

Innovation and Competition in the Digital Economy:

Implications for Internet Platforms, Telecommunications Networks
and Data Sharing Initiatives

Dissertation submitted to attain the degree of

Doctor rerum politicarum (Dr. rer. pol.)

at the School of Business, Economics and Information Systems
of the University of Passau

Submitted by: Oliver Zierke, M.Sc.
First Reviewer: Prof. Dr. Jan Krämer
Second Reviewer: Prof. Dr. Andreas König

May 2021



Accepted as dissertation

at the School of Business, Economics and Information Systems of the University of Passau

Date of the disputation: July 22, 2021

Chair of the examining committee: Prof. Dr. Johann Graf Lambsdorff

First reviewer: Prof. Dr. Jan Krämer

Second reviewer: Prof. Dr. Andreas König

Abstract

Governments around the world currently focus on shaping the digital economy. Particular attention is paid to Internet platforms, Internet infrastructure and data as essential components of the digital economy. The three studies in this thesis contribute to the understanding of the behavior of firms in each of these domains and derive insights for future regulations and business projects.

The first study deals with the ranking of content on Internet platforms and how it affects the incentives of content providers to invest in content quality. The focus of the study is on sponsored ranking and organic ranking, but the case that a vertically integrated content provider is favored by an Internet platform is also taken into account. Using a game theoretic model, it is shown that there is no ranking design that strictly leads to more investment compared to the other designs. It is also shown that the Internet platform usually chooses the type of ranking that, from the perspective of the Internet platform and consumers, yields the best expected overall content quality. The second study deals with the incentive of Internet service providers to throttle specific Internet content. The key finding is that Internet service providers use this instrument to utilize the capacity of their telecommunications network more efficiently. This leads not only to more benefits for Internet users, but also to a higher incentive to invest in network capacity due to better monetization. The third study examines the circumstances under which firms are willing to share data with other firms. By means of an economic laboratory experiment, it is shown that more data is shared if the firms have control over who exactly they share data with. Thus, for example, data pools that grant unrestricted data access to all participating firms can be expected to perform worse than data pools that give their participating firms control over with whom their uploaded data is shared. In addition, the third study finds that established relationships are characterized by more data sharing and less volatility in the amount of shared data than new relationships. The study concludes that data sharing projects should not be expected to work optimally right away.

In summary, the studies in this thesis identify a number of costs that may arise when digital firms' choice is restricted by regulation or design. The ability of Internet service providers to throttle certain content and the ability of Internet platforms to choose the ranking design are usually used in the best interests of consumers. Data sharing also works best when firms are free to decide who gets their data.

Acknowledgments

First and foremost, I would like to thank my supervisor Prof. Dr. Jan Krämer who guided me throughout my doctoral studies and whose door was always open. Even in difficult phases without results in sight, he pointed me to the right direction. Furthermore, I would like to thank my second examiner Prof. Dr. Andreas König and the chairperson of the examination committee, Prof. Dr. Johann Graf Lambsdorff, for agreeing to join the doctoral committee.

My thanks also go to my colleagues and friends. The academic exchange and various social activities after work have enriched my time in Passau. My special thanks go to my office colleague and co-author Nadine Stüdlein, because of the very good collaboration in joint academic projects, the fruitful academic exchange, and because she greeted me every morning with a grin. No less grateful am I to Dr. Michael Wohlfarth, with whom I shared the office for the first two years and who helped me to quickly find my way around the chair and the city of Passau. Also, I thank Dr. Torben Stoffer who broadened my methodological horizons by giving me an insight into the world of questionnaires and who was also willing to weekly join my journey to Regensburg to enjoy the free time in the nearest bouldering hall. I would also like to thank Dr. Victoria Fast, who proofread large parts of this thesis and with whom I successfully completed various PhD courses and many other challenges together. Moreover, I thank Dr. Bastian Haberer, among others, for several night shifts in which we discussed various game-theoretical modelling approaches, Janina Hofmann for introducing me to the world of yoga and other crazy things, Dr. Daniel Schnurr for the good collaboration in teaching and research and several comfortable evenings on the balcony of his apartment in Passau, Katrin Gleixner for her organizational support, not only related to lectures, and Prof. Dr. Thomas Widjaja for every instance when he exceptionally had time for lunch together.

Finally, I would like to thank my parents. They supported me before and during my doctoral studies and still do so today. Without their support, this thesis would not exist.

Contents

Introduction	1
Paying for Prominence	9
Introduction	10
Related Literature	14
Baseline Scenario: Organic Ranking	16
Sponsored Ranking	19
The Effect of Rankings on Quality Investment Incentives	21
Hybrid Sponsored Ranking with Quality Factor	27
Vertical Integration and Self-Preferencing	31
Welfare Analysis	35
Conclusions	38
References	42
Appendix	47
Raising Total Capacity Provision of Mobile Network Operators Through Throttling	61
Introduction	62
Related Literature	64
A Model on the Impact of Bandwidth Throttling on Capacity Investment Incentives in Mobile Networks	66
Conclusion	71
References	73
Data Sharing between Firms	77
Introduction	78
Investigations on Control in the Experimental Literature	81
Experimental Design	84
Procedures	86
Results	88
Discussion and Conclusion	97
References	98
Appendix	104

Introduction

Shaping the digital economy is currently one of the most important issues for governments and firms around the world. This becomes evident when considering digital policy strategies such as Australia's Tech Future (Australian Government 2019), Canada's Digital Charter (Government of Canada 2019) or the European Data Strategy of the European Commission (European Commission 2020b), which have all been introduced in recent years. Key elements of such strategies typically include measures to limit the abuse of market power by large Internet platforms, measures to improve the provision of broadband capacity in telecommunications networks, and measures to promote new technologies and services through better access to data. Each of the three studies in this thesis is closely connected to one of these topics. All studies contribute to a better understanding of the behavior of firms in a digital economy and provide insights that should be taken into account by regulators and firms.

The first study (Krämer and Zierke 2020) of this thesis investigates the impact of Internet platforms' ranking systems on content providers' incentives to invest in content quality. It contributes to the broad and intensive debate on a sensible design and regulation of powerful Internet platforms. Not only since the start of the Covid-19 pandemic Internet platforms have become an essential facility for large parts of the entire economy. Google search, for example, which is the dominant search engine in most countries around the world, controls the access of website owners to the majority of Internet users. Specifically, website owners can usually only get high traffic rates if they have a good listing on Google. A similar situation is present at Apple's App Store or Google's Play Store as it is very difficult for app developers to reach iOS or Google Android users without being listed in the corresponding App Store. Further, even if circumventing the platform seems to be easy as in the case of YouTube, one can consider this platform as an essential facility, as a significant portion of the potential viewership will only be accessible to video producers through this platform. Due to this strong position, the market behavior of Internet platforms is being critically scrutinized by regulators all over the world. A current example from the United States of America is the ongoing debate about the non-transparent requirements that Apple places on apps in its App Store (Kastrenakes 2020). In the same context, the European Commission has already opened an antitrust investigation against Apple in 2020 (European Commission 2020a).

However, while this investigation is mainly focused on the requirements to be listed in the App Store at all, it is of equal importance for content providers where they are listed. In other words, the ranking algorithms used by Internet platforms are of major importance. For this reason, the European

Commission issued a regulation for platforms in 2019 with the obligation to make their ranking criteria transparent (EU Regulation 2019/1150 Article 5). The first study of this theses investigates such rankings of Internet platforms and evaluates how different ranking designs affect content providers' incentives to invest in costly content quality. At its core, the study distinguishes between two types of ranking. The first type, organic ranking, is purely based on the quality of the content observed by the platform. The second type, sponsored ranking, is based on a second price auction. Surprisingly, the results of the study indicate that a ranking based on quality does not necessarily provide higher incentives for content providers to invest in quality than a sponsored ranking. Instead, the study identifies three effects, the *competition effect*, the *uncertainty effect* and the *revenue-share effect*, whose strengths vary among the different ranking systems.

The effects are identified by means of a game-theoretic model. This model is based on a set-up in which two content providers offering their content to Internet consumers through a monopoly Internet platform. Content providers can invest in the quality of their content, whereby one content provider is more efficient at generating quality than the other content provider. The Internet platform offers consumers free access to the content of the content providers. Under organic ranking, the content providers only have to pay a revenue share for the use of the Internet platform. Consequently, the Internet platform tries to maximize the total revenue of the content providers in order to receive the maximum profit from charging the revenue share. Note that the revenue of a content provider is positively related to its content quality and its position in the platform's ranking. Based on that, a content of higher quality generates more revenue in the better ranking position than a content of lower quality. Therefore, it is in the best interest of the platform to rank the higher quality content first and the lower quality content last. However, Internet platforms usually cannot perfectly detect the quality of the content. Consequently, there is a likelihood that the platform will erroneously assign the better position to the content of lower quality. Under sponsored ranking, the platform not only charges a revenue share, but receives an additional revenue stream from content providers by auctioning the better ranking position by means of a second price auction. Consequently, the content quality does not matter when awarding positions under sponsored ranking.

This difference leads directly to the first effect, the *competition effect*. Under sponsored ranking, the content providers compete by bidding for the best position in the platform's ranking. This has no direct effect on the content providers' incentives to invest in quality. Under organic ranking content providers compete in quality for the best position in the platform's ranking, i.e., a better quality results in a higher probability to reach the better position in the ranking of the platform. This, of

course, has a positive effect on the content providers incentives to invest in quality. Consequently, the *competition effect* works in favor of the incentives to invest in quality under organic ranking. This effect is stronger the more similar content providers are in their efficiency to improve the quality of their content through further investment. If the difference is very large, it disappears completely.

The second effect, the *uncertainty effect*, is related to the limited ability of the platform to assess quality. This ability is especially important if the platform uses organic ranking. Usually, the more efficient content provider offers the better quality more often and therefore receives the better ranking position more often. This leads to a higher expected demand, which creates the incentive to boost this demand further by providing higher quality content. However, the expected demand of the more efficient content provider decreases the more often the Internet platform erroneously assigns the better position to the content provider with the worse quality. As a consequence, the incentive of the more efficient content provider to invest in quality also decreases. The reverse argumentation applies to the content provider with the lower efficiency. If uncertainty increases, it gets more often the better position and therefore in expectation more demand. Consequently, the incentive of the less efficient content provider to invest in quality increases. However, from the Internet platform's and the Internet users' perspective, the negative impact on the efficient content provider outweighs the positive impact on the less efficient content provider. Under sponsored ranking, there is no *uncertainty effect*, as the platform can perfectly observe the content providers' bids. In particular, the more efficient content provider always wins the auction and gets the better ranking position. The result is comparable to organic ranking with no uncertainty. Thus, a lower uncertainty works in favor of the overall incentives to invest in quality under sponsored ranking.

The third effect, the *revenue-share effect*, is a result of platform's second revenue source from the auction under sponsored ranking. In particular, there is a trade-off between the two revenue sources. The bids in the auction strictly increase when the revenue share decreases. Consequently, the Internet platform weighs the gain from an increase in the revenue share against the associated decrease in auction revenue. As a result, the Internet platform charges a strictly lower revenue share from content providers under sponsored ranking compared to organic ranking. Note that a lower revenue share increases content providers' incentive to invest in content quality, i.e., the *revenue-share effect* works in favor of content providers' incentives to invest in quality under sponsored ranking.

Overall, the first study shows that an organic ranking only creates higher incentives for content providers to invest in quality if the content providers are very similar in terms of their efficiency in producing quality and if the ability of the Internet platform to assess quality is very high. Besides

this, the study considers two extensions. The first extension investigates hybrid forms of organic and sponsored rankings, where the influence of the ranking criteria bid and quality are weighted. The results show that the three effects mentioned above carry over. Further, a fourth effect is identified, the *payment-reduction effect*. This effect works in favor of hybrid rankings. It reflects that in a hybrid ranking, the content provider that wins the auction has to pay less the better its content quality is compared to the content of the other content provider. However, the result that no ranking system strictly dominates the other ranking system persists, i.e., it depends on the similarity of content providers and platform's ability to assess quality if a system with more weight on sponsored ranking or a system with more weight on organic ranking yields more investment incentives. The second extension deals with vertical integration, where the platform always gives the better position in the ranking to the content provider with which it is vertically integrated. Here, the *revenue-share effect* is of particular importance, because a vertically integrated content provider does not have to pay a revenue share. This has a positive effect on its incentive to invest in content quality. However, the non-integrated content provider does not benefit from this revenue share reduction, but only suffers from the fact that it always gets the lower position in the ranking. Eventually, the overall incentives of content providers to invest in quality may be greater under vertical integration compared to other situations. However, this depends on whether the more efficient or less efficient content provider is vertically integrated, whether the difference in efficiency between the two content providers is rather large or rather small and whether the revenue share effect is of high importance.

In summary, the results of the first study show that there is no ranking system that provides strictly higher incentives to invest in quality. The welfare analysis also shows that the ranking algorithm that is most profitable for the platform is in many cases also the algorithm that leads to the highest expected quality from the consumer's perspective. In other words, the study provides evidence that it is usually in the interest of Internet platforms to design their ranking algorithm in the best interest of consumers.

The second study (Zierke 2020) is about the incentives of Internet Service Providers (ISP) to throttle specific types of Internet content in their telecommunications network. This study is closely linked to the debate on net neutrality. Since the seminal article by Wu (2003), net neutrality has been a continuously and controversially discussed topic not only in science but also in the public. Although the ideal of an "(...) Internet that does not favor one application (say, the world wide web), over others (say, email)" (Wu 2003, p. 145) is far from being realized, regulators around the world enforce neutrality obligations on telecommunications infrastructure referring to the principle of net neutrality.

Neutrality regulations are currently in force, for example, in Brazil since 2014 (Presidência da República 2014), the European Union since 2015 (EU regulation 2015/2120) and India since 2018 (Ministry of Communications 2018). In other countries like the United States of America, the debate on net neutrality is not yet over. Recently, a senator affiliated with the Democratic Party announced plans to introduce a bill that restores the FCC's rights to enforce net neutrality, which have been withdrawn by the former Republican government (Kelly 2021). One of the major goals of net neutrality regulations is to prevent that Internet Service Providers (ISP) discriminate specific content in their network, because such discrimination could hurt static competition between established Internet content providers or dynamic competition between established and new Internet content providers (Wu 2003). On this basis, practices such as content-type-based throttling of data streams are among the core practices to be prevented by net neutrality regulations (Easley, Guo, and Kraemer 2018).

However, the second study of this thesis shows that a prohibition of content-type-based throttling may entail costs not only for ISPs but also for Internet users. In particular, the game theoretic model shows that one of the reasons for ISPs' use of throttling is to improve the overall quality of experience (QoE) of Internet users. The QoE, especially of media content like video streams, depends on two major factors. The first factor is resolution (i.e., the bit rate). Audio as well as video streams can have a high resolution or a low resolution. Throttling allows ISPs to reduce resolution by reducing the bandwidth available for media content, i.e., media content takes up less of the ISPs' network capacity. The second factor is the frequency of congestion events in the telecommunications network when consuming media content. Each of these congestion events can interrupt or even completely prevent Internet users' media consumption. There is a trade-off between both factors. By using throttling, an ISP lowers the resolution and thus the QoE of the media stream, but it also lowers the frequency of congestion events and thus simultaneously improves the QoE of the stream. One finding of the study is that ISPs resolve this trade-off in the best interests of Internet users. In this way, an ISP can offer more utility per unit of network capacity, which is associated with a higher willingness to pay of Internet users per unit of network capacity and thus a better monetization of each unit of network capacity. Further, as monetization of each unit of network capacity is improved, the incentives of ISPs to invest in network capacity also increases. In summary, the second study shows that allowing ISPs to throttle certain types of content can lead to more overall capacity in telecommunications networks and more utility per unit of network capacity for Internet users.

The third study (Krämer, Stüdlein, and Zierke 2020) is about how to encourage firms to share their valuable data. Data is one of the most important resources in the digital economy. The data possessed

by a firm is generally useful not only for that firm but also for other firms. In this sense, each firm is a bottleneck, as it is usually the only one that can provide access to a specific set of data. For this reason, governments and business initiatives around the world are trying to stimulate data sharing among firms. In India, for example, a panel of experts appointed by the Indian government recently recommended in its draft report to establish a new regulatory authority for non-personal data that would force firms to share their data (Rai 2020, Mehrotra 2020). A less radical approach was taken by the European Commission in 2019 with the launch of a "Data Sharing Support Centre", which provides expertise and experience on data sharing (Fraunhofer Institute for Open Communication Systems 2019). This is in line with the 2020 published European Data Strategy. According to this strategy, the main objectives of the European Commission are to provide expertise and legal certainty for data sharing (European Commission 2020b). In this way, the voluntary sharing of data between firms shall be promoted (European Commission 2020b).

Accordingly, the aim of the third study of this thesis is to show how data sharing initiatives should be designed to encourage voluntary data sharing between firms. For this purpose, the study investigates *control* over how dedicated a firm can decide which other firms gets how much of its data as an important design parameter. More precisely, three regimes are differentiated. First, a regime in which a firm can control how much data it makes accessible but not which firms get access (low control), second, a regime in which a firm can decide how much data it makes accessible and whether it excludes some firms from access to this data (medium control), and third, a regime in which a firm can decide specifically for each potential sharing partner how much data it makes accessible (high control). These scenarios reflect that, for example, a data sharing platform can be designed in such a way that either all participating firms have access to all uploaded data or that each firm can decide who has access to its uploaded data or that firms can exclusively share data bilaterally on the platform. Based on a literature review, the study expects more control to encourage data sharing. This expectation is tested by means of an economic experiment. The results confirm that control over who has access to shared data fosters overall data sharing. However, the importance to specifically control who receives how much data could not be confirmed. Moreover, it is shown that established relationships are characterized by more data sharing and less volatility compared to new relationships. The study concludes that any system that gives control to firms over who has access to their shared data is weakly dominating a system where firms have no such control and that data sharing projects should not be expected to perform optimally immediately. .

In summary, all three studies show that restricting the free choice of firms in the digital economy may

come at costs. The first two studies show that Internet platforms, as well as ISPs, tend to design their services in the interest of consumers if they are not restricted in their freedom of choice by regulation. The third study also shows that restricting the freedom of choice of firms can be detrimental, i.e., in data sharing initiatives where firms are deprived of their control over who exactly has access to their shared data, less data may be shared, overall. These costs should be considered when designing future regulations or projects in the digital economy.

Bibliography

- Australian Government, Department of Industry, Science, Energy and Resources. 2019. Setting the agenda for Australia’s tech future. Retrieved from <https://www.minister.industry.gov.au/ministers/karenandrews/media-releases/setting-agenda-australias-tech-future>, accessed on 10/07/2020.
- Easley, Robert F, Hong Guo, Jan Kraemer. 2018. Research commentary — From net neutrality to data neutrality: A techno-economic framework and research agenda. *Information Systems Research* **29**(2) 253–272.
- European Commission. 2020a. Antitrust: Commission opens investigations into Apple’s App Store rules. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1073, accessed on 11/07/2020.
- European Commission. 2020b. A European strategy for data. COM(2020) 66 final.
- Fraunhofer Institute for Open Communication Systems. 2019. European “support centre for data sharing” released as a beta version. Retrieved from https://www.fokus.fraunhofer.de/en/fokus/news/scds_07_2019, accessed on 13/07/2020.
- Government of Canada. 2019. Minister bains announces Canada’s digital charter. Retrieved from <https://www.canada.ca/en/innovation-science-economic-development/news/2019/05/minister-bains-announces-canadas-digital-charter.html>, accessed on 10/07/2020.
- Kastrenakes, Jacob. 2020. Hey.com exec says Apple is acting like ‘gangsters,’ rejecting App Store updates and demanding cut of sales. Retrieved from <https://www.theverge.com/2020/6/16/21293419/hey-apple-rejection-ios-app-store-dhh-gangsters-antitrust>, accessed on 11/07/2020.
- Kelly, Makena. 2021. Democrats are gearing up to fight for net neutrality. Retrieved from <https://www.theverge.com/2021/3/9/22321995/net-neutrality-ed-markey-save-the-internet-open-ajit-pai-rosenworcel>, accessed on 26/03/2021.
- Krämer, Jan, Nadine Stüdlein, Oliver Zierke. 2020. Data sharing between firms. Working Paper.
- Krämer, Jan, Oliver Zierke. 2020. Paying for prominence. Working Paper.

- Mehrotra, Karishma. 2020. It ministry panel for sharing non-personal data, wants dedicated authority. Retrieved from <https://indianexpress.com/article/technology/tech-news-technology/it-ministry-data-sharing-6503927/>, accessed on 10/07/2020.
- Ministry of Communications, Department of Telecommunications Networks & Technologies (NT) Cell. 2018. No 12-30/nt/2015/ott (pt): Regulatory framework on 'net neutrality'. Retrieved from https://dot.gov.in/sites/default/files/DoT%20Letter%20on%20Net%20Neutrality%20Regulatory%20Framework%20dated%2031%2007%202018_0.pdf, accessed on 10/07/2020.
- Presidência da República, Casa Civil Subchefia para Assuntos Jurídicos. 2014. Law no. 12.965, april 23rd 2014 (the comitê gestor da internet no brasil, trans.). Retrieved from <https://www.cgi.br/pagina/marco-civil-law-of-the-internet-in-brazil/180>, accessed on 10/07/2020.
- Rai, Saritha. 2020. India seeks to limit Facebook, Google dominance over online data. Retrieved from <https://www.bloomberg.com/news/articles/2020-07-13/india-seeks-to-limit-facebook-google-dominance-over-online-data>, accessed on 13/07/2020.
- Wu, Tim. 2003. Network neutrality, broadband discrimination. *J. on Telecomm. & High Tech. L.* **2** 141–179.
- Zierke, Oliver. 2020. Raising total capacity provision of mobile network operators through throttling. *PTC 2020 Conference Proceedings* Available at https://www.ptc.org/PTC20/Proceedings/Paper_YS_1_21_Zierke_Oliver.pdf.

Paying for Prominence

The effect of sponsored rankings on the incentives to invest in the quality of free content on dominant online platforms

Authors Jan Krämer, University of Passau, Germany.
Oliver Zierke, University of Passau, Germany.

Presented at¹

- IS Brown-Bag Seminar, 2017, Passau, Germany.
- 28th European Regional Conference of the International Telecommunications Society (ITS), 2017, Passau, Germany.
- The 45th Research Conference on Communications, Information and Internet Policy (TPRC), 2017, Arlington, USA.
- Colloquium for doctoral students, 2019, Passau, Germany.
- 46th Annual Conference of the European Association for Research in Industrial Economics (EARIE), 2019, Barcelona, Spainien.
- 17th ZEW Conference on the Economics of Information and Communication Technologies, 2019, Mannheim, Germany.

Abstract

Many independent content providers (CPs) offer free content through a dominant platform (e.g., an app store) and obtain revenues later from the usage of their content (e.g., through ads or in-app purchases). Attaining a more prominent placement on the platform is important for CPs, because it gives them more demand, and thus more revenues. We consider CPs' investment incentives to produce content quality under various ranking algorithms when the dominant platform demands a share of the CPs' revenues. We consider a game-theoretic model in which two CPs that differ in their efficiency to produce content quality compete for prominence on the platform. The platform commits to a revenue share and to a ranking algorithm that is either based on content quality (organic ranking) or bids (sponsored ranking), or a combination thereof. We highlight that CPs' incentives to invest in content quality under a given ranking regime can be explained by the interaction of four effects, that we denote as competition, uncertainty, revenue-share, and payment-reduction effect. We highlight how the strengths of these effects vary between different rankings and that it depends crucially on the CPs' efficiency differences and the platform's uncertainty to assess content quality which ranking yields the highest content quality, consumer surplus and platform profit. For instance, we show that under a pure sponsored ranking the platform demands the lowest revenue share, and that this ranking provides the highest investment incentives when CPs' efficiency differences and the platform's uncertainty are relatively high. We also study the case where the platform is vertically integrated with one CP and biases the ranking towards its own CP. We show that the platform's self-preferencing lowers the content quality of the non-integrated CP, but that vertical integration yields a significant increase of the integrated CP's content quality, and can result in an overall higher consumer surplus. Therefore, our results bear important implications for the regulation of digital content platforms.

¹The list does *not* include presentations held by co-authors.

Paying for prominence: The effect of sponsored rankings on the incentives to invest in the quality of free content on dominant online platforms

Jan Krämer, Oliver Zierke

Chair of Internet and Telecommunications Business, University of Passau, Dr.-Hans-Kapfinger-Str. 12, 94032 Passau, Germany, jan.kraemer@uni-passau.de, o.zierke@yahoo.de, <http://www.ibusiness.uni-passau.de>

Many independent content providers (CPs) offer free content through a dominant platform (e.g., an app store) and obtain revenues later from the usage of their content (e.g., through ads or in-app purchases). Attaining a more prominent placement on the platform is important for CPs, because it gives them more demand, and thus more revenues. We consider CPs' investment incentives to produce content quality under various ranking algorithms when the dominant platform demands a share of the CPs' revenues. We consider a game-theoretic model in which two CPs that differ in their efficiency to produce content quality compete for prominence on the platform. The platform commits to a revenue share and to a ranking algorithm that is either based on content quality (organic ranking) or bids (sponsored ranking), or a combination thereof. We highlight that CPs' incentives to invest in content quality under a given ranking regime can be explained by the interaction of four effects, that we denote as competition, uncertainty, revenue-share, and payment-reduction effect. We highlight how the strengths of these effects vary between different rankings and that it depends crucially on the CPs efficiency differences and the platforms uncertainty to assess content quality which ranking yields the highest content quality, consumer surplus and platform profit. For instance, we show that under a pure sponsored ranking the platform demands the lowest revenue share, and that this ranking provides the highest investment incentives when CPs efficiency differences and the platforms uncertainty are relatively high. We also study the case where the platform is vertically integrated with one CP and biases the ranking towards its own CP. We show that the platform's self-preferencing lowers the content quality of the non-integrated CP, but that vertical integration yields a significant increase of the integrated CP's content quality, and can result in an overall higher consumer surplus. Therefore, our results bear important implications for the regulation of digital content platforms.

Key words: online platform; ranking algorithm; content quality; sponsored search, platform regulation; self-preferencing

1. Introduction

Several online platforms have become the gatekeepers for digital content. For example, in order to be noticed by consumers, mobile apps must either appear on Apple's App Store or on Google's Play

Store, and are most often present on both platforms (Bresnahan et al. 2014)¹; videos must be posted on Youtube, and websites must appear on Google Search. In all of these platforms the majority of the content providers (CPs) offer their content for free to consumers (free-to-download apps, free-to-play games, free-to-view videos, free-to-access websites) and derive revenues predominantly through advertisements or in-app purchases. For instance, in the first quarter of 2018, more than 94% of all apps on Google’s Play Store, and more than 88% of all apps on Apple’s App Store were free to download (Statista 2018).

When consumers search for content on a platform, they are presented with a results list that ranks the CPs in some order. Empirical research has found that such rankings have a significant causal effect on consumer demand, i.e., consumers tend to demand more of the content that is listed higher, because it is listed higher (see, e.g., Baye et al. 2009, Glick et al. 2014, Ursu 2018). For example, in 2018 a slight change in the Play Store’s algorithm has altered the ranking in which apps appeared, and in turn led to a drop in downloads of some apps in the range of 70-90% (Lanier 2018). Consequently, a dominant platform is a powerful attention broker: By granting a CP more or less prominent position in the ranking, they can directly influence the demand that this CP receives, and thus the revenue that it makes.

Platforms usually employ either an organic ranking algorithm (based on observable ranking factors) or a sponsored ranking algorithm (based on bids), or a combination thereof, to determine the most prominent position in the results list. For example, before 2015, the search results lists, both in the App Store and the Play Store would be purely based on an organic ranking. However, in 2015 Google (Siliski 2015), and in 2016 Apple (Vincent 2016) have introduced a sponsored ranking in their respective app stores, allowing CPs to place bids for a more prominent placement, shown above the organic results (Villanueva 2016). Google Search introduced a sponsored ranking (AdWords) already in 2000, and Youtube in 2008 (Zafra 2008). The role of sponsored rankings has become increasingly important in recent years, as organic search results were moved further and further down the results lists. Especially on mobile devices with limited screen size, often the first page of the search results list is completely occupied by sponsored search results. This effect is further exacerbated in the context of the growing popularity of voice assistants and voice search, where it is just the top position in the ranking that counts.

Next to the ranking, and in the absence of a headline price, consumers’ demand for free content depends on the quality of the content. By ‘quality’ we mean all objective characteristics of the

¹ Indeed, developers are typically present on both platforms because there is little overlap in consumers between the two dominant app stores. In a recent antitrust case against Google, the European Commission has found that the Android and iOS mobile operating systems operate in different relevant markets (European Commission 2018). Similarly, in the US, the supreme court has found in the case *Apple vs. Pepper* (Docket No. 17-204) that there is little competition between platforms and that Apple’s App Store is a monopoly (Goldman and Fung 2019)

content that are valued by most consumers, as opposed to horizontal characteristics or individual tastes, which users may regularly disagree on. For example, everything else being equal, an app with higher quality will be used more, and thus, content of higher quality creates more opportunities to create revenues through advertisements or in-app purchases; but, of course, content of higher quality is also costlier to produce. In equilibrium CPs invest in content quality only until the expected marginal increase in revenue from the quality improvement equals the marginal cost of the quality improvement. Platforms have a significant impact on this trade-off, because they can affect CPs' revenues through two channels. First, as highlighted above, platforms have an indirect impact on CPs' revenues through their ranking algorithm. Second, platforms have a direct impact on CPs' revenues because they demand a revenue share. For example, both Apple and Google demand a 30% revenue share from in-app purchases of any app in their respective app store (Bergen 2016). Similarly, YouTube demands a share of 32-49% of the revenues generated by the ads played in videos (Google n.d.), because every ad is displayed through Google's own ad network AdSense. AdSense is also the dominant ad network for display and search advertising on independent websites,² and Google likewise demands a 32-49% revenue share here.

In this paper, our main research question is to study how the increasing importance of sponsored search on dominant online platforms affects the quality of free content offered by independent content providers when the platform demands a revenue share. Understanding the trade-offs faced by the platform and content providers, as well as their welfare consequences is important, both from a business and from a public policy perspective. Platforms make significant profits in these content markets. For example, it was estimated that Apple made USD 11.5 billion in annual revenue in its app store in 2017 (Jones 2018). Moreover, revenues are likely to increase further, as app store purchases have increased by roughly 40% annually in the years from 2013 to 2016 (Ovide 2017), currently amounting to USD 38.5 billion (Nelson 2018). Realizing the importance of platform markets, policymakers around the world scrutinize dominant platforms' ranking algorithms and their impact on content, considering to impose stricter regulations (Crémer et al. 2019, Furman et al. 2019, Stigler Committee on Digital Platforms 2019). Moreover, following a wave of net neutrality regulations around the world, which prohibit Internet service providers to offer pay-for-priority schemes to content providers (Krämer et al. 2013), some observers have noted that, likewise, dominant online platforms should be scrutinized for offering pay-for-prominence schemes in the form of sponsored rankings (Krämer and Schnurr 2018), or should be banned from self-preferencing their own content on the platform (Crémer et al. 2019, p.66), and that more research is needed to investigate the role of 'neutrality' regulation also in this context (Easley et al. 2018).

²Recent research suggests that Google's ad serving and tracking technology runs on about 75% of the top one million websites (Englehardt and Narayanan 2016). Moreover, in the European Union, Google has been fined 1.49 billion euros in a high profile antitrust case, because it leveraged its dominant position in search to online advertising (European Commission 2019a).

To this end, we consider a game-theoretic model with a dominant (monopolistic) platform and two CPs that offer free content and compete for users' attention on the platform. Everything else being equal, the top listed CP receives more demand, and the CP of higher quality receives more demand. However, CPs differ in their efficiency to produce content quality. The platform commits to a revenue share and either employs an organic ranking (based purely on content quality) or a sponsored ranking (based purely on bids). We also consider a hybrid sponsored ranking algorithm, which adjust bids by a 'quality factor', but we can show that our analyses of a pure organic ranking and a pure sponsored ranking provide important bounds for the spectrum of feasible outcomes that may emerge under any hybrid sponsored ranking.

We show that a platform's ranking algorithm can have a significant impact on content quality through four different channels: (1) First, under an organic ranking, CPs have a higher incentive to invest in content quality than under a sponsored ranking, everything else being equal, because they need to win the quality-based competition in order to attain the top position. This is denoted as the *competition effect*. It is strongest when CPs are of similar efficiency. (2) Second, an organic ranking depends crucially on the platform's ability to assess content quality accurately. If the platform faces uncertainty in determining content quality it may erroneously list the lower quality content in the top position. This leaves the higher quality content with a relatively lower demand and therefore leads to a reduction in quality at the high end. In reverse, it gives the lower quality content relatively more demand, which leads to an increase in quality at the low end, everything else being equal. This is what we denote as the *uncertainty effect*. (3) Third, under a sponsored ranking the platform generally demands a lower revenue share than under an organic ranking, because it additionally receives revenues from the position auction. As mentioned above, a lower revenue share provides CPs with a higher incentive to invest in content quality. This gives rise to a *revenue-share effect*, which is stronger under a sponsored ranking. (4) In hybrid sponsored rankings with quality factor, the payment of the CP that wins the position auction is the lower, the higher its quality. This provides an additional incentive to invest in quality, denoted as the *payment-reduction effect*, that is not present under either a pure organic or a pure sponsored ranking.

We highlight that the overall effect of a sponsored ranking on content quality depends on the relative strengths of these four effects. Depending on parameters, either ranking can lead to higher content quality, higher welfare and higher profits for the platform, and we characterize the conditions for this to occur. For example, an organic ranking can lead to higher content qualities of both CPs only if CPs are similar in efficiency and the platform's ability to accurately assess content quality is high. In all other cases, a sponsored ranking induces a higher content quality for at least one of the two CPs. In many cases, the more efficient CP may actually prefer a sponsored ranking,

because this allows it to attain the top position with more certainty and with a lower revenue share, despite the fact that it has to additionally invest in bids.³

We also investigate the case where the platform is vertically integrated with one of the CPs, and always promotes its own CP in the ranking, known as ‘self-preferencing’. Generally, due to the self-bias in the ranking, the CPs do not compete for the first position anymore, neither in quality, nor in bids. This results in a lower competition effect, and a lower revenue-share effect for the non-integrated CP than under an organic ranking or sponsored ranking. However, vertical integration yields a boost in the revenue-share effect for the integrated CP, which can raise its quality significantly. Self-preferencing will therefore reduce the content quality of the non-integrated CP, but vertical integration is likely to significantly rise the content quality of the integrated CP. Overall, this can result in a higher consumer surplus and industry profit. This casts a somewhat different light on recent cases of self-preferencing by large platforms, such as Google, which was fined by the European Commission because it favored its own comparison shopping site on Google Search (European Commission 2017), and Apple, which is being accused of favoring its own apps in the App Store (Mickle 2019, Nicas and Collins 2019). Our results therefore also have important implications for the ongoing policy debate on the appropriate regulation of dominant online platforms.

2. Related literature

Our model relates to several streams of the literature. First, there exists a large body of literature spanning over computer science, information systems, operations management, marketing, and economics that considers keyword auctions and sponsored rankings. This stream of the literature usually focuses on the effect of various auction mechanisms and rankings on consumer search, the advertisers’ bidding behavior or the auctioneer’s revenue. For example, Abhishek et al. (2019) and Long et al. (2019) show that under a sponsored ranking the platform and consumers can overcome some of the information asymmetry and infer the quality of the content providers better than under an organic ranking, and that this yields better search results. We make a related finding and show that sponsored rankings increase the incentives to invest of high quality content providers, because under an organic ranking they face greater uncertainty to be ranked erroneously in a lower position. This is what we denote as the *uncertainty effect*.

However, in the extant literature on sponsored rankings, quality or other vertical ‘performance’ measures (e.g., click-through-rates or ‘popularity’) are typically assumed to be exogenously given (e.g., Feng et al. 2007, Weber and Zheng 2007, Armstrong et al. 2009, Athey and Ellison 2011, Berman and Katona 2013, Hagiu and Jullien 2014, Long et al. 2019, Abhishek et al. 2019), and

³This is consistent with remarks by some app developers in response to the app stores’ introduction of a sponsored ranking (see e.g., D’Onfro (2015) and Balakrishnan (2016)).

therefore this literature cannot study how the auction design or ranking algorithm affects investment incentives, although some papers consider how the ranking affects pricing (e.g., Chen and He 2011, Dellarocas 2012).⁴ We are only aware of two notable exceptions: Chen et al. (2010) consider a first-price auction with quality factor in which bidders are first exogenously assigned either a low or a high quality level, but the low quality bidders are allowed to invest in a performance upgrade. They identify a unique parameter threshold at which the low quality bidder wishes to upgrade and to compete more aggressively with the high quality bidder. This relates to what we denote as the *competition effect*, and for which we also identify a unique parameter threshold. Second, Katona and Zhu (2018) consider a sponsored ranking with quality factor and also show, in independent work, that there exists both a pure strategy equilibrium and a mixed strategy equilibrium depending on parameters. In the pure strategy equilibrium, the ‘low cost’ bidder has higher incentives to invest in quality than under a pure sponsored ranking, and this incentive is strongest for intermediate weighting factors (β in our case). This corresponds to our *payment-reduction effect*. However, these papers do not consider situations in which the platform demands a revenue share from the bidders, or in which the platform has uncertainty to determine quality, and thus they do not identify the same comprehensive set of investment incentive effects that arise under the various ranking algorithms, including biased rankings.

Second, our model is related to the literature on biased intermediation, e.g., in the context of booking or search platforms. The paper that is closest to ours is by De Cornière and Taylor (2019), who study in a general setting how biased intermediation affects firms’ investment incentives—predominantly in the context of a vertically integrated product or service. Our set up is related to their *congruent payoffs* model, in which they find that the vertically integrated firm, which is recommended by the platform, has a stronger incentive to invest (e.g., in quality) than the non-recommended firm. This is similar to our finding in the case of vertical integration and a biased ranking. However, unlike De Cornière and Taylor (2019), we assume that content providers (CPs) are heterogeneous with respect to their costs to invest in quality, and we focus on the case where none of the CPs is vertically integrated with the platform. Moreover, in contrast to their model, we consider the platform as a strategic player who can influence the CPs’ investment incentives also through its strategic choice of the revenue share.

Third, our investigation is related to the policy debate on the regulation of dominant online platforms, and particularly the notion of ‘platform neutrality’, which is currently very topical among EU policy makers (see, e.g., CNNum 2014, European Commission 2017, Krämer and Schnurr 2018,

⁴ In some models, content providers can invest in non-quality related aspects, such as search engine optimization, to improve their ranking position with some probability in an organic ranking (see, e.g., Berman and Katona (2013)). This relates to the parameter γ in our model.

Dillet 2018, ARCEP 2018). In this context, ‘neutrality’ means any unreasonable discrimination among content providers, that is, discrimination that is not based on objective (quality) standards, but purely on commercial considerations. From a scientific view, Easley et al. (2018) note that the same issues that have been debated in the context on the net neutrality debate at the infrastructure level, can also be debated at the software level, i.e., in the context of dominant software platforms. Easley et al. highlight app stores as an example for such dominant platforms and argue that the pay-for-prominence scheme that is implemented by a sponsored ranking is akin to a pay-for-priority scheme that has been condemned under net neutrality regulation. Evidently, selling the top ranked position to the CP that bids the most to be in this position, or simply granting the top position to one’s own content, is a form of non-quality-based discrimination.

3. Baseline scenario: organic ranking

3.1. The model

Consider a market with two CPs, each of which provides exactly one unit of content. The CPs differ in their efficiency to produce content quality. We assume that the cost function of the less efficient CP, denoted by L , to produce content of quality q_L , is $C_L = q_L^2$; and that the cost function of the more efficient CP, denoted by H , to produce content of quality q_H , is $C_H = k q_H^2$ with $k \in (0, 1]$. Thus, we obtain that CP L incurs (weakly) higher costs for producing a certain quality level than CP H . Thereby we make the standard assumption of a convex cost function, i.e., producing quality improvements becomes more difficult for higher levels of $q_i, i \in \{L, H\}$, formally $\frac{\partial C_i}{\partial q_i} > 0$ and $\frac{\partial^2 C_i}{\partial q_i^2} > 0$. Both CPs offer their content through a monopolistic platform (i.e., a platform with significant market power) that ranks the respective content according to some ranking algorithm. We assume that all users have to use the platform in order to retrieve content. Users may use the content of either none, one or both CPs, i.e., we allow for both single- or multihoming. However, we abstract from explicitly modelling users’ search process for content, and instead build our model solely on two assumptions about the total demand for content. These assumptions are corroborated by the literature on ordered search as well as the empirical literature on rankings, which we have discussed above:

ASSUMPTION 1. Holding content quality constant, for each content provider total demand increases with a higher ranking position.

More formally, let the total demand of CP i be $d_1(q_i)$ in the top position, and $d_2(q_i)$ in the second highest position. Then we assume that $d_1(q_i) > d_2(q_i) > 0$ for any given quality level q_i .

ASSUMPTION 2. Holding the ranking position constant, for each content provider total demand increases with a higher content quality. For any given quality level the marginal increase in demand is weakly higher in a higher ranking position.

More formally this means that $\frac{\partial d_1}{\partial q_i} > 0$ and $\frac{\partial d_2}{\partial q_i} > 0$ with $\frac{\partial d_1}{\partial q_i} \geq \frac{\partial d_2}{\partial q_i}$ for any given level of q_i .

In order to fix ideas and to make the exposition of the model more tractable, we focus on the family of linear demand functions, where $d_1(q_i)$ and $d_2(q_i)$ increase linearly in quality, i.e. $d_1 = n_1 q_i$ and $d_2 = n_2 q_i$, with $n_1 > n_2 > 0$ (see, e.g., Chen et al. 2010, Liu et al. 2010, for a similar assumption). Our insights are not limited to this class of demand functions, however, but depend more generally only on Assumptions 1 and 2. Therefore, in the Online Appendix, we provide proofs for our main propositions also for more general demand functions $d_1(q_i)$ and $d_2(q_i)$ that are either more concave or more convex than in the case of linear demand.

We further assume that the platform may not be able to observe content quality perfectly, and may err when applying an organic ranking algorithm. As a result the content of lower quality may sometimes be ranked at the top under an organic ranking. We denote the platform's probability to identify the CP with the highest content quality correctly by $\gamma \in [1/2, 1]$.⁵ Thus, under an organic ranking the content of CP i receives a demand of

$$D_{i,1} = \gamma d_1(q_i) + (1 - \gamma) d_2(q_i) \quad (1)$$

if it provides a higher quality than its competitor, and

$$D_{i,2} = (1 - \gamma) d_1(q_i) + \gamma d_2(q_i) \quad (2)$$

if it provides a lower quality than its competitor. In order to break ties, we assume that the platform promotes CP H with probability γ when both content qualities are equal.

As motivated in the introduction, we assume that content is free to download for users, but that CPs receive a gross revenue per usage of their content, e.g., through advertisements or in-app purchases. Therefore, gross revenues of the CPs can be written as

$$R_i = F_{-i}(q_i) D_{i,1} + (1 - F_{-i}(q_i)) D_{i,2}, \quad (3)$$

where $F_{-i}(q_i) = \text{prob}(q_{-i} < q_i)$, i.e., F_{-i} is the cumulative distribution function (CDF) of the competitor's quality choice, which we denote by q_{-i} . Furthermore, the platform charges a share s of each CP's gross revenue. Thus, we can write a CP's profit function as

$$\Pi_i = (1 - s)R_i - C_i, \quad (4)$$

and the platform's profit function as

$$\Gamma = s \sum_i R_i. \quad (5)$$

⁵ In practice the platform may use a variety of cues, such as consumers' reviews, views or likes, to infer content quality in a dynamic learning process. We abstract from this learning process and instead model all possible states of that process by γ , ranging from $\gamma = 1/2$ (where the platform has no cues and picks a CP at random) to $\gamma = 1$ (where the platform can observe quality perfectly and always identifies the CP with the highest content quality correctly).

The timing of the game is then as follows: *Stage 1*: The platform announces its revenue share s . *Stage 2*: The CPs H and L set their qualities q_L and q_H , and the platform ranks the CP with the highest quality first with probability γ . *Stage 3*: Content is demanded depending on its ranking position and quality, and revenues are realized.

3.2. Equilibrium derivation

We solve for the subgame perfect equilibrium of the game by backwards induction, beginning in Stage 3. We only sketch the proof here and relegate the detailed proof to Appendix A.

Stage 3: Given the CPs' quality choices in Stage 2 and the revenue share in Stage 1, the CPs' revenues and profits are readily given by (3) and (4), respectively.

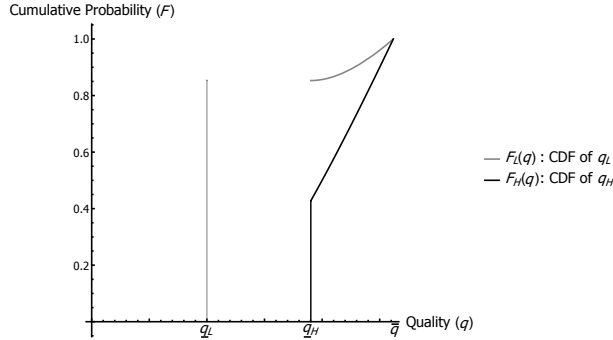
Stage 2: The analysis of Stage 2 entails two cases. First, if the CPs' efficiency is relatively similar, both compete in qualities for the top position. In this case there does not exist a pure strategy equilibrium in qualities, but a unique mixed strategy equilibrium. Second, if CPs' efficiency is very different, then CPs effectively do not compete for the first position anymore and there exists a unique pure strategy equilibrium in qualities. This yields the following lemma.

LEMMA 1 (Equilibrium Quality Choice under Organic Ranking). *Under an organic ranking, there exists a unique pure strategy equilibrium when $k \leq \tilde{k}$, and a unique mixed strategy equilibrium when $k > \tilde{k}$, where $\tilde{k} = \frac{N_1}{N_1 + \sqrt{N_1^2 - N_2^2}}$, with $N_1 = \gamma n_1 + (1 - \gamma)n_2$ and $N_2 = (1 - \gamma)n_1 + \gamma n_2$. In the pure strategy equilibrium, CPs L and H choose quality levels of $q_L = \frac{(1-s)N_2}{2}$ and $q_H = \frac{(1-s)N_1}{2k}$, respectively. In the mixed strategy equilibrium, CPs L and H choose their respective minimum quality \underline{q}_L and \underline{q}_H with positive probability $F_L(\underline{q}_H)$ and $F_H(\underline{q}_H)$, and a quality level between \underline{q}_H and \bar{q} according to their CDF $F_L(q_L)$ and $F_H(q_H)$. The equilibrium values are as follows:*

$$\begin{aligned}
 F_L &= \frac{(1-s)\bar{q}_L N_1 - k\bar{q}_L^2 - ((1-s)q_L N_2 - kq_L^2)}{(1-s)(N_1 - N_2)q_L}, & F_H &= \frac{(q_H - \underline{q}_L)^2}{(1-s)(N_1 - N_2)q_H}, \\
 \underline{q}_L &= \frac{(1-s)N_2}{2}, & \underline{q}_H &= \sqrt{\frac{(1-s)\bar{q}_L N_1 - k\bar{q}_L^2}{k}} \\
 \bar{q} = \bar{q}_L = \bar{q}_H &= \frac{(1-s)(N_1 + \sqrt{N_1^2 - N_2^2})}{2}.
 \end{aligned}$$

Proof: See Appendix A. \square

In order to exemplify Lemma 1, Figure 1 shows the CDFs, $F_L(q)$ and $F_H(q)$, with their respective positive mass points, \underline{q}_L and \underline{q}_H , in the mixed strategy equilibrium ($k > \tilde{k}$). It can be seen that the quality choices of CP H first-order stochastically dominate those of CP L . In Appendix B we show that this is a general result, and holds not only for linear demand functions, but also for more convex or more concave demand functions.

Figure 1 CDF of q_L and q_H in the mixed strategy equilibrium under an organic ranking.


Note. Numerical example derived for $n_1 = 6$, $n_2 = 4$, $s = 0.5$, $\gamma = 1$, $k = 0.75$.

The mixed equilibrium outcome reflects that an organic ranking creates a relatively dynamic environment, where a CP would like to constantly adapt its quality level, depending on the rival CP's quality. In practice, such a mixed equilibrium may be implemented, e.g., through frequent updates of apps, which indeed is very common in app stores and has been found to be positively correlated with app quality (see, e.g., McIlroy et al. 2016). In the following, we therefore consider the expected qualities under the mixed strategy equilibrium as the relevant proxy for the long-run average content qualities of CP L and CP H , which we can then compare to the quality levels set under a pure strategy equilibrium.

Stage 1: In the first stage the platform selects the revenue share, s , which it commands from the CPs. Thereby, it anticipates how its choice of s will affect the CPs' subsequent quality choice. Since the platform is a monopolist, s is determined by maximization of the platform's profit and results to

$$s = \arg \max_s \Gamma = 1/2. \quad (6)$$

for the case of linear demand functions. Note that in the case of more general demand functions, the equilibrium revenue share will not be exactly at $s = 1/2$ and depend more generally on the shape of the demand function (see Appendix A).

4. Sponsored ranking

4.1. The model

We now assume that the first position in the platform's ranking is awarded through a second-price auction. In reality, with n sponsored slots to be awarded, a generalized second-price auction (Edelman et al. 2007), also known as a position auction (Varian 2007), is used. However, with $n = 1$ sponsored slots, as we have here, the position auction coincides with a standard second-price auction. That is, CPs submit their bids, b_H and b_L , and the CP that has submitted the highest bid is awarded the top position at the price of the second-highest bid. Assume i is the CP that has

submitted the highest bid, and $-i$ the other CP, i.e., $b_i > b_{-i}$. Hence, in the second-price auction the high bidder has to pay $r_i = b_{-i}$ and receives the top position, whereas the low bidder pays $r_{-i} = 0$ and receives the second position. Moreover, in contrast to the organic ranking in Section 3, there exists no uncertainty in the determination of the first position, because payments are always perfectly observable by the platform. Consequently, $D_{i,1} = d_1(q_i)$ and $D_{i,2} = d_2(q_i)$, such that under a sponsored ranking algorithm a CP's profit is

$$\Pi_i = \begin{cases} \Pi_{i,1} = (1-s)d_1(q_i) - C_i - b_{-i} & \text{if } b_i > b_{-i}, \\ \Pi_{i,2} = (1-s)d_2(q_i) - C_i & \text{if } b_i < b_{-i}. \end{cases} \quad (7)$$

If both CPs have submitted the same bid, we use the same tie breaking rule as under an organic ranking and consider CP H as the highest bidder.

The timing is the same as before, but with an additional stage in which the auction takes place:⁶ *Stage 1*: The platform announces the revenue share s . *Stage 2*: The CPs H and L simultaneously choose their qualities q_L and q_H . *Stage 3*: The CPs simultaneously submit their bids b_L and b_H , and the platform ranks the CP with the highest bid first. *Stage 4*: Content is demanded depending on its ranking position and quality, and revenues are realized.

4.2. Equilibrium derivation

We again use backwards induction to solve for the subgame perfect equilibrium. The detailed proof for more general demand functions is relegated to Appendix C.

Stage 4: Stage 4 corresponds to Stage 3 in Section 3.2.

Stage 3: In this stage the CPs submit bids given their quality choice in Stage 2. In a second-price auction it is a weakly dominant strategy to bid one's true valuation (see, e.g., Krishna 2010, Proposition 2.1). A CP's value for the first position is determined by the additional profit that can be attained by being in this position. That is, in equilibrium $b_i = \Pi_{i,1} - \Pi_{i,2} = (1-s)(d_1(q_i) - d_2(q_i))$.

Stage 2: Consequently, the CP that has a larger incremental profit from being in the first position will always win the auction in Stage 3. As investment costs in quality are sunk at this stage, the CP that has chosen the higher quality level in Stage 2 will therefore always win the auction in Stage 3. Appendix C shows that CP H can choose its profit maximizing quality level for all feasible values of k without risking that CP L might choose a higher quality. Therefore, in equilibrium H chooses its quality such that it maximizes its profit in the first position and L chooses a quality that maximizes its profit in the second position. This immediately yields the following result.

LEMMA 2 (Equilibrium Quality Choice under Sponsored Ranking). *Under a sponsored ranking, there exists a unique pure strategy equilibrium, where CPs L and H choose quality levels of $q_L = \frac{(1-s)n_2}{2}$, and $q_H = \frac{(1-s)n_1}{2k}$, respectively.*

⁶ We note that we would obtain the same results if bids and qualities were chosen by CPs simultaneously in the same stage.

Proof: See Appendix C. \square

Stage 1: Finally, in the first stage, the platform determines the revenue share, s , so that $s = \arg \max_s \Gamma = \arg \max_s s(d_1(q_H) + d_2(q_L)) + b_L$. This yields

$$s = \frac{n_1^2 + k(3n_2^2 - 2n_1n_2)}{2n_1^2 + k(4n_2^2 - 2n_1n_2)} \quad (8)$$

for the case of linear demand functions (see Appendix C for the case of more general demand functions).

5. The effect of rankings on quality investment incentives

We now investigate the impact of a sponsored ranking in comparison to an organic ranking on the CP's incentives to invest in the quality of their content. To this end, it is instructive to compare the expected quality levels as detailed in Lemmas 1 and 2. In the following, we will highlight three effects that explain why and how content qualities may differ between the two ranking algorithms. Where necessary, we will differentiate between the variables that relate to the organic ranking algorithm and the variables that relate to the sponsored ranking algorithm by superscript o and superscript s , respectively.

5.1. The competition effect

First, note that the qualities chosen under sponsored ranking are identical to those chosen in the pure strategy equilibrium under organic ranking (i.e., for $k \leq \tilde{k}$), provided the platform chooses the same revenue share, $s = s^o = s^s$, and $\gamma = 1$. That is, whenever CP L chooses *not* to engage in the quality-based competition for the first position, both rankings will yield the same content qualities, everything else being equal. However, when CPs do compete for the first position under an organic ranking (i.e., $k > \tilde{k}$), then content qualities will be higher under the organic ranking than under the sponsored ranking, everything else being equal (i.e., $s = s^o = s^s$, and $\gamma = 1$). To see this, notice first that for $k > \tilde{k}$ it holds that $\underline{q}_L \leq \underline{q}_H < \bar{q}$ (see Appendix D). Second, recall that under an organic ranking CP L chooses the minimum quality level, \underline{q}_L , with probability $F_L(\underline{q}_H)$, but with probability $1 - F_L(\underline{q}_H)$ it chooses a higher quality between \underline{q}_H and \bar{q} . Third, see from Lemmas 1 and 2 that \underline{q}_L coincides with the quality level that CP L would choose under a sponsored ranking, everything else being equal. Consequently, when CPs are not too different in terms of efficiency (i.e., $k > \tilde{k}$), the expected quality of L is higher under an organic ranking, everything else being equal. In similar vein, also the expected quality of CP H is higher under an organic ranking for $k > \tilde{k}$, everything else being equal. This result also holds for more general demand functions and is shown formally in Appendix E.

The intuition for this finding is as follows. While CPs compete for the first position under both ranking algorithms, they do so with different strategic instruments. Under a sponsored ranking

CPs compete with their bids, but not with their content qualities. In comparison, under an organic ranking CPs compete directly in content quality, which gives them an additional incentive to set higher quality levels. Therefore, we denote this quality investment incentive as the *competition effect*. In reverse, this means that when the quality-based competition breaks down, then the competition effect vanishes and CPs will choose the same quality levels under both algorithms.

Figure 2a illustrates the competition effect by holding the revenue share ($s = s^o = s^s = 1/2$) constant across ranking algorithms, but varying the strength of the competition effect through H 's efficiency advantage k . The figure shows that the expected quality levels coincide between both ranking regimes for $k \leq \tilde{k}$. This is because CP L 's efficiency disadvantage is so large that it does not compete anymore for the first position. However, when CP L competes for the first position, i.e., for $k > \tilde{k}$, each CP will choose a higher quality under an organic ranking. In fact, notice that when CPs are of very similar efficiency ($k \rightarrow 1$), then both compete very fiercely for the top position under an organic ranking, such that the expected quality of CP H and CP L approach each other, and *both* CPs offer a higher quality than CP H under a sponsored ranking.

This result also holds for more general demand functions, because F_H first order stochastically dominates F_L , but for $k = 1$ it follows that $F_L = F_H$ (see Appendix B). Hence, the CDFs of the quality choices of CPs L and H , as well as their respective expected profits, converge as $k \rightarrow 1$. Eventually, both CPs choose exactly the same mixed strategy when they are of equal efficiency. Under a sponsored ranking there is no competition in the quality-setting stage, and therefore, the CPs' qualities and profits do not converge to the same value for $k \rightarrow 1$. This latter insights immediately yields the following corollary.

REMARK 1. When CPs are almost equally efficient ($k \rightarrow 1$), a sponsored ranking amplifies the small competitive disadvantage of CP L and yields very different expected quality levels for CPs L and H , respectively. Under an organic ranking this is not the case, such that, when CPs are equally efficient, they will also set the same qualities.

The insights from this subsection are summarized in the following proposition, which also holds for more general demand functions (see Appendix E).

PROPOSITION 1 (Competition Effect). *The more similar CPs become in their efficiency to produce content quality (k increases, $k > \tilde{k}$), the higher is a CP's quality investment incentive under an organic ranking. This 'competition effect' does not exist under a sponsored ranking.*

Proof: See Appendix E. \square

5.2. The uncertainty effect

Next, see that the CP that expects a larger demand is more likely to invest in quality. This is due to the economies of scale that stem from the fixed investment costs in quality. This scale effect comes

in two flavors. First, even if a CP would not compete for the first position, it faces a minimum demand D_2 . The larger D_2 , the higher are the investments in quality, everything else being equal. This is what constitutes a *minimum-scale* effect that ensures a certain minimum content quality. Second, when CPs compete for the top position, they do so, because this secures them additional demand of $D_1 - D_2$. This is what constitutes the *additional-scale effect*. For CP L , who wins the second position more often, the minimum-scale effect is relatively more important, whereas for CP H , the additional-scale effect is relatively more important.

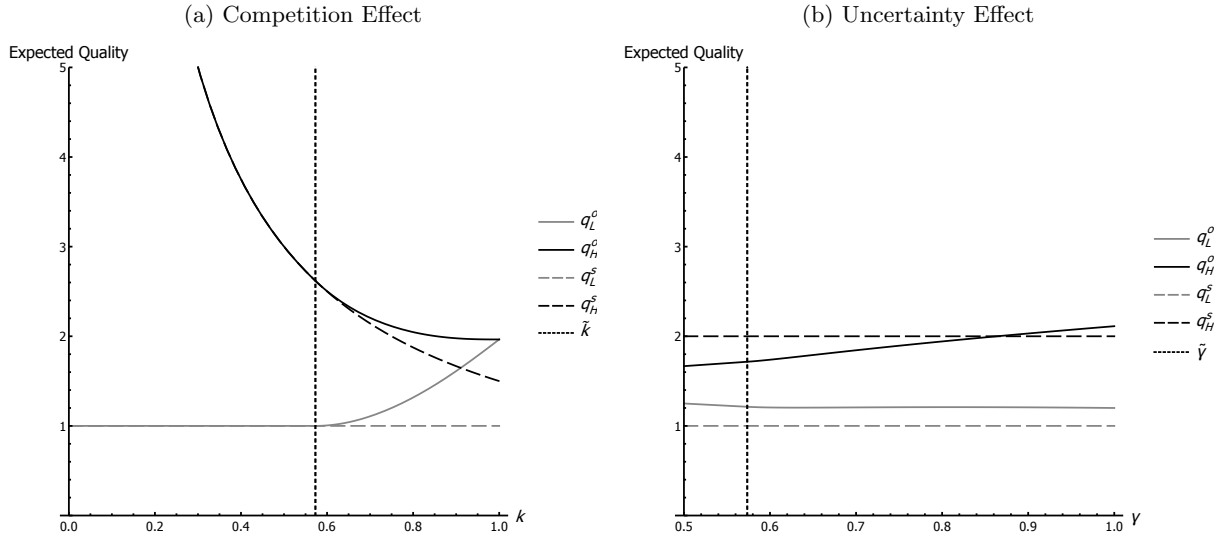
Against this backdrop, we can now introduce the uncertainty effect. An important difference between the two ranking algorithms is that under a sponsored ranking the platform has no uncertainty in identifying the CP with the highest bid, whereas under an organic ranking the platform can only identify the CP with the highest content quality with probability $\gamma \in [1/2, 1]$. Therefore, under a sponsored ranking the CP with the highest bid enjoys an additional scale of $D_1^s - D_2^s = d_1 - d_2$, but under an organic ranking the CP with the highest quality enjoys only a additional scale of $D_1^o - D_2^o = (d_1 - d_2)(2\gamma - 1) \leq D_1^s - D_2^s$. In reverse, under a sponsored ranking the CP with the lower bid enjoys a minimum scale of $D_2^s = d_2$, but under an organic ranking the CP with the lower quality enjoys a minimum scale of $D_2^o = (1 - \gamma)d_1 + \gamma d_2 \geq D_2^s$. Consequently, an increase in the uncertainty to attain the top ranking position despite having the highest quality (i.e., γ decreases), induces CP H , who offers a higher content quality more often, to *reduce* its quality (due to a decrease in the additional scale effect), and CP L , who offers a lower content quality more often, to *increase* its content quality (due to an increase in the minimum scale effect). In other words, the uncertainty effect is the reason why it is relatively less attractive to invest in quality under an organic ranking compared to a sponsored ranking for CP H , whereas the reverse is true for CP L .

PROPOSITION 2 (Uncertainty Effect). *The higher the platform's uncertainty in determining content quality (γ decreases), the lower (higher) is the quality investment incentive of the more (less) efficient CP under an organic ranking. This 'uncertainty effect' does not exist under a sponsored ranking.*

Proof: See Appendix F. \square

The uncertainty effect is exemplified in Figure 2b, which holds the revenue share ($s = s^o = s^s = 1/2$) and H 's efficiency advantage ($k = 0.75$) constant across ranking algorithms, but varies the strength of the uncertainty effect through γ . It can be noticed that, as the uncertainty increases ($\gamma \rightarrow \frac{1}{2}$), the expected qualities of H and L become more similar under the organic ranking, whereas γ has, of course, no effect under the sponsored ranking.

REMARK 2. Even if the platform can observe content quality perfectly ($\gamma = 1$), under an organic ranking CP H does not obtain the top ranking position with certainty, but only with probability

Figure 2 Competition Effect and Uncertainty Effect


Note. (a) Competition effect: Due to the CPs' competition in content quality under an organic ranking, expected quality levels are higher for $k > \tilde{k}$ under an organic ranking, everything else being equal. (b) Uncertainty effect: An increase in the platform's probability to rank the CPs according to their quality, γ , tends to increase (decrease) CP H 's (L 's) expected quality, but only under an organic ranking. Note that for values of $\gamma > \tilde{\gamma}$ the competition effect has a countervailing effect on L 's quality choice. Figures derived for $n_1 = 6$, $n_2 = 4$, $s = 0.5$, (a) $\gamma = 1$, (b) $k = 0.75$.

$F_L(q_H) < 1$ when $k > \tilde{k}$, and thus less often than under a sponsored ranking. Thus, even for $\gamma = 1$, in expectation CP H enjoys a lower additional-scale effect, and CP L a higher minimum-scale effect under an organic ranking than under a sponsored ranking.

5.3. The revenue-share effect

Finally, notice from Lemmas 1 and 2 that the quality levels under both ranking regimes are immediately affected by the revenue share s . The larger s , the lower the expected quality level. Next, we show that the platform will set a lower revenue share under a sponsored ranking. This means that, everything else being equal, due to this *revenue-share effect*, CPs have a stronger incentive to invest in quality under a sponsored ranking. Importantly, we prove in Appendix G that the platform's revenue share is generally lower under a sponsored ranking. For the case of a linear demand function this can be immediately derived by comparing (6), which shows that the revenue share under the organic ranking is $s^o = 1/2$, with (8), which yields a revenue share of $s^s < 1/2$ under the sponsored ranking for all feasible parameter values. This is because under a sponsored ranking the platform additionally considers the impact of s on its profits from the position auction. Whereas, under an organic ranking CPs will react to an increase in s only by lowering their quality, under a sponsored ranking, CPs will react to an increase in s by lowering their quality *and* their bid. This induces the platform to demand a lower revenue share under sponsored ranking.

PROPOSITION 3 (Revenue-Share Effect). *The lower a platform’s revenue share (s decreases), the higher is a CP’s quality investment incentive. This ‘revenue-share effect’ is stronger under a sponsored ranking ($s^s < s^o$).*

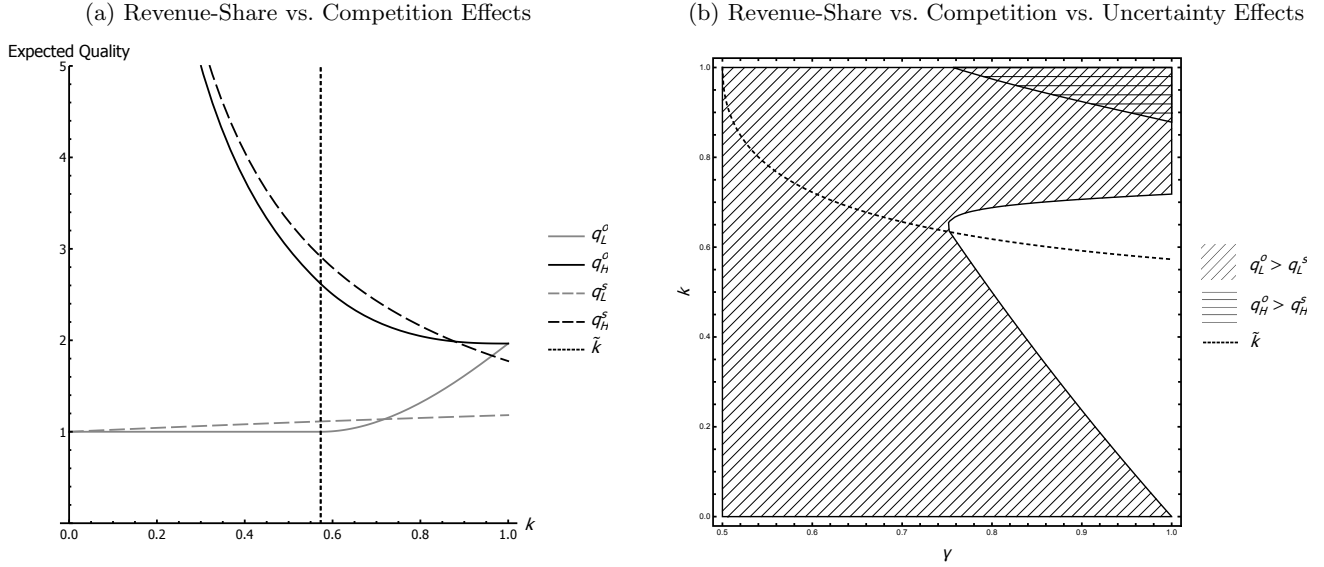
Proof: See above and Appendix G. \square

5.4. Interaction of quality investment effects

The overall impact of the ranking algorithm on content quality is governed by the interaction of the competition, uncertainty and revenue-share effects introduced above. While in the previous section, we proved the three effects analytically and showed that they act differently for the different ranking algorithms, we now investigate the interaction of these effects and proof by means of a numerical example that it will depend on parameters which ranking algorithm yields the highest expected content quality for one of the CPs, or even both CPs. In doing so, we also discuss and demonstrate that the complex trade-offs that content providers need to manage when determining their optimal content quality under a given ranking regime can be fully understood and explained in terms of the three quality investment effects highlighted in Propositions 1 to 3.

At first, let us consider the interaction of the competition and revenue-share effect, but without the uncertainty effect (i.e., $\gamma = 1$). Whereas the competition effect is stronger under an organic ranking, the revenue-share effect is stronger under a sponsored ranking. Thus, it will depend on the relative strength of these two effects which ranking algorithm will overall lead to higher expected qualities by each type of CP. Whereas in Figure 2a we have deliberately neglected the revenue-share effect by holding “everything else equal”, i.e., by fixing s across all ranking regimes, in Figure 3a we now apply the actual, regime specific equilibrium value of s at $\gamma = 1$. It can be seen that for low levels of k , higher qualities are expected under a sponsored ranking, because the competition effect is relatively weak (in fact zero for $k \leq \tilde{k}$) and thus the revenue-share effect dominates. On the contrary, if k is high, expected content qualities are higher under the organic ranking, because the competition effect now dominates the revenue-share effect.

Next, we consider how the revenue-share effect and the competition effect act in concert with the uncertainty effect. As we have noted above, varying γ directly affects the (additional and minimum) scale that CPs can expect in the first and second ranking position. However, γ also indirectly affects the strength of the competition effect and revenue-share effect. Figure 3b explores which ranking algorithm overall yields the highest content qualities for CPs H and L , respectively, under the assumption of the same demand functions as the previous figures, but for all possible combinations of k and γ . It can be seen that no ranking clearly dominates the other with respect to the CPs’ content quality. In the example, the organic ranking leads to higher qualities of *both* CPs only if i) both CPs are relatively similar in their efficiency to produce quality content (k is large, i.e., the

Figure 3 Comparison of Content Qualities between Organic Ranking and Sponsored Ranking


Note. (a) Expected quality levels when s is endogenous and $\gamma = 1$. CPs choose higher (expected) qualities under sponsored ranking when k is low due to the dominance of the revenue-share effect over the countervailing competition effect. In reverse, at high levels of k the competition effect can dominate the revenue-share effect for both CPs.

(b) Parameter regions of k and γ where an organic ranking yields a higher expected quality, $q_i^o > q_i^s$ than sponsored ranking. In the cross-hatched area, an organic ranking yields higher expected qualities for both CPs. In the remaining diagonally-shaded area, an organic ranking yields a higher expected quality only for CP L , but not for CP H . In the non-shaded area, an organic ranking yields a lower expected quality for both CPs. All figures derived for $n_1 = 6$, $n_2 = 4$.

competition effect is large), and ii) the platform has low uncertainty in determining content quality (γ is large). In all other cases, the sponsored ranking leads to higher content qualities for at least one CP. However, for a wide range of parameter values (except for low values of k combined with high values of γ), CP L will offer a higher quality under an organic ranking. This is the result of the uncertainty effect (which grants CP L a large minimum scale) and the competition effect, both of which are absent under a sponsored ranking. In reverse, when the competition effect is weak (k is low) and the uncertainty effect is weak (γ is high), then the revenue-share effect is the dominant effect. Since this effect is stronger under a sponsored ranking, in this parameter range both CPs will provide higher content qualities under a sponsored ranking.

Finally, it is important to note that the strength of each of the three effects will also be determined by the shape of the specific demand functions that are considered. In Appendix E to G we discuss this with respect to each of the three effects. In Appendix H we explore this in more detail with respect to the case of linear demand functions.

PROPOSITION 4 (No dominance of a single ranking algorithm). *No single ranking algorithm always yields the highest content qualities: Depending on the strengths of the competition,*

uncertainty and revenue-share effects, either the organic ranking or the sponsored ranking can yield the highest expected content quality for one of the CPs, or both CPs.

Proof: See Propositions 1 to 3, and Figure 3 for existence. \square

6. Hybrid sponsored ranking with quality factor

6.1. The model

In practice platforms often employ a hybrid sponsored ranking algorithm that determines the CPs' ranking positions according to a combination of quality and bid (e.g., in Google's or Apple's search ads). In this case CPs are assigned a ranking score, and the CP that has the highest ranking score is awarded the first ranking position, and has to pay the minimum amount that is necessary to achieve the same ranking score as the CP in the second position. We refer to this type of hybrid sponsored ranking algorithm as a sponsored ranking with quality factor, denoted by superscript sq . As we will show, this ranking combines the effects already identified under a pure organic and a pure sponsored ranking, but also yields an additional effect that is not present under either a pure sponsored or a pure organic ranking. Let the ranking score of CP i be given by $rs_i = q_i^\beta b_i^{1-\beta}$, where b_i is the bid submitted by CP i and $\beta \in [0, 1]$ is a weighting factor. See that $\beta = 0$ yields a pure sponsored ranking, identical to that studied in Section 4, and $\beta = 1$ yields a pure organic ranking, identical to that studied in Section 3. Let $rs_i > rs_{-i}$, then CP i has achieved the highest ranking score, is awarded the first position, and has to pay an amount r_i such that $q_i^\beta r_i^{1-\beta} = q_{-i}^\beta b_{-i}^{1-\beta}$. Rearranging yields the payment function

$$r_i = \left(\frac{q_{-i}}{q_i} \right)^{\frac{\beta}{1-\beta}} b_{-i}, \quad (9)$$

whereas CP $-i$ pays $r_{-i} = 0$ and achieves the second ranking position.⁷ The timing is the same as under the sponsored ranking without quality factor.

6.2. Equilibrium derivation and quality investment effects

The analysis of Stage 4 corresponds to Stage 3 in Section 3.2. In Stage 3, see that auction is again strategically equivalent to a second-price auction, because there is only one position to be auctioned off and qualities levels are sunk at this stage.⁸ Consequently, it is a weakly dominant strategy for each CP to bid the true incremental value of being in the first position, i.e., $b_i = (1-s)(d_1(q_i) - d_2(q_i))$. Notice that this holds irrespective of β . Thus, the CP with the higher quality will also submit the higher bid, and is consequently awarded the first position. In this context, see that for $\beta \rightarrow 1$, the payment of the firm with the highest quality, who wins the auction and is

⁷ Formally, the payment function (9) is only defined for $\beta < 1$, and in case $\beta = 1$ we set $r_i = r_{-i} = 0$.

⁸ Note that this generally not true if there is more than one slot to be auctioned off and the number of bidders exceeds the number of slots (Edelman et al. 2007).

awarded the first position converges to zero. This highlights again that for $\beta = 1$ we indeed obtain a pure organic ranking. In Stage 2, the CPs anticipate that whoever sets the higher quality also wins the auction. Because there is a direct correspondence between bids and quality, the platform has no uncertainty to determine content quality and the correct ranking order for any $\beta < 1$. Consequently, there is no uncertainty effect (i.e., $\gamma = 1$), just like under a pure sponsored ranking, and a CP's profit is

$$\Pi_i = \begin{cases} \Pi_{i,1} = (1-s) d_1(q_i) - C_i - r_i & \text{if } q_i > q_{-i}, \\ \Pi_{i,2} = (1-s) d_2(q_i) - C_i & \text{if } q_i < q_{-i}. \end{cases} \quad (10)$$

In general, the derivation of the equilibrium quality levels in Stage 2 is not tractable for arbitrary values of β . Thus, also the equilibrium revenue share in Stage 1 cannot be determined directly. However, based on our insights from pure organic rankings and pure sponsored rankings, we can nevertheless derive some important insights that allow us to understand the quality investment incentives under a hybrid sponsored ranking with quality factor, and how these relate to the investment incentives under a pure sponsored ranking or a pure organic ranking. In the following, we organize this discussion along the four quality investment incentives effects that are relevant under a hybrid sponsored ranking.

Payment-reduction effect: First, notice that the profit function (10) is similar to that under sponsored ranking (7), but with one important difference. See from (9) that a higher own quality *reduces* the payment that the winning bidder has to make. This provides the winning CP with an additional incentive to invest in quality that is not present under either a pure organic ranking or a pure sponsored ranking. We denote this as the *payment-reduction effect*, and highlight that it is maximized for some intermediate value of β . To see this, recall that $\beta = 0$ corresponds to a pure sponsored ranking, where the payment-reduction effect does not exist because the payment only depends on the quality of the other CP. Likewise, $\beta = 1$ corresponds to a pure organic ranking, where the payment-reduction effect does not exist because payments are zero anyway.

Uncertainty effect: Second, a decisive advantage of the sponsored ranking with quality factor over a pure organic ranking is that it removes the uncertainty effect, because the CPs' bids uniquely reveal their quality level. That is, CPs can expect to be ranked in the right order of their qualities for all values of $\beta < 1$. As highlighted by Proposition 2, this provides the highest incentives to invest in quality for CP H .

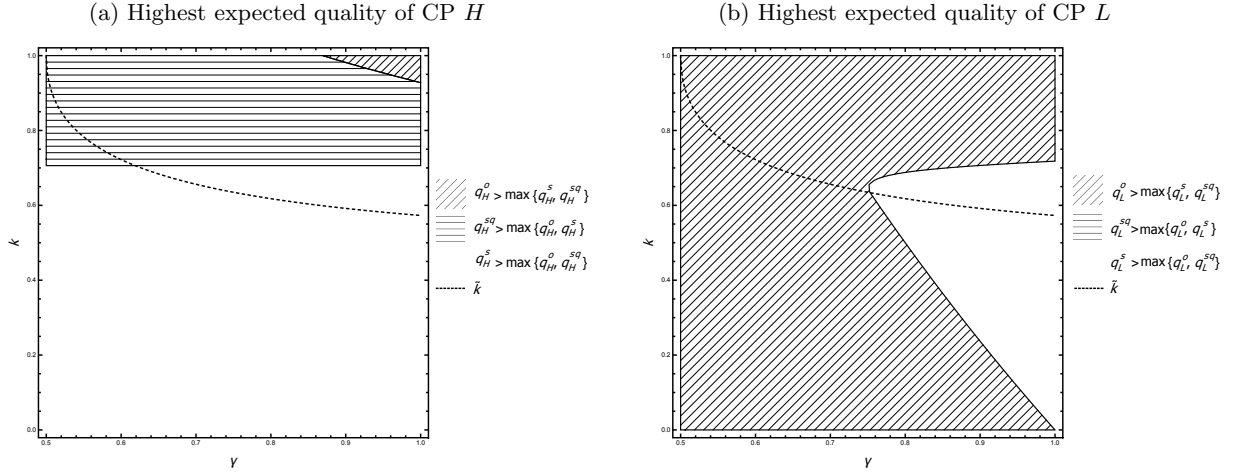
Competition effect: Third, a decisive advantage of the sponsored ranking with quality factor over a pure sponsored ranking is that, depending on β and k , it may also give rise to a competition effect, albeit to a lesser degree than under a pure organic ranking. Specifically, in Appendix I we show that for large enough k there exists a threshold $\tilde{\beta} < 1$ which is strictly smaller than $\beta \rightarrow 1$, above

which CP L wants to compete with CP H in qualities (and thus no pure strategy equilibrium, but only a mixed strategy equilibrium exists), and below which CP L prefers not to compete with CP H (and thus only a pure strategy equilibrium exists). For the competition effect to exist under a sponsored ranking with quality factor, $k > \tilde{k}$ is a necessary condition. However, in this ranking the actual threshold is higher than under a pure organic ranking (the more so, the lower β), because the first position is less attractive than under a pure organic ranking due to the additional payment that has to be made. The lower β , the higher is the payment under a sponsored ranking with quality factor, and the higher is the threshold for k . Indeed, if β is too low, the competition effect may not exist at all under a sponsored ranking with quality factor, even if $k = 1$. In particular, we can show numerically that this is the case for $\beta = 1/2$. This means that for $\beta \leq 1/2$ no competition effect (only a pure strategy equilibrium) exists, irrespective of parameters. Consequently, when the platform wants to stimulate investment incentives via the the competition effect, it must choose $\beta > 1/2$ in a sponsored ranking with quality factor, which means that it must place more emphasis on quality than on the bids.

Revenue-share effect: Finally, notice that under a sponsored ranking with quality factor the strength of the revenue-share effect must be in between that under a pure organic and under a pure sponsored ranking. This is because the revenue share, s increases as β increases, and bids converge to zero for $\beta \rightarrow 1$. For $\beta = 0$, $s < 1/2$ is the same as under a pure sponsored ranking, and for $\beta = 1$, $s = 1/2$ is the same as under a pure organic ranking.

In summary, our analysis highlights that a sponsored ranking with quality factor is not necessarily the ranking that leads to the highest investment incentives, although it combines the strengths of both a pure organic ranking (competition effect) and a pure sponsored ranking (no uncertainty effect), and provides an additional investment incentive not present under either pure ranking regime (payment-reduction effect). The reason is that the competition effect is lower than under a pure organic ranking, and the revenue-share effect is lower than under a pure sponsored ranking. The interaction of these effects has been overlooked in previous literature, or investment incentives were not taken into consideration at all, when designing a hybrid ranking algorithm. In fact, it is ambiguous which ranking regime ($\beta \in [0, 1]$) yields the highest expected quality levels and profits for the platform. In general, it will depend on parameters, because for $\beta = 0$ the revenue share effect is maximized, for $\beta = 1$ the competition effect is maximized, and for intermediate β the payment-reduction effect is maximized. If CPs are relatively different in their efficiency (low k), the competition effect is less important, and thus a low β is likely to be preferable—and vice versa if CPs are of similar efficiency.

This is also highlighted in Figure 4, which compares the quality level of CP H and CP L , respectively across the three ranking regimes (i) pure sponsored ranking ($\beta = 0$), (ii) sponsored

Figure 4 Comparison of Content Qualities between Organic Ranking, Sponsored Ranking, and Sponsored Ranking with Quality Factor


Note. Parameter regions of k and γ where a sponsored ranking with equal weight on quality and bids ($\beta = 1/2$) yields a higher expected quality than either a pure sponsored ranking ($\beta = 0$) or a pure organic ranking ($\beta = 1$) for CP H and CP L , respectively. While CP H 's quality level under the sponsored ranking with quality factor is higher than under any of the two other ranking regimes for intermediate values of k (horizontally-shaded area), CP L 's quality under a sponsored ranking with quality factor is strictly lower than under a pure sponsored ranking. All figures derived for $n_1 = 6$, $n_2 = 4$.

ranking with quality factor ($\beta = 1/2$) and (iii) pure organic ranking ($\beta = 1$). Clearly, for $\gamma = 1$, the highest expected quality of CP H is attained under a pure sponsored ranking for low values of k , and under a pure organic ranking for high values of k , whereas the hybrid sponsored ranking performs best for intermediate values of k . However, for CP L a sponsored ranking with quality factor never yields a higher expected quality than any of the other regimes. This is because for CP L all of the identified investment incentive effects are either absent or stronger in one of the other rankings: (1) for $\beta = 1/2$ no competition effect exists, (2) therefore also no payment-reduction effect exists for CP L , (3) the uncertainty effect is absent, and (4) the revenue-share effect is lower than under a sponsored ranking. Thus, q_L must be strictly lower than under a sponsored ranking for $\beta = 1/2$.

Taken together, this shows that the investment incentives and outcomes identified under the pure sponsored ranking and the pure organic ranking provide important and insightful bounds for the spectrum of feasible outcomes and investment incentives under any hybrid ranking algorithm. We will therefore use the pure ranking algorithms as benchmarks in the subsequent analysis.

PROPOSITION 5 (Payment-Reduction Effect). *Under a sponsored ranking with quality factor, the payment of the CP that wins the position auction is the lower, the higher its content quality.*

This ‘payment-reduction effect’ provides an additional quality investment incentive that does not exist under a pure sponsored ranking or a pure organic ranking.

Proposition 4 continues to hold correspondingly. No single ranking algorithm always yields the highest content qualities. Either the pure organic ranking, or the pure sponsored ranking, or the sponsored ranking with quality factor can yield the highest expected content quality for one of the CPs, or both CPs.

Proof: See above, and Figure 4 for existence. \square

7. Vertical integration and self-preferencing

So far we have considered *unbiased* ranking algorithms, i.e., algorithms that rank the CPs according to some predetermined and (possibly imperfectly) observable measure (either q or b , or a combination thereof). We now consider the special case where the platform is vertically integrated with one of the CPs and will always promote its own, vertically integrated content to the top position. Such ‘self-preferencing’ is at the heart of the ongoing policy debate on the regulation of online platforms and concerns have been raised, e.g., in the context of Google Shopping (European Commission 2017) or Apple Music (Mickle 2019, Nicas and Collins 2019). In the following, we consider the case where the integrated CP is of higher efficiency (CP H) than the competing CP, as well as the case where the integrated CP is of lower efficiency (CP L).

7.1. The model

The joint profit function of the vertically integrated platform and CP is

$$\Gamma = R_{\Gamma} - C_{\Gamma} + s R_{-\Gamma}, \quad (11)$$

where R_{Γ} is the revenue, and C_{Γ} is the cost function of the integrated CP/platform, and $s R_{-\Gamma}$ the revenue share that is collected from the non-integrated CP. Note that, due to vertical integration, R_{Γ} is independent of s . Also note that, from the perspective of the integrated CP, this is equivalent to the case where it is offered a revenue share of $s = 0$ by the platform.

The timing is now as follows: *Stage 1:* The platform announces its revenue share s for the non-integrated CP. *Stage 2:* The platform and the non-integrated CP simultaneously set their content qualities q_{Γ} and $q_{-\Gamma}$, and the platform ranks the integrated CP always in the first position. *Stage 3:* Content is demanded depending on its ranking position and quality, and revenues are realized.

7.2. Equilibrium derivation

We derive the subgame perfect equilibrium separately for the case where the integrated CP is H , i.e., of higher efficiency than the non-integrated rival, or CP L , i.e., of lower efficiency. Detailed proofs and results for more general demand functions are relegated to Appendix K.

Integrated CP is of higher efficiency: The analysis of Stage 3 is identical to the analysis of Stage 3 in Section 3.2. In Stage 2 the CPs know with certainty that CP H will always obtain the first, and that CP L will always achieve the second ranking position. Thus, there is no competition for the first position, neither in qualities nor in bids. Therefore, the optimal qualities are readily given by profit maximization and are, in principle, identical to those under a sponsored ranking (see Lemma 2). The only difference is that for the integrated CP (here H) the revenue share is set to zero. Thus optimal qualities are readily given by:

$$\begin{aligned} q_L^{v,H} &= \arg \max_{q_L} \Pi_L^{v,H} = \frac{(1-s)n_2}{2}, \\ q_\Gamma^{v,H} &= \arg \max_{q_\Gamma} \Pi_\Gamma^{v,H} = \frac{n_1}{2k}. \end{aligned}$$

In Stage 1, anticipating its impact on the quality q_L , the integrated platform maximizes (11) by setting the revenue share for the non-integrated CP to

$$s^{v,H} = \arg \max_s \Gamma = 1/2. \quad (12)$$

Consequently, for the non-integrated CP, the platform sets the same revenue-share as under an organic ranking. This is because, unlike under a sponsored ranking, the platform does not need to take into account the impact of the revenue-share on its revenue from bids.

Integrated CP is of lower efficiency: The same logic can be applied when the platform is integrated with CP L , i.e., the integrated CP is of lower efficiency. In this case the optimal content qualities are given by

$$\begin{aligned} q_H^{v,L} &= \arg \max_{q_H} \Pi_H^{v,L} = \frac{(1-s)n_2}{2k}, \\ q_\Gamma^{v,L} &= \arg \max_{q_\Gamma} \Pi_\Gamma^{v,L} = \frac{n_1}{2}. \end{aligned}$$

Here again, for the same reason as above, the optimal revenue-share for the non-integrated CP is given by

$$s^{v,L} = \arg \max_s \Gamma = 1/2. \quad (13)$$

LEMMA 3 (Equilibrium Quality Choice under Self-Preferencing). *With vertical integration and self-preferencing there exists a unique pure strategy equilibrium. If the integrated CP is of higher efficiency, then it chooses a quality level of $q_\Gamma^{v,H} = \frac{n_1}{2k}$, and the non-integrated CP a quality level of $q_L^{v,H} = \frac{(1-s)n_2}{2}$. If the integrated CP is of lower efficiency, then it chooses a quality level of $q_\Gamma^{v,L} = \frac{n_1}{2}$, and the non-integrated CP a quality level of $q_H^{v,L} = \frac{(1-s)n_2}{2k}$.*

Proof: See Appendix K. \square

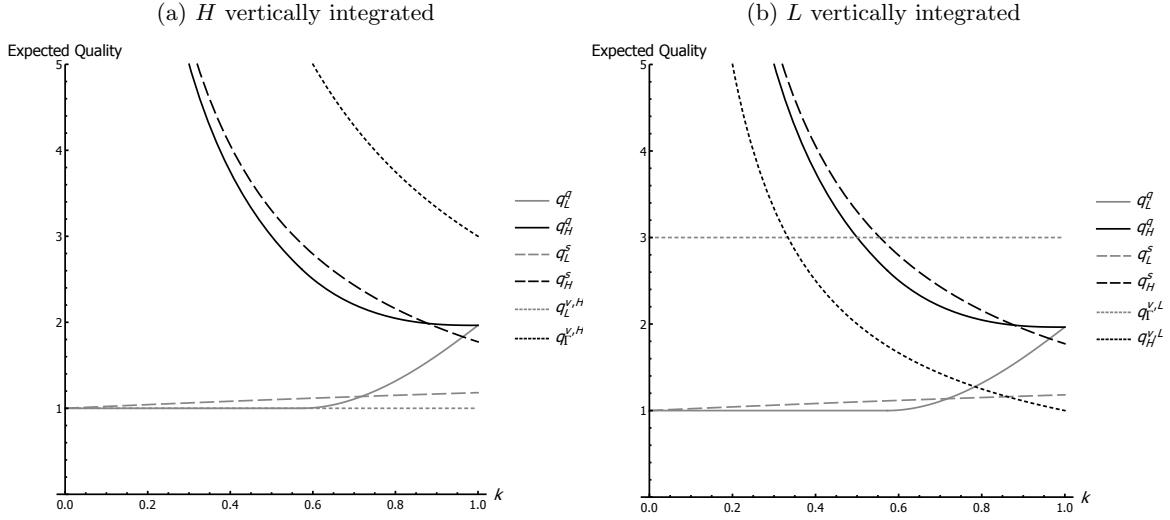
7.3. Quality investment incentives

In order to understand how vertical integration with self-preferencing impacts content quality, it is once again instructive to consider the four investment incentive effects. Thereby, one needs to differentiate between the impact of the fact that the ranking is biased (i.e., CPs cannot compete for the first ranking position through content qualities or bids), and the impact of the fact that one CP is vertically integrated (i.e., the revenue-share is asymmetric and set to zero for the integrated CP). As we will show next, a biased ranking generally has a negative impact on investment incentives, whereas vertical integration has a positive impact on the content quality of the integrated CP.

To see this, consider first what would happen if the ranking was ‘just’ biased (promoted either H or L unconditionally), but none of the CPs were vertically integrated with the platform. In this case the competition effect, the uncertainty effect and the payment-reduction effect are absent. Therefore, in case CP H is promoted, the content qualities are exactly the same as under a sponsored ranking, provided the revenue share s is the same. However, under a biased ranking, in which the CPs do not compete in bids, the platform would set a lower revenue share than under a sponsored ranking, thus yielding a lower revenue-share effect. Consequently, even when the platform is biased towards the more efficient CP, the content qualities of each CP would be strictly lower than the respective qualities under a sponsored ranking. This means that when comparing a biased ranking where H is always promoted with an organic ranking, then the same intuitions derived when comparing a sponsored ranking with an organic ranking carry over, but the parameter regions in which an organic ranking yields the higher expected quality will increase, for both H and L .

In case the ranking would promote the less efficient CP L , then outcomes would be even worse. This is because the less efficient CP is now enjoying the larger scale due to being in the first position, which creates a negative quality distortion at the top. One can show that in this case, the average quality set by CP H and CP L must be strictly lower than in a ranking that is biased towards the more efficient CP. For a full analysis of the case of a biased ranking without vertical integration we refer to Appendix J.

Now consider how vertical integration alters the above argumentation. In this case, the revenue-share effect acts differently on the integrated developer. Whereas the platform demands a positive revenue share from the non-integrated CP, it demands a revenue share of $s = 0$ from the integrated CP. This provides the integrated CP with strong incentives to invest in quality under vertical integration; indeed much stronger than under either an organic ranking or a sponsored ranking. Consequently, vertical integration with the platform will yield lower incentives to invest in quality for the non-integrated CP than under either organic ranking or a sponsored ranking; but vertical integration can yield a significantly higher incentive to invest in quality for the integrated CP than any other ranking regime.

Figure 5 Comparison of Content Qualities under Vertical Integration with Self-Preferencing


Note. Vertical integration results in a higher revenue-share effect for the integrated CP and may therefore lead to a higher content quality of the integrated CP than under either an organic ranking or a sponsored ranking. However, due to self-preferencing the content quality of the non-integrated CP is weakly lower than under either an organic ranking or a sponsored ranking. All figures derived for $n_1 = 6$, $n_2 = 4$, $\gamma = 1$.

This is exemplified in Figure 5. In case self-preferencing occurs with respect to CP H , Figure 5a shows that the boost in the revenue-share effect leads to a higher content quality of CP H than under an organic ranking or sponsored ranking. On the contrary, CP L will offer a weakly lower content quality than under an organic ranking or sponsored ranking due to the weakly lower competition and revenue-share effects. In case self-preferencing occurs with respect to CP L , Figure 5b shows that the boost in the revenue-share effect can compensate the weakly lower competition effect, so that CP L offers a higher content quality than under an organic ranking or sponsored ranking. In reverse, CP H experiences a weakly lower competition and revenue-share effect, and thus offers a lower content quality than under an organic ranking or sponsored ranking.⁹ Taken together, this yields the following proposition.

PROPOSITION 6 (Vertical Integration with Self-Preferencing). *Vertical integration with self-preferencing leads to (a) weakly lower competition and revenue-share effects for the non-integrated CP, and (b) a weakly lower competition effect, but a strictly higher revenue-share effect for the integrated CP, compared to either an organic ranking or a sponsored ranking. Hence, the expected content quality of the non-integrated CP is weakly lower, but the content quality of the integrated CP can be higher than under either an organic ranking or a sponsored ranking.*

Proof: See Appendix K and Figure 5 for existence. \square

⁹ In case self-preferencing occurs with respect to CP L and additional positive (negative) quality distortion occurs for CP L (CP H) due to the larger (smaller) expected scale compared to an unbiased ranking (see Appendix J).

8. Welfare analysis

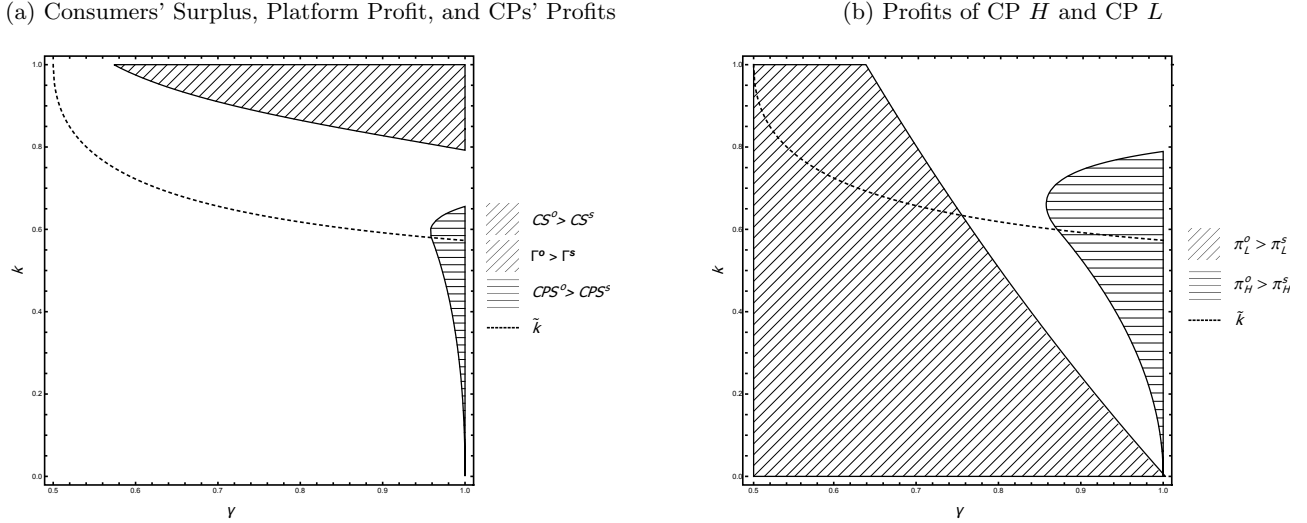
Finally, we compare the sponsored ranking and organic ranking algorithms also from a welfare perspective. We measure welfare by (i) consumer surplus (CS), (ii) platform profit, and (iii) CPs' profits/surplus (CPS). From an antitrust perspective, consumer surplus will be of paramount importance, especially in the US. However, it can be argued that CPs' profits are more important with respect to the long-run economic well-being of content producers. The platform's profit is important for determining which ranking regime would be preferred by the platform, and, when contrasted with consumer surplus, market failures can be identified.

Consumer surplus (CS): In our model consumer surplus is given by the sum of the (expected) demand of CPs H and L , i.e., $CS = D = D_L + D_H$. D_L and D_H are given in the Appendix by (27) and (28) for an organic ranking, and under a sponsored ranking we have simply $D_L = D_{L,2}$ and $D_H = D_{H,1}$. Figure 6a compares the consumer surplus between an organic ranking and sponsored ranking for the same demand function as in Figure 3b and over the whole feasible range of the parameters γ and k . It can be seen again that no ranking clearly dominates. In the example, organic ranking yields a higher consumer surplus only if the platform can assess content quality relatively well (γ is sufficiently large), and CPs differ not too much in efficiency (k is sufficiently high). In the light of our previous analysis, this means that an organic ranking is likely to lead to a higher consumer surplus only if the competition effect is relatively strong and the uncertainty effect is relatively weak.

CPs' surplus/profits (CPS): The CPs' surplus is given by $CPS = \Pi_H + \Pi_L$, where Π_i is given in Equations (4) and (7) for the organic and the sponsored ranking algorithm, respectively. Figure 6a also compares the CP surplus across the two algorithms. Again, no ranking dominates. In the example, the organic ranking yields a higher surplus for CPs only if the uncertainty effect is weak (γ is large), and the competition effect is weak, or even absent (k is small). This is because, when the competition effect is strong, CPs invest much more in quality, which is good for consumers, but bad for CPs' surplus, as it incurs large fixed costs. Consequently, CPs prefer an organic ranking over a sponsored ranking in about the same parameter region where an organic ranking results in lower content qualities for both CPs (cp. Figure 3b).

However, by just considering the combined CP surplus, we would mask the fact that the combined surplus is indeed mostly driven by CP H . Figure 6b highlights that CP L would prefer an organic ranking for a much larger range of parameters than CP H . This is mostly due to the uncertainty effect, which leaves CP L relatively better off when γ is low. In fact, notice that in the example CP L would never prefer an organic ranking over a sponsored ranking if there were no ranking uncertainty ($\gamma = 1$).

Figure 6 Comparison of Consumers' Surplus, Platform Profit, and CPs' Profits between Organic Ranking and Sponsored Ranking



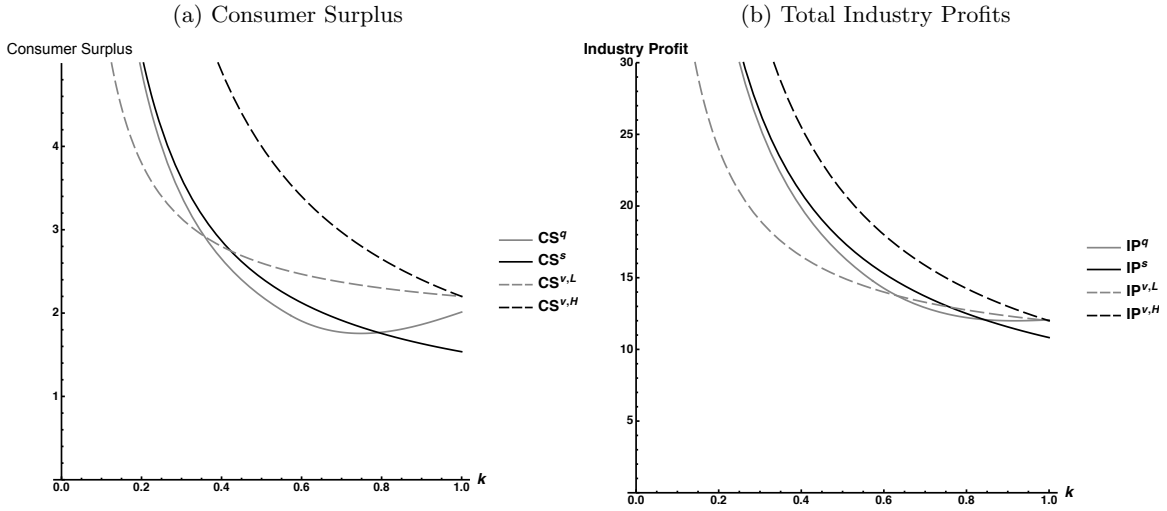
Note. (a) Comparison of consumers' surplus and CPs' surplus between the organic and the sponsored ranking over the full range of parameter values for k and γ . In the diagonally-shaded area, the organic ranking yields higher surplus for consumers. In exactly the same area, the organic ranking also yields higher profits for the platform. In the horizontally-shaded area, organic ranking yields a higher surplus for the CPs. In the non-shaded area, sponsored ranking yields a higher surplus for consumers (and the platform) and CPs. (b) Comparison of CP H 's and CP L 's preferred ranking algorithm for all parameter values of k and γ . In the diagonally-shaded area, organic ranking yields higher expected profit for L . The horizontally-shaded area yields higher expected profit for H under organic ranking. Figures derived for $n_1 = 6, n_2 = 4$.

Platform Profit (Γ): The platform surplus is readily given by Γ in (5). We can show numerically that, for all possible parameterizations of the linear demand functions, the platform makes higher revenues under a sponsored ranking exactly in the same parameter regions where the consumers receive a higher surplus under a sponsored ranking (see Figure 6a). This holds not only in the example but for linear demand functions more generally.¹⁰ Thus, if the platform were to choose a ranking algorithm, say in Stage 0 of an extended game, then it would always choose exactly the algorithm that would maximize consumer surplus. This is because the platform makes revenues based on the CPs' expected gross revenues, but not based on CPs' profits (i.e., not including the investment costs). CP's gross revenues are maximized when the total demand is maximized, i.e., when consumers' surplus is maximized. Therefore, when choosing the ranking algorithm, the platform's incentives and the consumers' incentives can be perfectly aligned.

Vertical Integration: Finally, it is insightful to consider the welfare outcomes also in the case of vertical integration with self-preferencing. From Proposition 6 we already know that vertical

¹⁰ However, this result does not necessarily hold for more general demand functions.

Figure 7 Consumer Surplus and Industry Profit under Vertical Integration with Self-Preferencing



Note. Vertical integration with self-preferencing can result (a) in a higher consumer surplus and (b) in a higher total industry surplus, both in case CP H or in case CP L is vertically integrated with the platform. All figures derived for $n_1 = 6, n_2 = 4, \gamma = 1$.

integration can yield higher content qualities for the integrated provider than any of the other unbiased ranking regimes. Figure 7 confirms that this also means that vertical integration can yield a higher consumer surplus and a higher total industry profit (i.e., the sum of platform profit and CP’s profit) than an unbiased ranking regime without vertical integration.¹¹ Vertical integration may even then lead to higher consumer surplus and industry profits than the unbiased rankings if the platform is integrated with the less efficient content provider.

PROPOSITION 7 (Welfare). *No single ranking algorithm always yields the highest social welfare: Depending on the strengths of the competition, uncertainty and revenue-share effects, either the organic ranking or the sponsored ranking can yield the highest consumer surplus, CPs’ surplus or platform profit.*

In case of a linear demand function the platform always prefers a sponsored ranking over an organic ranking exactly when this maximizes consumer surplus.

Vertical integration with self-preferencing can yield a higher consumer surplus and industry profit than an either the organic ranking or the sponsored ranking.

Proof: For the first part and third part of the proposition see Figure 6a and Figure 7 for existence, respectively. The second part is shown numerically for the full range of the parameters k and γ and arbitrary values of n_1 and n_2 with $n_1 > n_2$. \square

¹¹ Note that it is not possible to separate the platform’s profit from the CPs’ profits anymore with vertical integration.

9. Conclusions

9.1. Summary and theoretical contribution

We developed a game-theoretic model with a monopoly platform and two content providers in order to investigate how a sponsored ranking algorithm affects content quality and welfare in comparison to an organic ranking algorithm. We identified four distinct strategic effects – the competition effect, the uncertainty effect, the revenue-share effect, and the payment-reduction effect – which jointly determine content providers’ incentives to invest in content quality, and hence welfare. Intuitively one would expect that an organic ranking provides content providers with the highest incentives to invest in quality, precisely because content providers compete for the top ranking position in quality. Indeed this is what we denote as the *(1) competition effect*, and we can show that it becomes stronger the more similar content providers are with respect to their ability to produce content quality. However, we could also show that a sponsored ranking algorithm can yet lead to higher overall content quality and welfare. This is due to two additional effects.

(2) Revenue-share effect: Under a sponsored ranking, content providers react more strongly to a marginal increase in the platform’s revenue share for two reasons. First, whereas under an organic ranking a reduction in quality corresponds to a lower probability of being in the first ranking position, this is not the case under a sponsored ranking, where the ranking position does not directly depend on content quality. Therefore, under an organic ranking content providers cannot afford to decrease their quality as much as under a sponsored ranking in response to an increase in the revenue share. Second, under a sponsored ranking content providers can react to an increase in the revenue share through two strategic channels: through a reduction in quality, and through lower bids in the position auction. On the contrary, under an organic ranking, content providers can only react through a reduction in quality. Taken together, content providers are more sensitive to an increase in the platform’s revenue share under a sponsored ranking, which induces the platform to set a lower revenue share under this ranking regime. This in turn, provides content providers with a higher incentive to invest in quality under a sponsored ranking compared to an organic ranking.

(3) Uncertainty effect: Under an organic ranking content quality is also affected by the platform’s uncertainty in identifying content quality. The higher the uncertainty, the lower the incentives of the more efficient content provider to invest in quality, but the higher the incentives of the less efficient content provider to invest in quality. In reverse, a sponsored ranking does not suffer from such investment distortions due to CPs’ uncertainty because the winner of the position auction can be determined with certainty.

Furthermore, the same three effects can also explain the outcomes and investment incentives under a hybrid sponsored ranking, which considers both quality and bids. However, in hybrid sponsored rankings an additional investment incentive arises, because the CP who expects to win

the auction, has an extra incentive to to increase its quality in order to lower its payment in the auction. This is what we denote as the (4) *payment-reduction effect*. This does not mean, however, that overall investment incentives are strongest under a hybrid sponsored ranking, because we can show that the other investment incentive effects are less pronounced than under either a pure sponsored or a pure organic ranking.

Finally, we also considered the case where the platform is vertically integrated with one content provider and always promotes its own content. Vertical integration yields a boost in the revenue-share effect for the integrated provider, and thereby provides it with strong incentives to invest in content quality. However, because the ranking is biased towards the integrated provider, the non-integrated provider now has lower investment incentives than under an organic ranking or sponsored ranking. Nevertheless, vertical integration with self-preferencing can overall yield a higher average content quality and consumer welfare than unbiased ranking regimes without vertical integration.

9.2. Managerial implications

From a managerial perspective, our model highlights that a sponsored ranking may indeed be beneficial for content providers on a dominant platform, especially those who can produce content quality with high efficiency. In an environment where it is difficult for the platform to assess quality accurately, or where the platform applies an intransparent editorial policy, efficient content providers prefer to bid for a top position, which is a reliable way to signal their quality, rather than to rely on the platform's noisy (and potentially biased) quality assessment process. Moreover, a sponsored ranking shields content providers from engaging in a costly quality-based competition. Competition in quality is comparably costly, because marginal investments in 'quality' become increasingly expensive, whereas competition in 'money' entails constant marginal costs. Thus, a sponsored ranking can yield higher profits for content providers, because it prevents them from excessive investments in quality, above and beyond the quality level that they would offer absent any competition.

However, a sponsored ranking may also pick winners and losers. Those content providers that are less efficient and therefore have a competitive disadvantage are likely to be worse off under a sponsored ranking. In particular, we have shown that under a sponsored ranking even a very small difference in the providers' inefficiencies will, in contrast to an organic ranking, yield a distinct differentiation in the the providers' content quality. This is because under a sponsored ranking the more efficient content provider can secure the top position with certainty. On the contrary, under an organic ranking the more efficient provider can never be certain to be in the top position, even when the platform can assess content quality with certainty. This effect becomes even more pronounced when the platform cannot assess content quality accurately.

9.3. Policy implications

From a policy perspective, an important observation is that the platform's and consumers' preferred ranking algorithm can be perfectly aligned. This is because the platform prefers the algorithm that yields the higher weighted average of the expected qualities, which is a proxy for the gross revenues that content providers receive, and which therefore corresponds to a higher profit of the platform. Thus, our model provides an example of an online market where the exercise of market power by a dominant platform does not translate into a loss of consumer welfare. Yet, it is important to point out that under a total welfare standard market failures may still occur, because consumer welfare ignores the content providers' investment costs. This supposed shortcoming of the consumer welfare standard, which is adopted, e.g., in US antitrust enforcement, has recently led to a lively debate regarding the appropriate welfare standard for policymakers (see, e.g., Khan 2017). Our model informs this debate by formalizing the conditions under which the consumer welfare standard is not an appropriate signal for market power in digital markets.

Furthermore, in Europe policymakers have proposed to impose 'neutrality' regulation for dominant online platforms, like app stores (CNNum 2014, ARCEP 2018). Extending the legal framework of net neutrality, which currently only applies to Internet Service Providers, would possibly prohibit online platforms to offer pay-for-prominence regimes, such as a sponsored ranking (cf. Easley et al. 2018). Our results are a critical assessment in this regard. The application of such a rule would crucially depend on the market environment and its intended purpose. In general terms, we identify many instances in which a sponsored ranking regime would yield better results (in terms of quality and welfare) than an organic ranking regime. Thus, our findings do not support ex-ante regulation of 'platform neutrality'. If, however, the purpose of the rule is to ensure competition between content providers, especially in order to protect 'small' and less efficient content providers, then a 'neutrality' rule may be warranted.

A neutrality rule has also been proposed in an effort to prevent self-preferencing of the platform's own, vertically integrated content or service. For example, this is the proposed remedy of the European Commission following an anti-trust investigation of Google, who favored its own shopping comparison service in the search results (European Commission 2017). Similar remedies are envisaged in the context of app stores, because the ranking algorithm in Apple's App Store seems to be biased towards Apple's own apps (Mickle 2019, Nicas and Collins 2019). We show that vertical integration per se need not be to the detriment of consumers, because it can boost the quality of the integrated content or service by a surge in the revenue-share effect. However, self-preferencing creates a bias in the ranking that destroys the competition for the more prominent position. This results in lower content quality, especially for the non-integrated provider. Therefore, measures that prevent ranking biases, e.g., more transparency of the platforms on the factors that

determine a ranking position, can ensure fair and effective competition, which will improve market outcomes. In Europe, a first step in this direction has recently been pursued by the so-called Platform-to-Business Regulation (European Commission 2019b). Moreover, European lawmakers are currently considering additional ex-ante regulation for dominant platforms in the context of the Digital Services Act, which may include a complete ban of self-preferencing.

9.4. Limitations and future research

In closing, we would also like to mention some limitations of our model that may prompt future research. First, we have highlighted in the introduction that large platforms, such as app stores or search engines, are considered to have significant market power by policymakers both in the US and in the EU. Therefore we have modeled the strategic trade-offs faced by a monopolistic platform. Yet, a possible policy agenda may be to introduce more competition among platforms. Therefore, it could be insightful to see how competition between platforms affects content quality and a platform's choice of the ranking algorithm. While a complete analysis of competition is beyond the scope of this paper, two arguments can already be made here: First, in such an environment content providers will usually multihome (are present on both platforms), while consumers singlehome (choose one platform exclusively). For example, Bresnahan et al. (2014) show in the context of app stores that those apps which "satisfy the bulk of user demand" are available on both Apple's App Store and Google's Play Store. This gives rise to a situation in which platforms would only compete for the singlehoming users, but not for the content providers. In fact, once users have signed up with a given platform, and can only be reached through it, this creates a 'competitive bottleneck' (Armstrong and Wright 2007), which gives each platform monopoly power over the multihoming content providers. In the context of our model, this means that, given a certain user market share, each platform faces the same strategic trade-offs when setting the revenue share as if it were a monopolist, and therefore, our main results would not change. Second, in the context of online markets, it is reasonable to assume that multihoming content providers would offer the same content quality on each platform (e.g., because they can develop platform independent apps in Java). Further assume singlehoming consumers choose the platform from which they expect the higher surplus (average weighted content quality). Then platforms would compete for consumers on the basis of the ranking algorithm, and all platforms would prefer a sponsored ranking over an organic ranking for all values of $\gamma < 1$. This is because a sponsored ranking always promotes the content provider with the highest quality, whereas an organic ranking promotes the content provider with the highest quality only with probability $\gamma < 1$, so that the average weighted content quality is higher under a sponsored ranking.

In our analysis, we have taken the platform's ranking uncertainty γ and the content providers' cost difference k to be exogenous parameters. Future research could seek to endogenize these parameters, for example through modelling a consumer review system which would allow the platform to improve its ranking uncertainty, or by specifying a technological distance between a platform and the content provider that would motivate differences in the cost function. However, since we compare the different ranking algorithms for all possible combinations of these parameters, this would not yield new market outcomes, but rather limit the set of possible outcomes. In particular, recall that even if the platform can identify content quality perfectly in the long run, e.g., due to user feedback or downloads, which corresponds to $\gamma = 1$ in our model, then a sponsored ranking can still provide higher incentives to invest in quality than an organic ranking if the competition effect is relatively weak.

Moreover, we have considered content providers that can only invest in quality that increases consumers' utility. Thus, investments in quality are beneficial for both consumers and content providers. We would expect different outcomes in a setting where content providers have a strategic instruments that are able to reduce consumers' welfare, for example, prices or investments in a more intrusive advertising technology or other privacy diminishing measures. A sponsored ranking may increase the usage of such malicious content, because it would be possible to attain a high ranking position despite being of low quality for users. Clearly, this can be attenuated by a hybrid sponsored ranking with quality factor, but even more so by a pure organic ranking. Thus, if investments in malicious content is of a dominant concern, this would change our results and policy implications, and may support the application of a neutrality rule after all.

References

- Abhishek V, Jerath K, Sharma S (2019) Advertising on online marketplaces: Information asymmetry and the relevance of sponsored listings. Working Paper, Available at <http://dx.doi.org/10.2139/ssrn.3013468>.
- ARCEP (2018) Smartphones, tablets, voice assistants...: Devices, the weak link in achieving an open internet. ARCEP report, Retrieved from https://www.arcep.fr/uploads/tx_gspublication/rapport-terminaux-fev2018-ENG.pdf, accessed on 09/24/2018.
- Armstrong M, Vickers J, Zhou J (2009) Prominence and consumer search. *The RAND Journal of Economics* 40(2):209–233.
- Armstrong M, Wright J (2007) Two-sided markets, competitive bottlenecks and exclusive contracts. *Economic Theory* 32(2):353–380.
- Athey S, Ellison G (2011) Position auctions with consumer search. *The Quarterly Journal of Economics* 126(3):1213–1270.

- Balakrishnan A (2016) Why apple is bringing advertisements to its app store now. CNBC, Retrieved from <https://www.cnbc.com/2016/10/04/why-apple-is-bringing-advertisements-to-its-app-store-now.html>, accessed on 09/21/2018.
- Baye MR, Gatti JRJ, Kattuman P, Morgan J (2009) Clicks, discontinuities, and firm demand online. *Journal of Economics & Management Strategy* 18(4):935–975.
- Bergen M (2016) Google is offering app developers the same revenue-sharing terms apple just announced with one big advantage. Recode, Retrieved from <https://www.recode.net/2016/6/8/11889298/google-apple-subscription-app-revenue-share>, accessed on 08/13/2017.
- Berman R, Katona Z (2013) The role of search engine optimization in search marketing. *Marketing Science* 32(4):644–651.
- Bresnahan T, Orsini J, Yin PL (2014) Platform choice by mobile app developers. NBER Working Paper, Retrieved from <https://conference.nber.org/conferences/2014/EoDs14/multihoming%20BOY.pdf>, Accessed on 07/17/2019.
- Chen J, Feng J, Whinston AB (2010) Keyword auctions, unit-price contracts, and the role of commitment. *Production and Operations Management* 19(3):305–321.
- Chen Y, He C (2011) Paid placement: Advertising and search on the internet. *The Economic Journal* 121(556):F309–F328.
- CNNuM (2014) Platform neutrality: Building an open and sustainable digital environment. Conseil National du Numérique, Retrieved from https://ec.europa.eu/futurium/en/system/files/ged/platformneutrality_va.pdf, accessed on 09/21/2018.
- Crémer J, de Montjoye YA, Schweitzer H (2019) Competition policy for the digital era. Retrieved from <https://publications.europa.eu/en/publication-detail/-/publication/21dc175c-7b76-11e9-9f05-01aa75ed71a1/language-en>, accessed on 10/02/2019.
- De Cornière A, Taylor G (2019) A model of biased intermediation. *The RAND Journal of Economics* 50(4):854–882.
- Dellarocas C (2012) Double marginalization in performance-based advertising: Implications and solutions. *Management Science* 58(6):1178–1195.
- Dillet R (2018) Frances telecom regulator thinks net neutrality should also apply to devices. TechCrunch, Retrieved from <https://techcrunch.com/2018/02/15/frances-telecom-regulator-thinks-net-neutrality-should-also-apply-to-devices/>, accessed on 09/21/2018.
- D’Onfro J (2015) Developer: Paid search in google play store is ‘like a shot to the gut’. Business Insider, Retrieved from <https://www.businessinsider.com/developer-reaction-to-google-play-store-search-ads-2015-3>, accessed on 09/21/2018.

- Easley RF, Guo H, Krämer J (2018) Research Commentary - From Net Neutrality to Data Neutrality: A Techno-Economic Framework and Research Agenda. *Information Systems Research* 29:253–272.
- Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review* 97(1):242–259.
- Englehardt S, Narayanan A (2016) Online tracking: A 1-million-site measurement and analysis. *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 1388–1401 (ACM).
- European Commission (2017) Antitrust: Commission fines google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service. EC Press Release, Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/IP_17_1784, accessed on 08/08/2019.
- European Commission (2018) Antitrust: Commission fines google €4.34 billion for illegal practices regarding android mobile devices to strengthen dominance of google’s search engine. European Commission Press Release, Retrieved from http://europa.eu/rapid/press-release_IP-18-4581_en.htm, accessed on 08/17/2018.
- European Commission (2019a) Antitrust: Commission fines google €1.49 billion for abusive practices in online advertising. EC Press Release, Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1770, accessed on 08/08/2019.
- European Commission (2019b) Platform-to-business trading practices. Retrieved from <https://ec.europa.eu/digital-single-market/en/business-business-trading-practices>, accessed on 07/17/2019.
- Feng J, Bhargava HK, Pennock DM (2007) Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms. *INFORMS Journal on Computing* 19(1):137–148.
- Furman J, Coyle D, Fletcher A, McAuley D, Marsden P (2019) Unlocking digital competition: Report of the digital competition expert panel. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf, accessed on 10/02/2019.
- Glick M, Richards G, Sapozhnikov M, Seabright P (2014) How does ranking affect user choice in online search? *Review of Industrial Organization* 45(2):99–119.
- Goldman D, Fung B (2019) Iphone owners can sue apple for monopolizing app store, supreme court rules. CNN, Retrieved from <https://edition.cnn.com/2019/05/13/tech/apple-app-store-supreme-court/index.html>, accessed on 22/04/2020.
- Google (n.d.) Adsense revenue share. Google AdSense Help, Retrieved from <https://support.google.com/adsense/answer/180195?hl=en>, accessed on 08/07/2019.
- Hagiu A, Jullien B (2014) Search diversion and platform competition. *International Journal of Industrial Organization* 33:48–60.

- Jones C (2018) Apple's app store generated over \$11 billion in revenue for the company last year. *Forbes*, Retrieved from <https://www.forbes.com/sites/chuckjones/2018/01/06/apples-app-store-generated-over-11-billion-in-revenue-for-the-company-last-year>, accessed on 09/19/2018.
- Katona Z, Zhu Y (2018) Quality score that makes you invest. Working Paper, Available at <http://dx.doi.org/10.2139/ssrn.2954707>.
- Khan LM (2017) Amazon's antitrust paradox. *Yale Law Journal* 126:710–805.
- Krämer J, Schnurr D (2018) Is there a need for platform neutrality regulation in the EU? *Telecommunications Policy* 42(7):514–529.
- Krämer J, Wiewiorra L, Weinhardt C (2013) Net neutrality: A progress report. *Telecommunications Policy* 37(9):794–813.
- Krishna V (2010) *Auction Theory* (Academic Press), 2nd edition.
- Lanier L (2018) Google changes discovery algorithm on google play, leading to panic from some devs. *Variety*, Retrieved from <https://variety.com/2018/gaming/news/developers-sudden-drop-downloads-on-google-play-store-1202861850/>, accessed on 08/17/2018.
- Liu D, Chen J, Whinston AB (2010) Ex ante information and the design of keyword auctions. *Information Systems Research* 21(1):133–153.
- Long F, Jerath K, Sarvary M (2019) Leveraging information from sponsored advertising at online retail marketplaces. Kenan Institute of Private Enterprise Research Paper No. 20-03, Available at <http://dx.doi.org/10.2139/ssrn.3516104>.
- McIlroy S, Ali N, Hassan AE (2016) Fresh apps: an empirical study of frequently-updated mobile apps in the google play store. *Empirical Software Engineering* 21(3):1346–1370.
- Mickle T (2019) Apple dominates app store search results, thwarting competitors. *The Wall Street Journal*, Retrieved from <https://www.wsj.com/articles/apple-dominates-app-store-search-results-thwarting-competitors-11563897221>, accessed on 09/09/2019.
- Nelson R (2018) Global app revenue grew 35% in 2017 to nearly \$60 billion. *Sensor Tower*, Retrieved from <https://sensortower.com/blog/app-revenue-and-downloads-2017>, accessed on 10/23/2018.
- Nicas J, Collins K (2019) How apples apps topped rivals in the app store it controls. *The New York Times*, Retrieved from <https://www.nytimes.com/interactive/2019/09/09/technology/apple-app-store-competition.html>, accessed on 13/09/2019.
- Ovide S (2017) Apple spruces up the app store. *Bloomberg*, Retrieved from <https://www.bloomberg.com/gadfly/articles/2017-06-06/apple-is-making-critical-updates-to-its-app-store>, accessed on 08/13/2017.

- Siliski M (2015) A new way to promote your app on google play. Android Developers Blog, Retrieved from <https://android-developers.googleblog.com/2015/02/a-new-way-to-promote-your-app-on-google.html>, accessed on 08/13/2017.
- Statista (2018) Distribution of free and paid apps in the apple app store and google play as of 1st quarter 2018. Retrieved from <https://www.statista.com/statistics/263797/number-of-applications-for-mobile-phones/>, accessed on 09/19/2018.
- Stigler Committee on Digital Platforms (2019) Stigler committee on digital platforms, final report september 2019. Retrieved from <https://research.chicagobooth.edu/stigler/media/news/committee-on-digital-platforms-final-report>, accessed on 10/02/2019.
- Ursu RM (2018) The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science* 37(4):530–552.
- Varian HR (2007) Position auctions. *International Journal of Industrial Organization* 25(6):1163–1178.
- Villanueva T (2016) Apple to introduce search ads to its app store. Search Engine Watch, Retrieved from <https://searchenginewatch.com/2016/06/13/apple-to-introduce-search-ads-to-its-app-store/>, accessed on 09/21/2018.
- Vincent J (2016) Apple starts showing ads in app store search results. The Verge, Retrieved from <https://www.theverge.com/2016/10/6/13184346/apple-app-store-ads>, accessed on 08/13/2017.
- Weber TA, Zheng Z (2007) A model of search intermediaries and paid referrals. *Information Systems Research* 18(4):414–436.
- Zafra A (2008) Google brings sponsored videos on youtube search. Search Engine Journal, Retrieved from <https://www.searchenginejournal.com/google-brings-sponsored-videos-on-youtube-search/8010/>, accessed on 10/10/2019.

Online appendix

Appendix A: Equilibrium derivation under an organic ranking with more general demand functions

Stage 3: Given the CPs' quality choices in Stage 2 and the revenue share in Stage 1, the CPs' revenues and profits are readily given by equations (3) and (4), respectively.

Stage 2: In order to derive the equilibrium it is necessary to determine the reasonable quality range for each CP first. To this end, note that L can choose not to compete in qualities, i.e., to serve only users accessing the content in the second position. L 's profit in this situation is

$$\Pi_L = (1-s)D_{L,2} - q_L^2. \quad (14)$$

Maximization with respect to q_L results to

$$q_L = \frac{(1-s)}{2} \frac{\partial D_{L,2}}{\partial q_L}. \quad (15)$$

By inserting equation (15) into (14) we obtain an outside profit of

$$\tilde{\Pi}_L = (1-s)\underline{D}_{L,2} - \underline{q}_L^2. \quad (16)$$

Suppose there is a $\bar{q}_L > q_H$, resulting in a profit of $\hat{\Pi}$ for which L obtains the top position with probability γ . \bar{q}_L is a reasonable choice if $\hat{\Pi}_L \stackrel{!}{\geq} \tilde{\Pi}$ holds. We can rearrange this to determine the maximum reasonable quality of CP L as

$$\bar{q}_L = \sqrt{(1-s)\bar{D}_{L,1} - \tilde{\Pi}_L}. \quad (17)$$

This quality level is sufficient to outbid CP L . Choosing $q_H = \bar{q}_L$ provides CP H with an outside profit of

$$\tilde{\Pi}_H = (1-s)\bar{D}_{L,1} - \bar{q}_L^2. \quad (18)$$

Note that H chooses $q_H \geq \bar{q}_L$ if $\bar{q}_L \leq \arg \max_{q_H} \Pi_H$. Thus, there is a corner case where quality competition does not occur that is characterized by $\bar{q}_L \stackrel{!}{=} \arg \max_{q_H} \Pi_H$. Solving for k yields

$$\tilde{k} = \frac{(1-s)}{2\bar{q}_L} \frac{\partial D_{H,1}^m}{\partial q_H^m}. \quad (19)$$

Consequently CP H 's maximum quality level is:

$$\bar{q}_H = \begin{cases} \arg \max_{q_H} \Pi_H = \frac{(1-s)}{2k} \frac{\partial D_{H,1}}{\partial q_H} & \text{for } k \leq \tilde{k} \\ \bar{q}_L & \text{for } k > \tilde{k} \end{cases} \quad (20)$$

This means that two types of equilibria can arise depending on k . First, a pure strategy equilibrium occurs if $k \leq \tilde{k}$. In this case H 's maximization problem results to

$$q_H = \arg \max_{q_H} \Pi_H = \frac{(1-s)}{2k} \frac{\partial D_{H,1}}{\partial q_H}, \quad (21)$$

and L 's maximization problem yields

$$q_L = \arg \max_{q_L} \Pi_L = \frac{(1-s)}{2} \frac{\partial D_{L,2}}{\partial q_L}. \quad (22)$$

This is an unique equilibrium since choosing $q_L > q_H$ yields lower profits for L due to $q_H \geq \bar{q}_L$.

Second, a mixed strategy equilibrium can arise for $k > \tilde{k}$. Suppose CP H chooses $q_H \in (\underline{q}_H, \overline{q}_H]$ with a positive probability p . If choosing the quality level q_H were part of L 's quality strategy, then there would be a positive probability for a tie. L 's profit in case of a tie is $(1-s)D_{L,2} - C_L$. However, for L setting a marginally higher quality would give larger profits. This is because a slightly higher quality yields a marginal increase of costs for L , but a significant increase of expected revenue, since the probability to get the first position and to receive a demand of $D_{L,1}$ increases by p . Consequently choosing the quality level q_H cannot be part of the equilibrium strategy of L if q_H is chosen by H with probability p . Consider further that if q_H is not played by L , a marginal lower quality level yields the same probability for H to get the first position and to receive a demand of $D_{H,1}$. Thus, choosing q_H minus an increment is strictly better than choosing q_H with probability p as long as the decrease in costs for quality is higher than the decrease in expected revenue. Choosing the lower bound \underline{q}_H with probability one is also not an equilibrium, since L would outbid H by choosing \underline{q}_H plus an increment. Consequently, the equilibrium quality strategy of H consists of a mass point at its lower bound, and an atomless distribution of q_H between \underline{q}_H and \overline{q}_H .

The equilibrium strategy of L looks similar. We already mentioned that L will never choose a quality that is played by H with probability p . Thus, it will never play \underline{q}_H as long as \underline{q}_H is a mass point in H 's equilibrium strategy. It will further choose no quality level between \underline{q}_H and \underline{q}_L , because the probability to receive $D_{L,1}$ remains zero in this range; but the profit from receiving $D_{L,2}$ is maximized at \underline{q}_L . Further, L will never choose a $q_L \in (\underline{q}_H, \overline{q}_L]$ with a positive probability p . To see this, suppose L chooses a quality q_L with probability p . Then H will never choose a q_H which is equal to q_L minus an increment, because the probability to receive $D_{H,1}$ is much higher by choosing $q_H = q_L$. But, if H never chooses q_L minus an increment, then choosing q_L minus an increment instead of q_L yields the same probability for L to receive $D_{L,1}$. Consequently, an equilibrium strategy of L can only have a mass point at \underline{q}_L . This mass point is always part of L 's equilibrium strategy. If L never chooses \underline{q}_L then H always loses the quality based competition by choosing \underline{q}_H . Note that in such a situation H gets $\tilde{\Pi}_H$ if it chooses \overline{q}_L , and $(1-s)D_{H,2} - k\underline{q}_H^2$ if it chooses \underline{q}_H . Note further that for $k=1$ both profits are equal, because the equality of both profits determines the maximum quality level that L chooses. For every lower k the revenue side of both profits remains equal, but the costs of choosing \overline{q}_L decrease by a factor \overline{q}_L^2 , and the costs of choosing \underline{q}_H decrease by a factor \underline{q}_H^2 . That is, for every $k < 1$ choosing \overline{q}_L yields a higher profit. Thus, H will never choose a quality level which is lower than the minimal quality level of L . Consequently, any strategy where L does not choose \underline{q}_L cannot be part of an equilibrium strategy. Finally, choosing \underline{q}_L with probability one cannot constitute an equilibrium strategy either, because the best response of H is to choose \underline{q}_H with probability one, which, however, also cannot be an equilibrium strategy, as already mentioned.

Thus, the equilibrium strategies of both CPs consist of a mass point at the respective lower bound, and an atomless distribution between \underline{q}_H and \overline{q}_L . Consequently, for every quality choice of L , the condition $\Pi_L \stackrel{!}{=} \tilde{\Pi}_L$ holds, and for every choice of H , the condition $\Pi_H \stackrel{!}{=} \tilde{\Pi}_H$ holds. Solving for CDFs yields:

$$F_H = \frac{\tilde{\Pi}_L - ((1-s)D_{H,2} - q_H^2)}{(1-s)(D_{H,1} - D_{H,2})}, \quad (23)$$

$$F_L = \frac{\tilde{\Pi}_H - ((1-s)D_{L,2} - kq_L^2)}{(1-s)(D_{L,1} - D_{L,2})}. \quad (24)$$

Finally, with the CDFs at hand, we can derive q_H . Let the corresponding probability density functions (PDF) of H 's and L 's quality choice be f_H and f_L , respectively. Values of q where one of the PDFs is negative cannot be reasonable. As both PDFs are increasing in q , we can solve for the minimum feasible q by solving for the maximum root of both PDFs. Solving $f_L \stackrel{!}{=} 0$ and $f_H \stackrel{!}{=} 0$ we obtain that the maximum root is

$$q_H = \max \left\{ q_L, \frac{(D_{H,1} - D_{H,2})}{\left(\frac{\partial D_{H,1}}{\partial q_H} - \frac{\partial D_{H,2}}{\partial q_H}\right)} + \sqrt{\frac{(D_{H,1} - D_{H,2})^2}{\left(\frac{\partial D_{H,1}}{\partial q_H} - \frac{\partial D_{H,2}}{\partial q_H}\right)^2} - \frac{\tilde{\Pi}_H}{k} - \frac{(1-s) \left(D_{H,1} \frac{\partial D_{H,2}}{\partial q_H} - D_{H,2} \frac{\partial D_{H,1}}{\partial q_H}\right)}{k \left(\frac{\partial D_{H,1}}{\partial q_H} - \frac{\partial D_{H,2}}{\partial q_H}\right)} \right\}. \quad (25)$$

Note that this minimum quality is chosen by CP H with a probability of $F_H(q_H)$. Likewise, the probability that CP L chooses its minimum quality q_L , and does not compete for the top position in the ranking, is $F_L(q_H)$.

Proof of Lemma 1: In the case of a linear demand function with $d_1 = n_1 q_H$ and $d_2 = n_2 q_H$, equations (1) and (2) simplify to $D_{i,1} = N_1 q_i$ and $D_{i,2} = N_2 q_i$, respectively, where N_1 and N_2 are defined as in Lemma 1. Substituting this into the above formulas for F_i , q_L and \bar{q}_i , with $i = \{L, H\}$ yields Lemma 1.

Stage 1 In the first stage the platform selects the revenue share, s , which it commands from the CPs. Thereby, it anticipates how its choice of s will affect the CPs' subsequent quality choice. Since the platform is a monopolist, s is determined by maximization of the platform's profit

$$s = \arg \max_s \Gamma = \frac{D_L + D_H}{-\left(\frac{dD_L}{ds} + \frac{dD_H}{ds}\right)} = \frac{D}{-\frac{dD}{ds}}, \quad (26)$$

where $D = D_L + D_H$ is the expected total demand of both CPs with

$$D_L = F_L(q_H) D_{L,2} + \left(\int_{q_H}^{\bar{q}_L} (F_H f_L D_{L,1}) dq_L \right) + \left(\int_{q_H}^{\bar{q}_L} ((1 - F_H) f_L D_{L,2}) dq_L \right), \quad (27)$$

$$D_H = F_H(q_H) F_L(q_H) D_{H,1} + F_H(q_H) (1 - F_L(q_H)) D_{H,2} + \left(\int_{q_H}^{\bar{q}_L} (F_L f_H D_{H,1}) dq_H \right) + \left(\int_{q_H}^{\bar{q}_L} ((1 - F_L) f_H D_{H,2}) dq_H \right), \quad (28)$$

and where D_i is the expected demand of CP i 's content and $\frac{dD_i}{ds}$ is the corresponding total differential with respect to s . We note that in the linear demand case the platform would demand a revenue share of $s = 1/2$ in equilibrium.

Appendix B: Properties of F_H and F_L

In the following we show that the quality choices of CP H first-order stochastically dominate the quality choices of CP L , i.e., $F_H(q) \leq F_L(q)$ for all $0 \leq k < 1$. Furthermore, we show that for $k = 1$ the quality choices of both CPs coincide, i.e. $F_H(q) = F_L(q)$ for $k = 1$.

Proof of stochastic dominance of F_H over F_L : First, note that for any given $q = q_H = q_L$, we have $D_1 = D_{H,1} = D_{L,1}$ and $D_2 = D_{H,2} = D_{L,2}$. Consequently we can write the CDFs as follows:

$$F_H = \frac{\tilde{\Pi}_L - ((1-s)D_2 - q^2)}{(1-s)(D_1 - D_2)}, \quad (29)$$

$$F_L = \frac{\tilde{\Pi}_H - ((1-s)D_2 - kq^2)}{(1-s)(D_1 - D_2)}. \quad (30)$$

Suppose first-order stochastic dominance holds, then $F_H \leq F_L$ holds for any given quality level $q = q_H = q_L \leq \bar{q}_L$ and $0 \leq k \leq 1$. Substituting (29) and (30) into this condition yields

$$(1-k)q^2 \leq \tilde{\Pi}_H - \tilde{\Pi}_L. \quad (31)$$

Recall that the expected profit of L when choosing \bar{q}_L is $\tilde{\Pi}_L$. Consequently, we can write L 's outside profit as $\tilde{\Pi}_L = (1-s)\bar{D}_1 - \bar{q}_L^2$. Inserting this result and $\tilde{\Pi}_H$ into inequality (31) yields

$$(1-k)q^2 \leq (1-k)\bar{q}_L^2, \quad (32)$$

which is true for all $0 \leq k \leq 1$ and $q \leq \bar{q}_L$.

Proof of identity of F_H and F_L for $k = 1$: See that for $k = 1$ formula (32) holds with equality, from which it can be immediately followed that $F_H = F_L$ holds in this case.

Appendix C: Equilibrium derivation under a sponsored ranking with more general demand functions

Stage 4: Stage 4 corresponds to Stage 3 in Section 3.

Stage 3: In this stage the CPs submit bids given their quality choice in Stage 2. In a second-price auction it is a weakly dominant strategy to bid one's true valuation (Krishna 2010, Proposition 2.1). A CP's value for the first position is determined by the additional profit that can be attained by being in the this position. That is, in equilibrium $b_i = \Pi_{i,1} - \Pi_{i,2} = (1-s)(d_1(q_i) - d_2(q_i))$.

Stage 2: Consequently, the CP that has a larger incremental profit from being in the first position will always win the auction in Stage 3. As investment costs in quality are sunk at this stage, the CP that has chosen the higher quality level in Stage 2 will therefore always win the auction in Stage 3. For all feasible values of k , it can be shown that CP H can choose its profit maximizing quality level, q_H^m without risking that CP L might choose a higher quality. To see this, suppose L would choose a higher quality, i.e., $q_L = q_H^m$ (plus a small increment) instead of its profit maximizing quality in the second position, \underline{q}_L . Then it would expect to win the auction and make a profit of $\Pi_{L,1} = (1-s)d_1(q_H^m) - C_L(q_H^m) - b_H$ with $b_H = (1-s)(d_1(q_H^m) - d_2(q_H^m))$. However, it is then easy to see that $\Pi_{L,1} = (1-s)d_2(q_H^m) - C_L(q_H^m) < (1-s)d_2(\underline{q}_L) - C_L(\underline{q}_L) = \Pi_{L,2}$, because L 's net profits are as if it were in the second position, and \underline{q}_L is the profit maximizing quality choice for the second position. Consequently, given the same revenue-share s and for $\gamma = 1$, the content providers would set the same quality levels as under the pure strategy equilibrium given by Equations (21) and (22), respectively. Thus, under a sponsored ranking, there exists a unique pure strategy equilibrium in Stage 2, where CPs L and H choose the quality levels

$$q_L = \frac{(1-s)}{2} \frac{\partial d_{L,2}}{\partial q_L} \quad (33)$$

$$q_H = \frac{(1-s)}{2k} \frac{\partial d_{H,1}}{\partial q_H} \quad (34)$$

Proof of Lemma 2: In the case of a linear demand function it follows that $D_{i,1} = d_1(q_i) = d_1 q_i$ and $D_{i,2} = d_2(q_i) = d_2 q_i$ with $i = \{L, H\}$. Substituting this into the above formulas immediately yields Lemma 2.

Stage 1: Finally, in the first stage, the platform determines the revenue share, s , so that $s = \arg \max_s \Gamma = \arg \max_s s(D_{H,1} + D_{L,2}) - b_L$. This yields:

$$s = \frac{D_{H,1} + D_{L,2} + \frac{db_L}{ds}}{-\left(\frac{dD_{H,1}}{ds} + \frac{dD_{L,2}}{ds}\right)} = \frac{D}{-\frac{dD}{ds}} + \frac{\frac{db_L}{ds}}{-\frac{dD}{ds}} \quad (35)$$

Appendix D: Proof of $\underline{q}_L \leq \underline{q}_H < \bar{q}_L = \bar{q}_H$

Notice that $\underline{q}_L \leq \underline{q}_H$ holds if H chooses quality levels of $q_H \leq \underline{q}_L$ with zero probability, i.e.,

$$F_H(\underline{q}_L) = \frac{\tilde{\Pi}_L - \left((1-s)\underline{D}_{L,2} - \underline{q}_L^2\right)}{(1-s)(\underline{D}_{L,1} - \underline{D}_{L,2})} \stackrel{!}{\leq} 0.$$

Rewriting results to

$$F_H(\underline{q}_L) = \frac{\tilde{\Pi}_L - \tilde{\Pi}_L}{(1-s)(\underline{D}_{L,1} - \underline{D}_{L,2})} = 0.$$

Thus, at $q_H = \underline{q}_L$ the CDF of H 's strategy is zero, i.e. the probability that H chooses \underline{q}_L or a lower quality is zero. To see that $\underline{q}_L < \bar{q}_L$, notice that the maximum quality of L is

$$\bar{q}_L = \sqrt{(1-s)\bar{D}_{L,1} - \tilde{\Pi}_L}.$$

Inserting $\tilde{\Pi}_L$ yields

$$\bar{q}_L = \sqrt{\underline{q}_L^2 + (1-s)(\bar{D}_{L,1} - \underline{D}_{L,2})}.$$

Note that $\bar{q}_L > \underline{q}_L$ holds as $\bar{D}_{L,1} > \underline{D}_{L,2}$. Next, see that $\bar{q}_L = \bar{q}_H$ holds if the CDF of H is one at $q_H = \bar{q}_L$. This results in

$$1 \stackrel{!}{=} \frac{\tilde{\Pi}_L - \left((1-s)\bar{D}_{L,2} - \bar{q}_L^2\right)}{(1-s)(\bar{D}_{L,1} - \bar{D}_{L,2})}$$

Inserting \bar{q}_L yields

$$1 \stackrel{!}{=} \frac{\tilde{\Pi}_L - \left((1-s)\bar{D}_{L,2} - \left((1-s)\bar{D}_{L,1} - \tilde{\Pi}_L\right)\right)}{(1-s)(\bar{D}_{L,1} - \bar{D}_{L,2})}.$$

Rearranging yields

$$1 = \frac{(1-s)(\bar{D}_{L,1} - \bar{D}_{L,2})}{(1-s)(\bar{D}_{L,1} - \bar{D}_{L,2})},$$

i.e., the CDF of H is one at $q_H = \bar{q}_L$ and thus the maximum quality level of H is $\bar{q}_L = \bar{q}_H$ for $k > \tilde{k}$.

Appendix E: Proof of Proposition 1 (Competition Effect)

Consider that $q_L^o = q_L^s$ for $k \leq \tilde{k}$ and equal s . Consider further that q_L^s is constant in k and that $q_L^s = \underline{q}_L^o$. From $\bar{q}_L^o > \underline{q}_L^o$ for $k > \tilde{k}$ it follows that $E[q_L^o] > q_L^s$ for $k > \tilde{k}$ and equal s , i.e., the competition effect induces CP L to choose higher qualities under organic ranking.

Consider that $q_H^o = q_H^s$ for $k \leq \tilde{k}$ and equal s . Note that $E[q_H^o] > q_H^s$ for $k > \tilde{k}$ holds if

$$\int_{q_H^s}^{\bar{q}_H^o} f_H(q_H) q_H dq_H > F_H(\underline{q}_H^o) \underline{q}_H^o + \int_{\underline{q}_H^o}^{\bar{q}_H^o} f_H(q_H) q_H dq_H.$$

Note that both sides of the inequality are equal for $k \leq \tilde{k}$. Consequently $E[q_H^o] > q_H^s$ for $k > \tilde{k}$ holds if the first derivative of the left side with respect to k is higher than the first derivative of the right side. Rewriting the inequality by means of partial integration yields

$$0 > 2 F_H(q_H^s) q_H^s - F_H(\bar{q}_H^o) \bar{q}_H^o + \int_{q_H^s}^{\bar{q}_H^o} F_H(q_H) dq_H - \int_{\underline{q}_H^o}^{q_H^s} F_H(q_H) dq_H.$$

Note that $F_H(q)$ and \bar{q}_H^o are independent of k , i.e. k affects q_H^s and \underline{q}_H^o . Calculating the first derivative of the inequality with respect to k results to the following condition for $E[q_H^o] > q_H^s$ under $k > \tilde{k}$:

$$0 > 2 q_H^s f_H(q_H^s) \frac{\partial q_H^s}{\partial k} + F_H(\underline{q}_H^o) \frac{\partial \underline{q}_H^o}{\partial k}$$

Note that $q_H^s > 0$, $f_H(q_H^s) > 0$, $F_H(\underline{q}_H^o) > 0$, $\frac{\partial q_H^s}{\partial k} < 0$ and $\frac{\partial \underline{q}_H^o}{\partial k} < 0$. Consequently $E[q_H^o] > q_H^s$ for $k > \tilde{k}$ holds.

Appendix F: Proof of Proposition 2 (Uncertainty Effect)

In order to proof the uncertainty effect, we have to show that under an organic ranking CP H has a lower incentive invest into q_H as γ falls, and that CP L has a higher incentive to invest in q_L as γ falls, everything else being equal. First consider the case, where the pure strategy equilibrium emerges ($k < \tilde{k}$). Here the uncertainty effect can be shown immediately. Since $D_{H,1} = \gamma d_1(q_H) + (1 - \gamma) d_2(q_H)$, it follows that $\frac{\partial D_{H,1}}{\partial q_H} = \gamma \frac{\partial d_1(q_H)}{\partial q_H} + (1 - \gamma) \frac{\partial d_2(q_H)}{\partial q_H}$, which increases in γ since $\frac{\partial d_1(q)}{\partial q} > \frac{\partial d_2(q)}{\partial q}$. Thus, by Equation (21) q_H must decrease as γ decreases. Likewise, it holds that $D_{L,2} = (1 - \gamma) d_1(q_L) + \gamma d_2(q_L)$ decreases in γ . Thus, by Equation (22) q_L must increase as γ decreases.

From this proof, it is also evident that the uncertainty effect must become stronger (i.e. γ has a larger effect on investment incentives) if the demand function becomes more convex, because this means that $\frac{\partial d_1}{\partial q_i} - \frac{\partial d_2}{\partial q_i}$ increases. The reverse holds for a more concave demand function.

For $k > \tilde{k}$ generally the same logic applies, but here the uncertainty effect interacts with the competition effect, which has a countervailing effect on CP L 's quality choice. However, the uncertainty effect follows from the minimum scale effect and the additional-scale effect, respectively. In Appendix B it is shown that F_H stochastically dominates F_L for $k < 1$, which means that in the mixed strategy equilibrium CP H chooses the highest content quality more often than CP L . Thus, when γ decreases, CP H suffers more often from the corresponding reduction in $D_{i,1}$ than from the corresponding increase in $D_{i,2}$, which makes it less attractive to invest in q_H . This corresponds to the additional scale effect. The reverse holds true for CP L , who sets the lower content quality more often, and thus benefits more from the increase in $D_{i,2}$ than the decrease in $D_{i,1}$ as γ decreases.

Appendix G: Proof of Proposition 3 (Revenue-Share Effect)

Below, we derive that the platform will set a lower revenue share under a sponsored ranking for linear demand curves and that this result holds also for a range of concave and convex demand curves which are sufficiently linear, i.e., not too convex or too concave. This means that, everything else being equal, due to this *revenue share effect*, CPs often have a stronger incentive to invest in quality under a sponsored ranking.

At first consider the case where the competition effect and the uncertainty effect are not present, i.e., $k < \tilde{k}$ and $\gamma = 1$. In this case, under both regimes, CP H will always attain the first position and CP L will always attain the second position. Consequently, CPs face the same strategic incentives to invest in quality under both regimes. This means that, in this case, both the total demand as well as the total differential of demand with respect to s must coincide, i.e., $D^o = D^s$ and $\frac{dD^o}{ds} = \frac{dD^s}{ds}$. Yet, by comparing formulas (6)

and (8) it is evident, that the platform will nevertheless select a lower revenue share in equilibrium under a sponsored ranking, due to the additional influence of s on the expected auction revenue. Therefore, in this case it must hold for all demand functions that $s^s < s^o$ in equilibrium.

Given this benchmark, we now consider what would happen if k were to increase from $k = \tilde{k}$, i.e., if additionally the competition effect is present. Three arguments are to be made here:

(i) From Proposition 1, we know that, given the same s , the expected content quality of each CP is higher under an organic ranking for $k > \tilde{k}$. Consequently, also total demand and total revenue must be higher under an organic ranking. This means that for any $k > \tilde{k}$ it must hold that $D^o > D^s$. This implies that the nominator of the first term in formula (6) is larger than in formula (8).

(ii) Due to the fact that CPs compete in qualities under an organic ranking, they cannot afford to reduce their qualities as much as under a sponsored ranking in response to an increase in s . This tends to lower the absolute value of $|\frac{dD^o}{ds}|$ relative to $|\frac{dD^s}{ds}|$.

(iii) Because CPs' qualities are different under the two regimes for a given level of s , the CPs' reactions to an increase in s now also depend on the shape of the demand curve.

In case of a linear demand curve, the slope is the same at any level of q_i , and then only arguments (i) and (ii) are relevant. This means that, for a linear demand curve, the first term in formula (6) must be larger than in formula (8), which means that also for $k > \tilde{k}$ it follows that s must be larger under an organic ranking than under a sponsored ranking.

A convex demand curve tends to decrease s^s further, because CPs' bids react more strongly to an increase in s . However, because qualities are higher under an organic ranking, the effect identified in argument (ii) is dampened and possibly reversed. It is not possible to say - for general convex demand curves - which effects dominate. However, it is safe to argue that $s^s < s^o$ should also hold for a large range of convex demand functions.

In reverse, a concave demand curve tends to reinforce the effect identified in argument (ii), which tends to increase s^o . Moreover, under a concave demand curve, CPs' reduction in quality is relatively stronger at lower quality levels. This tends to decrease s^s , because qualities levels are lower under a sponsored ranking, given the same s . On the other hand, CPs' bids react less strongly to an increase in s under a concave demand curve, which tends to increase s^s . Again, it is not possible to derive more general insights for general concave demand curves, but the result $s^s < s^o$ prevails as long as the demand curves are not too concave.

Finally, consider how an increase in ranking uncertainty would alter these results. A lower γ makes it less attractive to attain the top ranking position under an organic ranking, and thereby the competition effect is softened. In the limit, for $\gamma = 1/2$, the CPs are randomly assigned the top position and the competition effect vanishes completely, i.e., for $\gamma \rightarrow 1/2$ it follows that $\tilde{k} \rightarrow 1$. This also means that, at lower values of γ , the total differential $\frac{dD^o}{ds}$ must be relatively larger than at higher values of γ , where the competition effect would be stronger. Hence, a higher ranking uncertainty tends to decrease the optimal s^o , but has no effect on s^s , of course. Consequently, lowering γ renders the revenue shares to be more similar under the two ranking regimes, everything else being equal.

However, for the case of a linear demand function, one can show that $s^s < s^o$ continues to hold over the whole range of γ . More precisely, in case of a linear demand function the revenue share under sponsored ranking is

$$s^s = \frac{2(n_1)^3 - (kn_2)^2(n_1 - n_2) - kn_1n_2(3n_1 - 5n_2)}{4(n_1)^3 - 4kn_1n_2(n_1 - 2n_2)}.$$

which, for all feasible parameter values, is less or equal to the revenue share under organic ranking with $s^o = 1/2$.

Appendix H: Sensitivity analysis of n_1

In the following we study how an increase in the demand in the first position, relative to the second ranking position influences the investment incentive effects that we have identified. In our model with linear demand this translates into increasing n_1 relative to n_2 . This, of course, means that attaining the first ranking position becomes relatively more attractive, which has several implications for the dynamics of the competition under each ranking that can be immediately derived from the existing analysis.

First, an increase in n_1 , everything else being equal, magnifies the *competition effect*. To see this, note that an increase in n_1 leads to a decrease in \tilde{k} (see Lemma 1). This means that the competition effect already occurs at lower levels of k under an organic ranking, and that the quality-based competition for the first position becomes more intense. This in turns means that the competition effect becomes stronger. In consequence, the cumulative distribution function representing the quality levels under an organic ranking is shifted upward, i.e., content providers expected qualities go up. This is easy to see, for example, in terms of \bar{q} (the highest quality level that any content provider would set) in Lemma 1, which clearly increases in n_1 .

Second, an increase in n_1 , everything else being equal, leads to a stronger increase in the quality of CP H than in the quality of CP L . An increase in n_1 benefits, predominantly CP H (the high efficiency provider), because it attains the first position more often. This is already discussed in the context of the *additional scale effect* in Section 5.2. When n_1 increases, the additional scale that can be achieved in the first position becomes larger, which gives the CP that is in the first position an additional investment incentive. Since CP H is more often in the first position, both under an organic ranking and under a sponsored ranking, CP H will increase its quality more than CP L .

Third, an increase in n_1 , everything else being equal, leads to a stronger increase in quality of CP H under a sponsored ranking, relative to CP H s expected quality increase under an organic ranking. Under a sponsored ranking CP H always attains the first position. Thus, it always benefits from the additional scale effect, as discussed above. By contrast, under an organic ranking, even absent any uncertainty by the platform, CP H does not always secure the first position, because both CP H and CP L choose their qualities according to some probability distribution. Thus, CP does not always (but still more often than CP L) benefit from the additional scale effect. Therefore, an increase in n_1 will, everything else being equal, raise the quality of CP H more strongly under a sponsored ranking than under an organic ranking.

It is easy to see that this effect of an increase in n_1 becomes more pronounced if the platform has uncertainty in determining the correct ranking order under an organic ranking. A reduction in γ has no effect on the outcomes under a sponsored ranking, of course. However, under an organic ranking a reduction in γ means

that CP H attains the first position even less often, and accordingly reduces its expected quality even more relative to its quality under a sponsored ranking. This is precisely due to the uncertainty effect highlighted in Proposition 2.

Fourth, an increase in n_1 monotonically increases the equilibrium bids, b_i , under a sponsored ranking, however, it does not monotonically decrease the revenue share, s^s , set by the platform under a sponsored ranking. One of the fundamental insights of our model is that the platform would generally set a lower revenue-share under a sponsored ranking than under an organic ranking, because it takes into account that the revenue share does not only affect its revenue from content, but also its revenue from the bids. Under a sponsored ranking, an increase in s would not only trigger CPs to set lower content qualities (as is also the case under an organic ranking), but also to submit lower bids (see Stage 2 in Section 4.2). This makes the platform softer with respect to its choice of s , and hence the platform selects a lower revenue share than under an organic ranking (i.e. in the absence of bids). How soft the platform is when setting s , depends on the size of n_1 relative to n_2 . When n_1 is very low, then bids are very low and hence the platform will not be much softer than under an organic ranking. Hence, it sets the revenue share similar as in the case of an organic ranking. To see this, consider the limiting case where $n_1 = n_2$. In this case there is no value in attaining the first position relative to the second position. Thus, firms will bid exactly zero, and hence the revenue-share effect disappears completely, i.e., the revenue-share is set to the same level as under an organic ranking, i.e., $s^s = 1/2$. In other words, as n_1 becomes very low, relative to n_2 , the platform becomes tougher in setting s , because the revenue from bids matters less than the revenue from content.

The reverse is also true. If n_1 is very large relative to n_2 , then for the platform revenue from content also matters more than revenue from bids under a sponsored ranking. To see this, notice that under our linear demand specification the platform's revenue from content increases quadratically in n_1 in the sponsored ranking equilibrium, because $s(R_H + R_L) = s(n_1 q_H + n_2 q_L) = s\left(\frac{(1-s)}{2k} n_1^2 + \frac{(1-s)}{2} n_2^2\right)$. However, the platform's equilibrium revenue from bids increases only linearly in n_1 , because $b_L = (1-s)(n_1 - n_2)q_L = \frac{(1-s)^2 n_2}{2}(n_1 - n_2)$. Hence, for very large n_1 , the revenue from bids becomes less important and therefore the platform becomes tougher again when choosing the revenue share.

Taken together, this leads to a non-monotonic effect of n_1 on the platform's choice of the revenue-share under a sponsored ranking, and hence on the revenue-share effect. Consequently, the revenue-share effect is the strongest for intermediate values of n_1 relative to n_2 , because then the platform's trade-off between revenue from bids and revenue from content is the strongest.

In summary, while the size of n_1 relative to n_2 moderates the strength of the competition, uncertainty and revenue-share effects in various ways, it does not affect the validity of the propositions.

Appendix I: Existence of the competition effect under a sponsored ranking with quality factor

In order to understand how the competition effect works under a sponsored ranking with quality factor, see that the outside profit of L remains the same as under organic ranking or sponsored ranking, given the same revenue share, s . As long as CP L does not win the auction the design of r has no effect on CP L 's profit. This fact implies that not only the outside option of L under a sponsored ranking with quality factor

remains unchanged, but also its behavior in a pure strategy equilibrium remains unchanged and is as under a pure organic or a pure sponsored ranking given the same revenue share. However, when CP L engages in a quality competition with H by choosing a quality $q_L > \underline{q}_H$ its expected profit is:

$$\Pi_L = (1 - s) (F_H d_1(q_L) - (1 - F_H) d_2(q_L)) - C_L(q_L) - F_H E(r_L(q_L)) \quad (36)$$

where $E(r_L(q_L))$ is the expected payment of L if it wins the auction. Note that $E(r_L) = 0$ under the organic ranking. This means that under a sponsored ranking with quality factor it is less attractive for CP L to compete with H in qualities than under a pure organic ranking. Consequently, the competition effect is weaker under the sponsored ranking with quality factor for any $\beta < 1$. However, the competition effect still exists, as long as β, k are sufficiently high. To see this, consider a situation with $k > \tilde{k}$. Assume further that CP H expects CP L to choose $q_L = \underline{q}_L$, as in a pure strategy equilibrium. Then CP L 's best response is $q_H = q_H^{br}$. In order to proof if L has an incentive to deviate in this situation by overbidding CP H consider that

$$\Pi_{L,2} = \tilde{\Pi}_L = (1 - s) d_2(\underline{q}_L) - C_L(\underline{q}_L) \quad (37)$$

is L 's profit from staying in the second position, while

$$\Pi_{L,1} = (1 - s) d_1(q_{L,1}) - C_L(q_{L,1}) - E(r_L(q_{L,1})) \quad (38)$$

is L 's profit from being in the first position. There is competition in quality as long as L has an incentive to outbid H , i.e., as long as $\Pi_{L,2} < \Pi_{L,1}$ for some $q_L > q_H^{br}$. For $\beta = 1$, i.e., under a pure organic ranking, we already know that $E(r_L(q_{L,1})) = 0$ and $\Pi_{L,2} < \Pi_{L,1}$ holds when $k > \tilde{k}$. Further for $\beta = 1$ and $k \rightarrow 1$ the difference in profits $\Pi_{L,1} - \Pi_{L,2}$ is strictly positive (non-marginal), because, by assumption $d_1(q) > d_2(q)$ for any given q . Contrast this with a situation in which $\beta < 1$. Then, when CP L wins the auction, it has to make a positive payment, which decreases the difference in profits $\Pi_{L,2} < \Pi_{L,1}$. However, as $\beta \rightarrow 1$ the expected payment, $E(r_L(q_{L,1}))$, becomes arbitrarily small. Because the difference in profits is strictly positive when $\beta = 1$, there must be a threshold $\tilde{\beta} < 1$ above which CP L still wants to compete in qualities, because it then has a chance for attaining a strictly greater profit than in its outside option, but must make a payment that is smaller than the incremental profit.

This means that above $\tilde{\beta}$ there exists a unique mixed strategy equilibrium and below $\tilde{\beta}$ there exists a unique pure strategy equilibrium. In the pure strategy equilibrium CP L chooses the same quality as under a pure sponsored ranking (given the same revenue share s). However, given the same revenue share s , CP H chooses a higher quality than under a pure sponsored ranking, because H can lower its payment by increasing q_H (payment-reduction effect).

Appendix J: Biased ranking without vertical integration

J.1. The model

We now study the case of a *biased* ranking algorithm, where the first ranking position always given either to CP H or CP L , but irrespective of content quality or bids. That is, the platform and all CPs know in advance which CP will receive the first position, and this is common knowledge. However, as we will show,

it does not matter whether or not the platform knows the CP's type (i.e., if it is of high or low efficiency), because it will set the same revenue-share in either case. The promoted CP always receives a demand of $D_1 = d_1$ and the other CP always a demand of $D_2 = d_2$. It is useful to study this case, because it eliminates any type of competition for the first rank between the two CPs. The timing is now as follows: *Stage 1*: The platform announces its revenue share s . *Stage 2*: The CPs H and L simultaneously choose their content qualities q_H and q_L , and the platform ranks the promoted CP (either H or L) always in the first position. *Stage 3*: Content is demanded depending on its ranking position and quality, and revenues are realized.

J.2. Equilibrium derivation

Bias towards the more efficient CP: The analysis of Stage 3 is identical to the analysis of Stage 3 in Section 3.2. In Stage 2 the CPs know with certainty that CP H will always obtain the first, and that CP L will always achieve the second ranking position. Thus, the optimal qualities are readily given by profit maximization.

$$q_H^{b,H} = \arg \max_{q_H} \Pi_H^{b,H} = \frac{(1-s)}{2k} \frac{\partial D_{H,1}}{\partial q_H},$$

$$q_L^{b,H} = \arg \max_{q_L} \Pi_L^{b,H} = \frac{(1-s)}{2} \frac{\partial D_{L,2}}{\partial q_L},$$

In Stage 3, anticipating its impact on the quality q_L , the integrated platform maximizes its profit by setting the revenue share to

$$s^{b,H} = \arg \max_s \Gamma = - \frac{D_{H,1} + D_{L,2}}{\left(\frac{dD_{H,1}}{ds} + \frac{dD_{L,2}}{ds} \right)}. \quad (39)$$

Bias towards the less efficient CP: In Stage 2 optimal qualities are given by profit maximization as follows:

$$q_H^{b,L} = \arg \max_{q_H} \Pi_H^{b,L} = \frac{(1-s)}{2k} \frac{\partial D_{H,2}}{\partial q_H},$$

$$q_L^{b,L} = \arg \max_{q_L} \Pi_L^{b,L} = \frac{(1-s)}{2} \frac{\partial D_{L,1}}{\partial q_L},$$

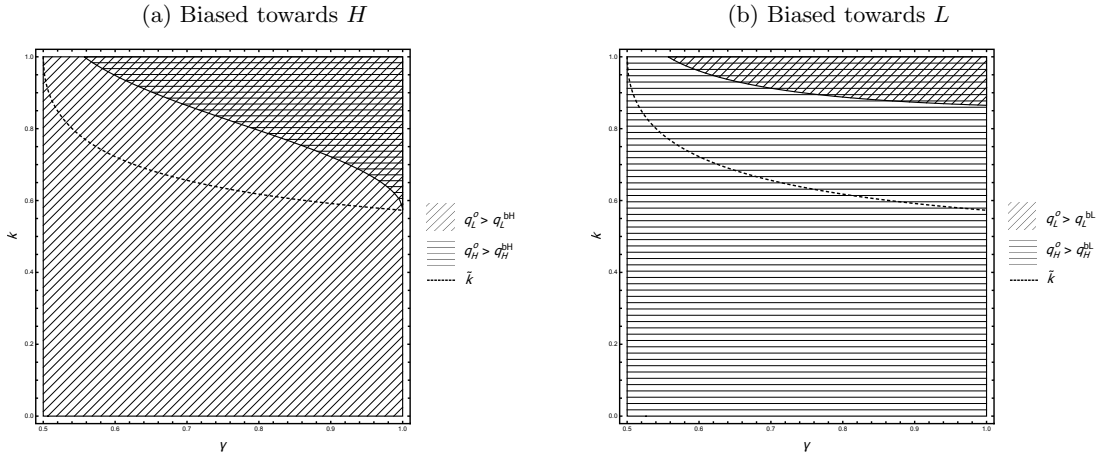
In Stage 3 the platform would then set a revenue share of

$$s^{b,L} = \arg \max_s \Gamma = - \frac{D_{H,2} + D_{L,1}}{\left(\frac{dD_{H,2}}{ds} + \frac{dD_{L,1}}{ds} \right)}. \quad (40)$$

J.3. Quality investment incentives

Bias towards the more efficient CP: The platform will generally set a lower revenue share under a sponsored ranking than under a biased ranking, i.e., $s^s < s^{b,H}$. By comparing $s^s = - \frac{D_{H,1} + D_{L,2} + \frac{dq_L}{ds}}{\left(\frac{dD_{H,1}}{ds} + \frac{dD_{L,2}}{ds} \right)}$ and $s^{b,H}$, it is immediately evident that under a biased ranking the platform does not need to consider the negative impact that an increase in s would have on its auction revenue. This makes the platform's revenue less elastic in s , and thus, the platform will select $s^{b,H} > s^s$. Otherwise, a biased ranking is governed by the same effects as a sponsored ranking. Thus, due to the lower revenue-share effect the content qualities of each CP will be lower under a biased ranking compared to a sponsored ranking. In reverse, this means that when comparing a biased ranking where H is always promoted with an organic ranking, then the same intuitions derived when comparing a sponsored ranking with an organic ranking carry over, but the parameter regions in which an organic ranking yields the higher expected quality will increase, for both H and L .

Figure 8 Comparison of Content Qualities between Organic and Biased Ranking



(a) When the ranking is biased towards H , then the quality of each CP is strictly lower than under a sponsored ranking, and thus, the parameter region in which an organic ranking leads to higher qualities becomes larger compared to Figure 3b. (b) When the ranking is biased towards L , the quality of CP H is strictly lower, but the quality of CP L is strictly higher than under a ranking biased towards H . Accordingly, the parameter region in which an organic ranking yields a higher quality for CP L (H) becomes smaller (larger) compared to Figure 3b. Figures derived for a linear demand function with $n_1 = 6$, $n_2 = 4$.

Bias towards the less efficient CP: Generally the same reasoning with respect to the competition, revenue-share and uncertainty effects applies in case CP L is promoted under a biased ranking. In particular, the platform will again set a higher revenue share than under a sponsored ranking for the reasons given above. However, now the less efficient CP L operates at a different quality level than CP H , and therefore it depends on the shape of the demand function whether the equilibrium revenue share is higher (for convex demand functions), or lower (for concave demand functions) than in the case when the ranking is biased towards CP H .

In any case, the difference in the revenue-share effect between the two biased rankings, if it exists, is only of second order importance for the chosen quality levels, because it stems solely from the fact that the quality levels at which CP L and H operate differ with respect to their respective slopes in the demand function (assuming a not too concave and not too convex demand function). However, it is of first order importance that the less efficient CP is now enjoying the larger scale due to it being in the first position, which creates a negative quality distortion at the top. To see this, note that in an unbiased ranking, or in a ranking that is biased towards H , CP H is awarded the top position more often than CP L and thus always enjoys the larger scale effect.¹² The relative inefficiency of CP L means that it cannot use the additional scale as efficiently as CP H and therefore produces a quality that is lower than the quality that CP H would have chosen in its place. In reverse, due to the additional scale, CP L now invests more than it would have in the second

¹² Recall that F_H stochastically dominates F_L , such that CP H attains the top position more often under an organic ranking; and that CP H will always attain the top position under a sponsored ranking.

position. Depending on the size of the scale advantage ($d_1 - d_2$) relative to the efficiency disadvantage (k), CP L may either offer a quality that is lower, or a quality that is higher than that of CP H . Finally, notice again that in any case a biased ranking would yield a lower quality level than a sponsored ranking even when $k = 1$ due to the lower revenue-share effect.

Appendix K: Proof of Proposition 6 (vertical integration)

K.1. Equilibrium derivation

Integrated CP is of higher efficiency: The analysis proceeds as in Section J and, thus optimal qualities are readily given by:

$$\begin{aligned} q_L^{v,H} &= \arg \max_{q_L} \Pi_L^{v,H} = \frac{(1-s)}{2} \frac{\partial D_{L,2}}{\partial q_L}, \\ q_\Gamma^{v,H} &= \arg \max_{q_\Gamma} \Pi_\Gamma^{v,H} = \frac{1}{2k} \frac{\partial D_{\Gamma,1}}{\partial q_\Gamma}. \end{aligned}$$

In Stage 1, anticipating its impact on the quality q_L , the integrated platform maximizes (11) by setting the revenue share for the non-integrated CP to

$$s^{v,H} = \arg \max_s \Gamma = \frac{-D_{L,2}}{\frac{dD_{L,2}}{ds}}. \quad (41)$$

Integrated CP is of lower efficiency: When the integrated CP is CP L , and the non-integrated CP is CP H , i.e., the integrated CP is of lower efficiency, optimal content qualities are given by

$$\begin{aligned} q_H^{v,L} &= \arg \max_{q_H} \Pi_H^{v,L} = \frac{(1-s)}{2k} \frac{\partial D_{H,2}}{\partial q_H}, \\ q_\Gamma^{v,L} &= \arg \max_{q_\Gamma} \Pi_\Gamma^{v,L} = \frac{1}{2} \frac{\partial D_{\Gamma,1}}{\partial q_\Gamma}. \end{aligned}$$

The optimal revenue-share for the non-integrated CP is given by

$$s^{v,L} = \arg \max_s \Gamma = \frac{-D_{H,2}}{\frac{dD_{H,2}}{ds}}. \quad (42)$$

K.2. Comparison of content qualities under vertical integration

The formulas for the revenue shares immediately show i) the two CPs operate under a different revenue-share, and that ii) the revenue share for the non-integrated CP will generally be different than under a biased ranking without vertical integration. Only the reaction of the non-integrated CP on s matters. Whether this leads to an increase or decrease in the revenue share depends, again, on the precise nature of the demand function and is ambiguous for the general case. However, in any case this difference in the revenue-share between a biased ranking with and without vertical integration is only of second order importance for the quality choice, as long as the demand function is not too concave and not too convex, and it vanishes completely for a linear demand function.

The former point on the asymmetric nature of the revenue-share effect is of first order importance for the quality choice and does not depend on the nature of the demand function. Whereas the platform demands a positive revenue share from the non-integrated CP, it demands a revenue share of $s = 0$ from the integrated CP. Consequently, the revenue-share effect acts on the non-integrated CP similarly as under a biased or organic ranking, but the absence of a revenue share provides the integrated CP with strong incentives to

invest in quality under vertical integration; indeed much stronger than under either organic, sponsored or biased ranking without vertical integration. Consequently, vertical integration with the platform will yield similar outcomes as under a biased ranking without vertical integration for the non-integrated CP, but it will yield a significantly higher quality for the integrated CP.

Raising Total Capacity Provision of Mobile Network Operators Through Throttling

Author Oliver Zierke, University of Passau, Germany.

Presented at 42th Pacific Telecommunications Council's Annual Conference (PTC), 2020, Honolulu, USA.
Awarded with the "2020 Young Scholar Program award"

Published in PTC 2020 Conference Proceedings.

Abstract

In recent years, mobile network operators worldwide introduced mobile plans that restrict data transmission rates of certain content types such as videos. However, this practice, known as throttling, is viewed critically as it violates so-called net neutrality and is suspected to be used as an instrument to avoid network capacity improvements. Thus, this paper investigates the impact of throttling on mobile network operators' incentives to invest in the capacity of their network. A stylized game-theoretic model is used with a monopoly mobile network operator providing access to the Internet for a homogeneous group of consumers. It is shown that throttling helps to use the capacity of mobile networks more efficiently, which, surprisingly, does not result in mobile operators offering less overall network capacity. Instead, the ability to throttle increases their incentives to improve network capacity.

Raising Total Capacity Provision of Mobile Network Operators Through Throttling

Oliver Zierke*

First version: Oktober 18, 2018

This version: May 18, 2021

Abstract

In recent years, mobile network operators worldwide introduced mobile plans that restrict data transmission rates of certain content types such as videos. However, this practice, known as throttling, is viewed critically as it violates so-called net neutrality and is suspected to be used as an instrument to avoid network capacity improvements. Thus, this paper investigates the impact of throttling on mobile network operators' incentives to invest in the capacity of their network. A stylized game-theoretic model is used with a monopoly mobile network operator providing access to the Internet for a homogeneous group of consumers. It is shown that throttling helps to use the capacity of mobile networks more efficiently, which, surprisingly, does not result in mobile operators offering less overall network capacity. Instead, the ability to throttle increases their incentives to improve network capacity.

1 Introduction

In recent years, mobile network operators (MNO), especially in the U.S., started to abolish data caps in their mobile tariffs. Instead, they introduced unlimited data tariffs with transmission rate restrictions for certain types of content (Welch, 2018). For example, T-Mobile USA offers a tariff in which all video streams are restricted to the use of a maximum data transmission rate of 1.5 Mbps (T-Mobile USA, 2018). Other examples for MNOs providing tariffs with throttling the data transmission rates of certain content types are Telkom LIT in South Africa (Saunders, 2018) or U Mobile in Malaysia (Qishin, 2018).

*Chair of Internet and Telecommunications Business, University of Passau, Dr.-Hans-Kapfing-Str. 12, 94032 Passau, Germany. Email: o.zierke@yahoo.de. Declarations of interest: None.

Deutsche Telekom’s attempt to introduce a tariff with a similar restriction on data transmission rates in Germany failed due to resistance from the national regulator with a signaling effect for the whole European Union (EU). The Bundesnetzagentur found that throttling practices violate the net neutrality regulation of the EU (Meyer, 2017). This regulation states that “providers of Internet access services shall treat all traffic equally” (EU regulation 2015/2120). The Academic literature on net neutrality has shown that such regulation can be detrimental for overall welfare as it prohibits practices such as charging for prioritization of data streams in telecommunications networks, although such a paid prioritization often results in a better resource allocation (Krämer et al., 2013). The analysis of this paper shows that there are also reasons of doubt whether a prohibition of throttling data streams is beneficial. The main concern regarding throttling in the context of net neutrality can be found in Recital 15 of the EU’s net neutrality regulation. It states that traffic management tools like throttling could be used as a substitute for network capacity expansion (EU regulation 2015/2120). This concern seems to be justified only at first glance. The overall effect of throttling is composed of two opposing effects. In practice, throttling, i.e., limiting the data transmission rate for some content, means on the one hand that videos, for example, are displayed to consumers with a lower resolution. This is detrimental to consumers’ quality of experience (QoE). On the other hand, the network load per video stream decreases, which in turn reduces congestion in mobile networks and is thus beneficial for consumers’ QoE. All in all, the impact of throttling on QoE is not obvious and as a consequence the influence on capacity investments is also not obvious, as QoE directly impacts consumers’ willingness to pay and thus the monetization of capacity which in turn is the key driver for capacity investments.

This paper investigates this relation of MNOs’ overall capacity investment incentives and throttling by means of a stylized game-theoretic model with a monopolistic MNO. The model shows that the ability of an MNO to throttle enables it to optimize the usage of its mobile network. Surprisingly, this optimization never results in lower incentives to invest in network

capacity but rather yields sometimes higher incentives, i.e., an MNO that is prohibited to throttle data streams (due to some form of net neutrality regulation) will never provide more, but sometimes less network capacity to its customers.

The paper is structured as follows. First, a brief overview of the literature is given, dealing primarily with the economic literature on net neutrality. Afterwards, a game-theoretic model is developed, and the results are analyzed and interpreted. Finally, an outlook for future research is given.

2 Related Literature

This paper is mainly connected to the literature on net neutrality. In 2003, a paper of Wu (2003) initiated a debate on the discrimination-free treatment of data in the Internet. The major focus of the economic literature on this topic is about pros and cons of allowing paid prioritization of data streams. In some of these papers a static setting is assumed in which the capacity of the telecommunications network is exogenous (e.g., Guo and Easley, 2016; Economides and Tåg, 2012). These papers mainly deal with the influence of net neutrality on the behavior of content providers and consumers. They show that net neutrality is advantageous for content providers and users if it removes barriers to access the telecommunications network for them. However, it is not considered that the overall capacity of the telecommunications network is also a barrier and that this barrier is controlled by the Internet Service Provider (ISP). Papers that take into account the endogeneity of network capacity include Bourreau et al. (2015), Choi and Byung-Cheol (2010), Economides and Hermalin (2012), Krämer and Wiewiorra (2012) and Peitz and Schütt (2016).¹ One of the major results of these papers is that paid prioritization yields an additional revenue stream that improves monetization of the mobile network capacity. This means that the incentives to invest in capacity are usually higher without net neutrality. This positive effect on capacity investments holds, as long as capacity is scarce. Capacity investments under net neutrality

¹For a detailed literature review on net neutrality see Easley et al. (2018) and Krämer et al. (2013).

are higher only if capacity expansion under paid prioritization would impair ISPs' ability to discriminate against content providers by charging fees for prioritization (see, e.g., Choi and Byung-Cheol, 2010; Baake and Sudaric, 2019).

However, the results of all these papers are based on an additional stream of revenue generated by paid prioritization. In contrast, throttling, which is the focus of the present paper, is a pure traffic management tool, i.e., the ISP grants prioritization without receiving any monetary reward. Economic papers on net neutrality that focus on neutrality violations where content providers do not pay for prioritization are rare. Some papers deal with the consequences of an ISP being vertically integrated with a content provider (see, e.g., Dewenter and Rösch, 2016; Guo et al., 2010), while other papers deal with models where the ISPs apply second degree price discrimination on the user side (see, e.g., Bandyopadhyay and Cheng, 2006; Inceoglu and Liu, 2019; Guo et al., 2013). The paper by Inceoglu and Liu (2019) is closely related to this paper. Their work focuses on price discrimination through zero-rating, a form of mobile plans where the data streams of certain content providers are not counted towards the customers' data cap. They find that zero-rating increases the incentive for ISPs to invest in network capacity. This is because zero-rating improves the ability of ISPs to extract the rent of their customers and consequently it improves the monetization of capacity. Similarly, the present paper shows that throttling improves the monetization of capacity and thus increases capacity investment. However, since this effect occurs in the absence of any price discrimination, this paper adds a new important aspect to the literature that has not been considered yet.

3 A Model on the Impact of Bandwidth Throttling on Capacity Investment Incentives in Mobile Networks

3.1 The Model Setup

Consider a market where consumers access the Internet and its content through a monopoly MNO that provides a mobile network with an overall capacity k for a fixed flat rate price p . Without regulation, the MNO is allowed to use throttling, which means that it can limit the bit rate consumption of some content units. The degree of throttling will be denoted as $t \in [0, \infty]$, with a higher t representing higher degrees of throttling and $t = 0$ representing a situation without throttling. The degree of throttling affects consumers' Quality of Experience (QoE) and thus their utility from accessing the Internet through the mobile network. The overall impact on QoE can be broken down into two effects.

First, throttling has an impact on congestion in peak traffic periods, in particular by lowering the average traffic consumption per content unit. For example, if each video stream uses a lower bit rate, this results in less frequent and less severe congestion events in the network. Note that the frequency of congestion events is affected not only by throttling, but also by the overall network capacity k , i.e., a higher network capacity as well as a higher degree of throttling is appropriate to prevent congestion. This relation reflects the concern that usage of throttling substitutes capacity expansion. In the model, the impact of congestion on QoE is denoted as $\Delta(t, k) \in [0, 1]$, where $\Delta = 1$ represents a network without any congestion and $\Delta = 0$ represents a network which allows no data transfers. The positive impacts of throttling and capacity on congestion imply $\Delta(t, k) \leq \Delta(\tilde{t}, k)$ for any $t < \tilde{t}$ and $\Delta(t, k) \leq \Delta(t, \tilde{k})$ for any $k < \tilde{k}$.

Second, in addition to the impact of throttling on congestion there is a direct effect of throttling on consumers' QoE. Throttling affects the average quality of content units not only during peak traffic periods but all the time. This means that with throttling consumers always have to be contented with, e.g., video streams having a lower resolution even though

the available network capacity would sometimes allow for higher bit rates. This direct effect of throttling on QoE is denoted as $\Theta_E(t) \in [0, 1]$ with Θ_E being twice differentiable in t and $\Theta_E(t) > \Theta_E(\tilde{t})$ for any $t < \tilde{t}$. Further, $\Theta_E(0) = 1$ addresses that there is no direct impact of throttling on QoE without throttling and $\Theta_E(\infty) = 0$ addresses consumers' inability to access Internet content if all content is fully throttled. Assuming Θ_E to be multiplicatively separable from Δ implies that QoE can be modeled by $\Delta(t, k) \Theta_E(t)$.

Changes in QoE affect consumers' behavior in two ways. First, consumers' utility per content unit changes and, consequently, their willingness to pay. Hence, in the following it is assumed that the willingness to pay per content unit is equal to the QoE. Second, it is reasonable to assume that QoE influences consumers' consumption profile. Let $x \in [0, \bar{x}]$ be the number of content units consumed by a consumer. The consumers' content consumption is the higher the less the QoE is affected by throttling or capacity constraints, i.e., $x(t, \Delta(t, k)) > x(\tilde{t}, \Delta(t, k))$ for $t < \tilde{t}$ and $x(t, \Delta(t, k)) \leq x(t, \Delta(t, \tilde{k}))$ for any $k < \tilde{k}$. The fact that there is no consumption if the network is fully throttled implies $x(\infty, \Delta) = 0$. The same applies to a network which allows no data transfer, i.e., $x(t, 0) = 0$. Finally, $x(0, 1) = \bar{x}$ reflects the maximum consumption in the absence of the negative effects of throttling and congestion.

The resulting utility function for consumers is given by:

$$U = x(t, \Delta(t, k)) \Delta(t, k) \Theta_E(t), \tag{1}$$

Note that in order to limit the complexity of the model, it is assumed that consumers are homogeneous. Consequently, given a certain parameter constellation, either all or none of the consumers would buy access to the mobile network. Without loss of generality, we normalize the mass of consumers to one.

From the network perspective, Δ is modeled as the ratio of mobile network's capacity k and total peak traffic T , i.e.,

$$\Delta = \min \left\{ \frac{k}{T(t, k)}, 1 \right\}, \quad (2)$$

which is a specification referring to Peitz and Schütt (2016). It reflects that if the available network capacity is sufficient to deliver content always as soon as necessary, then congestion is not an issue and $\Delta = 1$. If congestion is an issue, i.e., if the total peak traffic T exceeds the available bandwidth, it is modeled as the probability that a packet is sent in time, which is specified as $\Delta = \frac{k}{T(t, k)} \in [0, 1]$ (Peitz and Schütt, 2016). The total peak traffic itself depends on consumers' data consumption. The MNO can influence this consumption by throttling. First, throttling makes content less attractive and thus decreases x , the number of content units consumed by a consumer. Second, throttling lowers the data rate consumption per content unit, denoted as $\delta(t) \in [0, 1]$ in the following. This results in

$$T = x(t, \Delta(t, k)) \delta(t), \quad (3)$$

where $\delta(t)$ is assumed to be decreasing and twice differentiable in t . In other words, more throttling decreases total peak traffic. Further, it is reasonable to assume $\delta(\infty) = 0$ in order to address the absence of traffic in a situation where all content is fully throttled and $\delta(0) = 1$ in order to address that there is no impact of throttling without throttling. Thus, $\bar{T} = \bar{x}$ is the unrestricted peak traffic that occurs if there is no throttling and congestion in the mobile network. Note that consumers do not anticipate their impact on total peak traffic. This assumption is in line with the related literature (Economides and Hermalin, 2015; Choi and Byung-Cheol, 2010) and with reality where consumers' information regarding their weight on mobile cells' capacity in relevant area is very limited.

The MNO is assumed to incur costs from providing network capacity represented by a quadratic cost function. Other costs are assumed to be zero. Further, the MNO charges an

access price p to each consumer. Consequently, the MNO's profit function is

$$\pi = \begin{cases} p - k^2 & \text{if } p \leq U \\ 0 - k^2 & \text{else.} \end{cases} \quad (4)$$

The timing of the model is as follows:

Stage 1: The MNO sets network capacity k , throttling degree t , and access price p .²

Stage 2: Consumers decide to buy or not to buy access to the mobile network and use it depending on the QoE offered.

3.2 Equilibrium Derivation

The game is solved through backwards induction. As consumers are homogeneous, all of them will buy at Stage 2 if and only if $U \geq p$. Hence, the MNO's profit maximizing price at Stage 1 is $p^* = U$. Inserting this result into the profit function yields the following maximization problem regarding the choice of t and k :

$$\max_{t,k} (\pi = (x(t, \Delta(t, k)) \Delta(t, k) \Theta_E(t) - k^2)) \quad (5)$$

First, note that the interior solution of Equation (5) is $t^* = \arg \max_t \left(k \frac{\Theta_E(t)}{\delta(t)} \right)$ and $k^* = \frac{1}{2} \frac{\Theta_E(t^*)}{\delta(t^*)}$. Second, note that $k > \bar{T}$ is never a reasonable choice which implies the existence of cases with a corner solution $k^* = \bar{T}$. In such a case the profit maximizing throttling level is $t^* = 0$, as a higher degree of throttling provides no advantage, but would strictly lower consumer's willingness to pay. In order to specify the appearance of cases where the MNO chooses this corner solution, assume that there exists a border case $\bar{T} = \tilde{T}$ where the interior solution of Equation (5) with respect to k is equal to the corner solution $k = \bar{T}$. Considering that $k^* = \frac{1}{2} \frac{\Theta_E(t^*)}{\delta(t^*)}$, $t^* = 0$ and $k^* = \bar{T}$ hold for $\bar{T} = \tilde{T}$, yields $\bar{T} = k^* = 1/2$,

²By the Envelope Theorem, the results do not change if capacity is chosen in a previous stage.

i.e., the border case is characterized by $\bar{T} = \tilde{T} = 1/2$. In summary, the equilibrium price, equilibrium throttling level and equilibrium capacity investments are given by:

Lemma 1.

$$p^* = U(t^*, k^*), \quad (6)$$

$$t^* = \begin{cases} \arg \max_t \left(k^* \frac{\Theta_E(t)}{\delta(t)} \right) & \text{if } \bar{T} > 1/2 \\ 0 & \text{else,} \end{cases} \quad (7)$$

$$k^* = \begin{cases} \frac{1}{2} \frac{\Theta_E(t^*)}{\delta(t^*)} & \text{if } \bar{T} > 1/2 \\ \bar{T} & \text{else.} \end{cases} \quad (8)$$

3.3 Interpretation and Results

The first insight is that there is a case $\bar{T} \leq 1/2$ where the MNO will never use throttling. Thus, if the unrestricted peak traffic is sufficiently low, throttling will not be an issue. However, the more interesting case is $\bar{T} > 1/2$. Equation (7) shows, that some positive degree of throttling is profit maximizing if and only if $\frac{\Theta_E}{\delta}$ has its global maximum at $t \in (0, \infty)$.³ Note that $t = 0$ results in $\frac{\Theta_E}{\delta} = 1$, i.e., the MNO will only choose a throttling degree other than $t = 0$ if this yields $\frac{\Theta_E}{\delta} \geq 1$. In other words, an MNO will use throttling if and only if the negative direct effect of throttling on customers' QoE, Θ_E , is outweighed by the positive effect on customers' QoE, δ , due to less congestion. This result is summarized in the following proposition:

Proposition 1. *An MNO uses throttling if and only if it can thereby alleviate consumers' disutility from congestion in its mobile network.*

³Note that $t = \infty$ results in $\bar{T} \leq k$, where $t = \infty$ is weakly dominated by $t = 0$.

Correspondingly, Equation (8) shows that the ability to throttle never yields lower but rather higher incentives to invest in capacity. In particular, as the MNO chooses t in a way that maximizes $\frac{\Theta_E}{\delta}$, its choice will also maximize its incentive to invest in the network capacity.

Proposition 2. *An unrestricted MNO will always choose the throttling level that yields the highest incentives to invest in network capacity.*

Restricting the MNO's choice to use throttling will therefore never lead to an improvement in the overall network capacity, but rather bears the risk of reducing the incentive to invest in network capacity. The intuition behind this at a first glance counter-intuitive result is that each unit of network capacity is used more efficiently through throttling and becomes therefore more valuable, i.e., throttling increases the utility per capacity unit and thus consumers' willingness to pay per capacity unit rises. Consequently, the marginal revenue per capacity unit increases which makes it profitable for the MNO to incur more costs from capacity investments.

4 Conclusion

Summary and Theoretical Contribution: This paper focuses on the incentives for an MNO to invest in network capacity, if it is allowed to throttle data streams. By developing a stylized game-theoretic model with an MNO providing access to the Internet for its customers, it has been shown that obligations like the EU's regulation of net neutrality that prohibit throttling practices do not yield higher incentives for MNOs to invest in mobile network capacity but rather bear the risk to decrease these incentives. This is because MNOs may be able to offer a better quality of experience for a given network capacity by throttling. Such a better quality of experience leads to a higher willingness to pay of consumers, so that the MNO can enforce higher prices and therefore, achieves a better monetization of its net-

work. This leads to a higher incentive to invest in additional capacity in order to monetize this capacity as well.

In addition, the results show that the MNO uses throttling to improve the overall quality of experience of consumers. This is done by balancing the positive impact on congestion against the negative impact of lower content resolution, with the MNO choosing a solution that generates the most profitable outcome in terms of consumers' willingness to pay.

Managerial Implications: The model shows that throttling should primarily be understood as an instrument to improve the overall quality of the customers' experience. In particular, the negative effects of throttling on experience due to lower quality, e.g., in video resolution, should be balanced against the positive effects of less severe and less frequent congestion in the mobile network. A higher quality of experience may create scope for higher customer prices. This is not only beneficial in terms of current prices but should also be considered when scaling future investments in network capacity. It is expected that investments in network capacity will be more profitable in networks where throttling is used properly than in networks where no throttling is used, although throttling would improve the overall quality of experience for customers.

Limitations and Future Work: Note that the model has a limitation in terms of not taking into account competition between MNOs. However, even in a monopoly situation throttling is used by the MNO in order to alleviate consumers' disutility from network congestion, i.e., to make the mobile network more attractive. It is obvious that this cause is even more important under competition. Thus, the motivation of MNOs to use throttling persists if competition is taken into account and consequently the qualitative effects remain the same.

A further limitation is the assumption that consumers' willingness to pay is homogeneous, which means that the MNO's revenue and the consumers' willingness to pay are perfectly aligned. However, this assumption has no qualitative impact on the results. A heterogeneous

willingness to pay would at worst lead to a situation where the MNO would be less effective in monetizing improvements in the consumers' quality of experience. This will not affect the result that throttling is used in a way that improves consumers' overall quality of experience. The only effect would be that the increase in the incentive to invest in capacity generated by throttling would be less significant.

A somehow related assumption of the model is that consumers are homogeneous in their sensitivity to the effects of throttling on quality of experience. This is of special importance for future research investigating MNOs' ability to use throttling as an instrument for price discrimination. A paper of Zhai et al. (2018) on throttling in the context of cloud computing implies that such a discriminatory intention can be the reason why capacity is artificially scarce by the MNO in order to sufficiently differentiate a lower quality offer for consumers with a lower willingness to pay from a higher quality offer for consumers with a higher willingness to pay. The same mechanism is identified by Baake and Sudaric (2019) in the context of net neutrality with a focus on paid prioritization. However, Baake and Sudaric (2019) find that the result of Zhai et al. (2018) only holds if capacity is not scarce. In contrast, when capacity investments are sufficiently costly, the ability to discriminate works in favor of capacity investments. This positive effect once again results from a better monetization of each capacity unit. One can therefore expect that the ability to discriminate through throttling will usually amplify the effects presented in this paper.

References

- Baake, P. and S. Sudaric (2019). Net neutrality and CDN intermediation. *Information Economics and Policy* 46, 55–67.
- Bandyopadhyay, S. and H. K. Cheng (2006). Liquid pricing for digital infrastructure services. *International Journal of Electronic Commerce* 10(4), 47–72.

- Bourreau, M., F. Kourandi, and T. Valletti (2015). Net neutrality with competing Internet platforms. *The Journal of Industrial Economics* 63(1), 30–73.
- Choi, P. J. and K. Byung-Cheol (2010). Net neutrality and investment incentives. *The RAND Journal of Economics* 41(3), 446–471.
- Dewenter, R. and J. Rösch (2016). Net neutrality and the incentives (not) to exclude competitors. *Review of Economics* 67(2), 209–229.
- Easley, R. F., H. Guo, and J. Krämer (2018). From net neutrality to data neutrality: A techno-economic framework and research agenda. *Information System Research* 29(2), 253–272.
- Economides, N. and B. E. Hermalin (2012). The economics of network neutrality. *Rand Journal of Economics* 43(4), 602–629.
- Economides, N. and B. E. Hermalin (2015). The strategic use of download limits by a monopoly platform. *Rand Journal of Economics* 46(2), 297–327.
- Economides, N. and J. Tåg (2012). Network neutrality on the Internet: A two-sided market analysis. *Information Economics and Policy* 24(2), 91–104.
- Guo, H., S. Bandyopadhyay, H. K. Cheng, and Y.-C. Yang (2010). Net neutrality and vertical integration of content and broadband services. *Journal of Management Information Systems* 27(2), 243–276.
- Guo, H., H. K. Cheng, and S. Bandyopadhyay (2013). Broadband network management and the net neutrality debate. *Production and Operations Management* 22(5), 1287–1298.
- Guo, H. and R. F. Easley (2016). Network neutrality versus paid prioritization: Analyzing the impact on content innovation. *Production and Operations Management* 25(7), 1261–1273.

- Inceoglu, F. and X. Liu (2019). Multiproduct price discrimination with quantity limits: An application to zero-rating. *Economics Letters* 180, 41–45.
- Krämer, J. and L. Wiewiorra (2012). Network neutrality and congestion sensitive content providers: Implications for content variety, broadband investment, and regulation. *Information Systems Research* 23(4), 1303–1321.
- Krämer, J., L. Wiewiorra, and C. Weinhardt (2013). Net neutrality: A progress report. *Telecommunications Policy* 37(9), 794–813.
- Meyer, D. (2017). Net neutrality: Key zero rating decision made by germany. ZDNet. Retrieved from <https://www.zdnet.com/article/net-neutrality-key-zero-rating-decision-made-by-germany/>, accessed on 21/11/2018.
- Peitz, M. and F. Schütt (2016). Net neutrality and inflation of traffic. *International Journal of Industrial Organization* 46, 16–62.
- Qishin, T. (2018). U mobile’s giler postpaid and prepaid plans offer unlimited mobile data. thestar.com.my. Retrieved from <https://www.thestar.com.my/tech/tech-news/2018/06/26/umobile-launches-unlimited-mobile-internet-plans/>, accessed on 21/11/2018.
- Saunders, E. (2018). Telkom mobile data – A lot more data behind the scenes. fortressofsolitude.co.za. Retrieved from <https://www.fortressofsolitude.co.za/telkom-mobile-data/>, accessed on 21/11/2018.
- T-Mobile USA (2018). Cell phone data plans. t-mobile.com. Retrieved from <https://www.t-mobile.com/cell-phone-plans>, accessed on 21/11/2018.

- Welch, C. (2018). Unlimited data plans are a mess: Here's how to pick the best one. The Verge. Retrieved from <https://www.theverge.com/2018/6/28/17507322/best-unlimited-data-plans-verizon-att-tmobile-sprint>, accessed on 21/11/2018.
- Wu, T. (2003). Network neutrality, broadband discrimination. *Journal on Telecommunications and High Technology Law* 2, 141–179.
- Zhai, Y., M. Stinchcombe, and A. Whinston (2018). Service throttling in cloud computing. Retrieved from <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1350&context=icis2018>, accessed on 21/11/2018.

Data Sharing between Firms

Experimental Evidence

Authors Jan Krämer, University of Passau, Germany.
Nadine Stüdlein, University of Passau, Germany.
Oliver Zierke, University of Passau, Germany.

Presented at² DIW-based internal seminar / reading group, 2019, Berlin, Germany.

Abstract

B2B data markets, such as the ‘Data Intelligence Hub’ by Deutsche Telekom or the ‘Open Data Initiative’ by Adobe, Microsoft, and SAP seek to facilitate the sharing of data between firms. Data sharing is believed to promote firms’ ability to innovate and compete. However, many of these initiatives struggle to get off the ground. Although firms find it advantageous to use the data of other firms, they are reluctant to share their own data because they fear that this data could be used to their disadvantage. Robust economic insights on how to design B2B data markets to facilitate data sharing in such an environment are scarce. To this end, we investigate control over who among potential sharing partners receives how much data as an important design parameter. By conducting an economic laboratory experiment, we verify our expectation on control and find that more control improves data sharing.

²The list does *not* include presentations held by co-authors.

Data Sharing between Firms: Experimental Evidence

Jan Krämer* Nadine Stüdlein† Oliver Zierke‡

May 19, 2021

Abstract

B2B data markets, such as the ‘Data Intelligence Hub’ by Deutsche Telekom or the ‘Open Data Initiative’ by Adobe, Microsoft, and SAP seek to facilitate the sharing of data between firms. Data sharing is believed to promote firms’ ability to innovate and compete. However, many of these initiatives struggle to get off the ground. Although firms find it advantageous to use the data of other firms, they are reluctant to share their own data because they fear that this data could be used to their disadvantage. Robust economic insights on how to design B2B data markets to facilitate data sharing in such an environment are scarce. To this end, we investigate control over who among potential sharing partners receives how much data as an important design parameter. By conducting an economic laboratory experiment, we verify our expectation on control and find that more control improves data sharing.

Keywords: Data sharing, B2B data markets, laboratory experiment

JEL Codes: C91, L10, L20

1 Introduction

Data emerges as the essential resource in the 21st century and is increasingly important to our economy and society. Data has become vital to many business models. It allows targeted and personalized advertisements in online retailing and is used to develop and enhance artificial intelligence as well as advancing and reshaping Internet of Things technologies decisively

*Chair of Internet and Telecommunications Business, University of Passau, Dr.-Hans-Kapfnger-Str. 12, 94032 Passau, Germany. Email: jan.kraemer@uni-passau.de. Declarations of interest: none

†Chair of Internet and Telecommunications Business, University of Passau, Dr.-Hans-Kapfnger-Str. 12, 94032 Passau, Germany. Email: nadine.stuedlein@uni-passau.de. Declarations of interest: none

‡Chair of Internet and Telecommunications Business, University of Passau, Dr.-Hans-Kapfnger-Str. 12, 94032 Passau, Germany. Email: o.zierke@yahoo.de. Declarations of interest: none

(Hall, 2020). In order to exploit the full potential of data, the accessibility of data is crucial (European Commission, 2020; Ministry of Economic Affairs and Climate Policy, 2019). To this end, the European Commission is planning legislative measures to improve access to and use the of data (European Commission, 2021). One of the major intentions of the proposed Data Act is to support data sharing between firms (European Commission, 2021). This is to be achieved by reviewing competition rules, clarifying the rights of use for jointly generated data and the legal liability of data use (European Commission, 2020). Besides, data initiatives are being developed to promote data sharing. One example is Gaia-X, an initiative supported by the German and French governments that aims to create a trusted and sovereign data infrastructure for Europe to ensure secure and reliable data sharing (Gaia-X, 2020).

However, although data sharing is essential in promoting firms' innovativeness and competitiveness (Ministry of Economic Affairs and Climate Policy, 2019; PWC, 2018; Telekom, 2019), many firms are reluctant to share their own data, as it makes firms potentially vulnerable through a loss of control over sensitive internal data (Norman, 2001; Otto, 2019; PWC, 2018) and their competitive edge (European Commission, 2020; Natwichai et al., 2005; Trkman and Desouza, 2012). In addition, there is uncertainty about sharing partner involvement and uncertainty about how the sharing partners will use the shared data, as well as other coordination issues that make finding qualified sharing partners difficult (Cabrera and Cabrera, 2002; European Commission, 2020; Loebbecke et al., 2016; Ministry of Economic Affairs and Climate Policy, 2019). Thus, many firms miss vital opportunities for their business activities and ultimately business performance.

In practice, there are already some data initiatives which seek to facilitate structured sharing of data among firms. While large international initiatives, such as Gaia-X, are still being developed, there are a number of smaller B2B data sharing platforms initiated by businesses. The Data Intelligence Hub, for example, is a marketplace for data and analysis tools by Deutsche Telekom which allows firms to conclude individual data sharing

contracts (Telekom, 2019). In other initiatives, such as the Open Data Initiative by Adobe, Microsoft, and SAP, firms share data on a collective cloud platform, that can be accessed by all participating firms (Microsoft, 2019).

Despite the varying designs, all these initiatives can be roughly distinguished by the degree to which a firm can control and individually limit other firms' data access. Drawing on the literature on public goods, n-person prisoner dilemmas and oligopoly competition, we formulate the expectation that more control will have a positive effect on the amount of shared data. We test this insight from literature by conducting an experiment and show that control is indeed positively correlated with overall data sharing.

Our study is related to a wide range of literature streams dealing with sharing of knowledge, information and data in the contexts of intra-firm knowledge sharing (see Wang and Noe (2010) for an overview), joint ventures (see Bogers (2011) for an overview), and supply chain optimization (see, e.g., Kembro et al. (2014), Lee and Whang (2000) and Zhou and Benton Jr (2007)). However, in all these scenarios, there is a dependency between the sharing partners working together in a firm, on a project or on a product. We abstract from the existence of such dependencies and consider scenarios in which every potential sharing partner can succeed even if its potential sharing partners do not. There are some closer related theoretical models on firms sharing consumer data. In the models of Jentzsch et al. (2013) and Krämer et al. (2019), firms can share data on consumers' taste and in the models of Zhao and Xue (2012) and Ghosh et al. (2015), firms can share data on consumers' value. In these models, firms share data only if their joint profit increases with data sharing and the profit is redistributed so that sharing partners are better off after data sharing. Moreover, these models assume that data sharing agreements are perfect contracts that leave no room for opportunistic behavior after the agreement is signed. In reality, however, it is possible to exploit sharing partners as data is usually not a homogeneous good whose quality and quantity can be clearly defined in advance. We will show that in a scenario without a binding agreement, the theoretical prediction changes to no data sharing. Nonetheless, our

experimental results show that data sharing occurs even in the absence of any agreements and that the amount of data shared depends on the control dimension from above.

The remainder of this paper is structured as follows: First, we develop a prediction about the impact of control based on experimental results in the areas of public good games, n-person prisoner dilemma games, and oligopoly competition games. Second, we develop an experiment to test our prediction that more control leads to more data sharing. Third, we present the results of our experiment and finally, we discuss the results and derive an overall conclusion.

2 Investigations on control in the experimental literature

There is no exact replication of control over data sharing in other experiments. However, control in the broader sense, i.e., control over the relationships to others, has been investigated in many experiments. We identify relevant literature by considering the fact that data sharing has several similarities with public good games, n-person prisoner's dilemma games and oligopoly competition games. In all these games, cooperation at the individual level is costly, since non-cooperation, while others cooperate, maximizes individual profit. Further, in all these games, cooperation at the industry level is beneficial, since the overall profit is higher if all entities cooperate compared to the situation where no one cooperates. Due to these similarities, we assume the empirical results of these games suitable for making predictions regarding the effect of more or less control.

One of the most common ways of investigating more control in such games is to enable costly targeted monetary punishment. There is a broad literature on public good games that shows that the ability to punish leads to more cooperation (see, e.g., Ostrom et al. (1992), Fehr and Gächter (2000), Fehr and Gächter (2002), Gürer et al. (2006), Gächter et al. (2008), Nikiforakis and Normann (2008), Reuben and Riedl (2013), Gächter et al. (2017)).

Further, there is similar evidence for prisoner's dilemma games (see, e.g., Dreber et al. (2008) and Bone et al. (2015)) and oligopoly competition games (see Roux and Thöni (2015)) indicating that more control, i.e., the ability to punish, generally increases cooperation. This connection breaks down only in a few exceptional cases. For example, Nikiforakis and Normann (2008) show that the effectiveness of the punishment tool is important, and Nikiforakis (2008) shows that the appearance of counter-punishments can cause a negative effect of punishment on cooperation. However, it is also important to point out that the overall effect of monetary punishment is ambivalent. Many studies show that the costs caused by punishment can outweigh the positive effect of enhanced cooperation, resulting in a negative net effect on overall profits (see, e.g., Ostrom et al. (1992), Fehr and Gächter (2002), Dreber et al. (2008), Gächter et al. (2008), Gächter et al. (2017)).

To avoid the negative effect of costly monetary punishment, it is worth looking at the literature on the costless non-monetary punishment tool of ostracism. Especially with regard to public good games, many studies have found that the possibility to ostracize non-cooperators is sufficient to increase contributions to the public good (see, e.g., Cinyabuguma et al. (2005), Güth et al. (2007) or Maier-Rigaud et al. (2010)). This is consistent with the result of Riedl and Ule (2002) in the context of an n-person prisoner dilemma experiment. In other words, to encourage cooperation, it seems sufficient to provide control over who to cooperate with. This is confirmed by the results of experiments comparing bilateral relationships with multilateral relationships. To see the connection, consider that bilateral relationships are characterized by a firm being able to selectively decide whether or not to cooperate with a particular other firm. In a multilateral relationship, a firm cannot discriminate, i.e., it cooperates with all other firms or with none. For n-entity prisoner's dilemma games Marwell and Schmitt (1972) find in their experiment that the cooperation is significantly higher in a game with a bilateral relationship than in a game with a trilateral relationship. In a more recent experiment, Nosenzo et al. (2015) find the same result for a public good game when comparing a game with a bilateral relationship with a game with three, four or eight entities. Further,

this is in line with evidence from oligopoly experiments with Cournot competition. Dolbear et al. (1968), Huck et al. (2004) and Orzen (2008) find markets with two firms to be more collusive than markets with more than two firms. This result is confirmed for Bertrand and Cournot competition by Horstmann et al. (2018) by means of a meta-analysis and additional experiments.

Moreover, the literature provides evidence that it is not only important whether a firm can control with whom it cooperates, but also whether it can respond in a graduated manner to the behavior of its cooperation partners. The reason is that such an even more differentiated type of control allows for graduated punishment. By comparing case studies, Ostrom (1990, pp. 97–99) finds that cooperation works better when it is possible to impose graduated punishment. In this way, the reasons and circumstances for deviation can be taken into account (Ostrom, 1990, pp. 98–99). This is confirmed by Wright (2013). In his experiment, he investigates which punishment strategies are successful under a Bertrand competition (Wright, 2013). He finds that tacit collusion between competitors works best in markets where firms punish a small price cut of a competitor proportionally with a similarly small price cut in the next round, rather than with a disproportionately large price cut (Wright, 2013). Accordingly, a disproportional strategy was used by only a few participants, while the majority used a proportional strategy (Wright, 2013).

In summary, the literature suggests that cooperation can be enhanced by giving firms control over who they cooperate with and further enhanced by allowing graduated responses based on the severity of cooperation partners' deviations. However, it cannot be assumed that the insights from the literature discussed simply carry over to the data context. Data sharing differs from public good games, n-person prisoner's dilemma games and oligopoly competition games in an important aspect. Cooperation in data sharing does not directly reduce profit at the individual level, but rather generates only indirect costs. In a prisoner's dilemma, the player directly receives a lower reward if he or she decides to cooperate. Similarly, an oligopoly firm directly reduces its profits if it opts for a collusive price or quantity rather

than choosing the best response. In a standard public good game, an entity receives a lower reward from the private good because it invests resources in the public good rather than the private good. However, in the case of data sharing, the profit of the firm is not directly affected by whether the data are shared or not, as the data remain available to the firm after sharing. Instead, data sharing incurs only indirect costs through the (potential) opportunistic use of the data by the sharing partners.¹

3 Experimental Design

3.1 Model

To capture the described specificity of data sharing, we develop a simple game-theoretical model. We consider a set of n firms, which can, in principle, share data with each other. Each firm is endowed with a set of (unique) data points. The firms can simultaneously decide how much of their data to share. We then make the following key assumptions on the value of shared data:

- The total profit of all firms is the higher, the more data the firms share.
- The profit of a firm is the higher, the more data the other firms share with it.
- The competitive advantage of a firm is greater, the more data the other firms have shared with it.

These assumptions reflect that each firm would profit from having access to more data (e.g., because it allows for data-driven innovations), but that sharing of data also bears indirect costs through the (potential) opportunistic use of the data by the sharing partners. Further, reflecting the believe of the European Commission, overall the economy (in terms of producers' surplus) would be better off if more data is shared between the firms (European

¹See, e.g., Trkman and Desouza (2012) for a more detailed analysis of the potential risks.

Commission, 2020). All of these assumptions are captured by the following profit function of a firm $i \in \{1, \dots, n\}$:

$$\Pi_i = m x_i + m_s \sum_{j=1; i \neq j}^n s_{ji} x_j - \frac{c}{n-1} \left[\sum_{j=1; i \neq j}^n \left(m x_j + m_s \sum_{k=1; k \neq j}^n s_{kj} x_k \right) \right] \quad (1)$$

where x_i is firm i 's data endowment, $s_{ij} \in [0, 1]$ is the percentage of data firm i shares with firm j , m is the revenue per unit from own data, m_s is the revenue per unit from data shared by other firms and c is the degree of competition with $c \in [0, 1]$. The first term in the profit function represents the revenue from the firm's own data; the second term represents the revenue from data shared by the other firms with firm i ; and the third term covers the competitive influence of other firms depending on the average amount of data to which they have access. Specifically, we model the profit of a firm as the revenue from the data available to it minus c times the average revenue of its competitors from the data available to them.

For the experiment we used the following parameter specifications: $n = 3$, $m = m_s = 1$, $x_i = 100 \forall i$ and $c = \frac{1}{3}$, which relate to a market with three firms and a medium level of competition.

Regarding the theoretical benchmark consider that in our model for $c > 0$ the unique subgame perfect Nash equilibrium for an n -period game is not to share any data. Thus, in terms of the theoretical solution, our model is again not much different from public good games where the entities in equilibrium do not contribute, from prisoner's dilemma games where in equilibrium all players defect, or from oligopoly competition games where the firms play the one-shot Nash equilibrium instead of colluding. For all these games, however, it is well known that experimenters usually find a certain degree of cooperation. Consequently, we also expect to observe a degree of cooperation in the case of data sharing that exceeds the Nash prediction. Moreover, we expect that the amount of data sharing will depend on the kind of control that is granted to the firm.

3.2 Treatments

We identify three important levels of control in the context of data sharing. At the lowest level of control, called *Collective Sharing*, a firm has no differentiated control over who accesses its shared data, i.e., each firm can only share the same percentage of its data with all other firms. This reflects situations where a firm uploads part of its data to a data pool that provides unrestricted data access to all participating firms. At the medium level with slightly more control, called *Excludable Sharing*, the firm may fully exclude certain firms from access to its data, i.e., the firm specifies how much data to be shared and decides which of the potential sharing partners gets access to it. This reflects situations where a data pool allows firms to decide who can access their uploaded data, or where the firm licenses interested parties before granting access to part of its database via an API. At the highest level of control, called *Individual Sharing*, a firm is able to decide independently with whom it shares how much of its data, i.e., firms can set the amount of data to be shared specifically for each firm. It reflects a situation where two firms share data bilaterally.

4 Procedures

The experiment includes twelve sessions, four sessions for each treatment, with a total of 177 participants.² All sessions have been conducted in a laboratory at the School of Business, Economics and Information Systems at the University of Passau, Germany in January and February 2020. Subjects were recruited from the student subject pool of the University of Passau via the ORSEE platform (Greiner, 2015). In a session 15 participants were matched randomly and anonymously to groups of three firms each, with each participant representing one firm. There were 30 rounds per session. The participants of a group played together in a market over the whole 30 rounds (fixed partner matching). In all treatments, subjects were fully informed about the timeline of the experiment and the consequences of their actions.

²In one session the usual number of 15 participants was reduced to twelve because not enough participants showed up.

A session comprised seven steps. First, participants were randomly assigned to a seat. Subjects could neither see nor communicate with each other. Throughout the experiment, participants could take notes on a notepad provided for them. Second, the experimental instructions were handed out in print and read aloud from recording. The recording ensures that the voice, accent and emphasis were identical in each session. Third, before starting the actual experiment, the participants had to answer a series of comprehension questions. Fourth, after all participants had successfully answered the questions on the instructions, they were shown a video explaining the user interface of the experimental software. The video clearly illustrates the processes in the experiment and familiarizes participants with the user interface through which they submit their decision on how much data to share with the other firms, and in which they are presented information on past actions, as well as a profit calculator. Fifth, participants proceeded to the first round of the game without prior practice rounds in order to prevent unobserved learning effects. However, through the game, participants were able to access the profit calculator where they had the opportunity to check for any possible constellation of data sharing between the three firms, what the available data and the profits of all firms would be. Once the participants had made their decision, they were redirected to a waiting page. On this page they waited until all participants had finished the round. During this time, it was still possible to use the profit calculator. After all firms had determined the data points they wanted to share, the participants saw the results, i.e., the decisions, profits and available data from the current round, illustrated in a graph. Across all treatments, it took about an hour to complete the 30 rounds of playing the game. Sixth, at the end of the game, the subjects had to answer some post-questions. Finally, the subjects were paid out the profit of his or her firm from exactly one of the 30 rounds. The payout of the profit from exactly one round avoids budget effects. Each participant determined the round to be paid out by throwing a 30-sided dice. The profit their firm had made at the end of this round (measured in Thalers) was converted into Euro; 10 Thalers correspond to one Euro. Finally, the profit from the selected round in addition to a participation fee of 10

Euro was paid to the subjects privately and in cash. The duration of a session from entering the laboratory to the end of payment was about two hours. The average payout was 24.53 Euro.

5 Results

Our analysis focuses on how much data is shared in a given market. Therefore, in the following we treat the average contribution in a group in a single round as an observation from a particular market. Consequently, each observation entails three firms, each of which decides how much data the other two firms in the market receive. Note that there is some evidence in our data of a learning or coordination phase in the first rounds and an endgame effect in the last rounds. In order to avoid distortions caused by such adjustment and endgame effects, we do not take the first and last two rounds into account.³ Figure 1 provides an overview of the resulting data. The histogram shows how much data was shared on average from period 3 to 28 in each particular market, ranging from 0 (no firm in the market has shared any data in any round) to 100 (all firms in the market have shared all their data in all rounds).

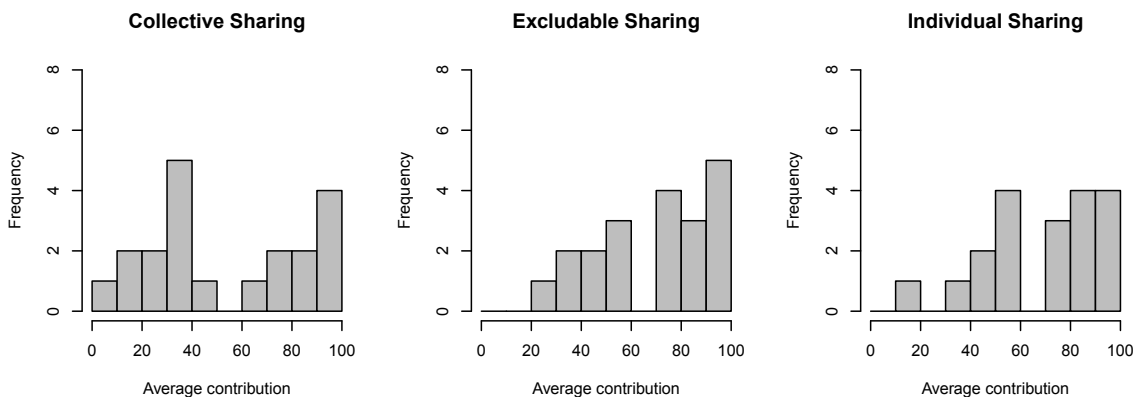


Figure 1: Histograms of the average data sharing of the markets.

³Table 5 in subsection 5.3 shows that the endgame-effect is strongest in the last two rounds.

Table 1: Summary statistics

Treatment	Obs.	Mean	SD
Collective	20	54.5	33.2
Excludable	20	69.5	23.5
Individual	19	67.1	23.7

One-Way ANOVA In order to test for treatment differences, we perform a one-way variance analysis (ANOVA). As in Figure 1, we use the average shared data from period 3 to 28 of each respective market. By doing so, we take into account that observations from the same market from different rounds are not independent. The result confirms differences between treatments on a 10% significance level ($F(2, 43) = 2.73, p = .08, \eta_p^2 = .11$). Thus, control has a significant impact on the amount of data shared in a market.

Result 1. *Control has an impact on the average amount of data shared in a market.*

Tukey post-hoc test However, a Tukey post-hoc test identifies no significant differences between the treatments (Collective Sharing - Excludable Sharing: $p = .12$, Collective Sharing - Individual Sharing: $p = .13$, Excludable Sharing - Individual Sharing: $p = 1$). This results from the data having some structural peculiarities. We will address this in the following sections by conducting a more detailed analysis based on the formation of reasonable sub-samples.

5.1 Developing sub-sample boundaries

A first notable observation is that the characteristics of data contributions in first half of the experiment differ from those in the second half. Contributions in the markets grew from an average of 58.7 data points in the entire first half to an average of 68.7 data points in the entire second half. A Mann-Whitney U-Test based on average market contributions over all rounds confirms this difference to be significant at a significance level of 5% ($U = 1330.5, Z = -1.922205, p = .03$). Furthermore, average market contributions are more

volatile in the first half than in the second half. By calculating the standard deviation of the average market contributions for each market from rounds 3 to 15 and 16 to 28, we obtain an average standard deviation of 11.95 for the first half and 7.45 for the second half. A Mann-Whitney U-Test confirms the difference again to be significant at a 1% significance level ($U = 2419, Z = -3.472727, p = .0003$). Thus, we observe that cooperation, i.e., higher and more stable data sharing, especially emerges in the second half of the experiment. This indicates that a distinction should be made between long-lasting business relationships and new business relationships.

Result 2. *On average, more data is shared in long-lasting business relationships than in new business relationships. Further, the amount of shared data is less volatile in long-lasting business relationships.*

Based on these insights, we conduct a round split in the middle of the game, i.e., in the following we analyze the first and second half of the game separately. Thus, we have the **early-rounds sub-sample** covering the played rounds 3-15 and the **late-rounds sub-sample** covering the other half of the played rounds minus the endgame-effect, i.e., rounds 16-28.

A first insight into the characteristics resulting sub-samples is provided by Figure 2, which shows histograms of the average contributions per market, separated by early and late rounds and separated by treatments. The values on the x-axis range again from 0 (no firm in the market has shared any data) to 100 (all firms in the market have shared all of their data). Note that, especially in the late round sub-sample, there are a large number of highly cooperative markets whose average data contributions are close to the maximum of 100. By classifying a market as supra-cooperative when the average contribution of firms in the market over all rounds exceeds 90 (out of 100) data points, we find 37.3% supra-cooperative markets in the late round sub-sample and 13.6% supra-cooperative markets in the early round sub-sample. A Fishers Exact Test confirms this difference to be significant at a 1% significance level ($p = .005$)

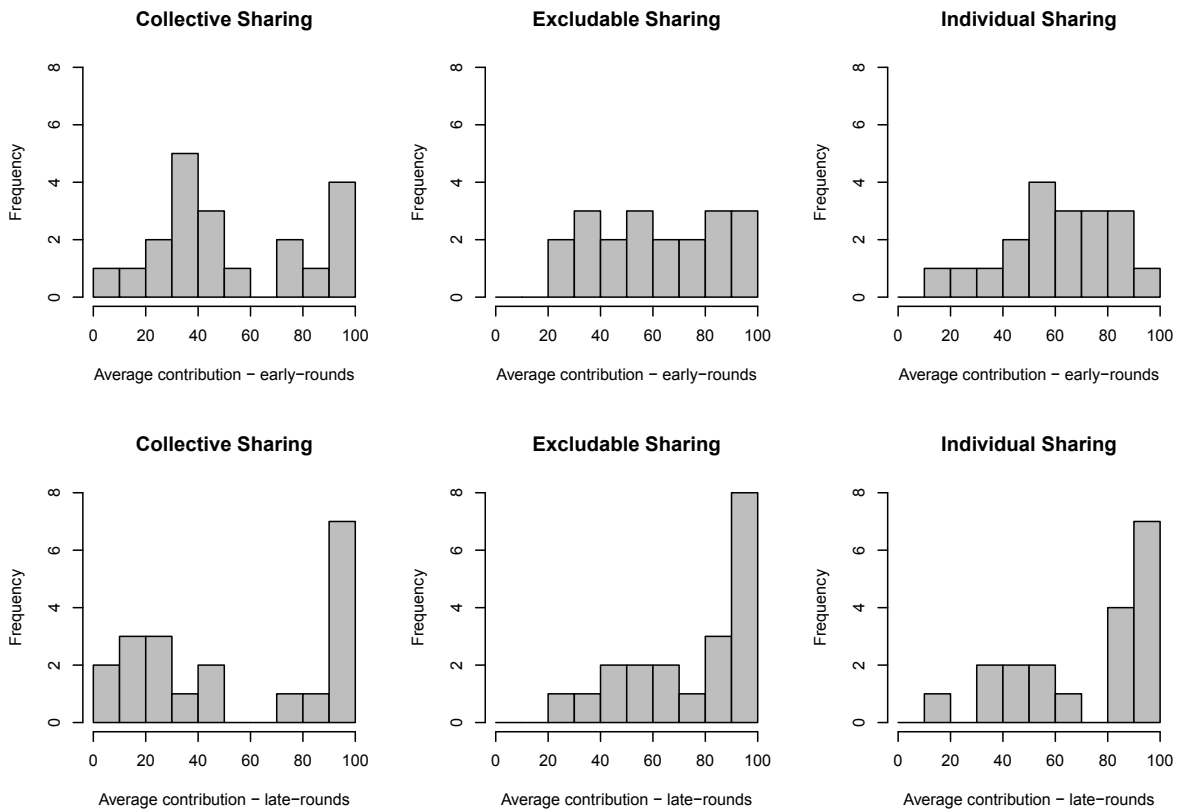


Figure 2: Histograms of the average data sharing of the markets in the early-rounds and late-rounds sub-sample.

Result 3. *Supra-cooperative markets are more likely to occur in long-lasting business relationships.*

Note, however, that there is no evidence for an impact of control on the probability of occurrence of supra-cooperative markets. A Fisher’s exact test⁴ shows that the null hypothesis that there is no relationship between the number supra-cooperative markets and control cannot be rejected (early-rounds sub-sample: $p = .51$, late-rounds sub-sample: $p = 1.00$).

Result 4. *There is no evidence for a relationship between the occurrence of supra-cooperative markets and control.*

⁴As the frequencies of the individual combinations are quite small, the Fisher’s exact test is used instead of the Chi-square test.

As supra-cooperative markets do not appear to depend on control, we focus on non-supra-cooperative markets in the remainder of our analysis. Table 2 shows the descriptive summary statistics of the resulting sub-samples.

Table 2: Summary statistics for non-supra-cooperative markets

Treatment	Round 3-15			Round 16-28		
	Obs.	Mean	SD	Obs.	Mean	SD
Collective	16	42.0	22.4	13	32.5	24.1
Excludable	17	55.8	21.1	12	61.9	19.7
Individual	18	58.9	21.2	12	59.1	24.6

5.2 Sub-Sample Analysis

One-Way ANOVA Based on the datasets from above, we again perform a One-Way ANOVA to test for differences in data sharing across treatments. To avoid dependencies between observations from the same market, we use the average contribution of a market from the sub-sample of early rounds and late rounds, respectively. We find an effect for the early-rounds sub-sample at a 10% significance level ($F(2, 48) = 2.89, p = .07, \eta_p^2 = .11$) and we find an effect for the late-rounds sub-sample at a 1% significance level ($F(2, 34) = 6.32, p = .005, \eta_p^2 = .27$), indicating a significant difference between the treatments.

Result 5. *In non-supra-cooperative markets, control has an impact on the amount of data shared in a market.*

Tukey post-hoc test In order to investigate which treatments differ significantly in the individual sub-samples, we perform a post-hoc test, using the Tukey-Test. The results in Table 3 show that only *Collective Sharing* differs significantly from *Individual Sharing* in the early-rounds sub-sample. In *Individual Sharing* on average of 16.9 more data points were shared than in *Collective Sharing*. In the late-rounds sub-sample, *Collective Sharing* differs significantly from *Excludable* and *Individual Sharing*, while there is no significant difference

between *Excludable* and *Individual Sharing* (see Table 3). In *Excludable Sharing* the data contributions were on average 29.4 data points and in *Individual Sharing* 26.6 data points higher than in *Collective Sharing*.⁵

Table 3: p-values from Tukey-Test to test for impact of control.

	Round 3-15		Round 16-28	
	Collective	Excludable	Collective	Excludable
Excludable	.17		.008***	
Individual	.07*	.90	.02**	.95

Notes: ***p < .01; **p < .05; *p < .1;

Result 6. *In new business relationships, significantly more data is shared in Individual Sharing than in Collective Sharing, while there is no significant difference between Collective Sharing and Excludable Sharing, nor between Excludable Sharing and Individual Sharing.*

Result 7. *In long-lasting business relationships, significantly more data is shared in Excludable Sharing and Individual Sharing than in Collective Sharing, while there is no significant difference between Excludable Sharing and Individual Sharing.*

5.3 Robustness Check with the Zero-Inflated Poisson Model

So far, the analyses only considered how much data in a market was shared on average over several rounds. This takes into account that different rounds played in the same market are not independent of each other. In the following analysis we use the so-called *Zero-Inflated Poisson Model* (see Lambert (1992) and Moffatt (2015, p. 257)). This more sophisticated approach allows us to incorporate the results of each round separately into the analysis and to take into account the interdependence between the rounds played in a market by using clustered standard errors. This analysis corroborates our results.

⁵In Appendix 7.1, we replicate the analysis by performing a Kruskal-Wallis test, followed by a Mann-Whitney U-post-hoc test (see Bortz (2005, p.287)), showing that these results are robust.

The Zero-Inflated Poisson Model is designed to analyze truncated data with excessive occurrence of zeros (Lambert, 1992). For this purpose, the data is separated into two sets. One set contains all non-zero observations plus the zero observations, which can be explained by the distribution of the non-zero observations (Lambert, 1992). The other set, the so-called "excessive" zeros, contains all zero observations that cannot be explained by the distribution of the non-zero observations (Lambert, 1992). In our data, however, we have an excessive number of observations at the upper bound where all data have been shared. Consequently, to use the Zero-Inflated Poisson Model, we invert the value of contributions. In particular, we calculate $Not\ Shared\ Data = 100 - Shared\ Data$, i.e., the dependent variable in our Zero-Inflated Poisson Model measures how much data has not been shared. Thus, the analysis reflects the previous analysis by addressing the problem of supra-cooperative markets through the separate consideration of the excessive zero observations.

Each regression (Model 1-5) in the resulting analysis in Table 4 and Table 5 entails two estimations. One estimation, the zero model (see Table 4), is based on a Logit model and estimates the probability that an observation is an excessive zero. Positive and significant coefficients indicate factors that are correlated with a higher probability that an observation is an excessive zero, and negative and significant coefficients indicate factors that are correlated with a lower probability that an observation is an excessive zero. The other estimation, the count model (see Table 5), is based on a Poisson regression and estimates the amount of shared data for all observations that are not excessive zeros. For the given subsample, a positive coefficient indicates a positive correlation with sharing less data, while a negative coefficient indicates a negative correlation with sharing less data.

In Table 4 Model 1-3, the coefficients *zero Excludable* and *zero Individual* measure whether excessive zeros are more likely to occur under the treatments *Excludable Sharing* and *Individual Sharing* than under *Collective Sharing*. In Model 4 and 5, we use *Individual Sharing* and *Excludable Sharing* instead of *Collective Sharing* as base treatment. *zero Round* and *zero Round²* control for round effects and the dummy variables *zero Endgame 26-30* con-

Table 4: Zero Models

	Excessive observation of full data sharing				
	(1)	(2)	(3)	(4)	(5)
zero (Intercept)	-1.190*** (0.450)	-3.759*** (0.808)	-2.726*** (0.603)	-1.584*** (0.404)	-0.995*** (0.368)
zero Excludable	0.195 (0.581)	0.211 (0.628)	0.214 (0.639)	0.590 (0.546)	
zero Individual	-0.395 (0.604)	-0.420 (0.642)	-0.427 (0.651)		-0.590 (0.546)
zero Collective				0.395 (0.604)	-0.195 (0.581)
zero Round		0.279*** (0.055)	0.099*** (0.015)		
zero Round ²		-0.006*** (0.001)			
zero Endgame Round 26			-0.170 (0.147)		
zero Endgame Round 27			-0.269* (0.153)		
zero Endgame Round 28			-0.440*** (0.149)		
zero Endgame Round 29			-0.920*** (0.247)		
zero Endgame Round 30			-2.832*** (0.544)		

Notes: ***p < .01; **p < .05; *p < .1;

trol for endgame effects in rounds 26-30. The result shows that we cannot reject the null hypothesis that the probability of observing an excessive zero under *Excludable Sharing* or *Individual Sharing* is equal to the probability of observing an excessive zero under *Collective Sharing*. This holds regardless of whether we control for round and end game effects (Model 2-3) or not (Model 1). Further, there is no significant difference between *Excludable Sharing* and *Individual Sharing* (Model 4-5). In other words, the model cannot explain the excessive occurrence of cases where all data in a market have been shared with differences in control. This is consistent with Result 4 where we find no effect of control on the occurrence of supra-cooperative markets.

However, Table 5 shows that differences in control can explain the amount of shared data in the subset of observations that are not part of the excessive zeros. Note that in Model 1-3 *count Excludable* and *count Individual* measure the difference of these treatments to

Table 5: Count Models

	Average data in a market that is not shared				
	(1)	(2)	(3)	(4)	(5)
count (Intercept)	4.116*** (0.083)	4.290*** (0.093)	4.162*** (0.083)	3.770*** (0.103)	3.832*** (0.077)
count Excludable	-0.284** (0.113)	-0.285** (0.113)	-0.292** (0.115)	0.062 (0.129)	
count Individual	-0.346*** (0.132)	-0.344*** (0.132)	-0.349*** (0.134)		-0.062 (0.129)
count Collective				0.346*** (0.132)	0.284** (0.113)
count Round		-0.039*** (0.014)	-0.006 (0.005)		
count Round ²		0.001*** (0.0004)			
count Endgame Round 26			0.020 (0.072)		
count Endgame Round 27			0.069 (0.075)		
count Endgame Round 28			0.102 (0.070)		
count Endgame Round 29			0.193** (0.083)		
count Endgame Round 30			0.611*** (0.118)		

Notes: ***p < .01; **p < .05; *p < .1;

Collective Sharing, respectively. In Model 4 and 5, we use *Individual Sharing* and *Excludable Sharing* as the base treatment instead. Again, *count Round* and *count Round²* control for round effects and the dummy variables *zero endgame 26-30* control for endgame effects. The results show at a significance level of 5% that firms in the *Collective Sharing* treatment share less data compared to the *Excludable Sharing* treatment and at a significance level of 1% that firms in the *Collective Sharing* treatment share less data compared to the *Individual Sharing* treatment. This holds true even if we control for round and endgame effects. We find no significant difference between *Excludable Sharing* and *Individual Sharing*.

In summary, the results confirm Results 6 and 7 that in non-supra-cooperative markets *Individual Sharing* and *Excludable Sharing* yield higher contributions than *Collective Sharing*. Furthermore, the existence of an increasing round effect (see Result 2) and the existence of an end game effect in the last two periods is confirmed.

6 Discussion and Conclusion

Although it is widely believed that more data sharing between firms would be beneficial for the industry as a whole, the success of projects to support such data sharing is currently limited. In this context, we identified control over the accessibility of shared data as one of the key design elements for such projects. Our experimental results confirm that control over who can access shared data promotes the success of data sharing. More specifically, *Collective Sharing* leads to less data sharing than *Individual Sharing* in new relationships, and *Collective Sharing* leads to less data sharing than *Individual Sharing* and *Excludable Sharing* in established relationships. Accordingly, a system where all firms have unrestricted access to shared data seems to be weakly dominated by systems where firms have more control over who exactly can access their shared data.

However, in contrast to results from comparable literature that does not explicitly address data sharing, we did not find that the ability to react gradually to small deviations of the cooperation partners is important, i.e., we found no significant difference between *Excludable Sharing* and *Individual Sharing*. Consequently, our experiment provides no evidence for the assumption that, e.g., a system in which firms can license other firms to access their shared data via a standardized API are worse than a system where firms grant individual access to a dedicated share of their data. With respect to data sharing platforms, our results do not indicate that it matters whether they allow firms to upload a data package and only give them control over who gets access to that data package, or whether the platform allows firms to control exactly what part of the data package which other firm can access. Further research is needed to determine whether this insight remains robust in other sharing contexts and to explore the underlying reasons.

In addition, we observed that sharing behavior in new relationships are significantly different from sharing behavior in established relationships. Higher volatility and lower sharing volumes, as well as the less frequent occurrence of supra-cooperative markets, indicate that it takes time to establish a stable data sharing cooperation, regardless of how much control is

given to the firms. Thus, in reality, data sharing projects should not be expected to perform optimally immediately. Further, we could not find a correlation between the occurrence of supra-cooperative markets and control. This indicates that control is particularly important in markets where firms are reluctant to share data.

Note that our results in the experiment apply to situations without information restrictions and with a medium degree of competition. Future research should investigate whether these results also apply to other environments. However, our main result in terms of the overall importance of control is backed by a broad body of literature on public goods, prisoner dilemmas and oligopoly competition experiments. Thus, it is very likely that this main result will persist and it is therefore advisable to give firms control over who gets access to the data they share in order to favor overall data sharing

References

- Bogers, M. (2011). The open innovation paradox: Knowledge sharing and protection in R&D collaborations. *European Journal of Innovation Management* 14(1), 93–117.
- Bone, J. E., B. Wallace, R. Bshary, and N. J. Raihani (2015). The effect of power asymmetries on cooperation and punishment in a prisoner’s dilemma game. *PloS one* 10(1), e0117183.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Cabrera, A. and E. F. Cabrera (2002). Knowledge-sharing dilemmas. *Organization Studies* 23(5), 687–710.
- Cinyabuguma, M., T. Page, and L. Putterman (2005). Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics* 89(8), 1421–1435.
- Dolbear, F. T., L. B. Lave, G. Bowman, A. Lieberman, E. Prescott, F. Rueter, and R. Sherman (1968). Collusion in oligopoly: An experiment on the effect of numbers and information. *The Quarterly Journal of Economics* 82(2), 240–259.

- Dreber, A., D. G. Rand, D. Fudenberg, and M. A. Nowak (2008). Winners don't punish. *Nature* 452(7185), 348–351.
- European Commission (2020). Commission communication (EU): A European strategy for data. *Com(2020) 66*.
- European Commission (2021). European data governance. <https://digital-strategy.ec.europa.eu/en/policies/data-governance>. Accessed: 24.03.2021.
- Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *American Economic Review* 90(4), 980–994.
- Fehr, E. and S. Gächter (2002). Altruistic punishment in humans. *Nature* 415(6868), 137–140.
- Gächter, S., F. Mengel, E. Tsakas, and A. Vostroknutov (2017). Growth and inequality in public good provision. *Journal of Public Economics* 150, 1–13.
- Gächter, S., E. Renner, and M. Sefton (2008). The long-run benefits of punishment. *Science* 322(5907), 1510–1510.
- Gaia-X (2020). Gaia-X: A federated data infrastructure for Europe. <https://www.data-infrastructure.eu/GAIX/Navigation/EN/Home/home.html>. Accessed: 18.11.2020.
- Ghosh, A., M. Mahdian, R. P. McAfee, and S. Vassilvitskii (2015). To match or not to match: Economics of cookie matching in online advertising. *ACM Transactions on Economics and Computation (TEAC)* 3(2), 1–18.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association* 1(1), 114–125.
- Gürerk, Ö., B. Irlenbusch, and B. Rockenbach (2006). The competitive advantage of sanctioning institutions. *Science* 312(5770), 108–111.

- Güth, W., M. V. Levati, M. Sutter, and E. Van Der Heijden (2007). Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics* 91(5-6), 1023–1042.
- Hall, D. W. (2020). Growing the artificial intelligence industry in the UK. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf. Accessed: 19.02.2020.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6(2), 65–70.
- Horstmann, N., J. Krämer, and D. Schnurr (2018). Number effects and tacit collusion in experimental oligopolies. *The Journal of Industrial Economics* 66(3), 650–700.
- Huck, S., H.-T. Normann, and J. Oechssler (2004). Two are few and four are many: number effects in experimental oligopolies. *Journal of Economic Behavior & Organization* 53(4), 435–446.
- Jentzsch, N., G. Sapi, and I. Suleymanova (2013). Targeted pricing and customer data sharing among rivals. *International Journal of Industrial Organization* 31(2), 131–144.
- Kembro, J., K. Selviaridis, and D. Näslund (2014). Theoretical perspectives on information sharing in supply chains: A systematic literature review and conceptual framework. *Supply Chain Management* 19(5/6), 609–625.
- Krämer, J., D. Schnurr, and M. Wohlfarth (2019). Winners, losers, and facebook: The role of social logins in the online advertising ecosystem. *Management Science* 65(4), 1678–1699.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 34(1), 1–14.

- Lee, H. L. and S. Whang (2000). Information sharing in a supply chain. *International Journal of Manufacturing Technology and Management* 1(1), 79–93.
- Loebbecke, C., P. C. van Fenema, and P. Powell (2016). Managing inter-organizational knowledge sharing. *The Journal of Strategic Information Systems* 25(1), 4–14.
- Maier-Rigaud, F. P., P. Martinsson, and G. Staffiero (2010). Ostracism and the provision of a public good: Experimental evidence. *Journal of Economic Behavior & Organization* 73(3), 387–395.
- Marwell, G. and D. R. Schmitt (1972). Cooperation in a three-person prisoner’s dilemma. *Journal of Personality and Social Psychology* 21(3), 376.
- Microsoft (2019). The Open Data Initiative. <http://ddi.uni-paderborn.de/>. Accessed: 26.07.2019.
- Ministry of Economic Affairs and Climate Policy (2019). Dutch digitalisation strategy: Dutch vision on data sharing between businesses. <https://www.government.nl/binaries/government/documents/reports/2019/02/01/dutch-vision-on-data-sharing-between-businesses/Dutch+vision+on+data+sharing+between+businesses.pdf>. Accessed: 26.07.2019.
- Moffatt, P. G. (2015). *Experimetrics: Econometrics for experimental economics*. Macmillan International Higher Education.
- Natwichai, J., X. Li, and M. Orłowska (2005). Hiding classification rules for data sharing with privacy preservation. In *International Conference on Data Warehousing and Knowledge Discovery*, pp. 468–477. Springer.
- Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* 92(1-2), 91–112.

- Nikiforakis, N. and H.-T. Normann (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics* 11(4), 358–369.
- Norman, P. M. (2001). Are your secrets safe? knowledge protection in strategic alliances. *Business Horizons* 44(6), 51–61.
- Nosenzo, D., S. Quercia, and M. Sefton (2015). Cooperation in small groups: The effect of group size. *Experimental Economics* 18(1), 4–14.
- Orzen, H. (2008). Counterintuitive number effects in experimental oligopolies. *Experimental Economics* 11(4), 390–401.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.
- Ostrom, E., J. Walker, and R. Gardner (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review* 86(2), 404–417.
- Otto, B. (2019). Interview with Reinhold Achatz on “data sovereignty and data ecosystems”. *Business & Information Systems Engineering* 61(5), 635–636.
- PWC (2018). Data exchange as a first step towards economy. <https://www.pwc.de/en/digitale-transformation/data-exchange-as-a-first-step-towards-data-economy.pdf>. Accessed: 12.02.2019.
- Reuben, E. and A. Riedl (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior* 77(1), 122–137.
- Riedl, A. and A. Ule (2002). Exclusion and cooperation in social network experiments. *Mimeo, University of Amsterdam*.
- Roux, C. and C. Thöni (2015). Collusion among many firms: The disciplinary power of targeted punishment. *Journal of Economic Behavior & Organization* 116, 83–93.

- Telekom (2019). Data Intelligence Hub. <https://dih.telekom.net/en/>. Accessed: 12.02.2019.
- Trkman, P. and K. C. Desouza (2012). Knowledge risks in organizational networks: An exploratory framework. *The Journal of Strategic Information Systems* 21(1), 1–17.
- Wang, S. and R. A. Noe (2010). Knowledge sharing: A review and directions for future research. *Human Resource Management Review* 20(2), 115–131.
- Wright, J. (2013). Punishment strategies in repeated games: Evidence from experimental markets. *Games and Economic Behavior* 82, 91–102.
- Zhao, X. and L. Xue (2012). Competitive target advertising and consumer data sharing. *Journal of Management Information Systems* 29(3), 189–222.
- Zhou, H. and W. Benton Jr (2007). Supply chain practice and information sharing. *Journal of Operations Management* 25(6), 1348–1365.

7 Appendix

7.1 Kruskal-Wallis-Test and Mann-Whitney U-Test (main effects)

To confirm the results of the One-Way ANOVA and the Tukey-Test, we additionally perform a Kruskal-Wallis-Test on the early-rounds and late-rounds sub-sample for the observations with an average contribution below 90 data points. The result from the early-rounds sub-sample indicates an impact of control at a significance level of 10% (Chi square = 5.93, $p = .06$, $df = 2$). The result from the late-rounds sub-sample yields a Chi square = 10.26, $p = .006$, $df = 2$, indicating a significant impact on a 5% level.

In order to investigate which treatment levels differ in the respective sub-samples, we perform a post-hoc test, by using the Mann-Whitney U-Test. Additionally, we adjust the p-values based on Holm (1979) to obtain robust p-values for the pairwise comparisons of treatments. The results in Table 6 show that at a significance level of 5% *Collective Sharing* differs significantly from *Excludable Sharing* and *Individual Sharing* in the late-rounds sub-sample. Further, given a significance level of 10% *Collective Sharing* differs significantly from *Individual Sharing* in the early-rounds sub-sample.

Table 6: p-values from
Mann-Whitney U-Test for impact of control

	Round 3-15		Round 16-28	
	Collective	Excludable	Collective	Excludable
Excludable	.14		.01**	
Individual	.07*	.57	.02**	.67

Notes: *** $p < .01$; ** $p < .05$; * $p < .1$;