Reputational Concerns in Online Platform Competition

University of Oxford

Department of Economics

St Hilda's College



by **Emmanouil-Spyridon Perdikakis**[1]

in partial fulfilment of the requirements

for the degree of

Master of Philosophy in Economics

Oxford, June 12, 2020

[1]Wordcount: 18468

Reputational Concerns in Online Platform Competition

Emmanouil-Spyridon Perdikakis

June 12, 2020

Abstract

This paper examines online platforms' incentives to protect their users' privacy, when users value the latter and the platforms' incentives arise through purely reputational concerns. The main question addressed is whether the introduction of a competing platform will increase effort incentives, relative to the monopoly case. When revenue is linear in market share and users' utilities are given by the Hotelling specification, I find that effort is reduced in the duopoly equilibrium. This is because in a duopoly setting, the expected increase in market share achieved through a high reputation is reduced. When the rewards-to-reputation function is sufficiently concave, the value of an increase in market share becomes larger for lower market shares and competition can increase effort incentives. I show that this "concavity" effect arises as the effect of platform popularity, when users' preferences favour one platform relative to the other and as the effect of decreasing marginal value of data, when users' data is used to estimate population parameters. Allowing platforms to respond to their posterior reputations through additional choice variables increases the convexity of the rewards-to-reputation mapping, thus worsening the impact of competition on effort incentives. I demonstrate this intuition by analyzing two-sided platform competition.

Acknowledgements

I would first like to express my deep gratitude to my supervisor, Dr Margaret Meyer, for her unparalleled academic guidance and support throughout the year. It has been a true privilege to learn from her and I have benefited greatly from our lengthy and detailed discussions. I would also like to thank Prof Greg Taylor, for the very helpful meetings we had.

I gratefully acknowledge the Scholarship I was awarded by the Onassis Foundation, in order to pursue an MPhil Thesis on the "Economics of Information". I would also like to acknowledge the generous Scholarship I was awarded by the George and Victoria Karellias Foundation, to pursue postgraduate studies in Economics. I am immensely grateful to both Foundations and it is my honor to be their Scholar.

Finally, I need to thank my family and friends for their continuous support throughout the past two years. They have provided me with all the necessary motivation to come through the most challenging parts of this research process.

Contents

1 Introduction **1**

2 Related Literature **5**

3 Monopoly Platform **8**

3.1 Users 10

3.2 Equilibrium 11

4 Introducing a Competitor **13**

4.1 Comparative Statics for $e_M - e_D$ 19

5 Asymmetric Users' Distribution **20**

5.1 Strategic Complementarity 26

6 Estimating Population Parameters **29**

7 Two-Sided Competition **34**

7.1 Monopoly 36

7.1.1 Equilibrium 37

7.2 Duopoly 38

7.3 Responding to Reputation 41

8 Policy Application: Data Portability **43**

8.1 Modified Model 44

8.2 Poaching Users 46

9 A Model of Price Discrimination **49**

9.1 Monopoly 50

9.1.1 Platform Revenue 51

9.1.2 Users' Expected Utility 52

9.2 Duopoly 54

10 Conclusion **57**

A Benchmark Model **61**

A.1 Propositions 1 and 2 61

A.2 Corollaries 1 and 2 62

B Asymmetric Users' Distribution **63**

B.1 Proposition 3 63

B.1.1 Monopoly Comparative Statics 63

B.2 Proposition 4 63

B.2.1 Corollary 3 65

C Estimating Population Parameters **66**

C.1 Proposition 5: Unique symmetric PBE 66

C.2 Lemma 2: Strategic Substitutes 66

C.3 Companion to Figure 5 67

D Two Sided Competitition **69**

D.1 Proposition 6 69

D.2 Proposition 7 70

E Policy Application: Data Portability **72**

E.1 Proposition 8 72

F A Model of Price Discrimination **73**

1 Introduction

Following the surfacing of major data breaches, most notably the Cambridge Analytica scandal, suspicion has arisen regarding the extent to which online platforms respect their users' privacy. In the Cambridge Analytica case, the data of almost 70 million Facebook users was accessed without their knowledge and was subsequently used to create tools for the manipulation of public opinion, prior to the 2016 US elections. This is just one example of how privacy breaches can negatively impact social welfare. In Mark Zuckerberg's own words, the Cambridge Analytica scandal represented a "breach of trust" between Facebook and its users[1]. In the Senate hearing that followed the scandal[2], he pointed at the crucial importance of "trust" for Facebook's business model, which depends on maintaining long-term relationships with users who share their most personal information with the platform[3]. This motivates me to investigate whether platforms' concerns to maintain users' trust provide them with sufficient incentives to adopt privacy-respecting practices.

The dominance of few (or even a single) platforms in domains such as social networking, search and online shopping implies that users' sensitive information is concentrated among very few corporations. Thus, as the Cambridge Analytica case highlights, the privacy-protection incentives of few key platforms, operating in monopolies or oligopolies, is critical to social welfare. At the same time, the market power of corporations like Facebook, Google and Amazon is under increasing scrutiny and a rising number of policy makers are calling for greater competition and antitrust enforcement in the aforementioned industries. Therefore, it becomes crucial to understand how a rise in competition will affect those platforms' incentives to protect their users' privacy and this is precisely what I seek to do in this paper. Finding that platform incentives deteriorate as competition increases would imply that a policy trade-off arises between higher concentration of valuable data and the safer handling of such data. On the other hand, finding that incentives to preserve users' privacy increase in more competitive settings, would make

[1] https://www.forbes.com/sites/kathleenchaykowski/2018/03/21/mark-zuckerberg-addresses-breach-of-trust-in-facebook-user-data-crisis/#3cc9c33d3e36, 2018

[2] https://www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing/

[3] Senator Gary Peters' quote from the Senate hearing conveys the same message: "Can I believe...who has access to this information about me? So, I think it's safe to say, very simply, that Facebook is losing the trust of an awful lot of Americans as a result of this incident".

the argument for increased competition in such markets even more compelling.

In order to capture "trust" in the interaction between platforms and users with privacy concerns, I employ a two-period model of reputational concerns, belonging to the class of "noisy-signalling" models analyzed in Matthews and Mirman [1983].

I am interested in platforms' incentives to exert costly effort towards the protection of users' privacy. This effort could be thought of as implementing costly screening procedures before granting data access to developers or other third-party websites. Similarly, it can be thought of as the platform undertaking random audits to check whether third-parties who gain access to data abide by privacy-respecting practices, or as the platform following up on reports that accuse third parties of mishandling data. Alternatively, this effort can be interpreted as keeping up to date with the latest advances in cyber-security, to protect databases from bad actors who attempt to get access[4]. Whichever interpretation of effort holds, I find it natural to assume that such effort is hard to verify in court and is, thus, non-contractible.

Using concepts introduced in Kreps and Wilson [1982] and discussed in a more recent survey by Bar-Isaac and Tadelis [2008], users entertain the possibility that the platform is a Commitment type, meaning it always exerts the highest possible level of effort. The platform's reputation is defined as the probability with which users believe it is of Commitment type. Platforms know their own types before they choose their actions, but users observe only imperfect signals of these actions, i.e. whether a data breach occurs or not[5]. Crucially, exerting effort reduces the probability of a data breach occurring. A breach in the first period, therefore, reduces the platform's reputation and causes it to attract a smaller fraction of users to join it in the second period. Through this "punishment", reputational concerns give rise to positive, equilibrium effort levels by Normal-type platforms.

[4] In the same Senate hearing referenced previously, Mr Zuckerberg refers to the "arms race" between bad actors, whose hacking technology evolves constantly, and platforms, who need to *continuously* invest in updating their security and expertise practices.

[5] Assuming that users observe whether a data breach occurs or not is already a substantial assumption. For instance, the Cambridge Analytica breach took three years before it became publicly known. Recently, Mozilla Foundation has developed the Firefox Monitor, which alerts users whose data has recently been compromised due to a data breach. Similar initiatives resulting in closer scrutiny of privacy policies make this assumption more realistic.

My choice to use a model of both incomplete information and hidden action is deliberate. In my view, incomplete information realistically captures the fact that users are uncertain about the true attitudes of platforms regarding their privacy: even when platforms are fully transparent with their privacy policies, users often do not read them thoroughly, due to the texts being significantly long or convoluted and making extensive use of legal terminology. I also find this assumption to be consistent with the observation that platforms, especially social networks, seem concerned with convincing users that they value their privacy[6], which would correspond to a Commitment type platform in my model. I find the assumption of unobserved effort to also be realistic; As the discussion of effort concepts reveals, platforms seem to have significant ex post discretion on how to implement their privacy policies and as demonstrated in Mark Zuckerberg's Senate hearing, it is particularly hard to verify the extent to which breaches occur due to platforms' low effort. Thus, users face two distinct kinds of uncertainty regarding the treatment of their privacy; this conclusion guides my modelling choices.

I choose not to incorporate switching costs in my analysis. It is often argued, that users of social networks become gradually "addicted", which could be interpreted as a form of switching costs. In a model with sufficiently high degree of addiction, privacy-concerned existing users could be unresponsive to data breaches, even in the presence of a high reputation competitor. Assuming away such switching costs means that I investigate the *best-case scenario* with respect to the impact of competition on effort incentives. At the same time, not including switching costs means that the core analysis of sections 3-7 is mostly applicable to competition between traditional media outlets (see Gentzkow et al. [2014]). In that setting, the model can be used to ask: how does competition between media outlets influence incentives for provision of high quality news reports, when audiences have uniform preferences in favour of high-quality news?

The rewards-to-reputation function remains the core object of the analysis throughout the paper. Using models in which the rewards-to-reputation function become gradually

[6]Using Facebook as an example, its business website mentions: "..we take data protection and privacy very seriously and are committed to complying to data protection legislation..", while Mark Zuckerberg recently outlined his "Privacy-focused vision for Social Networking".

more complex, Sections 4, 5 and 6 introduce three effects that capture the entire impact of competition on effort incentives, when competition influences incentives only through the market share. In the benchmark model, platform revenues are assumed to be linear in the fraction of users they attract in a particular period[7]. In this setup, it is shown that competition weakly reduces effort incentives for any parameter configuration. The intuition behind this result is that, when a Hotelling specification is used to determine market shares in every state, the expected marginal increase in market share achieved through a high posterior reputation is always lower under the presence of a competitor. I refer to this as the "incremental location" effect of competition.

In Section 5, I explore how an asymmetric distribution of users along the Hotelling line, interpreted as platform "popularity", induces non-linearity in the rewards mapping. I find that sufficiently concave mappings in combination with a high-reputation competitor can overturn the benchmark result and cause effort incentives to rise in the duopoly equilibrium, which I refer to as the "concavity" effect of competition. This effect captures the intuition that a smaller expected increase in market share might be more valuable under a concave revenue mapping, since expected market shares in duopoly are smaller than their monopoly counterparts. Lemma 1 follows, where I discuss the appropriate concept of strategic complements for my model of duopoly and explain why strategic independence of efforts arises in the models of Sections 4 and 5. In Section 6, I examine how the "concavity" effect appears due to the decreasing marginal value of data, when platforms use data to estimate population parameters. Effort levels are shown to be strategic substitutes in that setting, which gives rise to the "strategic uncertainty" effect of competition.

In the sections discussed above, I make the following crucial assumption: that platforms' revenue per market share remains unchanged between monopoly and duopoly. In Section 7, I revoke this assumption to study a model of two-sided platform competition, in which platforms also compete with each other in attracting advertisers. Both "incremental location" and "concavity" effects are again present. However, the analysis of the previous paragraph must be altered to account for the fact that platforms can respond to their posterior reputations by changing prices in the advertising market. I argue that in this model, the capacity of platforms to respond induces convexity in the rewards-to-reputation

[7]This follows the way Gentzkow and Shapiro [2006] model the advertising revenues of media publishers.

mappings. Thus, the "concavity" and "incremental location" effects impact effort incentives in the same direction, both favouring monopoly. This intuition is confirmed by the results derived in this section, showing equilibrium effort is weakly higher in monopoly, for a wide range of parameters.

As I have highlighted already, I have developed a "noisy-signalling" model where both hidden action and hidden information are present. This characteristic is shared with signal-jamming models, like the career-concerns model of Holmström [1999]. However, in contrast with signal-jamming models, the agent (platform) does know its type, hence equilibrium effort levels are a function of the agent's type. This latter characteristic is shared with Spence's signalling model (Spence [1973]), which, however, is a model of pure hidden information. As Matthews and Mirman [1983] point out in their classic article, Perfect Bayesian Equilibria in noisy-signalling games are often unique (without resorting to refinements), they depend on the agent's prior reputation and the informativeness of different separating equilibria varies. All three characteristics differentiate this class of models from signalling models and all three are present in the solutions of the various models presented in this paper.

2 Related Literature

In terms of the research question asked, this paper is linked to the recently emerging literature on data and privacy. The following papers model users' incentives to be active in online platforms, when the former value their privacy and platform activity entails privacy costs. A recurring theme in this quickly expanding literature is reconciling high levels of activity on online platforms with users having high valuations for privacy. In Acemoglu et al. [2019], Bergemann and Bonatti [2019] and Choi et al. [2019], the key mechanism is the informational externalities of personal data. In Ichihashi [2020a], it is the sunk nature of privacy cost. This quickly expanding literature assumes perfect observability of privacy policies and focuses on how rational users' incentives to share their data are shaped in response to publicly known privacy policies. To the best of my knowledge, there is no work that examines how users' incentives are shaped, as they learn more about platforms' privacy practices. My model, instead, takes a more reduced form approach to modelling "privacy costs" and focuses on learning about platforms' characteristics and platforms'

incentives. Therefore, my work is complementary to the above strand of literature on privacy and information sharing.

The main premise of Bergemann and Bonatti [2019] is that an individual selling their data induces informational externalities to other individuals with correlated types, in the sense that the data buyer can infer information about these (potentially non data-selling) agents as well. Given that a large number of users are already using the platform, the latter can infer non-users' types. Hence, the incremental change in an individual's welfare due to the acquisition of their information is small. When online platforms command substantial market power, they need not reimburse individuals above this marginal drop in welfare resulting from the disclosure of their own information. On the other hand, the online platform/data broker can benefit from the externality through better inferences on the rest of the population. Thus, a gap is created between the data's marginal value for the firm and the required marginal compensation to users for providing it.

Ichihashi's [2020a] benchmark model deals with an infinite-period interaction between a monopoly platform and a single user. The platform accumulates information over time; if the user maintains positive activity level in every period, the platform will learn her type perfectly. The privacy cost is defined just like in Acemoglu et al. [2019][8] as the precision with which the platform can predict the user's type, given accumulated data. Crucially, a consumer will incur cost in the current period due to the information shared through her past activity, even if she becomes inactive. Thus, she regards privacy cost (due on activity so far) as sunk and activity decisions are made based on the marginal privacy cost. The key point is that the marginal privacy cost decreases as the platform accumulates more information on the consumer. Based on this observation, it is shown that a user that joins the platform becomes fully exposed, and in the limit, behaves as if a privacy cost doesn't exist.

In the recent work of de Corniere and Taylor [2020], data can be used in various forms by the firms that collect it (e.g. for better recommendations, price discrimination etc.). The authors define data as pro-competitive when an increase in the amount of data available

[8]Acemoglu et al. [2019] present a static, model with many users and informational externalities. The privacy cost is defined as the precision with which the platform can infer a user's type, given other users' signals. Users only select whether they join the platform or not.

to a firm, raises the equilibrium level of utility its consumers derive. The paper's key message is that different uses of data have different implications on whether data is pro- or anti-competitive. I examine a question of the opposite direction, i.e. how competition impacts platforms' privacy-protection efforts. However, I make a point of similar nature, i.e. that the way platforms monetize users' attention determines whether competition is conducive, or not, to effort incentives.

In a different strand of literature around privacy and consumer data, Taylor [2004] and Acquisti and Varian [2005] show that when information disclosure is endogenous, ratcheting forces make it optimal for sellers to commit against employing price discrimination. Ichihashi [2020b] derives conditions under which the same result holds for the case of a multi-product monopolist, who can use data for both better recommendations and for price discrimination. Conitzer et al. [2012] point out that the commitment outcome is also achieved when consumers can freely anonymize their purchases. In Section 10, I show how the model can be modified to account for price discrimination concerns and relate to this literature.

The rest of the paper is structured as follows: Sections 3 and 4 analyze the benchmark model in the monopoly and duopoly cases, respectively. Section 5 introduces an asymmetric distribution of users' preferences and examines how a non-linear returns-to-reputation mapping changes the impact of competition on effort incentives. In the same section, I discuss what the relevant concept of strategic complementarity in my model of duopoly is. Section 6 builds on that intuition and discusses effort incentives when data is used to estimate population parameters. Section 7 expands platforms' set of choice variables and shows how returns-to-reputation become convex in a model of two-sided competition. Section 8 uses a modified version of the benchmark model, to analyze the effects of data-portability legislation, inspired by the European Union's GDPR. Finally, Section 9 introduces a significantly modified version of the benchmark model that micro-founds privacy costs via price discrimination concerns and Section 10 concludes the paper.

3 Monopoly Platform

A single online platform, referred to as platform A, is positioned at $A = 0$ of a Hotelling line. Platform A offers a service and wishes to attract users to register; Registration of users lasts 1 period, while A and users live for 2 periods. Users make their participation decisions at the beginning of each period. Platform A charges users with no registration or usage fees, but gets its entire revenue from advertisers (more on A's revenue later). Platform A has some ex post discretion on how to treat the data of users it has attracted in a given period. As earlier discussed, this can be thought of as unobserved effort that the platform exerts to better filter third parties that get access to consumer data. Crucially, I assume that this effort is non-contractible. As mentioned already, I find this assumption reasonable; even if a data breach is made publicly known, it could be quite costly, if at all feasible, to prove that it was due to lack of due diligence by the platform.

In attracting privacy-concerned potential users, the platform faces potential gains from establishing a reputation of caring about users' privacy. I capture this intuition by introducing incomplete information; a platform can be of either of two types, $\{N, C\}$, which stand for Normal and Commitment type, respectively and the type is the platform's private information. The Commitment type is non-strategic and always chooses the same action. In particular, the Commitment type is the "good" type and always chooses the highest level of effort[9], $e = 1$.

Users of an online platform care about the level of effort it exerts, because this effort level determines the probability with which they experience a breach of their personal data and suffer privacy disutility. I think of this privacy cost as mainly psychological but it could also be motivated by price or non-price discrimination concerns. I define the "outcome" binary random variable s_t^A, which can take values $\{b, n\}$, standing for *breach* and *no breach*. The value of this random variable becomes publicly known at the end of each period. I also define $P(s_t^A = breach|type)$ as the probability that platform A will suffer a *breach*, given its type and this probability is a function of the effort level exerted in period $t \in \{1, 2\}$. Throughout this paper, I use the following state-transition technology:

[9]This is in contrast to Mailath and Samuelson [2001], who model the non-strategic type as the one that always produces low quality outcomes. Following the classification of Bar-Isaac and Tadelis [2008], their paper uses a "Separation" model, while I use an "Imitation" model.

$$P(s_t^A = breach|type) = \begin{cases} 0 & type = Commitment \\ 1 - e_t & type = Normal \end{cases}$$

As the above specification suggests, breach probabilities in period 2 are independent of period-1 outcomes. I call the platform's "reputation", the probability with which users believe that the platform's type is C. The platform has prior reputation μ_A and I assume it is aware of this prior reputation. After observing the 1st period realization of s_t^A, fully rational, Bayesian users, update their beliefs on the platform's type using Bayes' Rule. Posterior reputation depends on the prior and also, crucially, on the effort level that users conjecture a Normal-type platform exerts in the first period, $\tilde{e_1}$. I assume all users share the same conjecture, so that there is a single posterior reputation following each s_1^A realization. The posterior reputation after observing a $\{b\}$ realization in period 1 is equal to zero, since C types never produce breaches under the specified transition technology. The posterior reputation a platform achieves, following a $s_1^A = n$ realization in period 1 is:

$$\mu'_A = \mu(C|s_1^A = n) = \frac{\mu_A}{\mu_A + (1 - \mu_A)\tilde{e_1}} \tag{1}$$

Posterior reputation following a good outcome takes values in $[\mu_A, 1]$. What is immediately apparent in expression (1), is that posterior reputation is decreasing in the effort conjecture, $\tilde{e_1}$. This is very intuitive; if users anticipate that a Normal type platform exerts a lot of effort to prevent breaches from occurring, it must be that a "no breach" outcome is less surprising, less indicative of a Commitment type and thus has a smaller effect of improving the platform's prior reputation[10]. When $\tilde{e} = 0$, then all good news indicate a Commitment type with certainty, so posterior reputation following $s_1 = "no\, breach"$ is equal to one. In contrast, when $\tilde{e} = 1$, the Normal type platform perfectly replicates the Commitment type's behaviour in period 1, hence posterior reputations are not updated and $\mu'_A = \mu_A$.

[10]This is in contrast with other models of effort provision: For instance, in models using the setup of Holmström [1999], posterior reputation is not a function of effort conjectures, hence the resulting first-order condition is linear in the simplest model.

3.1 Users

Before deriving a Normal-type platform's optimal effort choice, I turn to users' payoffs and participation decisions. These will ultimately determine the platform's revenues as a function of their second-period reputation and thus platform incentives. As mentioned already, users make their participation decisions $\alpha_t \in \{\varnothing, A\}$ at the beginning of each period, meaning that users choose between joining platform A and staying idle. There is a unit-measure continuum of massless users, spread uniformly along the [0,1] Hotelling line. Each user is characterized by location θ, which also determines the utility each user derives from using the platform's services. I assume that users suffer disutility equal to (-1) in any period in which they suffer a breach. Thus, expected utility for a user of type θ, as a function of their participation decision, is given by:

$$EU_t(\alpha_t = A, \theta) = \beta(1-\theta) - P(s_t^A = b|I_t) \tag{2}$$

$$EU_t(\alpha_t = \varnothing, \theta) = 0 \tag{3}$$

Implicit in this utility specification is the following assumption: a user that has suffered a breach in the first period will incur additional disutility of (-1) if their data is breached again in the second period. The disutility of an active user that experiences a breach in the second period is independent[11] of both α_1 and s_1^A. Last, but not least, I abstract from network effects and informational externalities by assuming that the utility users derive from using the platform is independent of other users' participation decisions (both past and present). If informational externalities, as they were discussed in the Related Literature section, were present, then the outside option would not yield $EU_t(\alpha_t = \varnothing) = 0$, but $EU_t(\alpha_t = \varnothing)$ would be negative and decreasing in the mass of active users.

The general utility specification used in Hotelling models is $EU(\alpha_t = A, \theta) = v - \beta\theta - P(b|I_t)$, where v is the standalone utility derived by using the platform's services and β is the transportation cost parameter. I am assuming $v = \beta$ to simplify the exposition and facilitate comparisons between monopoly and duopoly [12]. Users are fully rational and use all available information to determine the probability of platform A suffering a breach. I_t

[11] This means, for example, that a user that joins platform A in both periods does not incur higher privacy cost in the event of a second-period breach than a user who just joins platform A in $t = 2$. Users suffer only because their current period data is exposed.

[12] The results stated hold for the general case too, for a range of values of v around β. The significance of v for incentives will be discussed later.

is their information set at the beginning of period t, when they choose α_t; $I_1 = \{\mu_A, \tilde{e_1}\}$ and $I_2 = \{\mu_A, \tilde{e_1}, s_1^A\}$. As is usual in Hotelling competition models, β is a measure of transportation costs. Since being idle yields $EU_t = 0$, a user of type θ will join in period t iff:

$$\theta \leq \frac{1}{\beta}(\beta - P(b|I_t)) := \theta^{ind} \tag{4}$$

In the first period, $P^A(b|I_1) = (1-\mu_A)(1-\tilde{e_1})$, hence the position of the indifferent user is not a function of the true effort level that a Normal type firm exerts in period 1, but only of the users' conjecture about that effort level. In period 2, $P^A(b|I_2) = P^A(N|I_2)$, since it is never optimal for a Normal type platform to exert effort in period 2. The platform only cares about its reputation to the extent that it influences participation decisions. Since e_2 is chosen after a_2 has been chosen by users, a Normal platform has no incentive to choose $e_2 > 0$. Thus, I will focus only on period 1 effort choices and drop the t subscript from effort levels and conjectures.

As (4) suggests, the location of the indifferent user is a function of the platform's reputation, i.e. in the second period it is a function of the first period outcome realization. For given effort conjecture and prior reputation, there are thus only two possible second period cutoffs, corresponding to the possible realizations of s_1^A. These are given by:

$$\theta_n = \frac{\beta - 1}{\beta} + \frac{\mu'_A}{\beta} \tag{5}$$

$$\theta_b = \frac{\beta - 1}{\beta} \tag{6}$$

In the above expressions, which simply follow from (4), the subscripts indicate the s_1^A realization to which the cutoffs correspond, and the "*ind*" superscript is omitted for elegance of exposition. I work with $\beta \geq 2$ so that even following a $s_1^A = b$, the monopolist will capture at least $\frac{1}{2}$ of the market. This implies full market coverage in every state of the duopoly examined later. It is also important to notice for any finitely large value of β, the monopolist will never enjoy full market coverage, even after a $s_1^A = n$ realization. This will ensure that cutoffs are always linear in posterior reputation levels.

3.2 Equilibrium

Given that users are uniformly spread along the Hotelling line, i.e. that the pdf of θ is $g(\theta) = 1$, the market share platform A gets, for a given level of reputation and thus a

given indifferent user location, is given by: $\int_0^{\theta_{ind}} g(\theta)d\theta = \theta^{ind}$. In this benchmark model, I assume that the platform earns advertising revenue as a fixed price, p, per fraction of the users market that it attracts. Hence, the total revenue platform A earns through advertising in a given period is simply $r(\theta^{ind}) = p\theta^{ind}$. This means that the total expected profit the Normal-type online platform earns is given by:

$$\pi(e, \tilde{e}) = r(\theta_1) + e\delta r(\theta_g) + (1 - e)\delta r(\theta_b) - C(e) \tag{7}$$

The above expected profit function takes as inputs both true period 1 effort and the conjecture, since all indifferent user cutoffs depend on the conjecture; δ is the factor at which the platform discounts period-2 revenues when it chooses period 1 effort. The cost function is increasing and convex in e, $C(e) = \frac{1}{2}\alpha e^2$, and the Normal-type platform chooses effort to maximize the above profit function, taking effort conjectures as given.

Now that I have specified the information structure, the type and strategy spaces and the payoffs for all parties, I can define the solution concept of the "noisy-signalling" game. The solution concept used will be Perfect Bayesian Equilibrium (henceforth referred to as "equilibrium"). In equilibrium, the following must hold: users choose α_t optimally in every period to maximize their expected utility, given I_t; following the first period realization, users update their beliefs on the platform's type using Bayes' Rule; Normal type platforms choose e_1 optimally to maximize their expected profit, taking users' conjecture as given and finally, users' conjecture is correct.

Returning to the Normal type's unconstrained maximization problem, i.e. $\max_e \pi(e, \tilde{e})$, the first-order condition with respect to effort that any interior solution must satisfy is the following:

$$[e] \quad \delta r(\theta_g) - \delta r(\theta_b) = \alpha e \tag{8}$$

Taking users' effort conjectures as fixed, the potential "prizes" that the platform is rewarded are also fixed and the platform's effort choice simply regulates the probability with which it gets each prize. A platform exerting zero effort would still get revenue $r(\theta_b)$ with certainty, so effort provision is motivated by the difference $r(\theta_g) - r(\theta_b)$, as equation (8) suggests. Equation (8) is alternatively written as:

$$\frac{\delta p}{\beta} \frac{\mu_A}{\mu_A + (1 - \mu_A)\tilde{e}} = \alpha e \tag{9}$$

Equation (8) defines the monopolist platform's optimal effort provision, taking conjectures as exogenously fixed. To turn it into an equilibrium defining equation, I must impose the equilibrium condition that conjectures are correct, i.e. $\tilde{e} = e$. A greater difference between revenue in the two potential states induces higher effort provision, however, higher effort provision also reduces the size of the prize by reducing posterior reputation following $s_1^A = n$ and thus decreasing θ_n. Imposing the equilibrium condition means also taking the second effect into account. Taking conjectures to be correct and rearranging terms yields a quadratic first-order condition that must be satisfied by any interior effort level played by the Normal-type monopolist in equilibrium.

$$e^2 + \frac{\mu_A}{1-\mu_A}e - \frac{\delta p}{\alpha\beta}\frac{\mu_A}{1-\mu_A} = 0 \tag{10}$$

In the Appendix, I prove the following Proposition.

Proposition 1 *If the cost function is sufficiently convex, there exists a unique Perfect Bayesian Equilibrium of the 'noisy signalling' game discussed above. Type C plays $e = 1$ in every period; type N plays $e_2 = 0$, while e_1 is given by the unique solution to* (10)*. I call that unique solution e^M. Users conjectures are correct, i.e. $\tilde{e_1} = e^M$.*

In the following Corollary, I provide some easily interpreted comparative statics regarding e^M.

Corollary 1 *Let conditions hold for the existence of a unique equilibrium effort, defined as e^M. Then, the following comparative statics hold: $\frac{\partial e_M}{\partial \delta p} > 0$, $\frac{\partial e_M}{\partial \mu_A} > 0$ and $\frac{\partial e_M}{\partial \alpha} < 0$.*

To interpret these comparative statics results, an argument under fixed conjectures suffices, i.e. they all match the signs of the partial derivatives taken if equilibrium effort were given by equation (8). This means that the adverse effect of increasing effort provision on the size of the potential prize to be earned is not enough to overturn the direction of the above effects on equilibrium effort provision.

4 Introducing a Competitor

Next, I analyze the case in which a competing online platform, B is also positioned at point $B = 1$ on the same Hotelling line. Both platforms are present in both periods and both live for two-periods only. My goal is to examine how effort provision of a Normal-type platform A differs between the monopoly and duopoly settings. The competitor platform

is of either Commitment or Normal type, types of A and B are determined independently of each other and each type produces breaches with the same technology that characterized platform A outcomes in the previous section. I define the random variable s_t^B as the outcome of platform B in period t. Furthermore, I assume that platform B has prior reputation (defined as the probability that platform B is Commitment type) equal to μ_B. Crucially, the platforms do not know each other's types[13]. Platforms hold the same prior distribution of their competitor's types as the users do and update their beliefs on each other's type in the same way that users do[14]. In other words, platforms do not have an informational advantage over users regarding the competitor platform's private information.

I assume that the competitor platform also earns solely advertising revenue and charges no subscription fees. I also assume away, at least for now, competition on the advertisers' side of the market, that would be reminiscent of the two-sided-markets literature and Armstrong [2006]. Finally, I assume that the exogenous revenue per user, p, that platform A earns is unaffected by the competitor's presence.

Users now choose which platform they endorse. Similar to Armstrong [2006], I assume that the users are single-homing, i.e. they only join one of the two platforms. This could be potentially unrealistic if one thinks about different individuals having multiple profiles in different social networks, like Pinterest, Twitter, Instagram, etc. although it becomes more plausible if one thinks about 2 online platforms that offer directly comparable services (e.g. it's less likely that someone would be an active user of both Facebook and Google Plus). At the beginning of each period, users choose $\alpha_t \in \{\varnothing, A, B\}$, but as I have mentioned already, $\beta \geq 2$ guarantees that $\varnothing$ is never chosen by any user. The expected utility users derive from joining platform B is given by an expression symmetric to (2) and a user of type θ joins platform A iff:

$$\beta(1-\theta) - P(s_t^A = b|I_t) \geq \beta\theta - P(s_t^B = b|I_t) \tag{11}$$

[13]This is an important assumption; if platform A knew the type of platform B, then a Normal-type platform A would have two different choice variables; period-1 effort for the case where it faces a Commitment type and period-1 effort for the case where it faces a Normal type competitor. In this case, users would be required to form two effort conjectures for each platform and the belief updating process would be more analytically complicated.

[14]Although platforms' posterior beliefs on their competitor's type are irrelevant for payoffs.

The realizations of s_1^A and s_1^B are publicly observed, hence the second period information set, $I_2 = \{\mu_A, \mu_B, \tilde{e_A}, \tilde{e_B}, s_1^A, s_1^B\}$ is the same for all users, regardless of which platform they joined in period 1. I_2 consists of both platforms' period 1 outcome realizations and the effort conjectures users entertain for each platform's Normal type. It is important to emphasize that the random variables s_1^A and s_1^B are assumed to be independent, hence there is no informational aspect to competition; realizations of s^A do not convey information on the effort or type of platform B and vice versa. Thus, each platform's posterior reputation in period 2 is only a function of that platform's own period 1 realization and expression (1) still applies. For given $\tilde{e_A}, \tilde{e_B}$, the relevant second period state is now defined as the pair of platforms' period 1 realizations. Thus, the second period state can take either of four different values $\{b, b\}, \{n, b\}, \{b, n\}, \{n, n\}$. Replacing the inequality symbol with an equality one, (11) defines the location of the indifferent user in each of the four period-2 states. The cutoff value (which is also platform A's market share) as a function of the state is given by:

$$\theta(s_1^A, s_1^B) = \frac{1}{2} - \frac{P^A(Normal|I_2)}{2\beta} + \frac{P^B(Normal|I_2)}{2\beta} \tag{12}$$

Remember that in period 2, $Pr(b|I_2) = Pr(N|I_2)$, since Commitment types never produce breaches and Normal types never exert any effort in equilibrium. Thus, platform A's expected[15] profit, as a function of its own conjecture of platform B's Normal-type effort[16], e_B, and as a function of users' conjectures, $\tilde{e_A}, \tilde{e_B}$, is the following:

$$\begin{aligned}\pi(e_A, \tilde{e_A}, e_B, \tilde{e_B}) = p\theta_1 + e_A\delta p(P^B(n)\theta_{nn} + P^B(b)\theta_{nb}) \\ + (1 - e_A)\delta p(P^B(n)\theta_{bn} + P^B(b)\theta_{bb}) - C(e)\end{aligned} \tag{13}$$

In equation (13), $P^B(b)$ is the probability with which platform B is going to suffer a breach in the first period, hence it is given by $P^B(b) = (1 - \mu_B)(1 - e_B)$. Similarly, $P^B(n) = 1 - P^B(b) = \mu_B + (1 - \mu_B)e_B$. θ_{nn} is platform A's market share in the event that both platforms' period 1 outcome realizations are $\{no\ breach\}$ and the other θ terms are defined accordingly; to simplify notation, I omit the fact that these cutoffs also depend on $\tilde{e_A}, \tilde{e_B}$ and prior reputations. Expected profit in (13) could also be equivalently expressed in terms of the distribution of platforms' posterior reputations (which are in turn induced

[15] Expectation taken over the probability distribution of period-2 states, which is fixed for given effort of the Normal-type competitor.

[16] Whenever I refer to platforms A/B, their maximization problems and their effort levels, I will always be referring to the respective Normal types.

by the outcome realizations).

Platform A is assumed to be risk neutral, hence it chooses e_A to maximize expected profit (13), taking $\tilde{e_B}$, $\tilde{e_A}$ and e_B as given. The corresponding first order condition an optimal, interior effort level must satisfy is given by:

$$(P^B(n)\theta_{nn} + P^B(b)\theta_{nb}) - (P^B(n)\theta_{bn} + P^B(b)\theta_{bb}) = \frac{\alpha}{\delta p}e \tag{14}$$

Once again, the above condition simply states that since platforms are risk neutral, in an interior optimum the expected marginal benefit from increasing effort provision must equal the (deterministic) marginal cost of doing so.

Focusing on the terms of the left-hand side of (14), $P^B(n)\theta_{nn} + P^B(b)\theta_{nb} = E(\theta_n)$ is the expected market share platform A attains in period 2, conditional on $s_1^A = n$, and $E(\theta_b)$ is defined similarly. Expectations are taken over the competitor's posterior reputation, i.e. over the distribution of s_1^B. Since θ terms are all linear in posterior reputations, the Law of Total Probability applies and $E(\theta_n)$ depends only on B's prior reputation, μ_B. It is important to realize that the above argument is valid only when $e_B = \tilde{e_B}$; platform A uses its own conjecture of B's effort to form expectations over the possible levels of B's posterior reputation. These posterior reputations are held by the users, who have conjectures of their own, hence the LoTP applies only when these two different conjectures coincide - which does happen in equilibrium, when they are also correct, equal to B's true effort level. This point will be investigated in more detail in Section 5.1. When $e_B \neq \tilde{e_B}$, A's expected market share, following $s_1^A = n$, is given by:

$$E(\theta_n) = \frac{1}{2} + \frac{\mu'_A}{2\beta} - \frac{E^A(\mu(C|s^B))}{2\beta} \tag{15}$$

where E^A signifies expectation taken according to A's perceived distribution of s^B. When $e_B = \tilde{e_B}$ and the perceived distributions of users and A coincide, $E^A(\mu^B(C|s^B)) = \mu_B$.

I present the first order condition for an interior optimal effort level in equilibrium, when all effort conjectures are correct:

$$e_A^2 + \frac{\mu_A}{1-\mu_A}e_A - \frac{\delta p}{2\alpha\beta}\frac{\mu_A}{1-\mu_A} = 0 \tag{16}$$

The only difference between (16) and the corresponding equation for the monopolist, (10) is a multiplicative term in the constant. As predicted by the preceding discussion, e_B does

not affect effort incentives in equilibrium. This means that in equilibrium, when $e_B = \tilde{e_B}$, platform A chooses its optimal effort level, as if it is facing a competitor whose period 2 reputation is certainly given by μ_B. However, B's prior reputation is also irrelevant to effort incentives in equilibrium; To see why, it is useful to compute the expected market shares defined above, in equilibrium:

$$E(\theta_n) = \frac{1}{2} + \frac{\mu'_A - \mu_B}{2\beta} \tag{17}$$

$$E(\theta_b) = \frac{1}{2} - \frac{\mu_B}{2\beta} \tag{18}$$

B's prior reputation enters these expressions identically, hence the revenue map's linearity makes it irrelevant to marginal incentives[17]. At this point, the paper's main question can finally be asked: Does the presence of a competing online platform increase effort incentives, relative to the monopoly case, or not? I summarize the above analysis and provide the answer in the following Proposition, whose proof is in the Appendix:

Proposition 2 *In the duopoly version of the noisy-signalling game, when platforms' cost functions are sufficiently convex, there exists a unique Perfect Bayesian Equilibrium; in equilibrium, period 1 effort of A's Normal type, e^D is given by the unique solution to* (14)*. Users' conjectures about the effort levels of both platforms' Normal types are correct. Platforms' conjectures about the effort levels of their competitor's Normal types are also correct. Furthermore, $e^D \leq e^M$ for any admissible parameter configuration and, in particular, $e^D < e^M$ for all parameter constellations that yield unique interior solutions in both cases.*

There is no specific mention of e_B in Proposition 2, but the above analysis can be applied to platform B, without any change[18]. The final part of the proposition might seem counter-intuitive, but the above analysis and Figure 1 shed some light on it. Figure 1 compares between the monopoly and duopoly scenarios. It contains all the information that platform A takes into account when making its effort decision in either the monopoly or the duopoly setting. In the arguments below, I am keeping $\tilde{e_A}$ fixed. The vertical axes measure utility, while the horizontal axis is the [0,1] Hotelling line. The downward sloping lines are the expected utilities users receive by joining platform A after each of the two possible first-period S_A realizations. In other words, they are the posterior demand curves

[17] Observing equation (15) reveals that this is also the case off-equilibrium.

[18] It is not required that e^B is interior for e^A to be interior in equilibrium.

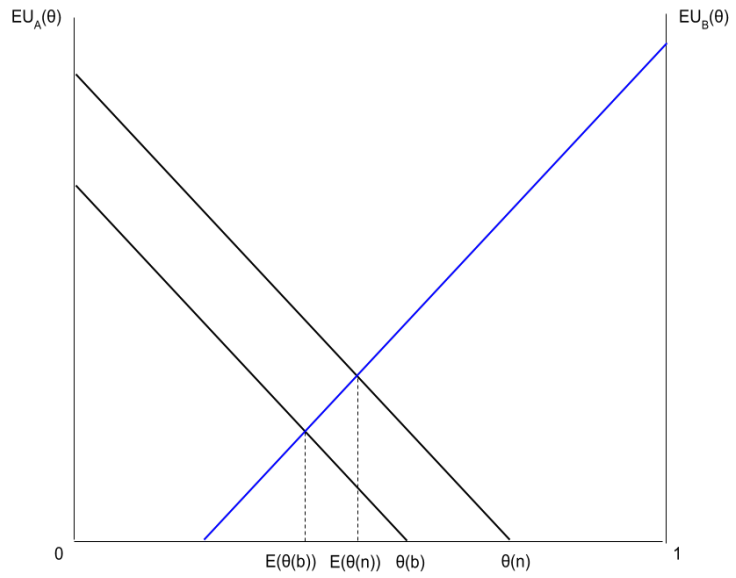


Figure 1: Hotelling competition uniformly reduces the marginal benefits of exerting effort.

the monopolist at A=0 faces and they determine θ_b and θ_g, the difference of which guides the monopolist's effort choice. The upward sloping line is the demand that platform A expects platform B to face in the second period. The intersection of this upward sloping line with A's posterior demand curves determine $E(\theta_b)$ and $E(\theta_n)$, i.e. the expected market shares platform A achieves after each possible period 1 realization of its own outcome.

To explain the reduction in effort incentives under duopoly, it is sufficient to observe that, as Figure 1 suggests, the expected gain in market share from achieving a high reputation in period 2 is lower under duopoly, i.e. $\theta_g - \theta_b \geq E(\theta_g) - E(\theta_b)$. Geometrically, whenever the competitor's posterior demand curve has slope above zero, this inequality is going to hold, hence the result holds for any parameter constellation that yields interior market shares in every state.[19] I refer to the above as the "incremental location" effect of competition, which pushes e^D below e^M. Note that this effect does not only arise in equilibrium; no restrictions were imposed on effort conjectures through the preceding argument.

[19] As mentioned already, assuming $\beta \geq 2$ suffices to achieve this.

4.1 Comparative Statics for $e_M - e_D$

The comparative statics of optimal effort in duopoly with respect to the model parameters $\mu_A, \delta p$ are of the same direction as in the monopoly case. The question of interest now becomes how these parameters influence the difference $e_M - e_D$. As Figure 2 suggests,

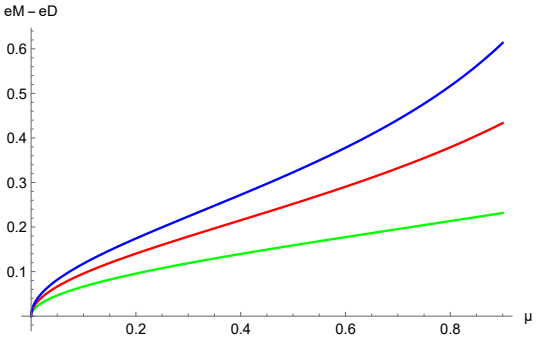


Figure 2: Difference between optimal effort levels as a function of prior reputation. Lower lines correspond to larger values of $\frac{a}{\delta p}$. Green = 1, red = $\frac{1}{2}$, blue = $\frac{1}{3}$.

in which I have fixed $\beta = 2$, one can show more generally that $e_M - e_D$ is increasing in the platform's prior reputation and in the ratio $\frac{\delta p}{\alpha}$. What is shown in Figure 2 is that increasing the ratio induces an upward shift of the difference, when the latter is expressed as a function of prior reputation. Taking a vertical cross section of the 3 lines (i.e. fixing μ_A) also suggests that the difference is increasing in $\frac{\delta p}{\alpha}$, ceteris paribus.

These comparative statics results are also relatively simple to interpret, if we resort to Figure 1 and the "incremental location" effect. For the following argument, I am, again, fixing users' conjectures $\tilde{e_A}$. In monopoly, effort is determined by $\delta p(\theta_n - \theta_b) = \delta p \frac{\mu'_A}{\beta}$, while in duopoly effort incentives arise from the difference $\delta p(E(\theta_n) - E(\theta_b)) = \delta p \frac{\mu'_A}{2\beta}$. Thus the effect of an increase in either δp or in μ_A on effort incentives is mitigated by the existence of the competitor - essentially the same intuition behind the $e_M > e_D$ result. What is shown analytically in Corollary 2 is that the adverse effect of an increase in equilibrium effort on the rewards to achieving a high reputation in period 2 is not enough to overturn the original intuition, so the difference grows even though higher monopoly effort puts greater downward pressure on the monopoly reward to high reputation.

Corollary 2 *Assume that parameter values $a, \mu_A, \mu_B, \beta, \delta, p$ are such that unique interior equilibria are guaranteed in both monopoly and duopoly; $e^M - e^D$ is the difference between*

A's optimal effort levels between the two cases. The following comparative statics hold: $\frac{\partial(e^M - e^D)}{\partial \mu_A} > 0$, $\frac{\partial(e^M - e^D)}{\partial \delta p} > 0$ *and* $\frac{\partial(e^M - e^D)}{\partial a} < 0$.

Corollary 2 concludes the examination of the benchmark model. In the next section, I seek to expand the base setup to further understand how the main result arises and under what conditions the effort of Normal-type platforms can increase as the result of introducing a competitor.

5 Asymmetric Users' Distribution

In the first extension examined, I use the benchmark framework exactly as described in sections 3 and 4, but with one difference: users are no longer uniformly distributed along the line. Instead, I assume that the probability density function of θ is given by a member of the following parametric family of distributions:

$$g_k(\theta) = 2(1-k)\theta + k \tag{19}$$

Each value of the parameter $k \in [0, 2]$ characterizes a single member of the above family of distributions. The support of θ is always [0,1]. One can verify that $k = 1$ yields the uniform distribution over that interval, which was used in the benchmark section. For values of $k \neq 1$, the density rotates around $g(0.5) = 1$; values $k = 0$ and $k = 2$ yield triangular distributions. High values of k mean that platform A is "more popular", i.e. there is a greater mass of consumers with preferences favouring platform A, while low values of k mean platform B is the relatively popular one. Fixing k, i.e. the distribution of users, A's market share in period 2, after reputations have been updated, is again given by the cdf of θ at $\theta = \theta_{ind}$. Thus, A's discounted period-2 revenues are given by:

$$\delta r(\theta_{ind}) = \delta p(\int_0^{\theta_{ind}} g(\theta)d\theta) = \delta p(\frac{1-k}{2}\theta_{ind}^2 + k\theta_{ind}) \tag{20}$$

where θ_{ind} is the indifferent user, given the updated beliefs. I proceed by discussing the monopoly and duopoly cases in sequence; indifferent user cutoffs are defined exactly as in the benchmark model, for both monopoly and duopoly and for every state and period. The distribution $g(\theta)$ determines how these cutoffs translate into market shares. In the following calculations, I will be using the normalization[20] $\delta p = 1$.

[20]Since in every maximization problem that I will analyze, δ enters only through the ratio a/δ, where a is the cost parameter, I will be omitting it in later sections.

The maximization problems are defined exactly equations (7) and (13) and I will skip to the monopolist's equilibrium first-order condition to avoid repetition. Defining $\gamma_m = 2(1-k)(\beta-1)+k\beta$, the latter equilibrium condition is given by:

$$\frac{1}{\beta^2}[(1-k)(\mu'_A)^2+\gamma_m\mu'_A] = \alpha e_A \tag{21}$$

When the equilibrium condition $\tilde{e_A} = e_A$ is imposed, the above first-order condition is revealed to be cubic in effort. However, it reduces to the corresponding, quadratic, version presented in section 3, when $k = 1$. The analysis of (21) yields the following result, which I prove in the Appendix:

Proposition 3 *In the monopoly model, when the distribution of users along the Hotelling line is given by $g_k(\theta) = 2(1-k)\theta + k$; if μ_A is bounded below by $\bar{\mu} > 0$ and a is sufficiently large, there exists a unique, interior equilibrium effort level, e^M. For $\beta \geq 2$, it holds that $\frac{\partial e_M}{\partial k} < 0$, for any $\mu_A \in (0,1)$ and any $k \in [0,2]$.*

This is an important finding, according to which in monopoly, more popular platforms will exert less effort in equilibrium. Figure 3 shows geometrically why that is always the case, when $\beta \geq 2$. Holding $\tilde{e_A}$ fixed, the marginal benefit of increasing effort for a monopolist, i.e. the left hand side of equation (21) is given by the area of trapezoid M[21]. $\beta \geq 2$ is exactly the condition required for $\theta(b) \geq \frac{1}{2}$, which guarantees that the area of M will decrease as the slope of $g(\theta)$ further decreases. Intuitively, when platform A faces a distribution of users with high k, i.e. when it is "popular", achieving a high reputation in the second period will only induce a small increase on its market share, relative to the market share under a low posterior reputation.

Before further analysing the effect of a non-linear revenue function, I briefly discuss the duopoly problem. Using the definition of $r(\theta)$ given above, the problem a Normal type duopolist at $A = 0$ faces can be expressed as follows:

$$\max_{e_A}[e_A Er(\theta(\mu'_A, \mu_B(C|s_B))) + (1-e_A)Er(\theta(0, \mu_B(C|s_B)))] - C(e_A) \tag{22}$$

In the above expression, I have expressed θ terms as functions of posterior reputations instead of outcome realizations, but the two formulations are equivalent. Expectations

[21] This is shown in the Appendix.

are taken over the realization of s_B in the first period, so the maximization program in (22) is actually identical to the one in (13), only difference being that (13) uses $r(\theta) = p\theta$. Similar to the monopoly analysis, defining $\gamma_d = \beta - (1-k)\mu_B$, the duopoly first order condition, in equilibrium, is given by:

$$\frac{1}{4\beta^2}[(1-k)(\mu'_A)^2 + 2\gamma_d \mu'_A] = \alpha e_A \tag{23}$$

Comparing effort levels between monopoly and duopoly won't be as straightforward as in the previous section, since analytical expressions are more complicated. Once again, the Law of Total Probability applies and, in equilibrium, platform A behaves as if it is facing a competitor whose posterior reputation is equal to μ_B with certainty[22]. Thus, e_B terms are absent from equilibrium condition (23).

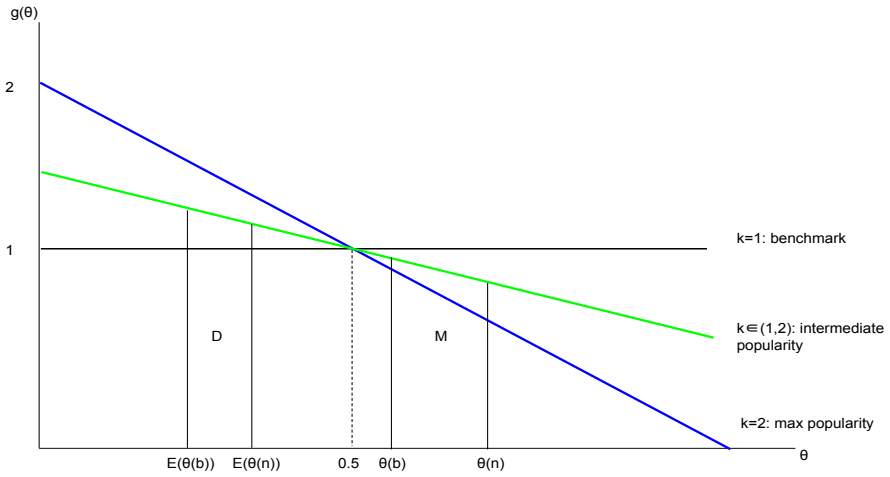


Figure 3: Highly popular platforms may face greater effort incentives under competition.

Proposition 4 *In the duopoly model, when the distribution of users along the Hotelling line is given by* $g_k(\theta) = 2(1-k)\theta + k$*; if* μ_A *is bounded below by* $\bar{\mu} > 0$ *and* a *is sufficiently large, there exists a unique, interior equilibrium effort level,* e_D*.*

After assuming conditions that guarantee uniqueness of interior equilibrium in duopoly, I derive intuitive comparative statics results, concerning the additional parameters that affect effort incentives in this model, namely k and μ_B. I present these results in the following Corollary:

[22]This point will be formally examined in Lemma 1.

Corollary 3 *Under the conditions described by Proposition 4, the following comparative statics results hold at the unique duopoly equilibrium:* $\frac{\partial e^D}{\mu_B} > 0$ *when* $k > 1$ *and* $\frac{\partial e^D}{\mu_B} < 0$ *when* $k < 1$*;* $\mu_B \geq \frac{1}{2}$ *is sufficient but not necessary to guarantee that* $\frac{\partial e^D}{\partial k} \geq 0$*, while* $\frac{\partial e^D}{\partial k} < 0$ *for* μ_B *sufficiently close to 0;*

The first comparative statics result of Corollary 3 states that duopoly effort incentives can be increasing in platform popularity. This is in contrast to the result obtained under monopoly. The second result states that the effect of a change in the competitor's prior reputation changes direction, depending on whether the platform is more or less popular relative to its competitor. As expected, the effect of μ_B is only relevant when $k \neq 1$; μ_B is not present in the benchmark first-order condition, which corresponds to $k = 1$.

In the following paragraphs, I explore the intuition behind the above results, using Figure 3. The change induced in this model is that for a given level of reputation, the location of the indifferent user is no longer equal to the demand that platform A faces. As I have mentioned, θ_{ind}, is given by the same expressions as before, i.e. the mapping from posterior reputation to θ_{ind} remains unchanged. However, the revenue function $r(\theta)$ is no longer linear in the location of the indifferent consumer, but is given by (20). The key point is that for a popular platform, i.e. for the platform at A=0 when $k > 1$, the revenue function is strictly concave in θ_{ind}.

For the following argument, I fix e_A (i.e I fix μ'_A too) and impose $e_B = \tilde{e_B}$. When the demand function is non-linear, the introduction of competition has two different effects on effort provision; on one hand, the "incremental location" effect is again present. It is the only effect present in the benchmark and it drives the result presented in Proposition 2. On the other hand, because both $E(\theta^D(b))$ and $E(\theta^D(g))$, as defined in (17) and (18), are smaller (in their levels) than their respective monopoly values, the concavity of $r(\theta)$ increases the potential gains from achieving a high reputation. I refer to the above as the "popularity" effect of competition. Essentially, I am comparing:

$$r(\frac{\beta - \mu_B}{2\beta} + \frac{\mu'_A}{2\beta}) - r(\frac{\beta - \mu_B}{2\beta}) \tag{24}$$

with the corresponding difference that monopolist faces

$$r(\frac{\beta - 1}{\beta} + \frac{\mu'_A}{\beta}) - r(\frac{\beta - 1}{\beta}) \tag{25}$$

In the model examined in this section, the platform at A=0 with $k = 2$ has the most concave demand function and thus I expect this additional effect to be the strongest for $k = 2$. Furthermore, I expect the "popularity" effect to provide the strongest incentives when the function arguments in (24) assume their lowest values, i.e. when competitor's prior reputation is 1 - i.e. when A faces a known Commitment type competitor. Put simply, I am trying to exploit the intuition behind Jensen's inequality. Figure 3 shows why the "popularity" effect is strongest for $k = 2$. Using the normalization $\delta p = 1$, the areas covered by the trapezoids M and D are exactly equal to the left hand sides of equations (21) and (23), respectively. Even though the base of D will always be smaller than the base of M (that is the "incremental location" effect), D gains in area relative to M, when $g(\theta)$ is downward sloping, which is precisely when $r(\theta)$ is concave. The relative magnitude of these two effects in equilibrium will determine the sign of $e^D - e^M$.[23]

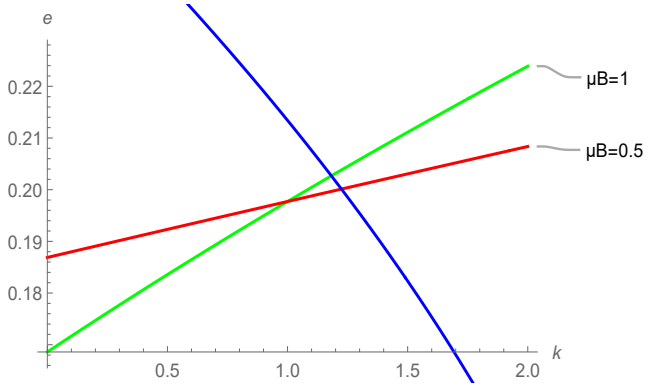


Figure 4: Parameters used are a=0.4, $\mu_A = 0.1$, and $\beta = 2$. The blue line plots e^M, while the upward sloping lines plot e^D for different levels of competitor's prior reputation. The red line corresponds to e^D when facing a competitor with prior reputation $\mu_B = 0.5$, while the green line corresponds to $\mu_B = 1$.

To illustrate the results discussed above, I provide some numerical output generated for particular parameter values. As Proposition 3 suggests, e^M is decreasing in the popularity parameter, k, ceteris paribus. Figure 4 shows that for the particular configuration used,

[23]When it comes to interpreting the sign of $\frac{\partial e^D}{\partial k}$, Figure 3 is not particularly helpful. It is drawn for large values of μ_B and relatively small values of μ'_A, to facilitate exposition; it always holds that $0.5 > E(\theta_b)$, but, whether $0.5 > E(\theta_n)$ or not, depends on the relative values of μ_B and μ'_A. Geometrically, when D lies at the center of Figure 3, it is not necessary that an increase in k will increase its area.

there exists a unique cutoff value k^* such that $e^D > e^M$ for any $k > k^*$. In other words, for sufficiently large k, the "popularity" effect dominates and competition becomes conducive to effort incentives. As shown in the Figure, this cutoff value is expected to be decreasing in the competitor's prior reputation, due to the interplay between the two variables discussed above. Both duopoly lines are upward sloping, reflecting the comparative statics result in Proposition 4, for $\mu_B \geq \frac{1}{2}$. The two intersect exactly at $k = 1$. This was expected; at $k = 1$, the model reduces to the benchmark version, where μ_B is irrelevant for effort incentives. The red line, corresponding to a lower μ_B crosses the green line from above and Proposition 4 also explains this; Since $r(\theta)$ becomes convex for $k < 1$, duopoly effort should in that case decrease with μ_B, ceteris paribus. The opposite should be obtained for $k > 1$ and this is exactly what Figure 4 depicts.

Comparative statics with respect to the parameter β will reveal little insight; that is because, throughout this paper, I am operating under the assumption that users' standalone utility, v, is equal to the transportation cost parameter. Instead, a brief discussion of how an increase in v affects monopoly incentives[24], when $v \neq \beta$, is helpful. The discussion will also illustrate why I assume $v = \beta$. Assuming $v = \beta$ implies that I focus on the interesting case, in which V is low enough so that the monopolist never achieves complete market coverage; this, importantly, ensures that market shares are always linear in posterior reputations. With respect to comparative statics, this means market shares are always linear in v and the market share differential between states is independent of v. In this case, an increase in v (within these bounds) affects effort incentives only through the convexity of r, thus increasing e^M, if $k < 1$ and decreasing it, otherwise.

Outside of this, two more cases must be distinguished: If v increases to such a value, so that both θ_b and θ_n are above 1, then users simply do not value their privacy highly enough and there are no effort incentives for a monopolist. If V increases and enters an intermediate range, in which $\theta_n \geq 1$ but $\theta_b < 1$, then the effect on e^M is ambiguous. Since the maximum market share is capped above, a larger V only increases θ_b and thus decreases effort incentives. On the other hand, if $k < 1$ and $r(\theta)$ is convex, a higher level of θ_b increases the value of an incremental increase in the market share achieved by having

[24] v is irrelevant for incentives under duopoly, since users derive equal standalone utility from both platforms.

a high posterior reputation.

This concludes my analysis of this particular extension. An asymmetric distribution of users' preferences introduces non-linearity in the mapping of platforms' posterior reputations into revenue and gives rise to a second effect of competition on effort incentives. For platforms that are "popular", a given expansion in the indifferent user's location is more valuable under the duopoly setting, in which cutoffs are lower. This effect can overturn the "incremental location" effect introduced in Section 4, in which case competition increases equilibrium effort incentives.

Before moving on to the next section, I briefly return to expand upon a point made previously; I examine exactly what conditions on the functional forms of $r(\theta)$ used in sections 4 and 5 lead to the derived irrelevance of e_B in determining the equilibrium values of e_A. Lemma 1 relates this finding to the notion of strategic complements that I find most appropriate for the models of duopoly presented in this paper.

5.1 Strategic Complementarity

Strategic complementarity typically concerns the sign of the mixed partial derivative $\frac{\partial^2 \pi^A}{\partial e_A \partial e_B}$. In the context of my model, where effort objects appear both as true effort terms and as conjectures, one should be careful as to how this derivative should be correctly defined. How does A's best response change, as its own conjecture of B's effort changes? e_B determines the probability with which A believes the various states are reached in period 2; it also affects the size of the rewards in each state, through B's posterior reputation. The second effect emerges, only when users' conjecture, $\tilde{e_B}$, varies as well.

However, I believe that platform A *should* take the second effect into account when calculating the impact of a change in B's effort on its own marginal incentives. Platform A realizes that B cannot systematically deceive A or the users in equilibrium, thus realizes that these two conjectures will have to be equal to each other. Put informally but hopefully intuitively, because A has no informational advantage over users regarding B's effort, if B attempts to deceive A and users by changing its effort, users and A will realize it at the same time. Hence, it is natural to assume that their conjectures, e_B and $\tilde{e_B}$, vary together and A takes into account both the effect on the distribution of states and on the

sizes of rewards.

Following this discussion, the question of interest becomes: What is the impact on marginal incentives for e_A, as both e_B and $\tilde{e_B}$ vary together? I believe that the relevant notion of strategic complements answers this particular question. The mechanics of this definition are more closely examined in the proof of Lemma 1.

I define the rewards-to-reputation function for platform A, $\phi(\mu'_A, \mu'_B)$, as the revenue platform A receives in period 2, as a function of both platforms' posterior reputations. This is exactly the object that has been driving effort incentives in every model presented so far. However, due to the assumptions made, the only channel through which posterior reputations influence second-period revenues is the market share. In other words, in the duopoly models examined so far, $\phi(\mu'_A, \mu'_B) = r(\theta(\mu'_A, \mu'_B))$. Hence, I have simply used $r(\theta)$ so far because this form facilitates exposition of the "incremental location" and "popularity" effects. However, proving the following result requires using the general form of the rewards-to-reputation function. I examine the case where $\phi(\mu'_A, \mu'_B)$ takes a particular functional form, which happens to nest the forms it has in the models of Sections 4 and 5, and prove the following Lemma:

Lemma 1 *Assume that under duopoly, the rewards-to-reputation function for platform A has the following form:* $\phi(\mu'_A, \mu'_B) = f(\mu'_B)g(\mu'_A) + h(\mu'_A) + p(\mu'_B)$, $\forall \mu'_A, \mu'_B \in [0,1]$, *where* f *is linear in platform B's posterior reputation. Then, effort levels,* e_A *and* e_B*, are neither strategic substitutes, nor complements.*

Crucially, f is linear in B's posterior reputation - there are no restrictions on the functions g, h, other than that they don't admit μ'_B as an argument and no restrictions on p either. This means that ϕ can be non-linear in both posterior reputation terms, just as it is in Section 5.

Proof: In duopoly, holding $e_B, \tilde{e_A}, \tilde{e_B}$ fixed, platform A wants to maximize expected profits with respect to e_A:

$$\pi^A = (P^B(n)\phi(\mu'_A, \mu'_B) + P^B(b)\phi(\mu'_A, 0))e_A + (P^B(n)\phi(0, \mu'_B) + P^B(b)\phi(0,0))(1 - e_A) - C(e_A) \quad (26)$$

where $P^B(n) = \mu_B + (1-\mu_B)e_B$, and prime reputation terms are the posterior reputation levels each platform achieves if its own outcome is "*no breach*" in period 1. Differentiating with respect to e_A yields:

$$\frac{\partial \pi^A}{\partial e_a} = (P^B(n)\phi(\mu'_A, \mu'_B) + P^B(b)\phi(\mu'_A, 0)) - (P^B(n)\phi(0, \mu'_B) + P^B(b)\phi(0,0)) - C'(e_A) \quad (27)$$

I emphasize again the question of interest that drives my notion of strategic complements: How does A's best response change, as B's effort changes, when A recognizes e_B and $\tilde{e_B}$ will vary together? The last part means that before differentiating (27) with respect to e_B, I impose that A's and users' conjectures of B's effort coincide, i.e. that $\tilde{e_B} = e_B$. Then, plugging the assumed form of ϕ in the above first derivative and keeping in mind that $P^B(b) = 1 - P^B(n)$, the product terms inside the first pair of parentheses evaluate to:

$$P^B(n)\phi(\mu'_A, \mu'_B) = P^B(n)f(\mu'_B)g(\mu'_A) + P^B(n)h(\mu'_A) + P^B(n)p(\mu'_B) \quad (28)$$

$$P^B(b)\phi(\mu'_A, 0) = P^B(b)f(0)g(\mu'_A) + P^B(b)h(\mu'_A) + P^B(b)p(0) \quad (29)$$

and, thus, their sum is equal to:

$$f(\mu_B)g(\mu'_A) + h(\mu'_A) + P^B(n)p(\mu'_B) + +P^B(b)p(0) \quad (30)$$

where I have used the linearity of f and the Law of Total Probability to get that, when $e = \tilde{e_B}$:

$$Pr^B(n)f(\mu'_B) + Pr^B(b)f(0) = f(\mu_B) \quad (31)$$

The product terms involving p in (30) appear in the exactly same form in the second pair of parentheses, thus (27) is given by:

$$\frac{\partial \pi^A}{\partial e_a} = f(\mu_B)g(\mu'_A) + h(\mu'_A) - f(\mu_B)g(0) - h(0) - C'(e) \quad (32)$$

$$= f(\mu_B)(g(\mu'_A) - g(0)) + (h(\mu'_A) - h(0)) - C'(e) \quad (33)$$

Since there are no e_B terms present in the above equation, the mixed partial derivative is trivially equal to zero, proving Lemma 1.

Furthermore, in the benchmark, μ'_B does not enter in a multiplicative manner so $g(\mu'_A) = 1, \forall \mu'_A$, hence B's prior reputation is also absent from A's equilibrium first-order condition. Importantly, the above argument does not require that $e_A = \tilde{e_A}$. It also does not require that the common conjecture of A and users regarding B's effort is correct, which

is required in equilibrium. Thus, the above result is not one of local nature. It holds more generally since the relevant strategic complements notion imposes that platform A's and the users' conjecture on B's effort vary together. Finally, to elaborate upon a point previously made, equation (33) shows that when ϕ belongs to this particular family of functions, A behaves as if it is facing a competitor of certain posterior reputation, equal to μ_B, i.e. as if ϕ is given by $\phi(\mu'_A) = f(\mu_B)g(\mu'_A) + h(\mu'_A) + p(\mu_B)$. Using this form of ϕ in A's maximization program, would immediately yield first-order condition (33).

The discussion of Lemma 1 completes Section 5. The model introduced in the section 6 will accomplish two goals: First, analyze an alternative economic mechanism, quite relevant to online platforms, which can give rise another instance of the "concavity" effect of competition. Second, expand the analysis to accommodate rewards-to-reputation functions which do not belong to the function family described by Lemma 1.

6 Estimating Population Parameters

I turn to the question of how the online platform turns the data of its users into revenue. The discussion of the concavity effect suggests that the implications of an increase in competition are dependent on the final use of users' data. For instance, consider the difference between using data to estimate fixed population parameters compared to using data to train learning algorithms; in the former case, the marginal contribution of an additional data point is decreasing, hence the revenue function would be concave in market share, while in the latter, returns to more data have been argued to be convex, or even given by step functions (e.g. Weyl and Posner [2018]). In this section, I attempt to explore such considerations by adding further structure to the platform's revenue mapping.

I adopt the following stylized setting, inspired by an example mentioned in de Corniere and Taylor [2020]: Assume that the platform earns its entire revenue from a single, monopolist firm M who offers every variety of an experience good it sells in the product market. The set of potential users of the online platform(s) coincides with the set of potential product consumers. M can sell to every consumer of the [0,1] continuum, not just the users of the platform it contracts with. However, M benefits from having access to data generated by users' interaction with the platform, as it can be used to make better product recommen-

dations. Assume that all users have the same true variety preference, t, that is unknown to firm M. The product seller has initial information on the consumers' true preference, that has the form of a signal with conditional distribution $z^0|t \sim \mathcal{N}(t, \sigma^0)$. Consumers who buy variety x receive utility given by:

$$U(x,t) = V - (x-t)^2 \tag{34}$$

Consumers know their (common) true preference s; they do not observe the type of product they are offered until they actually consume it, but they can form expectations on their derived utility, based on the information available to the seller, M.[25] When M has no access to platform data, it is assumed to offer recommendation z^0 and completely extract consumer surplus given by $EU(z^0, s) = V - Var(z^0|s)$. By contracting with the platform, M gains access to an improved signal $z|t \sim \mathcal{N}(t, 1/\theta)$, where θ is the market share of the platform; i.e. the signal's precision is increasing in the market share of the platform. Since M has access to every consumer, the revenue the former gets by having access to signal $z|t \sim \mathcal{N}(t, 1/\theta)$ is equal to:

$$r(\theta) = \int_0^1 (EU(z) - EU(z^0))di = \sigma^0 - 1/\theta \tag{35}$$

When the platform has full bargaining power against M, $r(\theta)$ defined above is the platform's revenue function. I assume that the platform deals with a different monopolist product seller in each period, hence period 1 is again irrelevant for effort incentives. Thus, using the expression for the location of the indifferent user in monopoly[26], (4), the rewards-to-reputation function for a monopolist online platform in the above setup, is defined as:

$$\phi(\mu'_A) = \sigma^0 - (\frac{\beta - 1 + \mu'_A}{\beta})^{-1} \tag{36}$$

where μ'_A is, with some abuse of notation, used to represent the level of A's posterior reputation (outside of function definitions, I am always using μ'_A as defined in (1)). Importantly, since $\mu \in [0,1]$, $\phi(\mu)$ is concave. I briefly state the monopolist Normal platform's maximization problem, given by:

$$\max_{e_A} \; e_A\phi(\mu'_A) + (1-e_A)\phi(0) - \frac{1}{2}ae_A^2 \tag{37}$$

[25] It is assumed that the information structure M has access to is public knowledge.

[26] A uniform distribution of users is used, so that A's market share is again equal to the indifferent user's location.

As usual, the Normal platform takes users' conjectures as fixed, when choosing optimal effort. The relevant first-order condition that any interior solution must satisfy is $\phi(\mu'_A) - \phi(0) = ae_A$ and proceeding identically to the previous monopoly analyses, I find the unique monopoly equilibrium effort, e^M. In this setup, $\beta \geq 2$ is sufficient to guarantee existence and uniqueness of solution to the equilibrium first-order condition, for all positive values of a and μ.

I proceed with the analysis of duopoly. Competition is identical to the setup described in Section 4. The only extra assumption that needs to be made, is that firm M can only contract with platform A, i.e. the presence of a competing platform does not alter firm M's outside option. Thus, $r(\theta)$ is again given by (35), but once again, competition affects the market shares platform A attracts when it achieves a given level of posterior reputation. The duopoly rewards-to-reputation function is defined by:

$$\phi(\mu'_A, \mu'_B) = \sigma^0 - (\frac{1}{2} + \frac{\mu'_A - \mu'_B}{2\beta})^{-1} \tag{38}$$

The discussion of Figure 3 applies again, however it captures only part of competition's impact in this case. The two effects of competition discussed in the previous section appear again: the "incremental location" effect is, as always, present. The expected increase in market share induced by achieving a high posterior reputation falls in the presence of a competitor, pushing effort incentives downwards. On the other hand, due to the decreasing marginal value of data in estimating the population mean, the extra profit generated by an increase in market share is higher, at lower market share levels. Similar to the "popularity" effect arising in Section 5 for platform A when $k > 1$, the "marginal value" effect arises in this setup, but both link to Jensen's Inequality.

Denoting as μ'_A and μ'_B the reputation levels platforms achieve after "*no breach*" realizations, the Normal type's maximization problem can be formulated as follows:

$$\begin{aligned}\max_e \ \pi^A = [\phi(\mu'_A, \mu'_B)P^B(n) + \phi(\mu'_A, 0)P^B(b)]e \\ + [\phi(0, \mu'_B)P^B(n) + \phi(0,0)P^B(b)](1-e) - c(e)\end{aligned} \tag{39}$$

Differentiating π^A with respect to effort and replacing $\phi(\mu'_A, \mu'_B)$ with the expression in (38), yields the first order condition that any interior optimum must satisfy, for given users' conjectures $\tilde{e_A}, \tilde{e_B}$ and given e_B:

$$2\beta P^B(n)(\frac{1}{\beta-\mu'_B}-\frac{1}{\beta+\mu'_A-\mu'_B}+\frac{1}{\beta+\mu'_A}-\frac{1}{\beta})-\frac{2\beta}{\beta+\mu'_A}+\frac{2\beta}{\beta}=ae \qquad (40)$$

where, $Pr^B(n)=\mu_B+(1-\mu_B)e_B$. Revenue function (35) does not yield a rewards-to-reputation function to which Lemma 1 applies, therefore, the Law of Total Probability cannot simplify the equilibrium first-order condition by cancelling out e_B and $\tilde{e_B}$ terms. However, focusing on symmetric competitors, both characterized by revenue function (35), same cost functions and equal prior reputations, I show in the Appendix that the following holds:

Proposition 5 *In a symmetric duopoly game, where competitors' revenue functions are given by* (35)*, if $\mu_A=\mu_B$ and the common cost parameter is sufficiently high, there exists a unique, symmetric Perfect Bayesian Equilibrium.*

In contrast to Sections 4 and 5, e_B becomes relevant in determining equilibrium effort incentives for platform A. An increase in e_B has two separate effects on e_A, whose combination I refer to as the "strategic uncertainty" effect: On one hand, holding $\tilde{e_B}$ fixed, it increases the probability with which the competitor attracts a high market share. The concavity of the revenue function becomes crucial again; when $r(\theta)$ is concave and A's market share becomes low in expectation, expected marginal rewards from achieving a high reputation increase. Thus, e_A tends to increase. On the other hand, in equilibrium, $\tilde{e_B}=e_B$ and an increase in e_B will decrease μ'_B, thus decreasing $B's$ market share after $s^B=\{no\,breach\}$. When $r(\theta)$ is concave, the increase in A's expected market share will tend to decrease effort incentives. The directions of these effects are always opposite of one another. When r is convex, the directions of both effects change.

The fact that the definition of complements discussed earlier considers varying both e_B and $\tilde{e_B}$ enables it to capture both these effects of strategic uncertainty. Thus, when the first effect dominates, efforts will be complements, according to the definition discussed earlier, and substitutes otherwise. When the rewards-to-reputation function belongs to the family described in Lemma 1, these two effects exactly cancel each other out and efforts are neither substitutes, not complements. The following result is derived, according to which the second effect dominates in the particular model examined:

Lemma 2 *In any PBE of the duopoly game characterized by the revenue function* (35)*,*

effort levels e_A *and* e_B *are strategic substitutes, i.e.* $\frac{\partial^2 \pi^A}{\partial e_A \partial e_B} \geq 0$.

The importance of strategic substitutability can be demonstrated with a simple comparative statics exercise; platform B's revenue mapping is given by $r(\theta_B) = \sigma^0 - \frac{1}{\theta}$, assume that, starting at the unique symmetric equilibrium, there is a small positive shock $\epsilon > 0$ in the informativeness of the signals B's users generate. This shock implies $r(\theta_B) = \sigma^0 - \frac{1-\epsilon}{\theta}$, shifting B's best response curve[27] upwards. Given Lemma 2, e_A will *decrease* as a result of this shock.

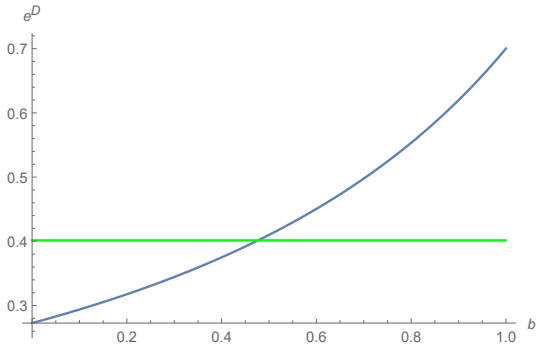


Figure 5: Model with fixed-probability competitor: Green line is $e = e^M$, while blue line is $e = e^D$, as a function of b. Parameter values used are $\mu = 0.45$, $\alpha = 2$, $\beta = 2$. There is a unique cutoff b^*, above which the "concavity" effect dominates and $e^D > e^M$.

Although both related to the convexity of $r(\theta)$, the "concavity" effect and the "strategic uncertainty" effects operate via different avenues. In contrast to the preceding discussion of the latter, the "decreasing marginal value" effect arises through the expected impact that the competitor's presence has on A's market shares, in equilibrium. This is exactly equal to $\frac{\mu_B}{2\beta}$, only depending on their prior reputation. Note that in the final expression of Lemma 1, μ_B is still present in the first-order condition, just as in the case of Section 5, meaning that μ_B is still relevant to incentives, even when "strategic uncertainty" is not.

In order to make the distinction between the "concavity" and "strategic uncertainty" effects clear, I employ a slightly modified version of the duopoly examined so far. Platform A remains unchanged and monopoly optimal effort is still e^M, defined as the unique solution to (37). I introduce a non-strategic competitor whose type is publicly known and given by $b \in [0, 1]$, i.e. the competitor's type is their fixed probability of a "no

[27] Where the best response is defined in analogy with the strategic substitutability concept.

breach" outcome in each period. The introduction of this change shuts down the "strategic uncertainty" effect of competition on $A's$ maximization problem, while it preserves the "incremental location" and "marginal value" effects. In this setup, since platform B's first-period outcome does not affect users' beliefs about $P^A(b)$ or $P^B(b)$, there are only two relevant period two states, corresponding only to the possible realizations of s_A. Fixing $\beta = 2$, the numerical output provided in Figure 5, shows that there exists a cutoff competitor type b^*, such that for $b > b^*$, the "marginal value" effect of duopoly dominates and effort provision for platform A increases with duopoly. The fact that effort incentives rise with the competitor's "no breach" probability is completely in line with the discussion of the two effects so far. A high value of b means that A's expected market shares are low, thus increasing the expected marginal value of collecting additional data. I provide the derivations behind this simpler model in the Appendix. With this discussion, I complete the analysis of Section 6.

Sections 4, 5 and 6 have all examined duopoly competition, in settings of gradually increasing complexity. In all models presented, the function mapping market shares to revenue remains unaffected by the introduction of a competitor platform. Under the linear revenue function of Section 4, competition only gives rise to the "incremental location" effect, which always favours incentives under monopoly. Non-linear rewards functions in Sections 5 and 6 give rise to different instances of the "concavity" effect of competition, which tends to favour duopoly incentives when the rewards function is concave. Lemma 1 explains why it is only in Section 6 that effort levels are strategically related. In the following section, I will revoke the underlying assumption of a fixed revenue function between monopoly and duopoly by examining a setting of two-sided competition.

7 Two-Sided Competition

I now introduce the model of two-sided competition. A monopoly platform can adjust prices it charges advertisers to maximize its revenue. In duopoly, platforms will also compete in the market for advertisers. The market for users is identical to the one described in the benchmark model, i.e. posterior reputation is still given by (1), indifferent user locations are again given by (5) and (6) under monopoly and the corresponding duopoly cutoffs are given by (11). However, instead of assuming platforms are paid an exogenously

determined amount, p, per fraction of the market they attract, I explicitly model their interaction with advertisers. There is a unit-mass continuum of advertisers, uniformly spread across a second Hotelling line. Advertisers are also assumed to be single-homing, i.e. each one advertises on a single (at most) platform. At the beginning of each period, platforms select the prices they charge to advertisers, who then decide which platform (if any) to advertise on, taking into account their relative preference between platforms, the prices set and the fraction of users each platform attracts. I use the following specification for the payoff an advertiser at location x gets from advertising on platform A.

$$\pi_t^{ad} = V + \theta_t - \beta_a x - p_t^A \tag{41}$$

V is the standalone revenue an advertiser receives by buying a platform's services, θ_t is platform A's user market share in period t, which is assumed to linearly increase the standalone revenue advertisers receive, regardless of their location. β_a is the "transportation cost" advertisers face when choosing between platforms. In a setting of targeted advertising, if one thinks of the advertisers as purchasing access to the platform's dataset, β_a can be interpreted as the match quality between the information sought by advertisers and the kind of data the platform extracts from its users. I assume that advertisers can choose what platform they endorse in each period and that no switching costs exist. The absence of switching costs and the assumed inability of platforms to commit to period 2 advertising prices in period 1 means that there are no intertemporal considerations in the platform's choice of the strategic variable p_1. Thus, users cannot make any inferences on the platform's type, based on the price it charges on that side of the market and no signalling game is induced by platforms' capacity to choose p_1. Finally, we note that advertisers do not care about the platform's reputation directly, but only instrumentally through the implied market share. The timing of the game when two platforms A and B are present in both markets is the following:

1. Nature randomly and independently chooses each platform's type.
2. Users make period 1 participation decisions based on prior reputations and effort conjectures.
3. Given period 1 market shares in the user market, platforms[28] choose period 1 advertising prices and advertisers make participation decisions based on period 1 platform market shares and prices.

[28]Both platform types choose prices, not just Normal types!

4. Platforms exert costly effort, $e1$.

5. At the end of the period, both platforms' outcomes are realized and observed by users.

6. At the beginning of period 2, users make participation decisions based on their posterior beliefs.

7. Given their period 2 market shares, platforms choose period 2 advertising prices and advertisers make participation decisions based on them.

8. Period 2 outcomes are realized.

7.1 Monopoly

I first examine a single, Normal type, platform positioned at A=0 of both Hotelling lines. As usual, the variable of interest is equilibrium effort of the Normal type, hence, I focus on the second period interactions that determine effort incentives. To find optimal first-period effort, I start by solving backwards, i.e. by finding optimal advertising prices as a function of posterior reputation, which I use to derive the rewards-to-reputation function the platform faces. Given the share of users, θ, that platform A attracts in the second period[29], given its posterior reputation and given the price it charges to advertisers, the indifferent advertiser's location is given by $x^* = \frac{V+\theta-p^A}{\beta_a}$. The optimal price is then found as the solution to:

$$\max_p (V + \theta - p^A)p^A$$

which yields:

$$p^M = \frac{1}{2}(\theta + V) \tag{42}$$

and

$$x^M = \frac{V + \theta}{2\beta_a} \tag{43}$$

Assuming $V \leq 2\beta_a - 1$ is required to ensure that even for $\theta_t = 1$, market share in the advertisers' market is interior. This is not an unreasonable assumption to make, given our current specification for advertiser profits; V is the utility they derive from advertising on a platform with no users, so it's presumably small. As usual, θ is given by (5) and (6) as a function of posterior reputation. Hence, both optimal price and the resulting market

[29] I drop the t subscript because I only analyze the second period.

share in the advertisers' market are functions of the posterior reputation. We denote these functions as $p^M(\mu'_A)$ and $x^M(\mu'_A)$ and define the rewards-to-reputation function as the platform's optimal revenue in the second period for a particular value of posterior reputation. In other words, $\Phi(p^M(\mu'_A), x^M(\mu'_A)) = p^M(\mu'_A)x^M(\mu'_A)$ and using expressions (42) and (43) yields:

$$\Phi^M(\mu'_A) = r^M(\theta(\mu'_A)) = \frac{(V+\theta)^2}{4\beta_a} \tag{44}$$

Because of the particular specification used for advertisers' profits, μ'_A enters Φ^M only through θ, hence $\Phi^M(\mu'_A)$ can be written as $r^M(\theta)$; this has also been the case in all previous sections. The revenue mapping, r^M, defined above is convex in the second period market share, which is in turn linear in posterior reputation, hence Φ is also convex in μ'_A. The convexity of r^M already suggests that competition between platforms will apply downward pressure on equilibrium effort incentives, through a negative "concavity" effect. I should emphasize that the non-linearity of r^M is not inherent in the revenue mechanism, in contrast to Section 6, but is generated by the fact that the platform can respond to its reputation by adjusting p^M. If price in the advertising market was fixed and higher reputation simply translated to attracting more advertisers, r^M would be linear in θ and the model would reduce to a marginally different version of the benchmark.

7.1.1 Equilibrium

Having derived the optimal rewards-to-reputation function, I can proceed with deriving optimal effort. The Normal-type monopolist platform solves the following problem:

$$max_e \pi^A = e_1\Phi^M(\mu'_A) + (1-e_1)\Phi^M(0) - C(e_1) \tag{45}$$

The solution concept employed again is Perfect Bayesian Equilibrium. In any equilibrium, prices are chosen optimally in the advertising market in both periods, by both platform types. Commitment type always chooses $e = 1$ in both periods. Normal type exerts, as usual, no effort in the second period and chooses first-period effort optimally by solving the above maximization problem. Users' conjecture on the effort of a Normal type platform must be correct. In order to facilitate the derivation of analytical solutions, I use a linear cost function, $C(e_1) = ae_1$. A brief discussion is required at this point of the process I follow to seek an interior equilibrium. The first-order condition of the maximization problem described above is:

$$\Phi^M(\mu'_A) - \Phi^M(0) = a \tag{46}$$

When users' conjectures are taken as fixed, the above equation is independent of the true effort level e_1 chosen by a Normal type platform[30]. This is because, given the linear cost function, the maximand is linear in e, hence an interior solution can only be optimal in the knife-edge case where the marginal cost and marginal benefit of e (which are both fixed for exogenous conjecture $\tilde{e}$) are exactly equal. In that case, the platform is actually indifferent between any effort level. However, I argue that the only interior effort conjectures, which are consistent with equilibrium play, are the ones that satisfy equation (46) and induce the kinfe-edge case. In the proof of the following proposition, I provide conditions such that corner solutions cannot be supported in equilibrium and there exists a unique solution to equation (46). Under these conditions, there exists a unique PBE of the monopoly platform game. I summarize the above analysis in the following proposition, whose proof is in the Appendix:

Proposition 6 *In the monopoly platform game whose timing is described above, there exist parameter constellations such that: price in the advertising market is uniquely defined as a function of each period's market share in the users' market; this function is the same for both Normal and Commitment-type firms; there exists a unique equilibrium effort level, e^M, played by Normal-type firms. Finally, user conjectures are correct and in equilibrium, the Normal-type firm is indifferent between any effort level.*

To maintain focus on the conceptual changes introduced by this model, I will not discuss comparative statics and instead proceed directly with the analysis of duopoly under two-sided competition.

7.2 Duopoly

Proceeding as usual, I introduce a competitor platform at B=1, but this time in both Hotelling lines, i.e. both markets. The timing of the game is exactly the one described at the beginning of this section; on the advertisers' side, platforms play a standard Hotelling game in every period, given their reputations and thus given their user market shares. The indifferent advertiser's location[31] is now given by:

$$x^D = \frac{1}{2} + \frac{1}{2\beta_a}((2\theta - 1) + p^B - p^A) \tag{47}$$

[30]The conjecture ,$\tilde{e}$, enters equation (7) through μ'_A

[31]$V \geq \beta_a$ is sufficient to guarantee that all advertisers receive positive utility in equilibrium.

In the equilibrium of the advertising market subgame, the corresponding optimal prices, given user market shares are given by:

$$P_a^D = \beta_a + \frac{1}{3}(2\theta - 1) \tag{48}$$

and[32] platform B's price is:

$$P_b^D = \beta_a - \frac{1}{3}(2\theta - 1) \tag{49}$$

Defining the rewards function similarly to the monopoly case, Φ^D is simply the revenue the platform receives in the second period, when posterior reputations are given by μ'_A, μ'_B, so that:

$$\Phi^D(\mu'_A, \mu'_B) = r^D(\theta(\mu'_A, \mu'_B)) = (\beta_a + \frac{1}{3}(2\theta - 1))(\frac{1}{2} + \frac{1}{6\beta_a}(2\theta - 1)) \tag{50}$$

Posterior reputations only influence Φ^D via the market share, hence Φ^D can also be expressed as $r^D(\theta)$. In sections 4, 5 and 6, this is the case by default, because of the following underlying assumptions: First, posterior reputations impact monopoly incentives only through θ. Second, competition only affects the mapping of reputation to market share and not the mapping from market share to revenue. By introducing competition in the second market, the second assumption is revoked, hence there appear two different mappings, r^D and r^M. By assuming that the platforms can respond to their posterior reputations, I also revoke the first assumption. It is simply because of the chosen functional forms that reputations' impact incentives only through their effect on θ. I chose to present such a model, so that the intuition developed so far can extend more naturally.

Little algebra reveals that ϕ belongs to the family of functions to which Lemma 1 applies. Thus, $\frac{\partial^2 \pi^A}{\partial e_A \partial e_B} = 0$ and the "strategic uncertainty" effect of competition does not arise. The maximization problem for the Normal type located at A=0 is now defined by:

$$max_e \quad eE(\Phi^D(\mu'_A, \mu_B(C|s_B)) + (1-e)E(\Phi^D(0, \mu_B(C|s_B))) - C(e) \tag{51}$$

where the expectation is taken over the distribution of the competing platform's period 1 outcome, s_B, and $\mu(C|s_B)$ is the posterior reputation of platform B, following a realization s_B in the first period. I derive the first-order condition in equilibrium, assuming again a linear cost function $C(e) = ae$. As it turns out, the intuition stemming from the analysis of the "incremental location" and "concavity" effects extends to this setting and I prove the following result in the Appendix, for the case of $\beta = 2$:

[32] Given full market coverage in the users' side, $\theta^B = 1 - \theta^A$.

Proposition 7 *For a sufficiently large, for any set of parameters that yields unique equilibria in both the monopoly and duopoly games, it holds that* $e^M \geq e^D$. *Under the same condition, the set of parameter configurations that induce a corner solution* $e^M = 1$ *is a superset of the configurations that induce* $e^D = 1$ *in the duopoly game.*

I provide the following numerical output to demonstrate Proposition 7. For sufficiently high prior reputation, both e^M and e^D attain corner solutions, $e^M = e^D = 1$. However, it always holds that $e^M \geq e^D$ and the set of parameter values that yield $e^D = 1$ is a strict subset of the set that yield $e^M = 1$. Note that $a = 0.1$ does not satisfy the condition of Proposition 7 derived in the Appendix. Proposition 7 is a result of restrictive, sufficient conditions and Figure 6 illustrates that the same results also arise for wider values of parameters. In explaining these results, the analysis of Section 5 is once again

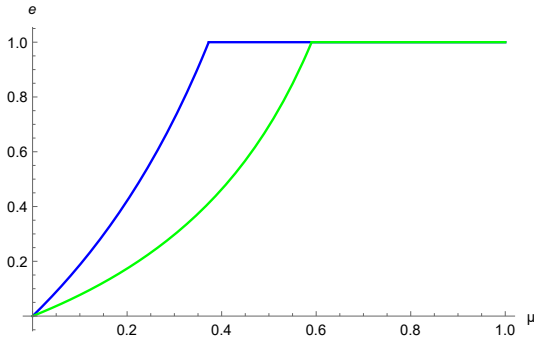


Figure 6: The blue line represents e^M, as a function of prior reputation, while the green line corresponds e^D. Parameters used were: $V = 2.5, \beta_a = 3, \beta = 2, \mu_B = 0, a = 0.1$.

useful, albeit not sufficient. The "incremental location" effect of competition is once again present; under duopoly, a given increase in posterior reputation yields a smaller expected increase in market share, on the users' side. The intuition behind the "concavity" applies again (this time though, as a "convexity" effect, favouring monopoly effort) since rewards to reputation are convex in both monopoly and duopoly. A brief discussion, however, is required to explain the difference from previous sections.

If under duopoly, platform A's interaction with advertisers remained the one described in Section 7.1, i.e. there was no competition with B in the second market, then $r^D(\theta) = r^M(\theta)$ would be given by (44) and competition would operate only through its effect on market shares. It becomes clear that in such a setting, the capacity of platform A to respond to its posterior reputation simply introduces another instance of the "concavity" effect,

mathematically equivalent to those examined in Sections 5 and 6. In the two-sided competition model, however, the presence of a competitor affects this mapping, making the identification of a "concavity" effect less straightforward.

I believe, however, that the following informal argument illustrates the intuition behind Proposition 7: I have argued that a convex rewards mapping arises because of the platforms' capacity to respond to posterior reputations by adjusting advertising fees. Because I have assumed that advertisers' outside option is always zero, the presence of a competitor *in both* markets implies a restriction of A's optimal response to μ'_A. In other words, according to the previous argument, the convexity of Φ^D with respect to market share θ should be lower than the convexity of Φ^M. Twice differentiating the revenue functions defined in (44) and (50) reveals that this is exactly the case. Therefore, not only does competition deteriorate effort incentives via the "incremental location" and "concavity" effects, but also through reducing the convexity of the revenue mapping.

The above discussion concludes my analysis of two-sided competition. Next, I formalize the argument that the convexity of ϕ arises due to platforms' capacity to respond to their reputation in period 2.

7.3 Responding to Reputation

I assume that a monopoly platform faces a known "fixed-q" rewards function, $\Phi(\mu, q)$, whose inputs are the level of posterior reputation and the value of a second choice variable, q, the value of which is selected by the platform in the second period, after the posterior, μ, is known. When q is optimally chosen in the second period as a response to posterior reputation, $q = q(\mu)$ and I define the "optimal" rewards-to-reputation function $\Phi(\mu, q(\mu))$. One can think of the difference between the two functions as the one between cost and cost function and my intuition stems from the producer theory result that the cost function is concave in input prices. I am interested in exploring what conditions on the *fixed-q* rewards function guarantee that the optimal rewards function $\Phi(\mu, q(\mu))$ becomes convex; to that end, I prove the following proposition:

Lemma 3 *Let $\Phi(\mu, q)$ and $\Phi(\mu, q(\mu))$ be the fixed-q and optimal reward functions as described above. The optimal reward function is convex at point $(\mu, q(\mu))$ if and only if: $\frac{\partial^2\Phi(\mu,q(\mu))}{(\partial\mu)^2} - \frac{\partial^2\Phi(\mu,q(\mu))}{(\partial q)^2}(\frac{dq(\mu)}{d\mu})^2 > 0$, where $\frac{\partial^2\Phi(\mu,q(\mu))}{(\partial\mu)^2}$ is the second partial derivative of*

$\Phi(\mu, q)$ *with respect to* μ*, evaluated at* $q = q(\mu)$.

Proof: Since q is optimally chosen in the second period, given posterior reputation, an interior optimum must satisfy the FOC:

$$\frac{\partial \Phi(\mu, q)}{\partial q} = 0, \quad at \quad q = q(\mu) \tag{52}$$

and the SOC for a local maximum:

$$\frac{\partial^2 \Phi(\mu, q)}{(\partial q)^2} < 0, \quad at \quad q = q(\mu) \tag{53}$$

The first order condition implicitly defines the constant function $g(\mu, q(\mu)) = 0$. Totally differentiating g with respect to μ yields:

$$\frac{d}{d\mu}\frac{\partial \Phi(\mu, q(\mu))}{\partial q} = \frac{\partial^2 \Phi(\mu, q(\mu))}{(\partial q)^2}\frac{\partial q(\mu)}{\partial \mu} + \frac{\partial^2 \Phi(\mu, q(\mu))}{\partial q \partial \mu} = 0 \tag{54}$$

which means:

$$\frac{\partial^2 \Phi(\mu, q(\mu))}{\partial q \partial \mu} = -\frac{\partial^2 \Phi(\mu, q(\mu))}{(\partial q)^2}\frac{\partial q(\mu)}{\partial \mu} \tag{55}$$

Turning to the optimal rewards function whose second-order total derivative I wish to evaluate:

$$\frac{d\Phi(\mu, q(\mu))}{d\mu} = \frac{\partial \Phi(\mu, q)}{\partial \mu} + \frac{\partial \Phi(\mu, q)}{\partial q}\frac{dq(\mu)}{d\mu} = \frac{\partial \Phi(\mu, q)}{\partial \mu} \tag{56}$$

where the second equality follows from the first-order condition, and totally differentiating again:

$$\frac{d^2 \Phi(\mu, q(\mu))}{(d\mu)^2} = \frac{\partial^2 \Phi(\mu, q)}{(\partial \mu)^2} + \frac{\partial^2 \Phi(\mu, q(\mu))}{\partial q \partial \mu}\frac{dq(\mu)}{d\mu} \tag{57}$$

Finally, using equation (55):

$$\frac{d^2 \Phi(\mu, q(\mu))}{(d\mu)^2} = \frac{\partial^2 \Phi(\mu, q(\mu))}{(\partial \mu)^2} - \frac{\partial^2 \Phi(\mu, q(\mu))}{(\partial q)^2}(\frac{dq(\mu)}{d\mu})^2 \tag{58}$$

where the second term is positive due to the second-order condition, (53). Hence, what is only required for the convexity of the optimal rewards function is that the corresponding function for *fixed-q* is not "too concave" with respect to μ, at $q = q(\mu)$. Alternatively, I have showed that even if the *fixed-q* rewards are linear, i.e. the first term of the RHS is zero, the optimal function is convex.

This, however, is not a sufficient argument to suggest that allowing the firm to choose q increases the difference $e_M - e_D$. Working with the general version of the rewards to

reputation (i.e. giving up on the structure acquired by assuming a revenue single function $\phi(\theta)$)) complicates the analysis significantly. However, parts of the intuition behind the "incremental location" and "concavity" effects should still carry through. Lemma 3 is a first step in enhancing the intuition formed through the analysis of Section 7, that allowing platforms to respond to their posterior reputation through additional variables deteriorates the impact of competition on effort incentives. As discussed already, in order to fully appreciate how the capacity to optimize $q(\mu)$ affects the difference $e^M - e^D$, the analysis must account for potential effects of competition on the optimal choice of q, holding μ fixed. Generalizing the analysis of platform competition in multiple dimensions would be an especially interesting direction for future research, but is outside the scope of this paper.

8 Policy Application: Data Portability

In this section, I introduce no additional effects of competition on effort incentives; instead, I employ a variation of the benchmark model of Section 4 to evaluate the effect of regulators introducing "data portability". My motivation stems, of course, from the introduction of GDPR by the European Union in 2016. My primary goal in this section is to identify what additional assumptions on users' payoffs and behaviour are required for such legislation to impact effort incentives in equilibrium. I abstract from the vast number of changes that the GDPR induces by focusing on data portability, inspired by the following suggestion made by Bergemann and Bonatti [2019].

"Data portability could however be a first and important element in introducing competition among platforms. Without it, any individual who is contemplating to leave one platform for another one, currently leaves his data on the previous platform, and starts from fresh at the new platform. With data portability, he could remove the data from the old platform and immediately transfer it to new platform."

The concept described above, matches very well with the "Right to Data Portability" described in Chapter 3, Article 20 of the General Data Protection Regulation[33], which states that data subjects (the users) should be able to retrieve their data from data con-

[33]https://gdpr-info.eu/art-20-gdpr/

trollers (the platforms) in a "machine-readable" form and have the right to transfer this data to a different controller. In a social-network context, this means users can migrate to a different platform, without having to set up a new profile from scratch. The "Right to Data Portability" however, does not suffice. It must be complemented with the "Right to Erasure" which is laid out in Article 17 of the same GDPR Chapter[34]. As Bergemann and Bonatti mention in the provided quote, it is important that data can be removed from the old platform, when the data subject wishes so. I now proceed to discuss under what additional assumptions, these two pieces of legislation, will impact effort incentives in my benchmark model.

8.1 Modified Model

So far I have explicitly assumed that users only suffer disutility from breaches that expose their current-period data. As mentioned in Section 3.1, the disutility a user suffers from a breach in period 2 is independent of their activity in period 1. This means that conditional on platform reputations and effort conjectures, users face identical problems in the two periods. I am now considering the case where users suffer additional disutility in period 2, of value c, when their period 1 data is breached. I will assume that the users incur the same privacy cost whether they have already been exposed already in period 1 or not[35]. A reasonable assumption would be that the disutility of having past data exposed in period 2 is weakly lower than the disutility of having current-period data exposed, i.e. $c \leq 1$. However, such an assumption has no impact on the analysis that follows.

Without "the right to data erasure", a user cannot influence the probability of his previous period data being exposed through her 2nd-period participation decision, hence the cost c is not taken into account when making that decision. In other words, without "the right to data erasure" the retroactive privacy cost is sunk in period 2. Under "the right to data erasure", I assume that users have the ability to delete their period 1 data at the

[34]https://gdpr-info.eu/art-17-gdpr/

[35]This assumption can be motivated by thinking that different bad actors can get hold of personal data in different times. If a certain group gains access to my data in the first period, it will hurt me again if this 1st period data is obtained by a different bad actor in period 2. This type of disutility is what I refer to as the "retroactive" privacy cost in period 2; users cannot influence whether they suffer this cost or not through their second period decision.

beginning of period 2 and two contrasting cases arise: If users do not need their past data to continue using a platform in period 2 and only face a potential privacy cost in the event of a second-period *breach*, then in any equilibrium, every user finds it optimal to delete their first-period data at the beginning of $t = 2$. This is the same result that arises in Ichihashi [2020a]. If that is the case, then this type of legislation will not impact effort incentives (but it will increase users' utility).

On the other hand, I will be focusing on the case in which users migrating from one platform to the other between periods will also transfer their first-period data[36]. It becomes immediately apparent, that for "data erasure" to be relevant to effort incentives in my model, it needs to be accompanied with "data portability". Data portability in this context means allowing users to transfer their past data to their 2nd period platform, if they decide to switch. Thus a user in period 2, when making her participation decision, will take into account the privacy cost of having her period 1 data exposed. Thus, under the assumptions made in this paragraph, users will be more sensitive to the platforms' posterior reputation, when making their period 2 participation decisions. Finally, under the "right to erasure", users' period 1 decisions are independent of the retroactive cost, c.

In the simplest case where introducing data portability has an impact on effort incentives, I further assume that:

1. Users' cost c is independent of whether their data has been breached already or not.
2. Platforms' revenue per user in period 2 is the same for re-attracting existing users and poaching users from the competitor's previous user base.

When the above hold, the change induced by portability legislation is summarized by the change in the expected utility a user receives by joining platform A in period 2:

$$EU(\alpha_2 = A, \theta) = \beta(1-\theta) - (1+c)P(s_2^A = b|I_2) \tag{59}$$

I examine only the duopoly setting; I solve platform A's maximization problem, assuming that the setup is otherwise identical to the benchmark duopoly presented in Section 4, the

[36]This is not such an obvious assumption; In a social media setup, it is certainly not the case that users would plausibly always want their entire profile transferred to a different platform. However, for a business-oriented platform, where users are firms and data concerns e.g. past transactions, it is a more plausible assumption to make.

equivalent of equation (15) now is:

$$e^2 + \frac{\mu_A}{1-\mu_A}e - \frac{p(1+c)}{2\alpha\beta}\frac{\mu_A}{1-\mu_A} = 0 \tag{60}$$

and it's easily seen that introducing portability raises duopoly equilibrium effort, for any $c > 0$. Note here that introducing portability means that the lower bound that β must be above of to ensure full market coverage in period 2 is raised. The above initial analysis can be augmented through the extensions presented in Sections 5,6 and 7, to further examine the impact of such legislation on duopoly incentives. The complementarity between such legislation and competition echoes arguments made in Bergemann and Bonatti [2019] and Ichihashi [2020a].

8.2 Poaching Users

In this subsection, I will be relaxing assumption 2, to explore a more interesting setup[37], in which poached users are more valuable to platforms in the second period. This is only sensible in a setting where the platforms' revenue depends on the amount of data users supply. If platforms (i.e. the advertisers that provide the platforms' revenues) are only interested in the number of users (i.e. the number of clicks or impressions) then it is harder to see why stolen users are more valuable than existing ones; however, even in this case, if ads are targeted, the targeting accuracy can realistically be assumed increasing in the total amount of data possessed by a platform, through a micro-foundation similar to the one provided in Section 6. If one thinks of the platform selling access to its users database, then poached users who carry both their current period and their previous period data, add more value to the database than existing users. Assuming that p is the revenue earned for attracting an existing user again and $(1+\delta)p$ is the revenue for stealing a user from platform B, I will now derive an expression for platform A's expected profit in period 2. Users' expected utilities are again given by (59).

Since the fraction of stolen users, σ, is now relevant for revenue, it must be identified. In period 2, given platforms' posterior reputations and users' effort conjectures, it is given by:

$$\sigma = max\{\theta_2(\tilde{e_A}, \tilde{e_B}, \mu'_A, \mu'_B) - \theta_1(\tilde{e_A}, \tilde{e_B}, \mu_A, \mu_B), 0\} \tag{61}$$

[37]The following idea was suggested to me by Prof Greg Taylor.

where θ_1 is platform A's first period market share and I am being more explicit than usual to emphasize how both market shares depend on effort conjectures and not true effort levels (except for in equilibrium).

$$\theta_1(\tilde{e_A}, \tilde{e_B}, \mu_A, \mu_B) = \frac{1}{2\beta}[\beta - (1-\mu_A)(1-\tilde{e_A}) + (1-\mu_B)(1-\tilde{e_B})] \tag{62}$$

$$\theta_2(\tilde{e_A}, \tilde{e_B}, \mu'_A, \mu'_B) = \frac{1}{2\beta}[\beta + (1+c)(\mu'_A - \mu'_B)] \tag{63}$$

Using $\sigma = \sigma(i,j)$, where $i,j \in \{b,n\}$, to suggest that the fraction of stolen users is a function of the period 2 state, I express platform A's expected profit as follows:

$$\begin{aligned} E\pi^A(e_A; e_B, \tilde{e_A}, \tilde{e_B}, \mu_A, \mu_B) = \\ e_A[P^B(n)(\theta_{nn}p + \sigma_{nn}\delta p) + P^B(b)(\theta_{nb}p + \sigma_{nb}\delta p] \\ + (1-e_A)[P^B(n)(\theta_{bn}p + \sigma_{bn}\delta p) + P^B(b)(\theta_{bb}\delta p + \sigma_{bb}\delta p)] - C(e) \end{aligned} \tag{64}$$

I am being more explicit with the derivations that in the rest of the paper to illustrate the problem that arises when attempting to solve the Normal type's maximization program. Expression (61) depends on effort conjecture levels and I cannot find any reasonable restrictions on user conjectures and/or prior reputations that free "poaching" terms of the max operator. The σ terms need not be differentiable for the usual analysis to carry on, but I would still need to find a way to evaluate the max expressions, after imposing the equilibrium conditions $\tilde{e} = e$.

In order to facilitate the analysis of this setup, I change the question I am asking; Instead of attempting to find a unique duopoly equilibrium, as I have done in the previous sections, I am now asking whether a symmetric equilibrium can exist in this setup (assuming equal prior reputations). In other words I am asking whether an interior effort level e_D exists, such that when users anticipate Normal types of both platforms to exert effort e_D and the Normal type of platform B indeed exerts effort e_D, it is optimal for platform A's Normal type to also exert effort e_D.

The benefit from switching to this simpler question is that when $\mu_A = \mu_B$ and users' conjectures satisfy $\tilde{e_A} = \tilde{e_B}$, then $\theta_1 = \frac{1}{2}$ and $\theta_{nn} = \theta_{bb} = \frac{1}{2}$, thus:

$$\begin{aligned} E\pi^A = e_A[P^B(n)\frac{1}{2}p + P^B(b)\theta_{nb}p + \sigma_{nb}p] \\ + (1-e_A)[P^B(n)(\theta_{bn}p) + P^B(b)\frac{1}{2}p] - \frac{1}{2}ae^2 \end{aligned} \tag{65}$$

where $\sigma_{nb} = \frac{1+c}{2\beta}\mu'_A$ and $\sigma_{bn} = 0$, since platform A takes conjectures as given. The FOC that any interior solution must satisfy becomes:

$$\frac{a}{p}e = \frac{(1+\delta)(1+c)}{2\beta}\frac{\mu}{\mu+(1-\mu)e_D} - \frac{\delta(1+c)}{2\beta}\mu \tag{66}$$

We can then say that there exists a unique symmetric equilibrium if there exists $e_D \in (0,1)$ such that equation (66) is uniquely satisfied by $e = e_D$. I summarize the above analysis in the following proposition:

Proposition 8 *For sufficiently high ratio $\frac{a}{p}$, there exists a unique value $e_D \in (0,1)$ that can be supported in a symmetric Perfect Bayesian Equilibrium of the platform competition game under the 'right to data portability' and the 'right to erasure'.*

Observing equation (66), it is easy to see that (as expected), the fraction of poached users in a given state is only relevant for effort incentives when $\delta \neq 0$. For $\delta = 0$, equation (66) reduces to (60), in equilibrium. One might have expected that $\frac{\partial e^D}{\partial \delta}$ would be positive in equilibrium, however that is not the case; the unique solution of (66) turns out to be decreasing in the parameter δ. The dynamics arising in this model are somewhat more complicated; due to the presence of (61), market shares become non-linear in posterior reputations, hence the analysis presented in previous sections does not apply readily.

What equation (66) does reveal, however, is that the effect of δ persists, even without "retroactive" privacy costs, i.e. even when $c = 0$. This means that a second, different set of assumptions has been identified, under which the discussed legislation impacts effort incentives in my model.

This discussion concludes the analysis of this policy application. The benchmark model was used to facilitate the discussion of legislation in the spirit of the "right to erasure" and "right to portability" of the GDPR. In both models presented, for such legislation to impact the privacy-protecting efforts of online platforms, it must be the case that users transfer their data along, when mitigating from one platform to the other. Assuming that this is the case, two setups have been identified in which introducing these two pieces of legislation will impact effort incentives under duopoly.

9 A Model of Price Discrimination

The final model that I present in this thesis constitutes the most significant departure from the benchmark presented in sections 3 and 4. I analyze an alternative version of the model presented so far, where online platforms have a different strategic choice to make, i.e. the intensity with which they use users' data to practice price discrimination. Various considerations justify developing this model: First, it allows for micro-foundation of users' privacy concerns; users are reluctant to give away their data (by participating in the platform) because it might make them worse-off in product markets. Second, by expressing the costs of preserving privacy in the same unit as the benefits of preserving privacy, this version of the model presents itself for welfare analysis; costs are now perceived by the platforms as loss of potential revenue when they decide not to practice price discrimination, while the benefit is measured by users' consumer surplus in the product market. In a related point, this model will also allow considering how data sharing can potentially increase total welfare by increasing the volume of trade in product markets. Third, the analysis presented in this section relates more clearly to previous literature on privacy concerns and endogenous information disclosure, briefly discussed in the Related Literature section and surveyed in Acquisti et al. [2016]. It thus allows complementing insights from this literature with of a model featuring incomplete information and reputational concerns. All the above reasons make the analysis introduced in this section a particularly attractive stepping stone for future research. I view the model introduced below as the first step in exploring additional questions around competition between online platforms with reputational concerns.

Most of the structure described in the benchmark model is present again. The model I develop below is also a noisy-signalling one, in which platforms are either Normal or Commitment types There are two periods, at the beginning of each users make participation decisions based on their expectations of the platforms' types and platforms benefit in the second period by having a high reputation, i.e. a high perceived probability of being Commitment type.

9.1 Monopoly

The major difference concerns the modelling of platforms' strategic choice and revenue mechanism. I begin with the analysis of a monopoly online platform. I assume that the monopoly platform is vertically integrated (or has full bargaining power against) a retail firm, F. Through users' activity on the platform, the latter learns each user's willingness to pay for good F[38]. I further assume that users' preferences for the online platform services (given by their location on the Hotelling line, θ) and their willingness to pay for good F, measured by the random variable w, are independently distributed. I define an online platform's choice variable for each period t, γ_t, as the probability with which the online platform respects users' privacy and does not share their willingness-to-pay data with the retail firm, F.

When the platform does share its users' data with the retail firm, the latter can employ first degree price discrimination and extract all consumer surplus that active platform users would otherwise enjoy in the product market. The importance of this mechanism is that I no longer have to resort to a reduced form privacy disutility suffered by users, but I can explicitly identify privacy disutility as the expected loss in consumer surplus that a user would suffer in the product market, if they were to join the online platform (more details on how consumers compute this loss later).

Similar to the model discussed in previous sections, I define the Commitment/non-strategic type as the "good" type, that always chooses $\gamma_t = 1$ while the Normal type always chooses $\gamma_2 = 0$ in the second period and chooses γ_1 optimally in the first period by solving a maximization problem. The following expressions summarize the function of the two types:

$$P(C) = \mu$$
$$P(s_t = 1|N) = \gamma_t$$
$$P(s_t = 1|C) = 1$$

where s is a binary random variable that takes value 1 when platform respects users' privacy, i.e. when platform does not share their data with firm F. The value of s_t is realized after users have made their participation decisions and is publicly observed. After a realization $s_1 = 0$, users are certain that they are facing a Normal type, hence posterior

[38] I use F to describe both the firm and the good it produces.

reputation is zero and when $s_1 = 1$, posterior reputation is given by $\mu(C|s = 1) = \frac{\mu}{\mu+(1-\mu)\tilde{\gamma_1}}$, where γ_1 is the policy conjecture users hold, so that the expression for the posterior reputation after the "good news" is identical to the one in the effort model. Thus, the information structure is identical to the one described in Section 3 and I omit further details to avoid further repetition.

9.1.1 Platform Revenue

The online platform receives all of its revenue by product firm F. The latter is an online retailer (e.g. a less ambitious version of Amazon that only sells a single, homogeneous product) that any consumer can visit, regardless of whether they use the platform or not. F is a monopoly over the supply of good F and in the absence of data over consumers, it would charge the monopoly optimal price, p^M and that is the price every customer would be offered upon entering the retailer's website. The demand function which is implied by the distribution of valuations $w \sim U[0,1]$ is $P = 1 - Q$, hence $\pi^M = max_p(1-P)P = \frac{1}{4}$. However, if data is available for a given consumer, that consumer would only be able to purchase the good at the personalized price, equal to their willingness to pay. When there is no consumer data available to F, its profits are given by π^M, the monopoly maximized profits, while if data is available on every consumer, i.e. when a platform that is used by every consumer shares data with firm F, profits are equal to the sum of valuations for good F (assuming production cost is zero). Given the distribution of valuations, that means $\pi^1 = \frac{1}{2}$, where the notation denotes profits under 1st degree price discrimination. When firm F gets data for a fraction θ of consumers, consumers are split between two segments and expected profits are equal to the sum of expected profits in the two segments (expectation with respect to the revealed valuations) are given by $\pi(\theta) = \theta\pi^1 + (1-\theta)\pi^M$.

Furthermore, I assume that each consumer demands one unit of good F in each period. Finally, I assume that the consumers delete their cookies after period 1. Hence they can ensure that they are not discriminated against by F if they do not use the platform in period 2, even if they have used it in period 1. This ensures that once again the only inter temporal link is the impact of the platform's period 1 policy on its period 2 reputation.

According to the above discussion, the expected surplus profit firm F earns if it gets access to data (= willingness to pay) for a fraction θ of the consumers is equal to $r(\theta) =$

$\pi(\theta) - \pi^M = \theta(\pi^1 - \pi^M)$. Having assumed vertical integration/full bargaining power to the platform, $r(\theta)$ is the revenue a platform earns if it shares data on its user base with F. Thus, the maximization problem a Normal type platform faces can be stated as follows[39]:

$$max_\gamma[(1-\gamma_1)r(\theta_1) + \gamma_1 r(\theta(1)) + (1-\gamma)r(\theta(0))] \tag{67}$$

Similar to the effort model, θ_1 is the fraction of users that participate in the platform in the first period and it is independent of the true policy γ_1 that Normal types use. $\theta(1)$ is the user market share after $s_1 = 1$ i.e. after no price discrimination has occurred in the first period and $\theta(1) \geq \theta(0)$. This last inequality is what drives the Normal platform's incentives to use $\gamma_1 < 1$; a higher reputation in the second period will lead to a higher users' market share, the data of whom will be certainly shared. In contrast to the effort model, the platform's choice is now relevant for period 1 profits too and thus it is not necessary for a stricter privacy policy to be directly costly through a cost function $C(\gamma)$, for a relevant economic trade-off to emerge.

9.1.2 Users' Expected Utility

When there is only a monopoly platform at point A=0 of the Hotelling line, users who join that platform get expected utility given by:

$$EU(\alpha_t = A, \theta) = v - \beta\theta - E(\Delta) \tag{68}$$

where Δ is the expected loss in consumer surplus a user faces in the product market if their willingness to pay is shared with firm F. All users derive the same standalone utility v from using the platform's services and in order to maintain a single-dimension of user heterogeneity, i.e. their location θ, I assume that consumers do not know their valuation, w, of firm F's product, prior to visiting firm $F's$ website. It holds that $E(\Delta) = (1-\mu)(1-\tilde{\gamma_1})E(\Delta|s_t = 0)$. The expectation in the loss term is now taken according to the distribution of w, which is assumed to also be U[0,1]. In order to compute the loss term, I must identify the counterfactual scenario relative to which the decrease in consumer surplus is defined. As mentioned before, consumers will face the monopoly price, when their data is not known to the platform. When a consumer computes the expected loss in consumer surplus conditional on their data being leaked to F, they know

[39]For fixed market shares, the maximand is linear in γ hence we would only get interior solutions in the knife-edge case of indifference. γ does however appear in the posterior reputation and period 1 market share, hence I can attempt to find a unique interior equilibrium, after imposing that conjectures are correct.

that with probability 0.5 their valuation is less than monopoly price, hence they suffer no loss between the two scenarios - they would get zero surplus in both cases. With probability 0.5, their w is more than 0.5, hence they suffer a decrease in CS when 1st degree price discrimination is employed and the latter is given by $E(w - p^M | w > 0.5) = 0.25$. Thus, $E(\Delta | s_t = 0) = 0.5 \cdot 0.25 = 0.125$ and $E(\Delta_1) = \frac{1}{8}(1-\mu)(1-\tilde{\gamma_1})$ and market shares/locations of indifferent users are given by the following expressions:

$$\theta_1 = \frac{1}{\beta}[v - \frac{1}{8}(1 - \tilde{\gamma_1})(1 - \mu)] \tag{69}$$

$$\theta_2(s = 1) = \frac{1}{\beta}[v - \frac{1}{8}(1 - \mu')] \tag{70}$$

$$\theta_2(s = 0) = \frac{1}{\beta}(v - \frac{1}{8}) \tag{71}$$

A comment before I present the solution of the Normal-type's maximization problem: consumer losses do not exactly correspond to increase in firm F revenue. When all users join the platform and the latter shares their valuation data with F, consumers lose on aggregate utility equal to 1/8. This is transferred to the firm alongside the monopoly dead-weight loss. As F learns consumers' valuations, it is able to make personalized offers and expand the volume of trade relative to the monopoly case. Through this very simple mechanism, price discrimination results in a loss of consumer surplus, but an increase in social welfare. The case of perfect price discrimination makes this point as straightforward as possible, but similar considerations should extend to more sophisticated methods.

To solve the maximization problem stated above, I proceed exactly as in section 7.1, where the linear cost function was used. I differentiate with respect to γ_1, holding user conjectures fixed and then impose that conjectures are correct in equilibrium. In equilibrium, the Normal type platform will be indifferent between any effort level, but the equilibrium is the only effort level at which profit-maximizing behaviour and correct beliefs can be sustained simultaneously. The resulting first order condition in equilibrium is:

$$(1 - \mu)^2\gamma^2 + [8v(1 - \mu) + \mu(1 - \mu) - (1 - \mu)^2]\gamma + (8v\mu - \mu(1 - \mu) - \mu) = 0 \tag{72}$$

I summarize the results of the above analysis in the following Proposition, which I prove in the Appendix.

Proposition 9 *For any $\mu \in (0,1)$, there exist pairs of values for v and β that guarantee the existence of a unique, interior, pure-strategy Perfect Bayesian Equilibrium. The Normal-type platform chooses $\gamma_1 = \gamma^M$ optimally, given user conjectures; users' conjecture is correct, $\tilde{\gamma_1} = \gamma^M$.*

Having identified the solution under monopoly, I immediately proceed with the analysis of duopoly.

9.2 Duopoly

To maintain tractability, I introduce a competitor of fixed type. Similar to what I have done in section 6, I assume that the platform at $B = 1$ is of fixed, publicly known type, $b \in [0,1]$, where $b = P(s^B = 1)$. Just like in the analysis of monopoly, $s^B = 1$ when platform B does not share users' data with a product firm. In the simple case that I examine, I assume that B shares data with a different retailer and that demands in the two retail markets (i.e. consumer valuations) are independent in every period. Assuming that Δ[40] terms are equal in both product markets, i.e. that consumers face equal expected "privacy costs" when their data is shared by either platform, cutoffs in each period and state are given by:

$$\theta_1 = \frac{1}{2} + \frac{\Delta}{2\beta}[(1-b) - (1-\tilde{\gamma})(1-\mu)] \tag{73}$$

$$\theta_n = \frac{1}{2} + \frac{\Delta}{2\beta}(\mu'_A - b) \tag{74}$$

$$\theta_s = \frac{1}{2} - \frac{\Delta}{2\beta}(b) \tag{75}$$

As shown in the monopoly model discussion, when valuations for the two products are distributed uniformly in $[0,1]$, $\Delta = \frac{1}{8}$. With independent retail markets, $r(\theta)$ terms are equal between periods and in every period 2 state, so $A's$ Normal type's maximization problem is:

$$max_\gamma\{\theta_1(1-\gamma) + \gamma\theta_n + (1-\gamma)\theta_s\} \tag{76}$$

so that the first order condition that must be satisfied in any equilibrium is: $\theta_1 = \theta_n - \theta_s$. Holding users' conjecture fixed, competition has two effects on platform A's incentives: by inducing a lower market share in period 1, it reduces profits earned by sharing consumers' data, thus increasing incentives towards a high γ. This effect is absent in the effort model,

[40] Using Δ as a shortcut for $E(\Delta|s_t) = 0$.

where period 1 revenue is irrelevant to effort incentives. On the other hand, by reducing the incremental increase in market share achieved by having a high reputation in the second period, competition reduces reputational premia and thus incentives for a strict privacy policy. When the equilibrium condition $\tilde{\gamma} = \gamma$ is imposed, the first order condition becomes:

$$\Delta\gamma + \frac{1}{(1-\mu)}[\beta + \Delta(1-b) - \Delta(1-\mu) + \Delta\mu]\gamma + \frac{1}{(1-\mu)^2}[(\beta-\Delta)\mu + (\mu-b)\Delta\mu] = 0 \quad (77)$$

In contrast to the benchmark effort model, because of the relevance of period 1, b matters for protection incentives. A very high b, i.e. a large prob that the competitor will protect users' privacy, means period 1 market share will be suppressed –but differential period 2 revenue is unaffected, just like in the effort model with linear revenue; this analysis is not nested in the return-to-reputation generalization employed before– hence gains from sharing users data drop and equilibrium γ should rise. Similarly, a lower initial reputation, i.e. a higher $(1-\mu)$ term, means that for any given level of conjecture, initial market share drops, hence incentives to share data drop. This should push towards a lower γ (more data sharing) when prior reputation rises. On the other hand, the "learning" effect of prior reputation is again present and the market share differential between period 2 states rises with prior reputation, which pushes incentives to preserve privacy up. These two conflicting effects will determine the impact of prior reputation on equilibrium privacy policies.

The relevance of period 1 makes this model quite different from the original one. Even this simple version, in which product markets are independent produces results starkly different , regarding duopoly effort. The above discussion shows that the role of reputation is more ambiguous in this setup (in contrast to Corollary 1 and the respective duopoly comparative statics). Figure 4 illustrates some of the dynamics between prior reputation μ, competitor's type, b, and equilibrium privacy policy in the duopoly model. While created under particular values of β, Δ, v, Figure 7 illustrates dynamics that arise for broader sets of parameters. First, it shows that a γ^D is increasing in the competitor's type , b and this seems to be robust across different values of μ_A. It is useful at this point to refer to the first order condition, $\theta_1 = \theta_n - \theta_s$. The RHS, which is the expected marginal benefit of increasing γ, evaluates to $\frac{\Delta\mu'_A}{2\beta}$, while the LHS, the marginal cost, is given by (86). Thus, a higher value of b only decreases the expected marginal cost of a higher γ, by reducing

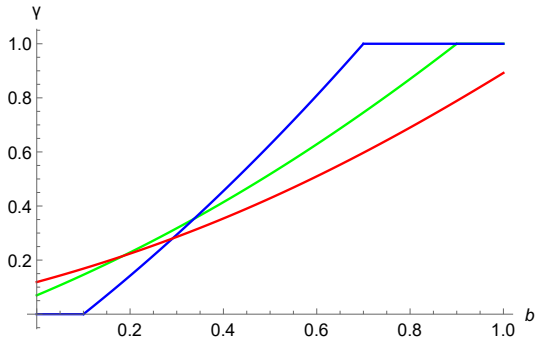


Figure 7: Equilibrium privacy policy plotted as a function of the competitor's type, b. Parameters used are $v = 0.15, \Delta = 0.125, \beta = 0.05$. The blue, green and red lines are drawn for prior reputation values 0.7, 0.5 and 0.3 respectively.

the profit platform A abandons when it does not share data in period 1. As a result, equilibrium γ^D is increasing in b.

Second, Figure 7 reveals that the effect of μ^A on γ^D depends on b. Looking at the vertical cross-section (i.e. holding b fixed and looking at the ranking of the three lines), it seems that $\frac{\partial \gamma^D}{\partial \mu_A} < 0$, for low values of b, while the opposite holds for intermediate and large values of b. To shed some light on this finding, I examine how a change in μ_A affects the expected marginal cost and marginal benefit of an increase in γ. Where does the effect of b on $\frac{\partial \gamma}{\partial \mu_A}$ come from? Following an increase in μ_A, there are:

- An increase in MB_γ equal to $B = \frac{\Delta}{2\beta}\frac{\partial \mu'_A}{\partial \mu_A}$, because of the higher reputation attained, following a good outcome in period 1.
- An increase in MC_γ equal to $C = \frac{\Delta}{2\beta}(1-\gamma)$, because an increase in μ_A increases 1st period profits from sharing data, hence the cost of not sharing.

These effects push sharing incentives to different directions, hence the overall effect depends on their relative magnitude. Both depend on b only indirectly, through γ; since, $\frac{\partial \gamma^D}{\partial b} > 0$, a higher b implies an increase in γ For values of γ such that $C > B$, $\frac{\partial \gamma}{\partial \mu_A} < 0$, while the opposite sign obtains when $C < B$. To match what is being shown in Figure 7, now I must argue that for sufficiently low values of γ, $C > B$, which is indeed the case, since $\lim_{\gamma \to 0} B = 0$ and $\lim_{\gamma \to 0} C = \frac{\Delta}{2\beta}$. Similarly, $B > C$, for sufficiently high values of γ, since $\lim_{\gamma \to 1} B > 0$ and $\lim_{\gamma \to 1} C = 0$.

The discussion of Figure 7 completes my introductory analysis of a model with price discrimination concerns. The relevance of period 1 for incentives introduces new dynamics in this model of micro-founded privacy costs. Further analysis is required to understand how the intuition discussed above carries through when variations similar to those of sections 5 and 6 are introduced. In this model, however, entirely new questions can be asked: Due to the micro-foundation employed, this model can be used to undertake welfare analysis and gauge the impact of platform competition on consumer surplus. By expanding the analysis to more general product demands, the model can potentially relate the impact of competition on sharing incentives, with primitives of the product demand.

In my view, a particularly interesting direction of analysis to be taken through this model would be to analyze multi-market competition between online platforms, that can lure users to give away their data *and* monetize that data in product markets. A natural starting point, using the model just introduced, would be to investigate the impact of competition when the platforms gather data on products with interrelated demands, in which case the Δ terms would be affected by the presence of a competitor too. In Section 7.3 I gave a suggestion of similar nature, regarding platform competition via multiple strategic variables. Given the increasingly expansive nature of online platforms, which are continuously trying to capture larger shares of users' time and attention, the analysis of effort incentives when platforms compete on multiple fronts or in multiple dimensions, seems a particularly relevant avenue of future research.

10 Conclusion

This paper explored the impact of competition on the incentives of online platforms to protect their users' privacy. I focused on the comparison between monopoly and duopoly to capture the oligopolistic nature of important digital-platform markets. When restricting the impact of competition on market shares, three effects capture its entire effect on effort incentives: the "incremental location", "concavity" and "strategic uncertainty" effects. The first effect of competition always arises and reduces the expected increase in market share achieved through a high reputation, hence decreasing incentives, relative to monopoly. In contrast, the "concavity" effect causes a platform whose services are preferred by the majority of users to face increased incentives in the presence of a competitor.

When the platform monetizes users' data via the estimation of population parameters, the "concavity" effect favours duopoly incentives and the impact of competition depends on the relative magnitude of the three effects. A particularly interesting instance of the "concavity" effect arises when a platform can respond to its posterior reputation via an additional strategic variable. This additional capacity of platforms seems to favour incentives under monopoly. I investigated this point by analyzing a model of two-sided competition, in which the first steps were taken in exploring the impact of competition when platforms antagonize on multiple fronts. The results presented and intuition developed through that model seem particularly worrying for policy-makers, as they strongly indicate a deterioration of effort incentives in duopoly. This points at the existence of a trade-off between the concentration of market power and sensitive data in the hands of few platforms, on one hand, and the protection of users' privacy, on the other. As corporations like Facebook, Amazon, Google, etc. keep growing in size and scope, they enter each others' traditional domains and compete in ever more dimensions; thus, the above model and results become especially relevant. These facts make the joint analysis of multi-market platform competition and privacy-respecting incentives a particularly important direction for future research and I hope that the models presented in the later sections present useful starting points for thinking about these issues.

References

Daron Acemoglu, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar. Too Much Data: Prices and Inefficiencies in Data Markets. NBER Working Papers 26296, National Bureau of Economic Research, Inc, September 2019. URL `https://ideas.repec.org/p/nbr/nberwo/26296.html`.

Alessandro Acquisti and Hal Varian. Conditioning prices on purchase history. *Marketing Science*, 24(3):367–381, 2005.

Alessandro Acquisti, Curtis Taylor, and Liad Wagman. The Economics of Privacy. *Journal of Economic Literature*, 54(2):442–492, June 2016.

Mark Armstrong. Competition in two-sided markets. *RAND Journal of Economics*, 37 (3):668–691, 2006.

Heski Bar-Isaac and Steven Tadelis. Seller reputation. *Foundations and Trends(R) in Microeconomics*, 4(4):273–351, 2008.

Dirk Bergemann and Alessandro Bonatti. The economics of social data: An introduction. *Cowles Foundation for Research in Economics, Yale University*, Cowles Foundation Discussion Papers(2171), 2019.

Jay Pil Choi, Doh-Shin Jeon, and Byung-Cheol Kim. Privacy and personal data collection with information externalities. *Journal of Public Economics*, 173:113–124, 2019.

Vincent Conitzer, Curtis R. Taylor, and Liad Wagman. Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases. *Marketing Science*, 31(2):277–292, March 2012.

Alexandre de Corniere and Greg Taylor. Data and Competition: A General Framework with Applications to Mergers, Market Structure, and Privacy Policy. *CEPR Discussion Paper DP 14446*, February 2020. URL `Available at SSRN: https://ssrn.com/abstract=3547379`.

Matthew Gentzkow and Jesse M. Shapiro. Media Bias and Reputation. *Journal of Political Economy*, 114(2):280–316, April 2006. doi: 10.1086/499414.

Matthew Gentzkow, Jesse M. Shapiro, and Daniel F. Stone. Media Bias in the Marketplace: Theory. NBER Working Papers 19880, National Bureau of Economic Research, Inc, February 2014. URL https://ideas.repec.org/p/nbr/nberwo/19880.html.

Bengt Holmström. Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, 66(1):169–182, 1999.

Shota Ichihashi. Dynamic Privacy Choices. *Working Paper*, 2020a. URL https://ssrn.com/abstract=3472151 or http://dx.doi.org/10.2139/ssrn.3472151.

Shota Ichihashi. Online Privacy and Information Disclosure by Consumers. *American Economic Review*, 110(2):569–595, February 2020b.

David Kreps and Robert Wilson. Reputation and imperfect information. *Journal of Economic Theory*, 27(2):253–279, 1982.

George J. Mailath and Larry Samuelson. Who Wants a Good Reputation? *Review of Economic Studies*, 68(2):415–441, 2001.

Steven Matthews and Leonard Mirman. Equilibrium limit pricing: The effects of private information and stochastic demand. *Econometrica*, 51(4):981–96, 1983.

A. Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374, 1973.

Curtis Taylor. Consumer privacy and the market for customer information. *RAND Journal of Economics*, 35(4):631–650, 2004.

A Benchmark Model

A.1 Propositions 1 and 2

The duopolist equilibrium first-order condition is:

$$e^2 + \frac{\mu_A}{1-\mu_A}e - \frac{\delta p}{2\alpha\beta}\frac{\mu_A}{1-\mu_A} = 0 \tag{78}$$

Define the LHS as $g(e)$, which has the same sign as $MC(e) - MB(e)$. Then, if $g(e) = 0$ has a unique solution which lies in (0,1), there exists a unique optimal effort level, which is interior. The constant term in the quadratic polynomial is negative, hence the product of solutions is negative, i.e. there is at most 1 positive solution. Thus, showing that a solution exists in (0,1) is sufficient and necessary for our purpose. Since g is continuous, I can appeal to the Intermediate Value Theorem. $g(0) = -\frac{\delta p}{2\alpha\beta}\frac{\mu_A}{1-\mu_A} < 0$, while $g(1) = 1 + \frac{\mu_A}{1-\mu_A} - \frac{\delta p}{2\alpha\beta}\frac{\mu_A}{1-\mu_A}$. Since $g'(e)$ is strictly positive for every non-negative e, $g(1) > 0$ is necessary and sufficient for the existence of an interior solution. The latter inequality requires $\frac{a}{\delta p} \geq \frac{\mu_A}{2\beta}$. Thus, assuming high enough cost parameter relative to the payoff parameter is sufficient for a unique interior optimum. If the above condition is not satisfied, that implies $g(e) < 0$, $\forall e \in [0,1]$ and optimization yields a corner solution $e_D^* = 1$. The second-order condition for the Normal type's maximization problem is very easy to check; Holding $\tilde{e_A}$ fixed, the second derivative of expected profits is simply $\frac{\partial^2 E\pi^A}{(\partial e)^2} = -a$, indicating a maximum. This will be the case in every model presented with quadratic costs, so checking second order conditions will be omitted in the rest of the proofs.

Note that since the corresponding monopoly FOC is very similar, the condition that guarantees unique interior monopoly effort is $\frac{a}{\delta p} \geq \frac{\mu_A}{\beta}$. The later is implies the former, hence any parameter constellation that yields an interior solution to the monopolist problem also does so for the duopolist's problem. Intuitively, since there is more to gain for the monopolist, it is easier for them to end up at a maximum effort corner solution. This discussion means that in Corollaries 1 and 2, I will be working with parameter constellations that ensure both equations admit unique interior solutions, i.e. $\frac{a}{\delta p} \geq \frac{\mu_A}{\beta}$.

A.2 Corollaries 1 and 2

When conditions are satisfied such that unique solutions e^M and e^D exist, I can find their analytical expressions, since FOCs in both monopoly and duopoly are quadratic equations.

$$e^M = \frac{\mu + \sqrt{4(1-\mu)\delta p\mu/(a\beta) + \mu^2}}{2(1-\mu)} \tag{79}$$

and

$$e^D = \frac{\mu + \sqrt{2(1-\mu)\delta p\mu/(a\beta) + \mu^2}}{2(1-\mu)} \tag{80}$$

The results of Corollary 1 obtain by simply differentiating expression (79). The signs of $\frac{\partial e^M}{\partial(\delta p)}$ and $\frac{\partial e^M}{\partial(a)}$ are trivial to derive; the analytical expression for $\frac{\partial e^M}{\partial(\mu^A)}$ is somewhat complicated, but Wolfram Mathematica confirms the result of Corollary 1, for all admissible parameter values.

Onto Corollary 2, I am first going to derive the sign of $\frac{\partial e^M - e^D}{\partial \delta p}$:

$$\frac{\partial e^M}{\partial \delta p} = \frac{1}{2(1-\mu)} \frac{1}{2\sqrt{4(1-\mu)\delta p\mu/(a\beta) + \mu^2}} \frac{4(1-\mu)\mu}{\beta}$$
$$\frac{\partial e^D}{\partial \delta p} = \frac{1}{2(1-\mu)} \frac{1}{2\sqrt{2(1-\mu)\delta p\mu/(a\beta) + \mu^2}} \frac{2(1-\mu)\mu}{\beta}$$

Thus, $\frac{\partial e^M - e^D}{\partial \delta p} > 0$ iff:

$$\frac{4}{2\sqrt{4(1-\mu)\delta p\mu/(a\beta) + \mu^2}} > \frac{2}{2\sqrt{2(1-\mu)\delta p\mu/(a\beta) + \mu^2}} \iff$$
$$4(1-\mu)\delta p\mu/(a\beta) + \mu^2 < 4(2(1-\mu)\delta p\mu/(a\beta) + \mu^2)$$

which holds, hence $\frac{\partial e^M - e^D}{\partial \delta p} > 0$. In the model of Sections 3 and 4, the cost parameter a enters the equilibrium condition only through the ratio $\frac{\delta p}{a}$, hence $\frac{\partial e^M - e^D}{\partial \delta p} < 0$. I move on to $\frac{\partial e^M - e^D}{\partial \mu}$. Taking the difference $e^M - e^D$ yields:

$$e^M - e^D = \frac{1}{2(1-\mu)}(\sqrt{4(1-\mu)\mu x + \mu^2} - \sqrt{2(1-\mu)\mu x + \mu^2})$$

where $x = \frac{\delta p}{a\beta}$ and uniqueness of solutions requires that $x \leq \frac{1}{\mu_A}$, according to the proof of Proposition 2. The denominator is decreasing in μ, while the numerator's monotonicity wrt μ is ambiguous. Evaluating the partial derivative of the above expression using Wolfram Mathematica, yields that $\frac{\partial e^M - e^D}{\partial \mu} > 0$, $\forall x \leq \frac{1}{\mu_A}$.

B Asymmetric Users' Distribution

B.1 Proposition 3

In equilibrium, the first order condition (20) becomes:

$$\alpha\beta^2 e^3 + 2\alpha\beta^2 \frac{\mu_A}{1-\mu_A} e^2 + \frac{\alpha\beta^2\mu_A^2 - \gamma(1-\mu_A)\mu_A}{(1-\mu_A)^2} e - (1-k+\gamma_m)(\frac{\mu_A}{1-\mu_A})^2 = 0 \quad (81)$$

where $\gamma_m = 2(1-k)(\beta-1) + k\beta$. I define the LHS of the above equation as $g(e)$ and find that the following inequalities hold:

$$g(0) < 0 \Leftrightarrow 1 + (2-k)(1-\beta) > 0$$
$$g(1) > 0 \Leftrightarrow a\beta^2 + m(2a\beta^2 - \gamma) + m^2(a\beta^2 - \gamma - (1-k))$$
$$g'(0) > 0 \Leftrightarrow a > \frac{\mu_A}{1-\mu_A}[\frac{2(1-k)(\beta-1)+k}{\beta}] > 0$$

Since all terms that appear are bounded above in absolute values, the above inequalities imply that for sufficiently convex cost function, there exists a unique interior equilibrium effort level (this is a sufficient, not necessary condition for either uniqueness and existence).

B.1.1 Monopoly Comparative Statics

Next, I provide comparative statics with respect to k, the popularity parameter. Assuming the conditions that ensure a unique equilibrium e^M, the Implicit Function Theorem ensures the existence of a local mapping $e^M(k)$, that is continuous and differentiable with respect to k in a neighbourhood of its equilibrium value, holding other parameters fixed. Having already assumed that $g'(e)$ is positive for every $e \in [0,1]$, I simply need to derive the sign of $\frac{\partial g}{\partial k}$ and appeal to the Implicit Function Theorem. The partial derivative is given by:

$$\frac{\partial g}{\partial k} = -\frac{\mu_A}{1-\mu_A}(2-\beta) - (\frac{\mu_A}{1-\mu_A})^2(2-\beta-1) \quad (82)$$

which is positive for any value of prior reputation $\mu_A \in (0,1)$, when $\beta \geq 2$. Hence, $\frac{\partial e_M}{\partial k} < 0$, for any $k \in [0,2]$ and any $\mu_A \in (0,1)$.

B.2 Proposition 4

First, I will show that the area of the trapezoid in Figure 3 that corresponds to the duopoly case is exactly equal to the left-hand side of equation (20). This is simply an illustration of the Fundamental Theorem of Integral Calculus. The area of a trapezoid is given by

$\frac{1}{2}(b+B)h$, which in this case translates to:

$$\begin{aligned}
D &= \frac{1}{2}[g(E(\theta(b)) + g(E(\theta(n))]\frac{\mu'_A}{2\beta} \\
&= \frac{\mu'_A}{4\beta}[2(1-k)(\frac{\beta+\mu'_A-\mu_B}{2\beta}) + 2(1-k)\frac{\beta-\mu_B}{2\beta} + 2k] \\
&= \frac{\mu'_A}{4\beta^2}[2(1-k)2(\frac{\beta-\mu_B}{2}) + 2(1-k)\frac{\mu'_A}{2} + 2k\beta] \\
&= \frac{\mu'_A}{4\beta^2}[2(1-k)(\beta-\mu_B) + (1-k)\mu'_A + 2k\beta] \\
&= \frac{1}{4\beta^2}[(1-k)(\mu'_A)^2 + 2[\beta-(1-k)\mu_B]\mu'_A]
\end{aligned}$$

which is exactly the LHS of equation (20). In equilibrium, i.e. when I impose that the conjecture terms present in μ'_A equal true effort levels, the above equation turns into:

$$\begin{aligned}
4\alpha\beta^2 e^3 + 8\alpha\beta^2\frac{\mu_A}{1-\mu_A}e^2 + [4\alpha\beta(\frac{\mu_A}{1-\mu_A})^2 - 2\gamma_d\frac{\mu_A}{1-\mu_A}]e \\
- [(1-k)(\frac{\mu_A}{1-\mu_A})^2 + 2\gamma_d(\frac{\mu_A}{1-\mu_A})^2] = 0
\end{aligned} \tag{83}$$

which is cubic in effort. The above equation does not reduce to its benchmark counterpart for $k=1$; this is not an error, these expressions were obtained after dividing by $(k-1)$, which would be invalid operation for $k=1$. Next, I use equation (83), to show existence of equilibrium in duopoly. As in the previous section, I define the left-hand side as $g(e)$ and appeal to the Intermediate Value Theorem (IVT). I find that the following inequalities hold:

$$\begin{aligned}
&g(0) < 0 \Leftrightarrow (1-k)(1-2\mu_B) + 2\beta > 0 \\
&g(1) > 0 \Leftrightarrow 4a\beta^2 + 8a\beta m + 4a\beta^2 m^2 - (2m^2+2m)(\beta-(1-k)\mu_B) > 0 \\
&g'(0) > 0 \Leftrightarrow 4a\beta^2 m^2 - 2\beta m + 2(1-k)\mu_B m > 0
\end{aligned}$$

where $m = \frac{\mu_A}{1-\mu_A}$. The condition on $g(0)$ holds for any $\beta \geq 2$, which is what I already assume to achieve full market coverage in duopoly. The condition for $g(1)$ is also easy to satisfy since the only negative term is bounded above by $4(\beta+2)$ which is smaller than $4a\beta^2$ for any $a \geq \frac{\beta+2}{\beta^2}$. These suffice to guarantee the existence of an interior solution thanks to the IVT. Getting sufficient conditions for a unique equilibrium is tougher for generic values of k, as the condition for $g'(0)$ suggests. In the worst case scenario, where $k=2$ and $\mu_B = 1$, the condition evaluates to $a > \frac{\beta+1}{2\beta^2 m}$, which is not satisfied for any reasonable a, for m too close to 0. However, if m, i.e. μ_A is bounded below by $\bar{\mu_A} > 0$, then I can argue that a sufficiently high will guarantee existence and uniqueness of equilibrium.

I emphasize that such conditions are not necessary but only sufficient, unlike in the proofs of Propositions 1 and 2, since the equilibrium conditions are now cubic.

B.2.1 Corollary 3

I operate under the assumptions stated in the previous paragraph that ensure existence of unique equilibrium, e^D. I appeal to the Implicit Function Theorem (IFT) to sign the partial derivatives: $\frac{\partial e_D}{\partial k}$ and $\frac{\partial e_D}{\partial \mu_B}$. By the IFT, it holds that:

$$\frac{\partial e_D}{\partial k} = -\frac{\frac{\partial g}{\partial k}}{\frac{\partial g}{\partial e}}$$

where both partial derivatives in the right-hand side are evaluated at the unique equilibrium. For the two results I seek, I need to compute the following partial derivatives:

$$\frac{\partial g}{\partial k} = -2m\mu_B e - 2m^2\mu_B + m^2 \tag{84}$$

$$\frac{\partial g}{\partial \mu_B} = 2m(1+e)(1-k) \tag{85}$$

Because I am operating under the previously stated assumptions, the sign of $g'(e)$ is positive, regardless of the true value of equilibrium effort - that's why I resorted to the very restrictive assumption of $g'(0) > 0$, to derive a result while remaining agnostic on the value of e^D, which is hard to analytically compute. The partial with respect to μ_B is positive if $k < 1$ and negative if $k > 1$. The partial derivative with respect to k is negative for any value of $e \in [0, 1]$, provided that $\mu_B \geq \frac{1}{2}$. Applying the IFT confirms the results stated in the Proposition.

C Estimating Population Parameters

C.1 Proposition 5: Unique symmetric PBE

As usual, I define $g(e)$ as the first order condition in equilibrium[41]; My argument is the following; I will show that there exists a unique value e^S, such that when $\mu_A = \mu_B$ and $\tilde{e_A} = \tilde{e_B} = e_A = e_B = e$, e^S is the unique interior solution to $g(e) = 0$. When these symmetry conditions hold, it also holds that $\mu'_A = \mu'_B = \mu'$.

$$g(e) = 2\beta P(n)(\frac{1}{\beta - \mu'} - \frac{1}{\beta + \mu' - \mu'} + \frac{1}{\beta + \mu'} - \frac{1}{\beta}) - \frac{2\beta}{\beta + \mu'} + \frac{2\beta}{\beta} - ae \tag{86}$$

where $P(n) = \mu + (1 - \mu)e$. Evaluating at $e = 0$ means $\mu' = 1$, while at $e = 1$, $\mu' = \mu$. Hence:

$$g(0) = 2\beta\mu(\frac{1}{\beta - 1} - \frac{1}{\beta} + \frac{1}{\beta + 1} - \frac{1}{\beta}) - \frac{2\beta}{\beta + 1} + 2 \tag{87}$$

where $\beta < \beta + 1$, hence $2 - \frac{2\beta}{\beta+1} > 0$; since $f(x) = \frac{1}{x}$ is decreasing and convex for positive x, it also holds that $f(\beta - 1) - f(\beta) > f(\beta) - f(\beta + 1)$; these two observations imply $g(0) > 0$.

$$g(1) = 2\beta(\frac{1}{\beta - \mu} - \frac{1}{\beta} + \frac{1}{\beta + \mu} - \frac{1}{\beta}) - \frac{2\beta}{\beta + \mu} + 2 - a \tag{88}$$

$2 - \frac{2\beta}{\beta+1} > 0$ is bounded above by 2. The same holds for the first term of $g(1)$; since $\beta \geq 2$ and $\frac{2\beta}{\beta-\mu}$ is decreasing in β, it is bounded above by 4, for $\beta = 2, \mu = 0$. Similarly, since $\frac{2\beta}{\beta+\mu}$ is increasing in β, it is bounded above by $\lim_{\beta\to\infty} \frac{2\beta}{\beta+\mu} = 2$. Thus, $g(1)$ is always negative for a sufficiently large. Differentiating $g'(e)$ yields an expression similar to (88) in which a is again subtracted and all positive terms that appear are bounded above. Thus, sufficiently high a guarantees that g is monotonically decreasing in $(0, 1)$. The Intermediate Value Theorem applies and Proposition 5 is proven.

C.2 Lemma 2: Strategic Substitutes

The first order condition for e_A is given by:

$$2\beta P^B(n)(\frac{1}{\beta - \mu'_B} - \frac{1}{\beta + \mu'_A - \mu'_B} + \frac{1}{\beta + \mu'_A} - \frac{1}{\beta}) - \frac{2\beta}{\beta + \mu'_A} + \frac{2\beta}{\beta} = ae \tag{89}$$

Differentiating with respect to both e_B and $\tilde{e_B}$, i.e. imposing that $\tilde{e_B} = e_B$ yields:

$$\frac{\partial^2 \pi^A}{\partial e_A \partial e_B} = (1 - \mu_B)(\frac{1}{\beta - \mu'_B} - \frac{1}{\beta + \mu'_A - \mu'_B} + \frac{1}{\beta + \mu'_A} - \frac{1}{\beta}) + P(n)(\frac{1}{(\beta - \mu'_B)^2}\frac{\partial \mu'_B}{\partial e_B} - \frac{1}{(\beta + \mu'_A - \mu'_B)2}\frac{\partial \mu'_B}{\partial e_B})$$

[41] With the slight difference that $g(e)$ in this section reflects the sign of $MB(e) - MC(e)$.

The derivative of μ'_B can be expressed as follows:

$$\frac{\partial \mu'_B}{\partial e_B} = \frac{\mu_B}{(\mu_B + (1-\mu_B)e_B)^2} = \frac{-(1-\mu_B)}{\mu_B + (1-\mu_B)e_B}\mu'_B$$

Thus $P(n)\frac{\partial \mu'_B}{\partial e_B} = -(1-\mu_B)\mu'_B$, thus:

$$\frac{\partial^2 \pi^A}{\partial e_A \partial e_B} = (1-\mu_B)(\frac{1}{\beta - \mu'_B} - \frac{1}{\beta + \mu'_A - \mu'_B} + \frac{1}{\beta + \mu'_A} - \frac{1}{\beta}) - (1-\mu_B)\mu'_B(\frac{1}{(\beta - \mu'_B)^2} - \frac{1}{(\beta + \mu'_A - \mu'_B)2}) \quad (90)$$

The sign of the mixed partial derivative depends on μ_B only through μ'_B. Evaluating the inequality using Wolfram Mathematica yields that the sign of the mixed partial derivative is negative for any values of $\beta \geq 2$ and $\mu'_A, \mu'_B \in [0,1]$.

C.3 Companion to Figure 5

In a duopoly of this nature, there are only two relevant period 2 states, depending on platform A's period 1 outcome. Platform B's outcome in period 1 is uninformative on either platform's period 2 outcome conditional on b. After a "breach" by A in period 1 the indifferent user is defined by:

$$\beta(1-\theta(b)) - 1 = \beta\theta(b) - (1-b) \quad (91)$$

which means $\theta(b) = \frac{\beta - b}{2\beta}$. Similarly, $\theta(n) = \frac{\beta + \mu'_A - b}{2\beta}$, where μ'_A is A's posterior reputation following 'no breach' in period 1. As usual, the FOC is given by:

$$\phi(\theta(n)) - \phi(\theta(b)) = ae \quad (92)$$

which in this setup means:

$$\sigma^0 - \frac{2\beta}{\beta + \mu'_A - b} - \sigma^0 + \frac{2\beta}{\beta - b} = ae$$

$$\frac{\mu'_A}{(\beta + \mu'_A - b)(\beta - b)} = \frac{a}{2\beta}e$$

$$\frac{a}{2\beta}(\beta - b)^2 e + \frac{a}{2\beta}(\beta - b)\mu'_A e = \mu'_A$$

multiplying by the denominator of μ'_A, i.e. $\mu + (1-\mu)e$:

$$\mu_A = \frac{a}{2\beta}(\beta - b)^2 \mu_A e + \frac{a}{2\beta}(\beta - b)^2(1-\mu_A)e^2 + \frac{a}{2\beta}(\beta - b)\mu_A e$$

$$\frac{2\beta}{a}\mu_A = [(\beta - b)^2 \mu_A + (\beta - b)\mu_A]e + (\beta - b)^2(1-\mu_A)e^2$$

Thus, the FOC in equilibrium is given by:

$$e^2 + \frac{\mu_A}{1-\mu_A}(1 + \frac{1}{\beta - b})e - \frac{2\beta}{a(\beta - b)^2}\frac{\mu_A}{1-\mu_A} = 0 \tag{93}$$

I've computed the unique solution to this equation for the parameters stated under Figure 5 and plotted it, as a function of the competitor's good outcome probability, b.

D Two Sided Competitition

D.1 Proposition 6

The first order condition that an interior effort must satisfy is:

$$\begin{aligned}
\phi(\mu'_A) - \phi(0) = a \Leftrightarrow \\
(V + \theta(\mu'_A))^2 - (V + \theta(0))^2 = 4\beta_a a \Leftrightarrow \\
2\theta(\mu'_A) + \theta^2(\mu'_A) - 2V\theta(0) - \theta^2(0) = 4\beta_a a \Leftrightarrow \\
\frac{2V\mu'_A}{\beta} + (\frac{\beta - 1 + \mu'_A}{\beta})^2 - (\frac{\beta - 1}{\beta})^2 = 4\beta_a a \Leftrightarrow \\
(\mu'_A)^2 + 2(\mu'_A)(\beta - 1) + 2\beta V\mu'_A = 4a\beta^2\beta_a := \gamma^M \Leftrightarrow \\
\mu_A^2 + 2\mu_A(\beta - 1)(\mu_A + (1 - \mu_A)\tilde{e}) + 2\beta V\mu_A(\mu_A + (1 - \mu_A)\tilde{e}) = \gamma^M(\mu_A + (1 - \mu_A)\tilde{e})^2
\end{aligned}$$

Imposing the equilibrium condition $\tilde{e} = e$ and expanding terms, yields:

$$\gamma^M e^2 + \frac{\mu}{1-\mu}[2\gamma^M - 2\beta V - 2(\beta - 1)]e + (\frac{\mu}{1-\mu})^2[\gamma^M - 2\beta V - 2(\beta - 1) - 1] = 0 \quad (94)$$

Proceeding as usual, I state three sufficient conditions that guarantee the existence of a unique, interior solution to (94):

$$\begin{aligned}
g(0) < 0 \Leftrightarrow \gamma^M < 2\beta V + 2(\beta - 1) + 1 \Leftrightarrow \\
a < \frac{2\beta V + 2(\beta - 1) + 1}{4\beta^2\beta_a} \quad (95)
\end{aligned}$$

$$g(1) > 0 \quad (96)$$

$$\begin{aligned}
g'(0) > 0 \Leftrightarrow 2\gamma^M - 2\beta V - 2(\beta - 1) > 0 \Leftrightarrow \\
a > \frac{\beta V + \beta - 1}{4\beta^2\beta_a} \quad (97)
\end{aligned}$$

Since $g'(e)$ is increasing in e, $g'(0) > 0$ is sufficient to guarantee $g'(e) > 0 \forall e \in (0, 1)$. Notice that since $\beta V + \beta - 1 < 2\beta V + 2(\beta - 1) + 1$, there always exist values of a that satisfy both conditions (95) and (97). Finally, the middle condition is satisfied when:

$$\gamma^M + \frac{\mu}{1-\mu}[2\gamma^M - 2\beta V - 2(\beta - 1)] + (\frac{\mu}{1-\mu})^2[\gamma^M - 2\beta V - 2(\beta - 1) - 1] > 0$$

When a satisfied the conditions stated, the second term of the RHS is positive and the third is negative. However, notice that for $\mu_A = 0$, the RHS is equal to $\gamma^M > 0$. Thus, $\exists \bar{\mu} > 0$, such that $\forall \mu_A < \bar{\mu}$, $g(1) > 0$. Thus for $\mu_A < \bar{\mu}$ and $a \in [\frac{\beta V + \beta - 1}{4\beta^2\beta_a}, \frac{2\beta V + 2(\beta - 1) + 1}{4\beta^2\beta_a}]$, all three sufficient conditions are satisfied and by the IVT, a unique interior solution to

$g^M(e) = 0$ exists. Notice that since for low values of μ, the negative term in $g(1)$ is multiplied by a factor of two orders of magnitude lower than γ^M, I expect the condition on $\bar{\mu}$ not to be particularly restrictive.

The interior market shares condition and the full market coverage condition that need be satisfied do not involve a or μ, hence can always be satisfied by appropriate values of β, β_a, V, without jeopardizing the above analysis.

D.2 Proposition 7

First, I present the equilibrium first-order conditions derived for the monopoly and duopoly settings of Section 7.1.

$$g^M(e) = \gamma^M e^2 + \frac{\mu}{1-\mu}[2\gamma^M - 2\beta V - 2(\beta-1)]e + (\frac{\mu}{1-\mu})^2[\gamma^M - 2\beta V - 2(\beta-1) - 1] = 0 \quad (98)$$

$$g^D(e) = \gamma^D e^2 + \frac{\mu}{1-\mu}[2\gamma^D - 6\beta_a\beta + 2\mu_B]e + (\frac{\mu}{1-\mu})^2[\gamma^D - 6\beta_a\beta + 2\mu_B - 1] = 0 \quad (99)$$

where $\gamma^M = 4\beta_a\beta^2 a$ and $\gamma^D = 18\beta_a\beta^2 a$. I will proceed with a slightly different argument of geometric nature. Since the equilibrium first order conditions derived both monopoly and duopoly are quadratic, the following are sufficient conditions for the proposed result to hold:

- the y-intercept of parabola $g^D(e)$, i.e. $g^D(0)$ is higher than the $g^M(0)$.
- the slope of $g^D(e)$ is greater than the slope of $g^M(e), \forall e \in (0,1)$, i.e. $g'_M(e) < g'_D(e), \forall e \in (0,1)$.
- The duopoly parabola is strictly increasing for positive values of e, i.e $g'(e) > 0, \forall e \in (0,1)$.

I will work with $\beta = 2$. I am going to derive sufficient conditions for the above to hold for any values of μ_B, β_a and any value of V that ensures interior market shares and full market coverage in the advertising market. I begin with the second condition, regarding the slopes.

$$g'_M(e) = 2\gamma^M e + \frac{\mu}{1-\mu}[2\gamma^M - 2\beta V - 2(\beta-1)]$$

and

$$g'_D(e) = 2\gamma^M e + \frac{\mu}{1-\mu}[2\gamma^D - 6\beta_a\beta + 2\mu_B]$$

Both are increasing in e, hence I need to derive conditions such that $g'_M(0) < g'_D(0)$. For $\beta = 2$, the latter inequality holds iff:

$$2\gamma^M - 4V - 2 < 2\gamma^D - 12\beta_a + 2\mu_B$$

This holds for every $\mu_B \in [0,1]$ if it holds for $\mu_B = 0$ i.e when:

$$2\gamma^M - 4V - 2 < 2\gamma^D - 12\beta_a \iff$$
$$28 \cdot 4a\beta_a - 12\beta_a + 4v - 2 > 0$$

The lowest value of V that ensures both full market clearing and interior monopoly market shares in every state is $V = \beta_a - \frac{1}{2}$. Using this value of V, the inequality becomes:

$$(28a - 2)\beta_a > 1 \iff a > \frac{1/\beta_a + 2}{28} \tag{100}$$

Condition (100) guarantees that $g'_D(e) > g'_M(e), \forall e \in (0,1)$ and $\forall(\beta_a, V)$ pairs that imply full market clearing and interior market shares. Next, I will derive a similar condition regarding the y-intercepts of the two parabolas.

$$g_M(0) < g_D(0) \iff$$
$$\gamma^M - 4V - 2 - 1 < \gamma^D - 12\beta_a + 2\mu_B - 1 \iff$$
$$16\beta_a a - 4v - 2 < 72\beta_a a - 12\beta_a + 2\mu_B \iff$$

Using the same "worst-case scenario" values for V and μ_B, the above inequality evaluates to:

$$56\beta_a a + 4\beta_a - 12\beta_a > 0 \iff$$
$$a > \frac{1}{7}$$

Condition (101) is more restricting than condition (100), (100) can be discarded. Finally, $a > 1/7$ is easily shown to suffice for $g'_D(e) > 0, \forall e \in (0,1)$, which is the final piece of the geometric argument made in this proof. Hence, it has been shown that when $a > 1/7$, for any combination of parameter values v, β_a, μ_A, μ_B that yields unique equilibria in both settings, the resulting equilibrium effort levels satisfy $e^M > e^D$.

E Policy Application: Data Portability

E.1 Proposition 8

Revenue for stealing a user from platform B is now: $p+\delta p$. The fraction of stolen users, σ is defined by equation (61) my goal is again to evaluate conditions under which platform A has no profitable deviation from a symmetric equilibrium. As stated in the main text, when users anticipate both platforms to exert e_D and platform B's Normal type indeed exerts e_D, then platform A's (Normal-type's) expected profit as a function of its own effort is given by:

$$\pi^A = e_A[P^B(n)\frac{1}{2}p + P^B(b)\theta_{nb}p + \sigma_{nb}p] + (1-e_A)[P^B(n)(\theta_{bn}p) + P^B(b)\frac{1}{2}p] - \frac{1}{2}ae^2 \quad (101)$$

where $\theta_{nb} = \frac{\beta+(1+c)\mu'}{2\beta}$, $\theta_{bn} = \frac{\beta-(1+c)\mu'}{2\beta}$. The first-order condition is given by:

$$\frac{a}{p}e_A = P^B(n)\frac{1}{2} + P^B(b)[\theta_{nb}+\delta\sigma_{nb}] - P^B(n)\theta_{bn} - P^B(b)\frac{1}{2}$$

$$= P^B(n)\frac{1}{2} + (1-P^B(n))[\theta_{nb}+\delta\sigma_{nb}] - P^B(n)\theta_{bn} - (1-P^B(n))\frac{1}{2}$$

$$= (P^B(n)-\frac{1}{2}) + (1-P^B(n))[\frac{\beta+(1+c)\mu'}{2\beta} + \delta\frac{(1+c)\mu'}{2\beta}] - P^B(n)\frac{\beta-(1+c)\mu'}{2\beta}$$

$$= (P^B(n)-\frac{1}{2}) - (P^B(n)-1)[\frac{1}{2} + \frac{(1+c)(1+\delta)\mu'}{2\beta}] - P^B(n)\frac{1}{2} + P^B(n)\frac{(1+c)\mu'}{2\beta}$$

And using that in the symmetric equilibrium $\mu'_A = \mu'_B = \mu' = \frac{\mu}{P^B(n)} = \frac{\mu}{\mu+(1-\mu)e_D}$:

$$\frac{a}{p}e_A = \frac{(1+c)(1+\delta)}{2\beta}\mu' - \frac{\delta(1+c)}{2\beta}\mu \quad (102)$$

which reduces to equation (60) for $\delta = 1$. It also (thankfully) reduces to the benchmark FOC for $\delta = c = 0$. Proving the stated result requires yet another appeal to the IVT. In the proposed equilibrium, when $e_A = e_D = e$, the first-order condition becomes:

$$\frac{a}{p}e + [\frac{a}{p} + \frac{\delta(1+c)}{2\beta}(1-\mu)]\mu e + [\frac{\delta(1+c)}{2\beta} - \frac{(1+c)(1+\delta)}{2\beta}] = 0 \quad (103)$$

Defining the LHS as $g(e)$, we observe that $g(0) < 0$ always holds; sufficiently large a can always guarantee that both $g(1) > 0$ and $g'(e) > 0 \forall e \in (0,1)$, yielding sufficient conditions for the existence and uniqueness of a symmetric, interior equilibrium, e_D.

F A Model of Price Discrimination

Beginning with the maximization problem, we got:

$$max_\gamma[(1-\gamma)\frac{1}{4}\theta_1 + \gamma\frac{1}{4}\theta(1) + (1-\gamma)\frac{1}{4}\theta(0)]$$

$$max_\gamma[(1-\gamma)\theta_1 + \gamma\theta(1) + (1-\gamma)\theta(0)]$$

$$max_\gamma[\theta_1 + \theta(0) + \gamma(\theta(1) - \theta(0) - \theta_1)]$$

Taking policy conjectures as fixed, i.e. the market share terms as fixed, an interior solution can only exist when the first order condition holds, which equates the MB of increasing γ, i.e. the benefit on enjoying a higher reputation in period 2, to the MC, i.e. the loss in period 1 revenue:

$$\begin{aligned}
\theta(1) - \theta(0) &= \theta_1 \\
v - \frac{1}{8}(1-\mu') - v + \frac{1}{8} &= v - \frac{1}{8}(1-\mu)(1-\tilde{\gamma}) \\
-1 + \mu' + 1 &= 8v - (1-\mu)(1-\tilde{\gamma}) \\
\frac{\mu}{\mu + (1-\mu)\tilde{\gamma}} &= 8v - (1-\mu)(1-\tilde{\gamma})
\end{aligned}$$

$$\begin{aligned}
\mu &= 8v\mu + 8v(1-\mu)\gamma - \mu(1-\mu) + \mu(1-\mu)\gamma - (1-\mu)^2\gamma + (1-\mu)^2\gamma^2 \\
0 &= (1-\mu)^2\gamma^2 + [8v(1-\mu) + \mu(1-\mu) - (1-\mu)^2]\gamma + (8v\mu - \mu(1-\mu) - \mu)
\end{aligned}$$

I define the RHS of the last equation as the function $g(\gamma)$. The sign of g is the same as the sign of (MC - MB). I will show that for every interior value of prior reputation $0 < \mu < 1$, there exists a range of values for v, such that a unique interior equilibrium exists.

$$g'(\gamma) = 2\gamma(1-\mu)^2 + 8v(1-\mu) + \mu(1-\mu) - (1-\mu)^2 \tag{104}$$

Since g' is increasing in γ, g' is positive across its support iff $g'(0)$ is positive, i.e iff $8v + \mu - (1-\mu) > 0$ or $v > \frac{1-2\mu}{8}$.

It holds that $g(0) = (8v\mu - \mu(1-\mu) - \mu) < 0$, iff $v < \frac{2-\mu}{8}$, while $g(1) > 0$ iff $v > \mu/8$. Going back to the definitions of user cutoffs (69), (70) and (71) and, we see that for

market shares to be always positive, it is required that $v > 1/8$ and assuming the last inequality also implies that g(1) is positive for any prior reputation level. Finally, $v > 1/8$ also implies that g is increasing for any level of prior reputation. Therefore, for a given reputation level $\mu \in (0,1)$, the equation $g(\gamma) = 0$ has a unique interior solution in (0,1), if $v \in (\frac{1}{8}, \frac{2-\mu}{8})$. For any $\mu \in (0,1)$, there exist values v that fall in the specified interval and these (μ, v) combinations guarantee the existence of a unique, interior, pure-strategy PBE.

A comment should be made on the requirement of full market coverage when duopoly is introduced. It is necessary that the lowest market share a platform can get is no smaller than $\frac{1}{2}$. Because of the absence of the transportation cost parameter from the FOC, we can always assume that β is small enough to guarantee that $\frac{1}{\beta}(v - \frac{1}{8}) \geq \frac{1}{2}$.