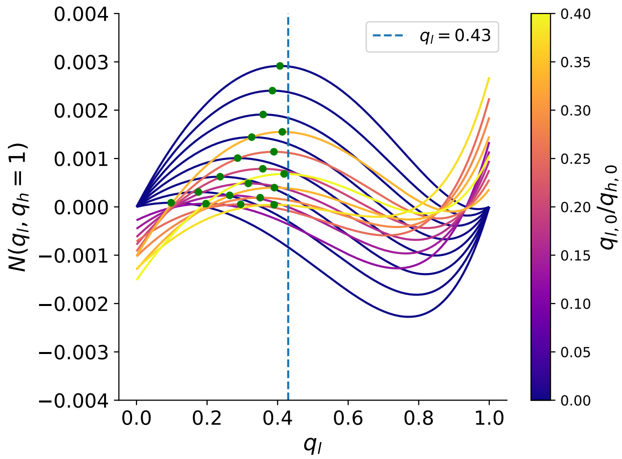


Figure 8: For a range of initial qualities and $q_h = q_h^* = 1$, the green dots mark the Nash bargaining solution. The x-values of these points are smaller than 0.43.

Case 2: $\rho_0 < \rho^*$. As the proof of this case in Proposition 5 directly shows, $\rho^* - \rho_T \le (4 - 5\rho^*) \log\left(\frac{q_h^*}{q_h^* - \epsilon}\right)$ if $T \ge \frac{\log(q_h^*/q_{h,0})}{\log((\tilde{b}+1)/2)} + \frac{L\|x_{h,0} - x_h^*\|^2}{2\epsilon}$.

Combining Cases 1 and 2, if $T \ge \max\left\{\frac{\log(\rho_0/\rho^*)}{\log(\tilde{b})}, \frac{\log(q_h^*/q_{h,0})}{\log((\tilde{b}+1)/2)} + \frac{L\|x_{h,0} - x_h^*\|^2}{2\epsilon}\right\}$, then $|\rho^* - \rho_T| \le (4 - 5\rho^*) \log\left(\frac{q_h^*}{q_h^* - \epsilon}\right)$, which completes the proof. □

The following lemma gives.

Lemma 5. *For all ρ_0 s.t. $\rho_0 \le \rho^*$, $\rho^* \le 0.43$.*

Proof of Lemma 5. The Nash bargaining objective evaluated at $q_h^* = 1$ is

$$N(q_l, q_h^*) = \left(\frac{q_l(1-q_l)}{(4-q_l)^2} - U_{l,0}\right)\left(\frac{4(1-q_l)}{(4-q_l)^2} - U_{h,0}\right), \tag{17}$$

where $U_{h,0} \stackrel{\text{def}}{=} U_h(q_{l,0}, q_{h,0})$ and $U_{l,0} \stackrel{\text{def}}{=} U_l(q_{l,0}, q_{h,0})$. Differentiating (17) with respect to q_l,

$$\frac{\partial N(q_l, q_h^*)}{\partial q_l} \tag{18}$$
$$= \frac{(7U_{h,0} + U_{h,0}\rho_0 + 4)q_l^3 + (-60U_{h,0} - 6U_{h,0}\rho_0 + 32)q_l^2 + (144U_{h,0} - 52)q_l + (-64U_{h,0} + 32U_{h,0}\rho_0 + 16)}{(4-q_l)^5}.$$

The roots of (18) correspond to the roots of the cubic numerator. It can be verified with graphing software that over all starting points $(q_{l,0}, q_{h,0})$ such that $\rho_0 \le \rho^*$, the roots q_l^* of this cubic are at most 0.43. (See Figure 8 for empirical evidence.) □

B Extension of Results and Proofs to n-firm Setting

We assume the N firms have an initial ranking of model qualities: $q_1 > \ldots > q_N$.

Definition 3 (Consumer Utility)**.** *A type-θ consumer has utility*

$$U_c(\theta) = \begin{cases} \theta q_n - p_n & \text{if it buys } n\text{'th-quality firm's model for } n \in [N], \\ 0 & \text{if it buys no model.} \end{cases} \tag{19}$$

Lemma 6 (Consumer Demands)**.** *Given the utilities in Definition 1,*

1. *consumer demand for the highest-quality firm is* $D_1 = 1 - \frac{p_1 - p_2}{q_1 - q_2}$
2. *consumer demand for firms* $n \in \{2, ..., N\}$ *is* $D_n = \frac{p_{n-1} - p_n}{q_{n-1} - q_n} - \frac{p_n}{q_n}$.

Lemma 7 (Equilibrium Prices and Utilities)**.** *The optimal prices for the firms are*

$$p_1^* = \frac{2q_1(q_1 - q_2)}{4q_1 - q_2}$$

for the highest-quality firm, and

$$p_n^* = \frac{q_n(q_{n-1} - q_n)}{4q_{n-1} - q_n}$$

for firms $n \in \{2, ..., N\}$*. These prices yield price-optimal utilities*

$$U_1(q_2, q_1, p_2^*, p_1^*) = \frac{q_1 q_2(q_2 - q_1)}{(4q_2 - q_1)^2} \tag{20}$$

and

$$U_n(q_n, q_{n-1}, p_n^*, p_{n-1}^*) = \frac{4q_n^2(q_n - q_{n-1})}{(4q_n - q_{n-1})^2}$$

for $n \in \{2, ..., N\}$.

Proposition B.1. 1. U_n *is increasing in* $q_n \forall n < N$,

2. U_n *is decreasing in* $q_{n-1} \forall n < N$,
3. U_N *is increasing in* $q_N - 1$*, and*
4. U_N *is increasing in* q_N *for* $q_N \leq \frac{4}{7} q_{N-1}$ *and decreasing in* q_N *otherwise.*

Definition 4. *(*N*-agent Nash bargaining objective)*

$$\begin{aligned}(q_1^*, ..., q_N^*) = \arg\max_{q \in [0,1]^N} \quad & \tilde{N}(q_2, q_1, q_{2,0}, q_{1,0})(\Pi_{n \in \{2,...,N\}} \tilde{N}(q_n, q_{n-1}, q_{n,0}, q_{n-1,0})) \\ \text{s.t.} \quad & U_1(q_2, q_1) \geq U_1(q_{2,0}, q_{1,0}) \\ & U_n(q_n, q_{n-1}) \geq U_n(q_{n,0}, q_{n-,0}), \ n \in \{2, ..., N\}\end{aligned}$$

where

$$\tilde{N}(q_n, q_{n-1}, q_{n,0}, q_{n-1,0}) \stackrel{def}{=} U_n(q_n, q_{n-1}) - U_n(q_{n,0}, q_{n-1,0}).$$

Proposition B.2 (Equivalence between maximal quality and the Nash bargaining solution)**.**

$$q_1^* = \max_{x \in \mathcal{X}} q(x).$$

Proposition B.3 (Non-decreasing revenues)**.** *There exist learning rate schedules* $\{\alpha_{n,t}\}_t$ *for* $n \in [N]$ *such that at no step of Algorithm 1 does any firm's revenue decrease.*

Proof. At round t, the highest quality firm can improve by any amount $q_{1,t-1} \to q_{1,t}$ without decreasing any other firm's utility. By the proof of the 2-firm case, firm 2 can then improve $q_{2,t-1} \to \hat{q}_{2,t}$ without decreasing any firm's utility. Following this logic then, firm n can improve $q_{n,t-1} \to \hat{q}_{n,t}$ without decreasing any firm's utility. As in the 2-firm proof, $\hat{q}_{n,t}$ is based on 3 quantities: $q_{n-1,t}$, $q_{n-1,t-1}$, and $\rho_{n,t-1} = \frac{q_{n,t-1}}{q_{n-1,t-1}}$. Given the sequential ordering of improvements (firm 1 improves, determining $\hat{q}_2$, then firm 2 improves based on determining $\hat{q}_2$, ..., then firm n,...) in Algorithm 2, $\hat{q}_{n,t}$ can be computed for each firm to determine their improvement threshold.

As in the 2-firm proof, firm 1 can set any learning rate $\alpha_{1,t} \leq \frac{1}{L}$. Then in order to not exceed their respective thresholds $\hat{q}_{n,t}$ firms $n \in \{2, ..., N\}$ must not exceed learning rates of $\alpha_{n,t} = \min\left\{\frac{\hat{q}_{n,t} - q_{n,t-1}}{\|\nabla_{x_{n,t-1}} f_{n,t-1}\|^2}, 1\right\}$. □

Proposition B.4 (Convergence to the Nash bargaining solution)**.** *If* $q_{n,0} \leq q_n^*$ *for all* $n \in \{2, ..., N\}$*, then there exist learning rate schedules* $\{\alpha_{n,t}\}_{t=1}^T$ *for all* $n \in [N]$ *such that after* T *rounds Algorithm 2 converges to* $(q_1^*, ..., q_N^*)$.

Proof. From the 2-firm proof, the highest-quality firm must adhere to a learning rate schedule $\alpha_{h,t} = \min\left\{\frac{(\tilde{b}-1)q_{1,t-1}}{\|\nabla_{x_{1,t-1}} f_{1,t-1}\|^2}, \frac{1}{L}\right\}$, and doing so, will converge to q_1^* in $T = \frac{\log(q_h^*/q_{h,0})}{\log((\tilde{b}+1)/2)} + \frac{L\|x_{h,0}-x_h^*\|^2}{2\epsilon}$ steps (within ϵ error). In order to not exceed $\hat{q}_{2,t}$ and violate the no-revenue-loss requirement, the second-highest-quality firm must adhere to $\alpha_{2,t} = \min\left\{\frac{\hat{q}_{2,t}-q_{2,t-1}}{\|\nabla_{x_{2,t-1}} f(x_{2,t-1})\|^2}, 1\right\}$. □

Proposition B.5 (Convergence to the Nash bargaining solution)**.** *If $q_{n,0} \le q_n^*$ for all $n \in \{2, ..., N\}$, then there exist learning rate schedules $\{\alpha_{n,t}\}_{t=1}^T$ for all n such that after T rounds Algorithm 1 converges to $(q_1^*, ..., q_N^*)$.*

Proof. We look at an arbitrary firm n and handle it cases as in the 2-firm proof.

Case 1: $q_{n,0} \le q_n^*$ and $\frac{q_{n,0}}{q_{n-1,0}} \ge \frac{q_n^*}{q_{n-1}^*}$.

The proof is identical to the 2-firm proof. Firm n should not update until $\frac{q_{n,t-1}}{q_{n-1,t}} = \frac{q_n^*}{q_{n-1}^*}$. At this point, for any learn rate schedule that firm $n-1$ maintains going forward, firm n can maintain a learning rate schedule such that $\frac{q_{n,T}}{q_{n-1,T}} = \frac{q_n^*}{q_{n-1}^*}$.

Case 2: $q_{n,0} \le q_n^*$ and $\frac{q_{n,0}}{q_{n-1,0}} < \frac{q_n^*}{q_{n-1}^*}$

We showed in 2-firm proof that there is a learning rate schedule $\{\alpha_{1,t}\}_t$ such that firms 1 and 2 converge to (q_1^*, q_2^*) in T rounds. Now we just have to ensure that the rate at which firm 2 converges to q_2^* makes it possible for firm 3 to converge to q_3^* without violating the no-revenue-loss constraint. Then extending this logic to the remaining firms completes the proof.

In the 2-firm proof, we showed that as long as, at every step $t \in [T]$, $\frac{\tilde{b}+1}{2} \le \frac{q_{1,t}}{q_{1,t-1}} \le \tilde{b}$ (where $\tilde{b} \approx 1.03$), then firm 2 will converge to q_2^* when firm 1 converges to q_1^* after T steps, simply by never exceeding $\hat{q}_{2,t}$. Therefore, we have to ensure that, at step t given firm 1's current quality $q_{1,t}$, firm 2 can improve $q_{2,t-1} \to q_{2,t}$ such that $\frac{\tilde{b}+1}{2} \le \frac{q_{2,t}}{q_{2,t-1}} \le \tilde{b}$. This in turn will ensure that firm 3 converges to q_3^* in T steps.

Note from earlier results in the paper that

$$\hat{q}_{2,t} = B\left(\frac{q_{2,t-1}}{q_{1,t-1}}, \frac{q_{1,t}}{q_{1,t-1}}\right)q_{1,t} \ge q_{2,t-1}\left(\frac{q_{1,t}}{q_{1,t-1}}\right) \ge q_{2,t-1}\left(\frac{\tilde{b}+1}{2}\right).$$

Therefore firm 2 should improve to $q_{2,t} = \min(\tilde{b}q_{2,t-1}, \hat{q}_{2,t})$. This ensures that $\frac{\tilde{b}+1}{2} \le \frac{q_{2,t}}{q_{2,t-1}} \le \tilde{b}$, which, by the same logic for firms 1 and 2, ensures that firm 3 converges to q_3^* in T steps by simply never exceeding $\hat{q}_{3,t}$ at every round. □

Different Consumer Distributions. For $\theta \sim U[0, \theta_{\max}]$, $p_l^* \to \theta_{\max} p_l^*$, $p_h^* \to \theta_{\max} p_h^*$, $U_l^* \to \theta_{\max} U_l^*$, and $U_h^* \to \theta_{\max} U_h^*$. With these changes, all other results in the paper carry through. For other distributions, it depends on the form of the pdf of θ. Let $p(\theta)$ be the pdf of θ. Then $D_l(p_l, p_h, q_l, q_h) = \int_{\hat{\theta}_h}^{\theta_{\max}} p(\theta)d\theta$, where $\theta_{\max}$ is the largest value that θ can take on, and $D_h(p_l, p_h, q_l, q_h) = \int_{\hat{\theta}_l}^{\hat{\theta}_h} p(\theta)d\theta$. These demands affect the optimal price and utilities, but we cannot calculate them unless we know $p(\theta)$.

C NeurIPS Main conference Reviews

C.1 Decision: Reject

The paper takes a theoretical modeling approach to study competition in a collaborative learning system. The paper establishes several theoretical insights; for example, full collaboration might lead to market collapse while one-sided collaboration coming from the lower-quality firm can improve revenue overall. The paper also proposes a more equitable, defection-free scheme in which both firms share but lose no revenue.

Overall, the paper studies an interesting theoretical problem, proposes an economic model of two firms, and provides a solid theoretical analysis. The review team found the above insights to be novel and interesting, although their validity might be limited by (i) the weak experimental evaluation, (ii) the stylized model and knowledge of model parameters, and (iii) the assumption of trust between firms. There is also some related literature on algorithmic monoculture (e.g., Kleinberg & Raghavan, PNAS 2021); it would be important for the paper to add a discussion on how these works compare to the present model and insights. Finally, reviewers had also raised concerns about the focus on a two-firm model; however, the authors have successfully addressed this by extending their results to N firms.

C.2 Review by Reviewer L4cP

Summary: This paper suggests a novel defection-free collaboration workflow. The suggested scheme considers two firms, with one (Firm h) having a better performing (ML) model than the other Firm (Firm l). Here, Firm h performs better, thereby "higher quality," because its dataset is more similarly distributed to the target dataset than Firm l, with data_h $\cup$ data_l $\sim$ data_target.

The considered setup is akin to the federated learning scheme, with zero training data transmission between the two firms (models), but only the evaluated outcomes, i.e., training loss or its gradient, can be shared. The caveat here is that in order to examine Model A's loss on Firm B's dataset, Firm B should be able to have full access to Firm A's model parameters. The paper gets away from this red flag by potentially introducing a "trusted central coordinator."

One of the key findings is Proposition 1, which suggests that the utilities of both Firms h and l increase as the quality of Model h increases, but the utility of Firm l only conditionally increases with respect to the quality of Model l. This leads to Algorithm 1, defection-free collaboration learning, which guarantees the increase of both firms at all times. The key functionality is to delicately tune the quality improvements of Firm l with respect to that of Firm h.

The work is tested on the MNIST dataset with LeNet-5 model structures, with each firm having 1,000 training samples but with different distributions.

Scores:

- **Soundness:** 3: good
- **Presentation:** 2: fair
- **Contribution:** 3: good

Strengths:

- The proposed work sets up a very interesting connection between operations management in economics and federated learning in machine learning. Simply put, the work tells us that naively allowing the competing firm (agent) to evaluate its model performance on my dataset can be detrimental, especially when the competing firm is already on higher ground.

Weaknesses: The paper is difficult to follow, especially for the common audiences in the ML community. It's not about all the theories from the economics, e.g., Nash bargaining and so on, but more about the notations. Section 2.1 (especially 2.1.1) needs to have more explanations. Also, the experimental setup significantly lacks details.

Rating: 6 (Weak Accept: Technically solid, moderate-to-high impact paper, with no major concerns with respect to evaluation, resources, reproducibility, ethical considerations.)

Confidence: 2 (You are willing to defend your assessment, but it is quite likely that you did not understand the central parts of the submission or that you are unfamiliar with some pieces of related work. Math/other details were not carefully checked.)

Author Rebuttal: We thank the reviewer for their detailed feedback and positive evaluation. We address each of the concerns raised:

Section 2.1 (especially 2.1.1) needs to have more explanations.

Thank you for bringing this to our notice. We have modified the notation in Section 2.1.1 (particularly bullet point 2) in our paper to hopefully make it more readable, and have expanded the explanation.

Why the same number of data points for Firms h and l?

This is for simplicity of setup - our conclusions are robust to the number of data points each firm holds. The main concerns/requirements of our experiments are that 1) firm h have a higher initial quality than firm l, and 2) the firms share data with each other in a way that decreases neither firm's utility over the course of the algorithm.

C.3 Review by Reviewer ucZC

Summary: The paper studies the dynamics of collaborative learning where participant incentives can lead to defection if not aligned with revenue goals. It uses a duopoly model where (two) firms collaborate to train a global model while maintaining or improving their revenue. Various collaboration schemes are evaluated, leading to the proposal of a defection-free algorithm that ensures both firms benefit without revenue loss, aiming for a Nash bargaining solution.

Scores:

- **Soundness:** 2: fair
- **Presentation:** 3: good
- **Contribution:** 2: fair

Strengths:

- The paper studies collaborative learning as a competitive market scenario, aligning with economic theory to ensure participation incentives. It shows that their model qualities improve maximally when both firms contribute fully to the collaboration.
- The paper introduces a defection-free algorithm that prevents revenue loss for participants, promoting sustained collaboration.
- The paper shows convergence to a solution that maximizes joint surplus, and their proposed algorithm converges to the Nash equilibrium, except in some trivial cases.

Weaknesses:

- The paper relies on simplified assumptions such as convex and smooth loss functions, which may not generalize to all real-world scenarios. There might be some data-privacy considerations as well.
- While extending results to an oligopoly is mentioned, the primary focus remains on a two-firm scenario.
- The paper emphasizes revenue preservation over model quality improvement, which might have a potential impact on accuracy for economy stability.

Rating: 3 (Reject: For instance, a paper with technical flaws, weak evaluation, inadequate reproducibility and/or incompletely addressed ethical considerations.)

Confidence: 2 (You are willing to defend your assessment, but it is quite likely that you did not fully understand central parts of the submission.)

Author Rebuttal: We thank the reviewer for their comments and feedback. We address the concerns raised below:

The paper relies on simplified assumptions such as convex and smooth loss functions, which may not generalize to all real-world scenarios.

Our analysis assumes smooth convex functions because this helps precisely control model-quality improvement during training, which is necessary to guarantee the no-revenue-decrease property of our algorithm. Current optimization theory reflects the practical performance on deep learning very poorly. Incorporating formal privacy guarantees (such as differential privacy) would also be excellent future directions.

The primary focus remains on a two-firm scenario.

All of our results and proofs carry through to the N-firm setting. We have added an appendix to the paper which states the algorithm for N firms, and restates and proves each result for this setting.

C.4 Review by Reviewer CKmX

Summary: This paper studies collaboration between owners of high- and low-quality model owners in a competitive setup using game theoretic tools. First, they showed complete collaboration leads to zero revenue. They then designed a defection-free algorithm that can provably converge to a Nash bargain solution in a multi-round regime. The analyses offer new insights to the field of economics and collaborative learning.

Scores:

- **Soundness:** 3: good
- **Presentation:** 3: good
- **Contribution:** 3: good

Strengths:

- The paper is well-written and the demonstration is clear.
- The problem setup is novel, and the authors modeled the relationship between utility and model quality through an economic lens. The analyses are neat and nice.

Weaknesses: I am not convinced by Line 229-230. I do not think q_l^* and ρ^* are reasonable to be assumed known in practice. There is a typo in Proposition 1. The 2nd item should be U_h is decreasing in q_l. Typo in Line 176, "have lower revenue that" should be "have lower revenue than".

Regarding the experimental setup, the distinction between low- and high-quality firms is based solely on the number of training epochs. With this approach, both firms could conduct local training and achieve models of the same quality (I would be curious to see what the revenues would be with local learning). I believe a more reasonable way to differentiate between low- and high-quality firms would be to base it on their target performance when they conduct local training until convergence.

Rating: 6 (Weak Accept: Technically solid, moderate-to-high impact paper, with no major concerns with respect to evaluation, resources, reproducibility, ethical considerations.)

Confidence: 4 (You are confident in your assessment, but not absolutely certain. It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work.)

Author Rebuttal: We thank the reviewer for their close reading of our work, the detailed feedback, and the positive evaluation. We address each of the concerns raised:

Regarding the experimental setup, the distinction between low- and high-quality firms is based solely on the number of training epochs.

This is an excellent point. We can achieve a differentiation between the quality of two firms setup in a variety of ways in practice: e.g. a) make firm h's data distribution closer to that of the target test distribution, b) make firm h's dataset larger than firm l's, or c) ensure firm h has a better initialization point or runs for longer training epochs than firm l, etc.

C.5 Review by Reviewer Bkmn

Summary: The paper investigates collaborative learning systems involving competitive participants who may defect if collaboration leads to revenue loss. The authors model the system as a duopoly where two firms train machine learning models and sell predictions to a market of consumers. The study explores various collaboration schemes, demonstrating that full collaboration leads to market collapse, while one-sided collaboration can improve both firms' revenues. The authors propose a defection-free algorithm where both firms share information without losing revenue, showing that it converges to the Nash bargaining solution.

Scores:

- **Soundness:** 3: good
- **Presentation:** 3: good
- **Contribution:** 3: good

Strengths:

- Relevance and Novelty: The paper addresses a significant and timely issue in collaborative learning, particularly in competitive environments. The proposed defection-free scheme is novel and provides valuable insights into ensuring sustained collaboration.
- Theoretical Foundation: The framework is grounded in economic theory, particularly the Nash bargaining solution, providing a robust theoretical basis for the proposed scheme.

Weaknesses: The primary issue with the paper is the potential lack of generalizability of the proposed model. The study focuses on a duopoly, and it remains unclear how the conclusions might change with more than two competitors.

Rating: 5 (Borderline Accept)

Confidence: 5 (Absolutely certain of the assessment)

Author Rebuttal: We thank the reviewer for their feedback and for the positive evaluation of our work. We address the questions and main concerns below:

How does the proposed defection-free algorithm scale with an increasing number of competitors?

All of our results and proofs carry through to the N-firm setting. We have added an appendix to the paper which states the algorithm for N firms, and restates and proves each result for this setting.