Strategic Data Sharing between Competitors

Nikita Tsoy [1] **Nikola Konstantinov** [1]

Abstract

Collaborative learning techniques have significantly advanced in recent years, enabling private model training across multiple organizations. Despite this opportunity, firms face a dilemma when considering data sharing with competitors—while collaboration can improve a company's machine learning model, it may also benefit competitors and hence reduce profits. In this work, we introduce a general framework for analyzing this data-sharing trade-off. The framework consists of three components, representing the firms' production decisions, the effect of additional data on model quality, and the data-sharing negotiation process, respectively. We then study an instantiation of the framework, based on a conventional market model from economic theory, to identify key factors that affect collaboration incentives. Our findings indicate a profound impact of market conditions on the data-sharing incentives. In particular, we find that reduced competition, in terms of the similarities between the firms' products, and harder learning tasks foster collaboration.

1. Introduction

Machine learning has become integral to numerous business functions, such as operations optimization and new products creation (Chui et al., 2022). Despite its power, its efficacy hinges significantly on the quality and quantity of the training data, making data a key asset for firms.

One way to enhance data access and machine learning models is collaboration via data sharing (Rieke et al., 2020; Durrant et al., 2022). However, at least two barriers exist to such collaborations. The first is privacy, which can be addressed by new collaborative learning techniques, such as federated learning (Kairouz et al., 2021). The second is incentives: if the entities have no collaboration incentives, they may not collaborate at all, free-ride (Blum et al., 2021), or attack the shared model (Blanchard et al., 2017).

Such conflicting incentives arise naturally between market competitors. While the competitors are an appealing data source since they operate on the same market, the collaboration could strengthen their models and hence intensify competition. This concern is especially important for big firms that can influence prices and act strategically. Strategic actions might increase profits in various ways and produce complex downstream effects. For example, firms might collude to capture more revenue[1] or engage in a price war to win a market.[2] Thus, data sharing between big firms can have complicated downstream effects.

Our contributions Despite the significance of the data-sharing trade-off, particularly for large companies, there are few works on the effects of competition on collaborative learning incentives. We aim to fill this gap by proposing a framework to investigate these effects. The framework is modular and consists of three parts: market model, data impact model, and collaboration scheme. These components represent the firms' production decisions, the impact of additional data on model quality, and data-sharing negotiation process, respectively.

We investigate the key factors of the trade-off, using a conventional market model from economic theory and a data impact model grounded in learning theory. We theoretically demonstrate the impact of market on data-sharing collaboration in the case of binary data sharing between two firms. In particular, we show that collaboration becomes more appealing as market competition, measured by the similarities of the firms' products, or learning task simplicity decreases.

2. Related work

Data sharing incentives in machine learning Collaboration incentives constitute an important research topic, especially in the context of federated learning (Kairouz et al., 2021). Notable lines of work concern incentives under data

[1]INSAIT, Sofia University, Sofia, Bulgaria. Correspondence to: Nikita Tsoy <nikita.tsoy@insait.ai>.

Federated Learning and Analytics in Practice: Algorithms, Systems, Applications, and Opportunities Workshop at 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA.

[1] https://www.nytimes.com/2021/01/17/technology/google-facebook-ad-deal-antitrust.html

[2] https://www.reuters.com/article/us-uber-results-breakingviews-idUSKCN1UY2X5

heterogeneity (Donahue & Kleinberg, 2021; Werner et al., 2022); compensation for costs related to training (Yu et al., 2020; Tu et al., 2022; Liu et al., 2022); fair distribution of collaboration gains (Lyu et al., 2020a;b; Blum et al., 2021); and free-riding (Richardson et al., 2020; Karimireddy et al., 2022). We refer to Zeng et al. (2021); Yang et al. (2020); Zhan et al. (2021) for a detailed overview.

To our awareness, only Wu & Yu (2022) consider the effect of competition on collaboration. However, the authors do not explain the mechanisms behind the effects of machine learning on the market. Thus, their framework can not predict coalition benefits but only indicate coalition benefits post factum. In addition, they use the marketing model of Rust & Zahorik (1993) instead of the classic economic models (Tirole, 1988) to model competition, potentially constraining applicability. Finally, their notion of the sustainability of coalitions does not arise from standard concepts in game theory. In contrast, we base our analysis on the standard Nash equilibrium concept (Nash, 1951).

3. Data-sharing problem

This section describes a general data-sharing problem, consisting of three components: market model, data impact model, and collaboration scheme. These componets outline consumers' and firms' consumption and production decisions, the impact of additional data on machine learning model, and data-sharing negotiation process among the competitors. Given a specific application, these components can be modeled by market research or sales teams, operation management or data scientists, and mediators, respectively. We begin with an example that will help us clarify the abstract concepts and then introduce the framework.

3.1. Running example

Consider a city with a taxi market dominated by a few firms. Each firm collects data (e.g., demand for taxis and traffic situation) to train a machine learning model for reducing costs or enhancing services independently or collaboratively with its competitors. Collaboration can improve the company's machine learning model but may also strengthen the competitors' models. Thus, the company must carefully evaluate the impact of data sharing on its profits.

3.2. Market model

The market model describes consumer actions (demand factors) and firm production actions (supply factors). We consider a market with m firms, $F_1, \dots, F_m$, each producing $q_i \in \mathbb{R}_+$ units of good G_i and offering them at price $p_i \in \mathbb{R}_+$, where $\mathbb{R}_+ = [0, \infty)$. In our example, $G_1, \dots, G_m$ represent taxi services from different companies, with consumers being city travelers, q_i are kilometers serviced by F_i, and p_i are the prices per kilometer.

In this setting, prices and quantities are $2m$ unknown variables. Hence we need $2m$ constraints to describe the consumers' and firms' market decisions, which we derive from the consumers' and the firms' rationality. In contrast to standard market models, our framework allows product utilities and costs to depend on the machine learning models of the firms $\boldsymbol{v} = (v_1, \dots, v_m)$.

Consumers' behavior Each consumer j optimizes their utility $u^j(\boldsymbol{g}^j, \boldsymbol{q}^j, \boldsymbol{v})$ by buying goods given the market prices. Their utility depends on three factors. The first is the consumed quantities $\boldsymbol{q}^j$ of products $G_1, \dots, G_m$. The second is the models $\boldsymbol{v}$, as these may impact the corresponding products' utilities. The last is consumed quantities $\boldsymbol{g}^j$ of goods outside the considered market (e.g., consumed food in our taxi example).

Assuming that each consumer j can only spend budget B^j and that consumers can not influence prices since there are many of them and they do not cooperate, we get the following consumption problem

$$\max_{\boldsymbol{g}^j, \boldsymbol{q}^j} u^j(\boldsymbol{g}^j, \boldsymbol{q}^j, \boldsymbol{v}) \text{ s.t. } \sum_{l=1}^{k} \tilde{p}_l g_l^j + \sum_{i=1}^{m} p_i q_i^j \leq B^j, \quad (1)$$

where $\tilde{\boldsymbol{p}}$ are the outside products' prices (which we consider fixed). The solution $q_i^{j,*}(\boldsymbol{p}, \boldsymbol{v})$ determines the aggregate demanded quantity of goods

$$q_i(\boldsymbol{p}, \boldsymbol{v}) \coloneqq \sum_j q_i^{j,*}(\boldsymbol{p}, \boldsymbol{v}). \quad (2)$$

The functions $q_i(\boldsymbol{p}, \boldsymbol{v})$ (demand equations) link $\boldsymbol{p}$ and $\boldsymbol{q}$ and constitute the first m restrictions in our setting.

Firms' behavior Firms maximize their expected profits, the difference between revenue and cost,

$$\Pi_i^e = \mathbf{E}_{\boldsymbol{v}}(p_i q_i - C_i(q_i, v_i)). \quad (3)$$

Here $C_i(q_i, v_i)$ is the cost of producing q_i units of G_i, and the expectation is taken over the randomness in the models' quality. Since the quality is often observed only after testing in production, we assume that the firms optimize expected profits when making production decisions. In our running example, C_i depends on driver wages, gasoline prices, and the scheduling quality of the machine learning model.

The firms may act by either deciding on their produced quantities (Cournot, 1838) or on their prices (Bertrand, 1883). As demand equations (2) may interrelate prices and quantities for various products, firms strategically consider their competitors' actions, resulting in a Nash equilibrium. The equilibrium conditions provide another m constraints, enabling us to solve the market model entirely.

3.3. The impact of data on the market

Each company F_i has a dataset D_i (e.g., trip data in our example) and may reciprocally share it with others, concluding in the final dataset D_i^c used to train its model. We *postulate two ways a model v_i can benefit the company*. The first is *reducing production costs* $C_i(q_i, v_i)$, for example, by minimizing drivers' time in traffic jams. The second is *increasing products quality* in $u^j(\boldsymbol{g}^j, \boldsymbol{q}^j, \boldsymbol{v})$, for example, by minimizing consumers' waiting time for taxi arrival.

3.4. Collaboration scheme

Following classic economic logic, we posit that *firms will share data if it increases their expected profits Π_i^e*. Since the firms can not evaluate the gains from unknown data, we assume that they know about each other dataset characteristics (e.g., size and distributional information). Although profit maximization determines individual data-sharing incentives, forming a coalition necessitates mutual agreement. Thus, the data-sharing problem crucially depends on the negotiation process details, such as the number of participants and full or partial data sharing.

4. Example market and data impact models

In this section, we instantiate the general framework using a conventional market model from economic theory and a natural data impact model justified by learning theory. These models allow us to reason quantitatively about the data-sharing problem, leading to the identification of several key factors in the data-sharing trade-off.

4.1. Market model

We use a utility and cost model standard in the theoretical industrial organization literature (Tirole, 1988; Carlton et al., 1990). Despite its simplicity, this model effectively captures the basic factors governing market equilibrium and is often used to obtain qualitative insights.

4.1.1. Demand

We assume that, in the aggregate, consumers (1) can be described by a representative consumer with quasi-linear quadratic utility (Dixit, 1979; Choné & Linnemer, 2019)

$$\max_{g,\boldsymbol{q}} u(g, \boldsymbol{q}) := \sum_{i=1}^{m} q_i - \Big(\sum_i q_i^2 + 2\gamma \sum_{i>j} q_i q_j\Big)/2 + g$$
$$= \boldsymbol{\iota}^\mathsf{T} \boldsymbol{q} - \boldsymbol{q}^\mathsf{T} \mathbf{G} \boldsymbol{q}/2 + g \text{ s.t. } g + \boldsymbol{p}^\mathsf{T} \boldsymbol{q} \le B. \quad (4)$$

Here, $\boldsymbol{\iota} = (1, \ldots, 1)^\mathsf{T}$, $\mathbf{G} = (1-\gamma)\mathbf{I} + \gamma \boldsymbol{\iota}\boldsymbol{\iota}^\mathsf{T}$, g is the quantity of a single outside good, and $\gamma \in \left(-\frac{1}{m-1}, 1\right)$ is a measure of substitutability between each pair of goods: higher γ corresponds to more similar goods. In our running example, γ describes the difference in service of two taxi companies, such as the difference in the cars' quality or the location coverage.

In this case, the demand equations (2) are well-known (see proof in Appendix A.1).

Lemma 4.1 (Amir et al. 2017)**.** *Assume that $\mathbf{G}^{-1}(\boldsymbol{\iota} - \boldsymbol{p}) > 0$ and $\boldsymbol{p}^\mathsf{T}\mathbf{G}^{-1}(\boldsymbol{\iota} - \boldsymbol{p}) \le B$. The solution to problem (4) is*

$$p_i = 1 - q_i - \gamma \sum_{j \ne i} q_j. \quad (5)$$

4.1.2. Supply

We assume linear cost functions in Equation (3)

$$\Pi_i^e = \mathbf{E}(p_i q_i - c_i q_i). \quad (6)$$

Here c_i depends on the machine learning model and $c_i^e := \mathbf{E}_{D_i^c}(c_i)$. Now, we describe the competition between firms.

Cournot competition Each firm chooses output level q_i, which determines the prices (5) and expected profits (6). The next standard lemma describes the Nash equilibrium of this game (see proof in Appendix A.2).

Lemma 4.2. *Assume that companies maximize their profits (6) in the Cournot competition game. If $\forall i\, (2-\gamma)(1-c_i^e) > \gamma \sum_j (c_i^e - c_j^e)$, equilibrium quantities and profits satisfy*

$$q_i^* = \frac{2 - \gamma - dc_i^e + \gamma \sum_{j \ne i} c_j^e}{(2-\gamma)d}, \Pi_i^e = (q_i^*)^2,$$

where $d := 2 + \gamma(m-2)$.

Bertrand competition Each firm sets price p_i, which determines the quantities (5) and expected profits (6). The following lemma describes the Nash equilibrium of this game (see proof in Appendix A.3).

Lemma 4.3. *Assume that companies maximize their profits (6) in the Bertrand competition game. If $\forall i\, d_1(1-c_i^e) > d_3 \sum_{j \ne i}(c_i^e - c_j^e)$, equilibrium prices and profits satisfy*

$$p_i^* = \frac{d_1 + d_2 c_i^e + d_3 \sum_{j \ne i} c_j^e}{d_4},\ \Pi_i^e = d_5(p_i^* - c_i^e)^2,$$

where $d_1, \ldots, d_5$ depend only on γ and m.

4.2. Data impact model

Next, we describe how firms can reason about machine learning impact on their costs. We assume all datasets are sampled from the same distribution, which gives the expected costs in form $c_i^e = c^e(n_i^c)$, where $n_i^c = |D_i^c|$. Using the examples below, we motivate that

$$c^e(n) = a + \frac{b}{n^\beta}, \beta > 0. \quad (7)$$

Here, β indicates *learning task simplicity*: higher β corresponds to a simpler task (Tsybakov, 2004). Additionally, we assume that $a < 1$ and $\frac{b}{1-a}$ is small enough to satisfy the technical requirements of Lemmas 4.1, 4.2, and 4.3.

Asymptotic normality Consider a company that needs to perform action $\boldsymbol{s} \in \mathbb{R}^n$ during production that will impact its costs. However, the production process is random because of noise $\boldsymbol{X} \in \mathbb{R}^m$, making the cost $c(\boldsymbol{s}, \boldsymbol{X})\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ a random function. If the firm knows a structural causal model of $\boldsymbol{X}$, it can use the maximum likelihood estimator to find structural parameters and choose optimal action $\boldsymbol{s}_{\text{fin}}$ based on this estimate. In this case, final cost $\mathbf{E}_{\boldsymbol{X}}(c(\boldsymbol{s}_{\text{fin}}, \boldsymbol{X}))$ will approximately have a generalized chi-square distribution (Jones 1983; see Appendix A.4), resulting in the following expected costs

$$c^e = \mathbf{E}_{D,\boldsymbol{X}}(c(\boldsymbol{s}_{\text{fin}}, \boldsymbol{X})) \approx a + \frac{b}{n} \tag{8}$$

and implying Equation (7) with $\beta = 1$.

Stochastic optimization A similar dependence arises when the company exploits stochastic optimization to optimize expected cost $c(\boldsymbol{s}) = \mathbf{E}_{\boldsymbol{X}}(c(\boldsymbol{s}, \boldsymbol{X}))$. If the algorithm generalizes (e.g., it is single pass SGD; Bubeck et al. 2015) and $c(\boldsymbol{s})$ is strongly convex, outcome $\boldsymbol{s}_{\text{fin}}$ will satisfy

$$\mathbf{E}_D(c(\boldsymbol{s}_{\text{fin}}) - c(\boldsymbol{s}^*)) = \mathrm{O}\left(\frac{1}{n}\right),$$

where $\boldsymbol{s}^*$ is the optimal action, resulting in the same dependence. Notice that β decreases for convex problems

$$\mathbf{E}_D(c(\boldsymbol{s}_{\text{fin}}) - c(\boldsymbol{s}^*)) = \mathrm{O}\left(\frac{1}{\sqrt{n}}\right),$$

justifying its use as a measure of task simplicity.

5. Data sharing between two firms

In this section, we aim to obtain qualitative insights about the impact of market parameters on collaboration by analyzing the case of two companies making a binary decision of whether to share all their data with each other. We consider other collaboration schemes: partial data sharing between two firms and full data sharing between many firms, in Appendices E and F.

Full data sharing between two firms According to our framework, both companies compare their expected profits for two cases, when they share their data fully and when they do not. Then *they collaborate if and only if they both expect an increase in profits*.

For Cournot competition, Lemma 4.2 and Equation (7) give the following collaboration criterion

$$\forall i\, \mathit{\Pi}^e_{\text{share}} > \mathit{\Pi}^e_{i,\text{ind}} \iff 2n_i^{-\beta} - (2-\gamma)n^{-\beta} - \gamma n_{-i}^{-\beta} > 0,$$

where n_{-i} is the size of the data of the player that is not i, $n := n_1 + n_2$, $\mathit{\Pi}^e_{i,\text{share}}$ is the expected profit in collaboration, and $\mathit{\Pi}^e_{i,\text{ind}}$ is the expected profit without collaboration.

For Berntrand competition, Lemma 4.3 gives the criterion

$$\forall i\ (2-\gamma^2)n_i^{-\beta} - (2-\gamma-\gamma^2)n^{-\beta} - \gamma n_{-i}^{-\beta} > 0.$$

The theorem below describes the properties of this criterion (see proof in Appendix A.5).

Theorem 5.1. *If $\gamma \le 0$, the firms will collaborate. If $\gamma > 0$, there exists a value $x_t(\gamma, \beta)$, where $t \in \{Bertrand, Cournot\}$ is the type of competition, such that*

$$\mathit{\Pi}^e_{share} > \mathit{\Pi}^e_{i,ind} \iff \frac{n_{-i}}{n} > x_t(\gamma, \beta).$$

The function x_t has the following properties:

1. *$x_t(\gamma, \beta)$ is increasing in γ.*
2. *$x_{Bertrand} \ge x_{Cournot}$.*
3. *$x_t(\gamma, \beta)$ is increasing in β.*

Discussion The theorem indicates that the *firms are more likely to collaborate when either the market is less competitive* (properties 1 and 2) *or the learning task is harder* (property 3). Indeed, the threshold x becomes smaller when the products are less similar, making the market less competitive. Also, x becomes smaller in the Cournot case since it is known to be less competitive than the Bertrand one (Shapiro, 1989). Finally, when the learning task is harder, x decreases, making collaboration more likely. Intuitively, it happens because the decrease in cost (7) from an additional data point is higher for smaller β.

We additionally explore several extensions of this setting and the welfare implications of data sharing between two firms in Appendicies C and D.

6. Conclusion

In this work, we introduced a modular framework for studying data sharing among market competitors that consists of three parts—market model, data impact model, and collaboration scheme. We instantiated it using a conventional market model and a data impact model grounded in learning theory and studied data sharing between two firms to examine the key market parameters that impact collaboration.

Our findings indicate a significant impact of market competition on data-sharing decisions. Specifically, we found that higher product differentiation generally increases the willingness for collaboration, as does learning task complexity. We hope our study will inspire further in-depth investigations into the nuanced trade-offs in data sharing, allowing competition and collaboration to coexist in data-driven environments.

Acknowledgments

The authors would like to thank Florian Dorner, Mark Vero, Nikola Jovanovic, Dimitar Dimitrov, and Sam Motamed for providing helpful feedback on earlier versions of this work.

References

Amir, R., Erickson, P., and Jin, J. On the microeconomic foundations of linear demand for differentiated products. *Journal of Economic Theory*, 169:641–665, 2017.

Bergemann, D. and Bonatti, A. Markets for information: An introduction. *Annual Review of Economics*, 2019.

Berry, S., Gaynor, M., and Scott Morton, F. Do increasing markups matter? Lessons from empirical industrial organization. *Journal of Economic Perspectives*, 33(3):44–68, 2019.

Bertrand, J. Book review of theorie mathematique de la richesse social and of recherches sur les principes mathematiques de la theorie des richesses. *Journal des savants*, 1883.

Binmore, K., Rubinstein, A., and Wolinsky, A. The Nash bargaining solution in economic modelling. *The RAND Journal of Economics*, pp. 176–188, 1986.

Blanchard, P., El Mhamdi, E. M., Guerraoui, R., and Stainer, J. Machine learning with adversaries: Byzantine tolerant gradient descent. *Advances in Neural Information Processing Systems*, 30, 2017.

Blum, A., Haghtalab, N., Phillips, R. L., and Shao, H. One for one, or all for all: Equilibria and optimality of collaboration in federated learning. In *International Conference on Machine Learning*, pp. 1005–1014. PMLR, 2021.

Bubeck, S. et al. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 2015.

Carlton, D. W., Perloff, J. M., and van't Veld, K. T. *Modern industrial organization*. Scott, Foresman/Little, Brown Higher Education Glenview, IL, 1990.

Choné, P. and Linnemer, L. The quasilinear quadratic utility model: An overview. *CESifo Working Paper*, 2019.

Chui, M., Hall, B., Mayhew, H., and Singla, A. The state of AI in 2022—and a half decade in review. QuantumBlack By McKinsey, 2022.

Cournot, A.-A. *Recherches sur les principes mathématiques de la théorie des richesses*. chez L. Hachette, 1838.

Dixit, A. A model of duopoly suggesting a theory of entry barriers. *The Bell Journal of Economics*, pp. 20–32, 1979.

Donahue, K. and Kleinberg, J. Model-sharing games: Analyzing federated learning under voluntary participation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(6):5303–5311, 2021.

Durrant, A., Markovic, M., Matthews, D., May, D., Enright, J., and Leontidis, G. The role of cross-silo federated learning in facilitating data sharing in the agri-food sector. *Computers and Electronics in Agriculture*, 193:106648, 2022.

Fried, D. Incentives for information production and disclosure in a duopolistic environment. *The Quarterly Journal of Economics*, 99(2):367–381, 1984.

Jones, D. Statistical analysis of empirical models fitted by optimization. *Biometrika*, 70(1):67–88, 1983.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., and Cummings, R. Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, 14(1–2):1–210, 2021.

Karimireddy, S. P., Guo, W., and Jordan, M. I. Mechanisms that Incentivize Data Sharing in Federated Learning. *arXiv preprint arXiv:2207.04557*, 2022.

Kóczy, L. Á. Partition function form games. *Theory and Decision Library C*, 48:312, 2018.

Liu, J., Zhang, G., Wang, K., and Yang, K. Task-load-aware game-theoretic framework for wireless federated learning. *IEEE Communications Letters*, 2022.

Lyu, L., Xu, X., Wang, Q., and Yu, H. Collaborative fairness in federated learning. In *Federated Learning*, pp. 189–204. Springer, 2020a.

Lyu, L., Yu, J., Nandakumar, K., Li, Y., Ma, X., Jin, J., Yu, H., and Ng, K. S. Towards fair and privacy-preserving federated deep models. *IEEE Transactions on Parallel and Distributed Systems*, 2020b.

Mas-Colell, A., Whinston, M. D., and Green, J. R. *Microeconomic theory*. Oxford university press New York, 1995.

Nash, J. Non-cooperative games. *Annals of Mathematics*, 54(2), 1951.

Raith, M. A general model of information sharing in oligopoly. *Journal of economic theory*, 71(1):260–288, 1996.

Reiss, P. C. and Wolak, F. A. Structural econometric modeling: Rationales and examples from industrial organization. *Handbook of econometrics*, 6:4277–4415, 2007.

Richardson, A., Filos-Ratsikas, A., and Faltings, B. Budget-bounded incentives for federated learning. In *Federated Learning*, pp. 176–188. Springer, 2020.

Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M. N., Landman, B. A., and Maier-Hein, K. The future of digital health with federated learning. *NPJ digital medicine*, 3(1):1–7, 2020.

Rust, R. T. and Zahorik, A. J. Customer satisfaction, customer retention, and market share. *Journal of retailing*, 69(2):193–215, 1993.

Shalev-Shwartz, S. and Ben-David, S. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

Shapiro, C. Theories of oligopoly behavior. *Handbook of industrial organization*, 1:329–414, 1989.

Tirole, J. *The theory of industrial organization*. MIT press, 1988.

Tsybakov, A. B. Optimal aggregation of classifiers in statistical learning. *The Annals of Statistics*, 32(1):135–166, 2004.

Tu, X., Zhu, K., Luong, N. C., Niyato, D., Zhang, Y., and Li, J. Incentive mechanisms for federated learning: From economic and game theoretic perspective. *IEEE Transactions on Cognitive Communications and Networking*, 2022.

Vives, X. Duopoly information equilibrium: Cournot and Bertrand. *Journal of economic theory*, 34(1):71–94, 1984.

Von Neumann, J. and Morgenstern, O. Theory of games and economic behavior. In *Theory of games and economic behavior*. Princeton university press, 1944.

Werner, M., He, L., Karimireddy, S. P., Jordan, M., and Jaggi, M. Towards provably personalized federated learning via threshold-clustering of similar clients. In *Workshop on Federated Learning: Recent Advances and New Challenges (in Conjunction with NeurIPS 2022)*, 2022.

Wu, X. and Yu, H. MarS-FL: Enabling Competitors to Collaborate in Federated Learning. *IEEE Transactions on Big Data*, 2022.

Yang, Q., Fan, L., and Yu, H. *Federated Learning: Privacy and Incentive*, volume 12500. Springer Nature, 2020.

Yu, H., Liu, Z., Liu, Y., Chen, T., Cong, M., Weng, X., Niyato, D., and Yang, Q. A sustainable incentive scheme for federated learning. *IEEE Intelligent Systems*, 2020.

Zeng, R., Zeng, C., Wang, X., Li, B., and Chu, X. A comprehensive survey of incentive mechanism for federated learning. *arXiv preprint arXiv:2106.15406*, 2021.

Zhan, Y., Zhang, J., Hong, Z., Wu, L., Li, P., and Guo, S. A survey of incentive mechanism design for federated learning. *IEEE Transactions on Emerging Topics in Computing*, 2021.