Pareto-Improving Data-Sharing*

Ronen Gradwohl
Ariel University
Department of Economics and Business Administration
Ariel, Israel
roneng@ariel.ac.il

Moshe Tennenholtz
Technion – Israel Institute of Technology
Faculty of Industrial Engineering and Management
Haifa, Israel
moshet@ie.technion.ac.il

CCS CONCEPTS

• **Applied computing** → **Electronic data interchange**; • **Theory of computation** → *Algorithmic game theory and mechanism design*; • **Social and professional topics** → Information system economics.

ACM Reference Format:
Ronen Gradwohl and Moshe Tennenholtz. 2022. Pareto-Improving Data-Sharing. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), June 21–24, 2022, Seoul, Republic of Korea.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3531146.3533085

Data-driven innovation promises to transform the economic landscape and bring tremendous benefits to individuals. One of the key ingredients to such innovation consists of data-sharing between firms. The importance of such sharing is underscored by the European Commission's Strategy for Data, which describes the Commission's plan to "invest in a High Impact Project on...infrastructures, data-sharing tools, architectures and governance mechanisms for thriving data-sharing and Artificial Intelligence ecosystems." Private industry is also rapidly developing such tools, including platforms such as Google Merchant and Azure Data Share that facilitate data sharing between firms. Finally, the growing area of federated machine learning focuses on designing algorithms that enable firms to share data used by their respective predictive analytics tools [Rasouli and Jordan 2021; Yang et al. 2019].

Despite the investment in and proliferation of data-sharing tools, the full realization of a data-driven transformation must overcome some hurdles. One of the major challenges is spelled out by the European Commission as follows: "In spite of the economic potential, data sharing between companies has not taken off at sufficient scale. This is due to a lack of economic incentives (including the fear of losing a competitive edge)," [European Commission 2020].

In this paper we formally study firms' incentives for data sharing, with a focus on the possible existence of data-sharing mechanisms that both benefit individuals and incentivize firms to participate. We show that firms' hesitancy may be justified, in that indiscriminate data-sharing can be mutually harmful.[1] However, we also show that data sharing need not be a zero sum endeavor—benefiting consumers at the expense of firms, or vice versa—and that more-carefully designed mechanisms for partial data-sharing can simultaneously benefit individuals as well as firms.

We undertake our study within the context of e-commerce, in which competing firms engage in imperfect competition over a set of consumers. Firms have heterogeneous data about consumers, and sharing all their data with one another leads to a more efficient outcome and benefits consumers. However, the increased competition brought about by data sharing lowers firm profits, and so firms do not willingly participate. On the other hand, when partial data can be shared, firms can do so in a way that increases their respective profits, but lowers consumer welfare. Most interestingly, we show that a fair, middle ground exists—that partial data can be shared in a way that increases firm profits, while at the same time also increasing market efficiency and benefiting consumers.

We demonstrate our ideas in the most-studied model of imperfect competition, namely, that of Hotelling [1929]. There are two firms, each located at a different endpoint of a unit interval, with a unit mass of consumers distributed across this interval. In order to study data sharing we depart from the standard model and suppose that the firms may have data about some of the consumers. We model data about a consumer as information about that consumer's location within the interval. Thus, we suppose that there are consumers whose locations are known only to the first firm, consumers whose locations are known only to the second firm, consumers whose locations are not known to either firm, and consumers whose locations are known to both firms.

Within this model, we first show that indiscriminate sharing of data is harmful to the firms. In particular, we compare firms' baseline profits—those attained in equilibrium with no data sharing—with their profits under full data-sharing, in which each firm shares all of its location data with the other firm. Relative to the baseline of no sharing, full data-sharing increases market efficiency and consumer welfare, but lowers firm profits. We then show that if firms share data about only some of the consumers, as opposed to all, then it is possible to attain *fair* outcomes—namely, ones that are strictly Pareto-improving, weakly increasing the utilities of all market participants, with at least one strict increase—relative to the baseline. In particular, our main result is the design of a mechanism that increases firm profits as well as each consumer's welfare, and, in particular, maximizes joint firm profits subject to being fair. In addition, we also design a mechanism for which consumer welfare is maximal, subject to being fair.

*The full version of this extended abstract is available at http://tinyurl.com/ydytkrfp.
[1]This is in line with a main insight of the literature on competitive price-discrimination.

FAccT '22, June 21–24, 2022, Seoul, Republic of Korea

ACM ISBN 978-1-4503-9352-2/22/06.
https://doi.org/10.1145/3531146.3533085

ACKNOWLEDGMENTS

Ronen Gradwohl gratefully acknowledges the support of National Science Foundation award number 1718670. The work by Moshe Tennenholtz was supported by funding from the European Research

Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 740435).

REFERENCES

European Commission. 2020. A European strategy for data. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX:52020DC0066. Accessed: 2021-05-13.

Harold Hotelling. 1929. Stability in Competition. *The Economic Journal* 39, 153 (1929), 41–57.

Mohammad Rasouli and Michael I Jordan. 2021. Data Sharing Markets. *arXiv preprint arXiv:2107.08630* (2021).

Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. 2019. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)* 10, 2 (2019), 1–19.