Discussion of “The Market for Reviews: Strategic Behavior of Online Product Reviewers with Monetary Incentives”

Anna Ressi¹

Received: 4 June 2020 / Accepted: 10 June 2020 / Published online: 2 July 2020
© The Author(s) 2020

Abstract In this note, I discuss the article of Dorner et al. who experimentally study the issue of generating high-quality online product reviews under two different incentive schemes. I first classify their work in the existing literature on marketing and information systems and highlight their contribution. I then argue that the existence of cognitive dissonance costs serve as a reasonable explanatory factor for their main results. Two further critical issues warn against the limited generalizability of the present study to the real review market and ability to provide practical implications.

Keywords Discussion · Online Product Reviews · Strategic Downvoting · Intrinsic Motivation · Crowding-Out · Helpfulness Ratings

In this note, I discuss the article of Dorner, Giamattei and Greiff, who study the issue of generating high-quality online product reviews when subjects’ payoffs are dependent on helpfulness ratings that can be strategically manipulated by others. The authors employ a laboratory experiment with a supplementary online questionnaire. In line with previous studies arguing that in online markets, reputation systems usually rely on voluntary feedback from customers, review writing is implemented as a contributing to a public good (see, e.g., Cabral and Li 2015; King et al. 2014). The authors introduce two different incentive schemes: a flat wage paid for each review irrespective of quality (FWT), and a tournament that awards a bonus for the review that received the highest helpfulness rating from the other reviewers in a given group (BT). They then compare review quality as measured by the review’s length, and the

✉ A. Ressi
anna.ressi@whu.edu

¹ Chair of Organization Theory, WHU – Otto Beisheim School of Management, Vallendar, Germany

assigned helpfulness ratings. As predicted by basic economic theory, they find that the tournament indeed triggers incentives to engage in strategic downvoting, i.e. to assign low helpfulness ratings to competitors' reviews. However, review quality is higher than under a flat wage without strategic downvoting. In contrast to intuition, this suggests that low helpfulness ratings do not seem to negatively affect individuals' motivation to write high-quality reviews. In this experimental setting, the tournament thus proves to be a superior incentive device to flat wages in terms of soliciting high-quality reviews.

In the following, I will classify the work of Dorner, Giamattei and Greiff in the existing literature on marketing and information systems and highlight their contribution in these fields. Next, I will comment on three critical issues in the paper. I will conclude with a brief summary and suggestions for future research.

1 Existing Literature

Over the past few years, it has become increasingly important for sellers operating in the online market to provide a well-functioning and reliable platform, where customers share their evaluations and experiences. According to a 2017 Podium study (Podium 2017), an astonishing 93% of buyers consult online reviews before making their purchase decision. Several companies have even begun to specialize on providing a platform for reviews and as such basically sell reviews as part of their product portfolio (e.g., TripAdvisor, Yelp). Hence, by examining drivers of reviewers' behavior, the authors address a topical and economically significant research question that is of interest to practitioners and academic scholars alike. Up to now, this issue has predominantly been scrutinized by scholars in the field of marketing and information systems. The sizeable literature in these two fields span a great variety of research methods, including interviews, lab, field experiments, as well as quasi-natural experiments. In the field of information systems, work on online reviews can be broadly classified according to three sub-streams: the impact of online reviews on economic outcomes, the factors that drive review generation including reviewing motivation or reviewer self-selection, and the determinants for helpfulness ratings (King et al. 2014; Gutt et al. 2019).¹ Research in marketing, viewing online reviews as a form of electronic word-of-mouth, is particularly interested in the second sub-stream, i.e. individuals' extrinsic and intrinsic motivation to share their experiences and opinions with others, and specifically focuses on the quantity and quality dimensions of review writing. The most salient extrinsic and intrinsic motivating forces of spreading electronic word-of-mouth that previous studies detect include economic incentives, altruism, social norms, self-enhancement, enjoyment of helping, reputation seeking, identity building, and status-seeking (King et al. 2014; Wu 2019). The article of Dorner et al. (2020) contributes to and also bridges the gap between the two latter sub-streams, as it analyzes the ramifications of extrinsic rewards and the resulting strategic considerations for the assignment of helpfulness ratings and individuals' motivation to write high-quality reviews.

¹ For an overview, see, for example King et al. (2014) and Dellarocas and Narayan (2006a).

As such, it also contributes to the scientific debate on whether and how monetary rewards crowd out intrinsic and other forms of extrinsic motivation for review writing. The following studies only provide a glimpse into the current state of knowledge. In an experiment on Amazon Mechanical Turk, Wang et al. (2012) show that reviews exhibit the same quality when people are not monetarily rewarded as when they earn a fixed fee. However, additional performance-contingent rewards in terms of helpfulness ratings tend to improve review quality. In contrast, Stephen et al. (2012) find that perceived effort does not increase when subjects are paid for their reviews, but are considered more helpful than when reviewers are not incentivized. In a lab experiment, Li and Xiao (2014) demonstrate that rebates by sellers incentivize more buyers to write reviews. In contrast, a field study on eBay by Cabral and Li (2015) shows that offering rebates only increase the propensity of buyers to provide feedback, when payments are relatively high (\$2). Moderate rebates (\$1), on the other hand, do not translate into a significant increase in feedback provision. Wang et al. (2016), as well as Khern-am-nuai et al. (2018), find an increase in review volume but not in review quality in response to an introduction of monetary incentives. The latter study, using a natural experiment, even finds a drastic decrease in overall review quality. Burtch et al. (2018) emphasize the role of social norms and demonstrate that monetary incentives are more effective in attracting people to write reviews, where social norms are more effective at motivating higher review quality in terms of length. One additional study, which may be considered as particularly close to the present paper of Dorner et al. (2020) in terms of stressing the interrelation of different forms of intrinsic and extrinsic incentives, is Wu (2019). Through in-depth interviews with top reviewers on Amazon, the study uncovers mutually reinforcing and countervailing effects of different incentive devices for the motivation to write reviews. In line with the findings of Dorner et al. (2020), the author, amongst others, documents anecdotal evidence for the existence of fierce competition for status recognition and manipulations of helpfulness ratings. Taken together, the lack of a clear picture of what motivates review writing and how different extrinsic and intrinsic motivations interact warrants further research in this direction.

2 The Explanatory Power of Cognitive Dissonance Costs

Previous literature, for example on subjective peer evaluations (e.g., Leibbrandt et al. 2018; Baliotti et al. 2016; Carpenter et al. 2010) and sabotage opportunities in competitive environments (e.g., Harbring and Irlenbusch 2005, 2008; Harbring et al. 2007), has already documented people's awareness and willingness to engage in strategic behavior that harms others but leads to own individual benefits. Of course, demonstrating the robustness of this finding in the context of online product review systems is a worthwhile endeavor. However, the study's main take-away rather seems to be the result that even though strategic downvoting occurs, review quality counterintuitively does not deteriorate in response. In this section, I will argue that this finding is not as surprising as it may seem at first sight.

In their paper, the authors argue that to maximize the chances of receiving the bonus, rational players will strategically assign the lowest possible helpfulness rating to their rivals. As everyone will anticipate this behavior, it is optimal to not contribute to the public good, that is, to write a review with the lowest possible quality. The bonus will then be split evenly. This line of argumentation is surely reasonable if individuals' preferences only comprise monetary payoffs. However, it is a well-established fact that people face so-called cognitive dissonance costs (see, e.g., Harmon-Jones and Harmon-Jones 2007). This concept dates back to Festinger (1957) and refers to a state of psychological discomfort people experience when they hold conflicting mental models of the world. In the context of the study, this implies that people suffer a loss in their utility when assigning a particular helpfulness rating that does not correspond to their actual evaluation. In their decision-making process, people will hence face a trade-off between increasing monetary gains and minimizing cognitive dissonance costs. Hence, it is reasonable to expect that people will indeed engage in strategic downvoting, but they will not do so to the fullest possible extent. Helpfulness ratings will then on average only experience a parallel downward shift, such that high-quality reviews will still receive better helpfulness ratings than low-quality reviews. Under these considerations, it is still rational to exert high effort, that is, to write high-quality reviews, as doing so maximizes the chances of receiving the bonus payment.

A second argument in the paper for why review quality is expected to decline in response to strategic downvoting is based on social approval theory. Helpfulness ratings usually serve as a signal of social approval or disapproval, which in turn boosts individuals' motivation to exert effort. However, as long as helpfulness ratings are not always at their minimum irrespective of the quality, they are still informative of social approval. This is because when interpreting the helpfulness rating as a signal of approval or disapproval, rational players are perfectly aware and factor in the incentives of others to assign lower ratings and can thus still deduce the informational content of the biased ratings.

3 Generalizability to Review Markets in Reality

Apart from the incentives for strategic downvoting and the associated negative side effects discussed above, there exist further general reasons for why one would expect review quality to be lower in BT than in FWT. First, according to (self-)signaling theory, people in the flat wage treatment can in principle derive utility by signaling to themselves and to others that they are willing to provide voluntary contributions, which they are unable to do in the bonus treatment. However, as in FWT quality declines over the four periods even though helpfulness ratings are constantly high, these signaling motives (as well as concerns for social approval) do not seem to crucially affect subjects' behavior in this experimental setting. At least, these motives do not seem to be sufficiently strong to overcome the usual boredom effects commonly observed in overly long lab sessions. Also, in line with the seminal work of Deci (1971), introducing incentives for quality instead of a fixed wage may crowd out intrinsic motivation to write high-quality reviews. However, it is very likely that

subjects in this experimental setting were influenced by demand effects to write a review, which is why intrinsic motivation, being a key factor in actual review markets, should have been vanishingly low (Khern-am-nuai et al. 2018).

Against this backdrop, it remains to be analyzed whether the paper's main result that a tournament will not crowd out intrinsic as well as other forms of extrinsic incentives (social approval, signaling) also persists in the field. It is reasonable to think that these three motivational forces are significantly more pronounced and impactful in reality, when people have the freedom to choose whether and for which products they want to write a review.² In particular, self-determination theory suggests that intrinsic motivation is higher when people's psychological need for autonomy, competence, and relatedness are satisfied (Deci and Ryan 1985; Ryan and Deci 2000).³ All in all, the implication that a bonus awarded to the reviewer who received the highest helpfulness rating is superior to a flat wage incentive scheme may not continue to hold when these additional motivating factors that are unlikely to be triggered in the lab, but have been shown to be pivotal in the field, are considered.

Finally, according to Khern-am-nuai et al. (2018), there generally exist three approaches on monetary incentives that platforms in practice use to generate reviews. The first is to not provide explicit monetary incentives and solely rely on intrinsic motivation or non-monetary incentives, such as Amazon. The second approach is to tie rewards to quality in terms of helpfulness ratings. This approach, however, is rarely implemented in practice due to operational difficulties. The third approach is to pay a fixed fee per review written, which has become increasingly popular among a number of review platforms such as Best Buy, or Kmart. As in FWT the present study of Dorner et al. (2020) rewards subjects with a flat wage, it is clearly related to the third approach. In addition, as the authors motivate their analysis by drawing upon the example of Amazon's top reviewer league table, one might conclude that they also speak to the first category. However, in the second treatment (BT), the experiment uses a tournament with a monetary bonus awarded to the subject with the highest helpfulness rating, which implies that BT rather reflects the second category. Hence, the paper's results cannot be directly transferred to the case of Amazon, because monetary incentives may lead to different crowding-in or crowding-out effects of motivation than non-monetary incentives. Also, Khern-am-nuai et al. (2018) argue that due to operational difficulties, performance-contingent monetary rewards are rarely implemented in practice. This further limits the generalization of the present lab experiment and its possibility to provide practical implications.

² Note that in the experimental setting of Dorner et al. (2020), subjects were asked to write a review about all four products, which were determined by the experimenters.

³ For experimental evidence in the realm of review markets see, e.g., Wang et al. (2012).

4 The Pitfalls of Implementing a Tournament Incentive Scheme for Review Markets

Even if in real review markets it does hold true that a tournament based on helpfulness ratings is not detrimental for promoting high-quality reviews, the implication that it is a superior incentive device to a flat wage may be misleading. This is because implementing a tournament for each product, as it is the case in the experimental setting of Dorner et al. (2020), would simply be too costly, not only in terms of total monetary payments, but also in terms of operational difficulties (Khern-am-nuai et al. 2018). At best, review platforms may rely on a tournament that spans a certain time period and aggregates the helpfulness ratings for reviews of multiple products. But then, even if strategic downvoting does not lead to a decline in review quality, other predicaments immediately follow that put the supposed superiority of the tournament into perspective. For example, reviewers will strategically focus on the products that promise the highest helpfulness ratings, which according to the “brag-and-moan” (Hu et al. 2006) will be the case for especially poor as well as especially outstanding products. Hu et al. (2006) empirically document this suggested behavior for product sold at Amazon and find that, in response, the average score does not necessarily reveal the product’s true quality.⁴ Also, reviewers will choose relatively popular products to increase the visibility of their reviews (Shen et al. 2015). Due to this self-selection and review concentration bias, potential buyers of mediocre or niche products will lack a reliable and informative signal of product quality to make an informed purchase decision. Moreover, whereas the study of Dorner et al. (2020) focuses on generating high-quality reviews, feedback platforms additionally face the upstream problem of attracting reviewers in the first place. As tournaments are inherently risky, incentives have to be significantly increased in order to attract the same number of reviewers than under the riskless flat wage incentive scheme.

Finally, in light of the comprehensive results found by previous studies, it may well be the case that relying on unincentivized reviews dominates both the flat wage as well as the tournament incentive scheme in terms of implementation costs, the quantity and quality of reviews, as well as the signaling power of reviews for product quality. For example, Cabral and Li (2015) find that increasing the likelihood of providing voluntary feedback requires relatively high monetary rewards. The experimental findings of Wang et al. (2012) and Wang et al. (2016) demonstrate that review quality is relatively insensitive to the introduction of monetary incentives, while Khern-am-nuai et al. (2018) even find a drastic decrease in overall review quality. In addition, unpaid reviews are likely to be most informative of true perceptions of product quality, as they are less prone to biases that distort reviewers’ evaluations and reduce buyers’ confidence and trust in the review system. For example, in a field experiment on eBay, Cabral and Li (2015) find that rebates for reviews offered by sellers tend to induce a positive bias due to reciprocity concerns. Similarly, the interview data of Wu (2019) documents reviewers’ feelings of obligation to write positive reviews on Amazon in response to receiving free

⁴ See also the study of Dellarocas and Narayan (2006b) for movie reviews. Further, Khern-am-nuai et al. (2018) provide evidence in favor of the latter approach.

product samples, which can equally well be considered a flat wage. The natural experiment of Khern-am-nuai et al. (2018) shows that a newly introduced payment for review writing attracts additional reviewers, who then assign inflated ratings for previously low-rated products. Another astonishing drawback of incentivizing reviewers is presented by Stephen et al. (2012). In multiple experiments, they find that customers' knowledge about reviewers' compensation influence the assessment of the focal product in the short- as well as the long-term. If customers know that the reviewer was paid, results suggest that they form less favorable expectations of product quality based on the review, reduce their willingness-to-pay, and also lower their product evaluations even after first-hand experience. Considering that Dorner et al. (2020) are particularly concerned about the loss of signaling power of helpfulness ratings, unincentivized reviews not only avert strategic downvoting, but also uphold the arguably more pivotal signaling power of product reviews per-se.

5 Conclusion

All in all, while this study contributes to our understanding of the influence of different incentive schemes on individuals' motivation mix to write high-quality reviews and determinants of helpfulness ratings, it remains to be verified whether the observed effects persist in more realistic settings. In this context, a major concern is that low helpfulness ratings may deplete intrinsic motivation to write reviews, which sellers and review platforms crucially rely on, but which is unlikely to have been a significant motivator in the experimental setting of Dorner et al. (2020). Also, providing practitioners with well-founded strategic prescriptions on the optimal reward system requires a more extensive study comparing not only the flat wage scheme and the (hardly practicable) tournament system awarding a monetary bonus, but also unpaid reviews relying on a non-monetary tournament (as it is the case for Amazon).⁵ In this context, pursuing a more comprehensive perspective in terms of potentially countervailing and mutually reinforcing effects on the number of reviews for different products, their quality and informativeness in terms of inherent distortions, as well as the informativeness of helpfulness ratings may be a particularly worthwhile and promising direction for future research.

Funding Open Access funding provided by Projekt DEAL.

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/>.

⁵ For evidence suggesting that readers' ratings of reviews improve review quality, see the field study of Chen et al. (2010).

References

2017. Podium. <http://learn.podium.com/rs/841-BRM-380/images/2017-SOOR-Infographic.jpg>. Accessed July 1, 2020.
- Baliotti, S., R.L. Goldstone, and D. Helbing. 2016. Peer review and competition in the Art Exhibitio game. *Proceedings of the National Academy of Sciences* 113(30):8414–8419.
- Burtch, G., Y. Hong, R. Bapna, and V. Griskevicius. 2018. Stimulating online reviews by combining financial incentives and social norms. *Management Science* 64(5):2065–2082.
- Cabral, L., and L.I. Li. 2015. A dollar for your thoughts: feedback-conditional rebates on eBay. *Management Science* 61(9):2052–2063.
- Carpenter, J., P.H. Matthews, and J. Schirm. 2010. Tournaments and office politics: evidence from a real effort experiment. *The American Economic Review* 100(1):504–517.
- Chen, Y., F.M. Harper, J. Konstan, and S.X. Li. 2010. Social comparisons and contributions to online communities: a field experiment on movielens. *American Economic Review* 100(4):1358–1398.
- Deci, E.L. 1971. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology* 18(1):20–28.
- Deci, E.L., and R.M. Ryan. 1985. *Intrinsic motivation and self-determination theory in human behavior*. New York: Plenum.
- Dellarocas, C., and R. Narayan. 2006a. What motivates consumers to review a product online? A study of the product-specific antecedents of online movie reviews. In *WISE*, 1–6.
- Dellarocas, C., and R. Narayan. 2006b. A statistical measure of a population's propensity to engage in post-purchase online word-of-mouth. *Statistical science* 21(2):277–285.
- Dorner, V., M. Giamattei, and M. Greiff. 2020. The market for review: strategic behavior of online product reviewers with monetary incentives. *Schmalenbach Business Review*. <https://doi.org/10.1007/s41464-020-00094-y>
- Festinger, L. 1957. *A theory of cognitive dissonance*. Evenstone: Row, Peterson.
- Gutt, D., J. Neumann, S. Zimmermann, D. Kundisch, and J. Chen. 2019. Design of review systems—a strategic instrument to shape online reviewing behavior and economic outcomes. *The Journal of Strategic Information Systems* 28(2):104–117.
- Harbring, C., and B. Irlenbusch. 2005. Incentives in tournaments with endogenous prize selection. *Journal Institutional Theoretical Economics* 161(4):1–28.
- Harbring, C., and B. Irlenbusch. 2008. How many winners are good to have? On tournaments with sabotage. *Journal of Economic Behavior & Organization* 65(3/4):682–702.
- Harbring, C., B. Irlenbusch, M. Kräkel, and R. Selten. 2007. Sabotage in corporate contests—an experimental analysis. *International Journal of the Economics of Business* 14(3):367–392.
- Harmon-Jones, E., and C. Harmon-Jones. 2007. Cognitive dissonance theory after 50 years of development. *Zeitschrift für Sozialpsychologie* 38(1):7–16.
- Hu, N., P.A. Pavlou, and J. Zhang. 2006. Can online reviews reveal a product's true quality? Empirical findings and analytical modeling of online word-of-mouth communication. In *Proceedings of the 7th ACM conference on Electronic commerce*, 324–330.
- Khern-am-nuai, W., K. Kannan, and H. Ghasemkhani. 2018. Extrinsic versus intrinsic rewards for contributing reviews in an online platform. *Information Systems Research* 29(4):871–892.
- King, R.A., P. Racherla, and V.D. Bush. 2014. What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of interactive marketing* 28(3):167–183.
- Leibbrandt, A., L.C. Wang, and C. Foo. 2018. Gender quotas, competitions, and peer review: experimental evidence on the backlash against women. *Management Science* 64(8):3469–3970.
- Li, L.I., and E. Xiao. 2014. Money talks: rebate mechanisms in reputation system design. *Management Science* 60(8):2054–2072.
- Ryan, R.M., and E.L. Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist* 55:68–78.
- Shen, W., Y.J. Hu, and J.R. Ulmer. 2015. Competing for attention: an empirical study of online reviewers' strategic behavior. *MIS Quarterly* 39(3):683–696.
- Stephen, A., Y. Bart, C. Du Plessis, and D. Goncalves. 2012. *Does paying for online product reviews pay off? The effects of monetary incentives on content creators and consumers*. In Zeynep Gürhan-Canli, Cele Otnes, and Rui (Juliet) Zhu (eds), *NA – Advances in Consumer Research*, Vol 40, Duluth, MN: Association for Consumer Research, 228–231.

- Wang, J., A. Ghose, and P. Ipeirotis. 2012. Bonus, disclosure, and choice: what motivates the creation of high-quality paid reviews? In *Proceedings of 2012 International Conference on Information Systems*. <http://aisel.aisnet.org/icis2012/proceedings/DigitalNetworks/9>.
- Wang, S., P. Pavlou, and J. Gong. 2016. Monetary incentives, online reviews, and product sales: an empirical investigation. In *ICIS 2016 Proceedings*. <http://aisel.aisnet.org/icis2016/Economics/Presentations/9>.
- Wu, P.F. 2019. Motivation crowding in online product reviewing: a qualitative study of amazon reviewers. *Information & Management* 56(8):103–163.

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