



Paying with your personal data: the insensitivity of private information provision to asymmetric benefits

Bettina Rockenbach^{1,2} · Abdolkarim Sadrieh³ · Anne Schielke¹

Received: 3 April 2020 / Revised: 13 July 2021 / Accepted: 27 July 2021 /
Published online: 5 August 2021
© The Author(s) 2021

Abstract

Internet services are often free of charge but ask for customers' personal data in exchange for usage. We experimentally study whether the provision of information-based public goods is susceptible to restraint when contributions not only make contributors better off but also enable a non-contributing “big player” to acquire substantial profits. We show that the presence of the big player crowds out the willingness to provide neutral tokens, but no such effect is observed for the provision of private information. Hence, collecting anonymized personal data instead of monetary fees can be more profitable to service providers and create greater benefits for customers.

Keywords Information public good · Sharing information · Personal data · Asymmetric benefits · Internet service providers

JEL Classifications C91 · D03 · D64 · D82 · H41 · M31

✉ Bettina Rockenbach
bettina.rockenbach@uni-koeln.de

¹ Experimental and Behavioral Economics, Department of Economics, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany

² Max Planck Institute for Research on Collective Goods, Research Group Experimental Economics, Kurt-Schumacher-Str. 10, 53113 Bonn, Germany

³ Faculty of Economics and Management, University of Magdeburg, 4120, 39016 Magdeburg, Germany

1 Introduction

Many economists believe that “data are becoming the new raw material of business” and that information on economic actors and activities have turned into “an economic input almost on a par with capital and labor.”¹ Data-driven businesses such as Google Search and Google Maps are amongst the most profitable ones.² In these businesses, asking customers to provide personal data in exchange for the service is ubiquitous. Usually, these services collect the users’ data automatically, without users incurring any cost, once they have given consent. In general, the amount of personal information shared correlates positively to the quality of the service.

Since all customers benefit from the improved quality of service, the provision of personal data constitutes a contribution to an information public good (Rockenbach & Sadrieh, 2012). While information providers’ benefits are usually moderate in these public goods, the data collection is often highly profitable for the service providers. They frequently use the aggregated information for other extremely profitable purposes (e.g. for targeted advertisements). Our central research question is whether the presence of a service provider who makes generous profits (i.e., the presence of a substantial payoff asymmetry to the advantage of a single non-providing party) crowds out the customers’ willingness to share personal information.

To avoid additional complexities that may generate confounds, we focus on information public goods for which contributions are (almost) costless and are oftentimes collected automatically (e.g. location data collected when using traffic apps, search data collected when using search engines, data collected when using software packages, or customer data originally collected for other purposes). Hence, our setup is not meant to model information public goods that incur high costs on providers (e.g. a high degree of cognitive effort, time, or creativity) such as knowledge collection sites (e.g. Wikipedia or other open science sites), creative entertainment (e.g. Youtube and similar social media sites), or collaborative production sites (e.g. SourceForge, GitHub, or other open-source sites). The analysis of the complex set of motives for contributions in those settings goes beyond the scope of our experiment. In our design, we avoid these complexity issues and focus on the pure trade-off between inequity and the desire to provide personal information for the benefit of others.

The parallelism of our setup to the information-sharing settings in the field results from the positive network effect that characterizes these settings, i.e. customers have additional utility gains as more and more other customers actively use the service. In the case of Google Search, Google Maps, Amazon, or Netflix, for example, it seems clear that the more customers contribute their usage information, the better the services can calibrate their recommendations. Often the accuracy of the services also increases as the network of active customers grows, due to the greater diversity of products and interactions. To implement these positive network effects, we model

¹ <http://www.economist.com/node/15557443>.

² E.g. see <https://www.investopedia.com/articles/personal-finance/030415/worlds-top-10-internet-companies.asp>.

the interaction as a public goods game, in which each customer receives an additional utility gain when others contribute to the service. In the field, most of these services are provided by firms that do not contribute information to the public good but earn profits from the collected data and from the access to the large network of active customers. The “big player” in our game resembles these service providers in the field.

We experimentally study how subjects’ willingness to provide personal information to an information public good is affected by the presence of a “big player.” The big player, resembling a data-driven business, cannot provide any own information, but benefits more from the collected data than any of the contributors does. Asking simple personal questions, we compare subjects’ information provision in the treatment with a big player to the provision of information in a control treatment without a big player. Additionally, we compare the difference between the contributions with and without a big player in the information-based setting to that difference in a public good setting with neutral tokens as units of provision. We design the token-based game to have an identical monetary payoff structure as the information-based game and a very similar cost and effort structure.

We find that the provision of information is less susceptible to the payoff asymmetry caused by the big player than the provision of tokens. The presence of the big player crowds out the willingness to contribute to the token-based public good, but surprisingly we observe no crowding-out effect in the information public good. In consequence, contributions and contributors’ payoffs are significantly higher in the information-based treatment with a big player than in the corresponding token-based treatment.

Our results demonstrate that personal involvement may interact in hitherto undocumented ways with the provision of public goods. Our study provides a first contribution to this area of research by documenting that the crowding out of contributions under payoff asymmetry is diminished in an information public good.

2 Experimental design and hypotheses

Subjects play a paper-and-pencil public goods game in groups of four active contributors and one passive big player. The big player cannot contribute to the public good, but profits from the contributions. In the information-based treatments, the four contributors receive 10 information sheets containing simple personal questions. To contribute a unit the subject adds a brief written answer to the question on the sheet. Sheets containing no answer do not count as contributions to the public good.³ The instructions clearly state that the personal information that subjects

³ The questions generally have non-verifiable answers. In this, our design resembles other experimental protocols using questions with non-verifiable answers, e.g. Ackfeld and Güth (2019) or Frik and Gaudeul (2020). Our questions, however, are simple and deliberately chosen to keep the cognitive and social cost of answering as low as possible. The questions are pretested and reported in Rockenbach et al. (2020). We asked subjects to reveal their gender, their eye color, their age, their subject of study, their shoe size, their study duration, their body height, their zip code, the frequency of returning clothes to a shop while falsely claiming that they were unused, and the frequency of overdrawing the bank account.

provide will not be revealed at any time or used for any other purposes except for the experiment. While our simple personal questions have a private component, breach of privacy is not part of our design. This sets our experimental design apart from most of the studies on privacy and breach of privacy issues.⁴

In the token-based treatments, the four contributors receive 10 token sheets and have to write the word “group” on the token sheet if they wish to contribute it to the public good. Hence, both in the information-based and in the token-based setting, each contributor can provide at most 10 units to the public good, amounting to a maximal total of 40 units.⁵

Contributions in both treatments involve no monetary cost. Since the effort of contributing is very similar, the only difference between contributions in the two settings results from the psychological difference between contributing a neutral token (token-based treatments) or a personal information (information-based treatments). To ensure that the psychological effort of providing the personal information is as low as possible, we chose the cognitively least demanding personal questions that we had tested in a previous study (Rockenbach et al., 2020). The chosen questions also exhibited the lowest degree of ethical conflict. With the choice of questions and of the experimental protocol, we intended to achieve the highest possible degree of comparability between the information-based and token-based treatments.⁶

Contributions benefit all other contributors as well as the big player in both information- and token-based settings. For each unit provided by a contributor i , each of the other contributors $j \neq i$ receives a return of $r = 0.24$ while the big player receives the tripled amount $3r = 0.72$.

⁴ In an early experimental study, Poindexter et al. (2006) use variations of scenario-type questionnaires to assess the preferences for different data safety measures. Tsai et al. (2011) experimentally test whether a more transparent presentation of privacy requirements and risks leads to more informed online-shop choices. Feri et al. (2016) simulate a breach of privacy in an experiment, showing that the failure to secure personal data leads to a decrease in information provision. Gaudeul and Giannetti (2017) allow subjects to reveal their true names in a partner selection experiment. Revelation leads to vulnerability, but also to partnerships that are more profitable. Ackfeld and Güth (2019) study how hiding personal information can be a signal for non-cooperative intentions. In their two-stage design, subjects first answer a series of non-verifiable personal questions, especially on the attitude towards cooperation. Next, they can decide to show or to hide their answers from potential partners. Frik and Gaudeul (2020) also use a set of non-verifiable questions on ethically difficult issues. They then measure the implicit value of privacy under risk by eliciting the price that subjects are willing to pay to decrease the probability of a breach of privacy.

⁵ The token-based setting resembles cases, in which contributions to the public good do not consist of personal data, but are fees for a service with network effects. To simplify things, we assume that the net effect of the own contribution is zero, i.e. the fee paid for the service equals the monetary equivalent of the utility gained from the service. However, due to the positive network effect, the own contributions increase the others customers' utilities (i.e. their monetary payoffs in the experiment). Note that the token-based treatments are framed in terms of money in the experiment, to avoid any additional terminology.

⁶ The type of personal information that we ask for is most comparable to the category of questions called “preferences” in Benndorf and Normann (2018). The information in that category was provided by almost all subjects in that study, resembling high willingness to provide similar information in Rockenbach et al. (2020).

Table 1 Treatments and number of independent observations

	Baseline	Big player
Token	TOKEN_BASE $N = 52$ (52 contributors, no big player)	TOKEN_BIG $N = 44$ (44 contributors, 11 big players)
Information	INFO_BASE $N = 52$ (52 contributors, no big player)	INFO_BIG $N = 44$ (44 contributors, 11 big players)

All group members, including the big player, start with the initial endowment $e = 6$. Each contributor i 's monetary payoff function π_i^A is given by:

$$\pi_i^A = e + r \sum_{j \neq i} \text{number of units provided by contributor } j,$$

i.e., contributor i 's payoff increases in the number of units provided by any of the *other* contributors $j \neq i$. The big player's monetary payoff increases in the number of units provided by *any* of the four contributors. The big player's payoff function π^B is given by:

$$\pi^B = e + 3r \sum_{j=1}^4 \text{number of units provided by contributor } j.$$

In a 2×2 experimental design, we vary both the unit of provision (*token* versus *information*) and the existence of the big player (*baseline* versus *big player*). All other aspects of the game are the same across all treatments.

Table 1 shows the number of subjects and the number of independent observations in each of the four treatments. Note that while the big players were actually present in the laboratory and this was common knowledge to all subjects, only the contributors were active decision makers.

2.1 Hypotheses

If subjects care about joint payoffs, it is likely that full provision is selected. The presence of the big player does not make a difference under the assumption of standard preferences. Predictions are different if we assume that players are inequity-averse with regard to monetary outcomes (Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999). Since all active group members and the big player start with the same initial endowment, any unit provided by a contributor increases the payoff difference between the contributors and the big player due to the latter's triple marginal return.⁷ Since the payoff disadvantage increases more for a contributor in the big

⁷ As an anonymous reviewer has pointed out, if expecting that all other players will contribute, increasing the contribution to a public good can actually be well in line with an expected decreased inequity. However, in our game, inequity aversion will always be greater in the treatment with the big player than without, since the big player benefits from others' contributions without providing own contributions.

player treatment than in the baseline treatment, inequity-averse individuals would contribute weakly fewer units to the public good in the former than in the latter, both in the information-based and in the token-based treatments. In a token-based setup, Engel and Rockenbach (2011) show that subjects contribute less to a public good that has positive externalities on wealthy third parties than to one that does not. Hence, we expect lower average provision levels in the big player treatments than in the baseline treatments.

Hypothesis 1. *Provision levels are lower in the big player than in the baseline treatments.*

Outcome-based theories (with and without other-regarding preferences) yield the same predictions in the information-based and in the token-based treatments because the monetary payoff functions are identical in both conditions and the cost of provision is similar. Hence, we do not expect treatment differences when comparing contributions in the information-based to those in the token-based treatments.

Hypothesis 2. *Provision levels do not differ between the information and token treatments.*

2.2 Experimental procedure

We collected one-shot paper-and-pencil decisions from a total of 192 subjects (57% female). All experimental sessions were conducted at the Cologne Laboratory for Economic Research (CLER), University of Cologne, Germany. We recruited our participants using ORSEE (Greiner, 2015). Written instructions informed participants about the course of the experiment. By using identical envelopes and unmarked questionnaires for all subjects, we guaranteed anonymity. We stated this clearly in the instructions. Furthermore, the instructions contained neutral terminology concerning the player types. We called contributors “group members A” and the big player “group member B.” Sessions lasted less than one hour. Including the €2.50 show-up fee per participant, contributors earned €14.95, on average, and big players €32.65.⁸

⁸ The Supplementary Material provides appendices with detailed descriptions of the experimental protocol. Appendix A.1. contains the experimental instructions, Appendix A.2. provides the post experimental questionnaire, Appendix A.3. displays the questions used in the information-based treatments, and finally Appendix A.4. provides photographs showing the experimental material. Data are available at: <http://dx.doi.org/10.7802/1899>.

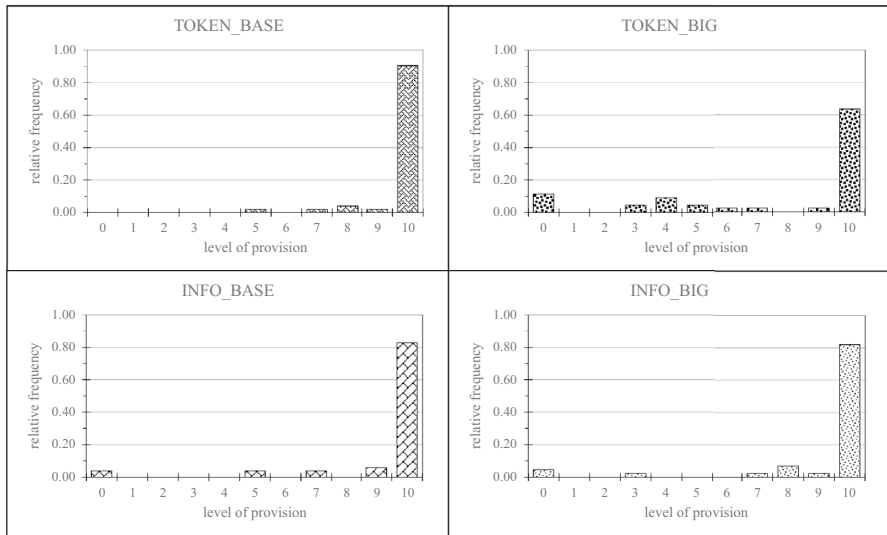


Fig. 1 Distribution of the provision levels

3 Results

Statistical inferences in this section are based on two-sided Mann–Whitney U tests. Each subject in our one-shot game represents an independent observation unit.

Figure 1 displays the distribution of provision levels by treatments. While the majority of subjects chooses full provision in all treatments, the minority refraining from full provision seems clearly larger in TOKEN_BIG than in the other three treatments. Comparing provision levels in the information-based treatments, we find no significant difference between the provision levels in INFO_BASE and INFO_BIG ($p=0.8989$).

Result 1a. *Provision levels do not differ significantly between the information-based big player and baseline treatment.*

Comparing provision levels in the two token-based treatments, we find significantly lower provision levels in TOKEN_BIG than in TOKEN_BASE ($p=0.0006$).

Result 1b. *Provision levels are significantly lower in the token-based big player than in the token-based baseline treatment.*

Taking results 1a and 1b together, we come to a mixed assessment of Hypothesis 1. While contributions to the public good are lower in the token-based treatment with a big player than without, the same does not hold for the information-based treatment comparison. Obviously, subjects' reactions to strong asymmetries in earnings depend on the type of contributions. A big player who benefits from other individuals' contributions seems to affect the willingness to contribute neutral tokens more

than the willingness to provide personal information. This holds true even though the monetary payoff structure is identical in both settings and the effort cost of contribution is similar.

Testing Hypothesis 2, we also find mixed results. While provision levels are significantly higher in INFO_BIG than in TOKEN_BIG ($p=0.0351$), they are not significantly different between INFO_BASE and TOKEN_BASE ($p=0.2322$).⁹

Result 2a. *In the presence of the big player, contributions in the token-based treatment are significantly lower than in the information-based treatment.*

Result 2b. *In the absence of the big player, contributions in the token-based and information-based treatments do not differ significantly.*

4 Discussion and conclusion

While the collection of personal data in online business models is ubiquitous, most of the behavioral research so far has focused on privacy issues. In this paper, we present a laboratory experiment that studies the aspect of payoff asymmetry in the relationship between the data collecting firms and their data providing customers.

The type of data collection that we have in mind typically takes place automatically with customers only incurring the effort cost of giving consent. The firms pool the collected data to create additional benefits for their customers. For example, maps applications use their customers' location data to enhance routing recommendations, while search engines and recommender applications use their customers' search terms and other characteristics to improve the quality of the feedback. Hence, the individuals providing personal information in such settings are contributing to an information public good that has mutual benefits, i.e. efficiency gains, for the customers (Rockenbach & Sadrieh, 2012).

The parallelism of our setup to the information sharing settings in the field results from the positive network effect that characterizes these settings. To implement these positive network effects, we model the interaction as a public goods game, in which each customer receives an additional utility gain when others contribute to the service. In the field, most of these services are provided by firms (our big player) that do not contribute to the public good but earn profits from the collected data and from access to the large network of active customers. Often these firms can monetize the collected customer data and the access to the network of active customers in multiple ways, e.g. by offering targeted advertisement or other services to third parties. Hence, whenever a for-profit firm supplies the services for an information public good, a substantial payoff asymmetry arises between the data-collecting firm and

⁹ When considering average earnings from the public good, the contributors' average earnings and the big players' earnings are higher in INFO_BIG (€6.50 and €26.40) than in TOKEN_BIG (€5.50 and €21.90), but earnings fail to differ significantly at the group level ($p=0.1085$).

its data-providing customers. Our central research question is whether and how this type of payoff asymmetry affects the customers' provision of personal information.

In our experiment, we model the data-collecting firm as a big player who benefits more than any other player from the collected information without providing any own information to the data pool. We study the effect of asymmetric payoffs on contributions by testing the difference between the information provision levels in a treatment with a big player (INFO_BIG) and in one without (INFO_BASE). We also compare the difference between the provision levels in the two information-based treatments to the difference between the provision levels in two analogous token-based treatments (TOKEN_BIG and TOKEN_BASE). This approach allows us to uncover any differential effects that asymmetric payoffs may have depending on the type of contributions, i.e. information-based versus token-based.

We find that while token-based contributions are significantly lower in a setting with asymmetric payoffs than in a setting without, the same is not true for information-based contributions. Since introducing a big player induces a substantial payoff asymmetry, we expected inequity-averse subjects to reduce their contributions in both settings. Hence, the surprising finding is not that payoff asymmetry leads to decreased contributions in the token-based treatments, but that it fails to do so in the information-based treatments.

Our experimental results seem well in line with the empirical observation that people are willing to provide their personal data in exchange for internet services. In fact, since in settings with a for-profit firm, individuals are more likely to provide personal information than monetary contributions, businesses that have a choice, should generally prefer collecting data on their network of active users and selling the insights or the access to the network to third parties. Introducing a fee-based model in some cases may even lead to an exodus of the active users, who were willing to accept payoff asymmetry in an information-based business model, but not in a fee-based business model.

By pointing out differences between information-based and token-based contributions in a public goods setting with payoff asymmetry, we contribute to the scant literature on the perception of fairness norms that vary with the type of input.¹⁰ As far as we can tell, this study is the first to point out that the domain-specific sensitivity and the insensitivity to benefit asymmetries, i.e., to inequity, may play an important role for a better understanding of the non-monetary drivers of information provision.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40881-021-00102-y>.

Acknowledgements Financial support from the German Research Foundation (DFG) through the research unit "Design & Behavior" (FOR 1371) is gratefully acknowledged. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1– 390838866.

Funding Open Access funding enabled and organized by Projekt DEAL.

¹⁰ Kube et al. (2012), e.g. find that the type of gift exchanged may affect perceived fairness norms.

Data availability The data is available at: <http://dx.doi.org/10.7802/1899>.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ackfeld, V., & Güth, W. (2019). Personal information disclosure under competition for benefits: Is sharing caring? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3350489>
- Benndorf, V., & Normann, H.-T. (2018). The willingness to sell personal data. *Scandinavian Journal of Economics*, 120, 1260–1278. <https://doi.org/10.1111/sjoe.12247>
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90, 166–193. <https://doi.org/10.1257/aer.90.1.166>
- Engel, C., & Rockenbach, B. (2011). We are not alone: The impact of externalities on public good provision. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1463259>
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114, 817–868. <https://doi.org/10.1162/003355399556151>
- Feri, F., Giannetti, C., & Jentzsch, N. (2016). Disclosure of personal information under risk of privacy shocks. *Journal of Economic Behavior & Organization*, 123, 138–148. <https://doi.org/10.1016/j.jebo.2015.12.001>
- Frik, A., & Gaudeul, A. (2020). A measure of the implicit value of privacy under risk. *Journal of Consumer Marketing*. <https://doi.org/10.1108/JCM-06-2019-3286>
- Gaudeul, A., & Giannetti, C. (2017). The effect of privacy concerns on social network formation. *Journal of Economic Behavior & Organization*, 141, 233–253. <https://doi.org/10.1016/j.jebo.2017.07.00>
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1, 114–125. <https://doi.org/10.1007/s40881-015-0004-4>
- Kube, S., Maréchal, M., & Puppe, C. (2012). The currency of reciprocity: Gift-exchange in the workplace. *American Economic Review*, 102(4), 1644–1662. <https://doi.org/10.1257/aer.102.4.1644>
- Poindexter, J. C., Earp, J. B., & Baumer, D. L. (2006). An experimental economics approach toward quantifying online privacy choices. *Information Systems Frontiers*, 8, 363–374. <https://doi.org/10.1007/s10796-006-9013-4>
- Rockenbach, B., & Sadrieh, A. (2012). Sharing information. *Journal of Economic and Organizational Behavior*, 81(2), 689–698. <https://doi.org/10.1016/j.jebo.2011.10.009>
- Rockenbach, B., Sadrieh, A., & Schielke, A. (2020). Providing personal information to the benefit of others. *PLoS ONE*, 15(8), e0237183. <https://doi.org/10.1371/journal.pone.0237183>
- Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, 22(2), 254–268. <https://doi.org/10.1287/isre.1090.0260>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.