

Wohlfarth, Michael

**Conference Paper**

## Data Portability on the Internet: An Economic Analysis

28th European Regional Conference of the International Telecommunications Society (ITS):  
"Competition and Regulation in the Information Age", Passau, Germany, 30th July - 2nd  
August, 2017

**Provided in Cooperation with:**

International Telecommunications Society (ITS)

*Suggested Citation:* Wohlfarth, Michael (2017) : Data Portability on the Internet: An Economic  
Analysis, 28th European Regional Conference of the International Telecommunications Society (ITS):  
"Competition and Regulation in the Information Age", Passau, Germany, 30th July - 2nd August,  
2017, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/169506>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen  
Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle  
Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich  
machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen  
(insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten,  
gelten abweichend von diesen Nutzungsbedingungen die in der dort  
genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal  
and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to  
exhibit the documents publicly, to make them publicly available on the  
internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content  
Licence (especially Creative Commons Licences), you may exercise  
further usage rights as specified in the indicated licence.*

# Data Portability on the Internet: An Economic Analysis

Michael Wohlfarth\*

June 7, 2017

## Abstract

Data portability allows users to transfer data from one online service to another service. As data gets increasingly valuable for online services and users alike, and as data portability will be politically enforced within the European Union by the General Data Protection Regulation (GDPR), we develop a game-theoretic model that incorporates the competitive effects arising from a right to data portability. We show, among others, that although data portability is designed to protect users, a right to data portability may hurt customers because new services increase the amount of collected data compared to the case without a right to data portability. However, profits for new services and total surplus unambiguously increase with a right to data portability, improving innovation and service variety. Consequently, policy makers should be aware that the decision to enforce data portability is far more complex than currently realized and should consider the case-specific effects.

*Keywords:* Data portability, Regulation of online services, Market entry and innovation, Consumer protection, Switching costs, GDPR

## 1. Introduction

In the digital ecosystem, data is considered to be the key ingredient for many of today's business models, crucially determining whether a service is successful. At the same time, the protection of (personal) data, users and competition becomes increasingly important for policy makers and competition authorities. For example, many of the European antitrust investigations against Google attribute either to the observation that consumers might be disadvantaged or that competition and innovation is hampered (c.f., Drozdiak and Schechner, 2016 for an overview of European antitrust probes against Google). In fact, personal data entered or

\* Chair of Internet and Telecommunications Business, University of Passau, e-mail: michael.wohlfarth@uni-passau.de

revealed<sup>1</sup> at a specific online service may lead to a lock-in effect for users as switching to competing services induces costs to re-enter the data required by the new online service (c.f., Klemperer, 1987 for related research). Hereby, innovation and service variety might be reduced as market entry is deterred. Illustrative examples are provided by online banking accounts (where switching leads to the necessity to re-enter recurring transfers), online mail or storage services (where switching leads to the necessity to re-enter general user information, and to re-upload files, photos, contacts or categories), or cloud computing environments (where preferences and adaptations have to be re-injected). Interestingly, these services suggest that a lock-in does not necessarily stem from network effects alone, i.e., the number of participating users or complementary provided services. Instead, as Chen and Hitt (2002) analyze empirically, there is a variety of factors (additionally) influencing a users' loyalty. We build on these observations and argue that the (amount of) already revealed (personal) data has a crucial role for (i) online services active in data-driven markets since it determines the service's profitability, and also (ii) users since they might be locked-in to a certain service.

It is well known that established systems designed to lock-in users may hamper the success of new services and lead to excessive rents of incumbent firms (c.f., Katz and Shapiro, 1994; Farrell and Klemperer, 2007) and so –eventually– to market failures. Accordingly, the European Commission has recently formulated a general “right to data portability” for personal data. Consequently, a standardized way of how information can be ported from one online service to another is required (c.f., European Commission, 2016b, p.45, Article 20); an issue most voluntarily provided abilities for users to export previously revealed data do not explicitly account for (c.f., Facebook, n.d.; Google, n.d.) and an issue also highlighted by the Deputy Chief Technology Officer of the United States (c.f., Macgillivray & Shambaugh, 2016). Ultimately, the European Commission's incentive aims to promote users' negotiation power vis-à-vis (dominant) online services by reducing lock-in effects, i.e., protecting the “fundamental rights and freedoms of natural persons” (European Commission, 2016b, p.32, Article 1). However, the economic effects of such an intervention on consumer's surplus, on the amount of data online services collect from their customers, on online service's profits, and on service variety are unclear to date. Albeit the regulation becomes binding for all European member states as of May 2018 and policy makers already discuss the possibility to port non-personal data (c.f., European Commission, 2017, p.11), academic analyses have so far been restricted primarily to the legal and technical dimensions of data portability. An analysis of strategic incentives, business strategies and economic outcomes is lacking, as Nobel prize laureate Jean Tirole outlined in a speech on competition and regulation of online platforms (c.f., Valero, 2015).

This paper tackles this research gap and analyses the competitive effects of a user's ability to port data from an incumbent online service / content provider (CP) to a joining one. Next to

---

<sup>1</sup> We will use the terms entered and revealed interchangeably.

firms' incentives (not) to promote data portability, we shed light on the ensuing effects on consumers as this is pivotal to the argumentation of the European Commission and the U.S. Deputy Chief Technology Officer, alike. In doing so, we develop a game-theoretic model that considers the economic effects arising from a right to data portability by considering two CPs generating revenues primarily through data revealed by users active at their platform. Thus, we abstract from any explicit business model (e.g., based on advertisements, or based on selling aggregated user-data to third parties) and simply assume that data revealed by users somehow transforms into revenues. Hence, additional data has a positive effect on a CP's profits. On the other hand, revealing data bears "costs" (i.e., a disutility) for users: Either as they have some effort revealing data as such (say, the time needed to enter the data), or –more general– as users give away data they attribute some value to. Consequently, whereas collecting more data is beneficial for CPs, users experiencing a higher disutility might switch to competing CPs or leave the market entirely. However, this opportunity is impeded by established switching-costs and lock-ins. The ability to port data by means of data portability arguably circumvents the established restrictions, but may also have impact on the data consumption of the CPs taken into consideration. These effects have to be taken into account when analyzing the competitive effects.

Our derived results show that data portability is not necessarily beneficial for users as joining CPs may demand for increased data revelation under this regime. Thus, the ultimate goal to protect users is not necessarily achieved. Conversely, the CPs' incentives (not) to promote data portability are unequivocally: whereas dominant CPs (incumbents) always suffer from data portability, emerging CPs (entrants) challenging incumbents are always better off. Nevertheless, as total surplus unambiguously increases under a data portability regime, predominantly due to the arising benefits for the entrant able to generate more revenues, the decision to enforce a right to data portability is far more complex than currently realized. These results are also valid if network effects are considered.

The remainder of this article is organized as follows. In Section 2, we briefly review previous literature related to data portability and highlight the multi-faceted character of this topic. Next, in Section 3, we lay out the game-theoretic model and assumptions. The model's main results are deduced and discussed in Section 4. Section 5 extends the model to account for network effects and shows the robustness of the deduced results. Finally, Section 6 concludes by highlighting policy as well as managerial implications, limitations, and avenues for future research.

## **2. Literature Review**

We refer to data portability as consumer's ability to transfer (personal) data revealed at one CP to another CP. To the best of our knowledge, the economic literature as well as literature in the field of information systems has so far not considered this concept explicitly in terms of

strategic incentives, business strategies, or economic outcomes. Albeit, the technical literature exposed the feasibility of that concept by proposing models to conveniently port data, e.g., between cloud computing vendors. In this vein, Ranabahu and Sheth (2010) propose semantic web techniques to achieve portability and Petcu and Vasilakos (2014) inter alia highlight open standards and open application programming interfaces (APIs) as technical solutions. Nevertheless, most technical studies are a proof of concept that data portability is technically feasible and do not explicitly discuss the occurring trade-offs for the involved parties.

In light of the already mentioned General Data Protection Regulation becoming effective in 2018, several legal investigations have been carried out. Graef (2015) conducts a legal analysis of data portability in social networks with respect to the (European) competition law and summarizes relevant cases. Swire and Lagos (2013) explicitly refer to consumer welfare and “express serious concerns about the RDP [right to data portability]” (Swire and Lagos, 2013, p.338) as, (i) the problems addressed by the regulation (e.g., monopoly power through lock-ins) are legally already covered by competition law, (ii) personal data can easily be exported, i.e., security problems arise, and (iii) it is unclear how a common standard can be achieved if a variety of different service providers are involved. The authors conclude that “the proposed RDP appears to reduce consumer welfare” (Swire and Lagos, 2013, p.379), but do not offer or discuss economic incentives or outcomes, which additionally highlights the necessity of economic backing in this context.

Moreover, this study is related to at least two streams of economic literature, which will be highlighted in the following. First, as we assume users to be locked-in when using a data-intensive online service and having costs to port these data, we draw on the literature investigating the role of switching costs. The results derived from this strand of literature show that an incumbent firm has an incentive to lower its price anticipating that an entrant enters the market (Klemperer, 1989). In essence, firms thus fiercely compete in early periods to gain market shares which can then be harvested in later periods (Klemperer, 1987a; Klemperer, 1987b). Hence, switching costs induce softened competition in later periods which allows the remaining firms to set higher prices. Indeed, as Gehrig and Stenbacka (2004) show analytically, competing firms have an incentive to establish high switching costs. The authors show that these can be achieved by (maximum) vertical differentiation (see additionally Hotelling, 1929 and d’Aspermont et al., 1979). Within the taxonomy introduced by Ray et al. (2012), our study deals with “user-related” switching costs as they include the effort a user needs to invest to “ensure a satisfactory switch of service and to recreate or transfer features” (Ray et al., 2012, p.199). More precisely, one may argue that within the framework provided by Ray et al. (2012), transfer costs are of particular importance. To demarcate our approach from previous literature related to the existence of switching costs and lock-ins, the fact that we assume data as being the considered good, which additionally inherently determines the degree of switching costs (c.f., Section 3 for details) is crucial and should be highlighted. Hence, the strategy derived from

the traditional switching cost literature would induce to set lower prices in early periods (i.e., collect less data) to deter entry and gain market shares which can thereupon be harvested. This, in turn, is not necessarily the equilibrium strategy of an incumbent within a data-driven market environment, as (i) switching costs would be lower in successive periods and (ii) profits in later periods from data already gained in early periods would be reduced. These specific aspects of the competitive environment further delineates our approach from, e.g., Caminal and Matutes (1989) who consider endogenous switching costs.

Second, our study on data portability is related to the strand on the (economic) effects stemming from interoperability. Within this strand, the literature on compatibility and standardization between different services, especially the ensuing effects of the availability of converters as considered by Farrell and Saloner (1992), should be highlighted. In their theoretical model, Farrell and Saloner show that the availability of (imperfect) converters allows users to benefit from other users using a competing technology, i.e., a converter induces benefits through compatibility. Thus, direct network effects resulting from interoperability are a central aspect of the depicted model. Another important view on interoperability is highlighted within the study conducted by Pollock (2009). Pollock evaluates the effects of controlling the possibility to convert “‘software’ or ‘services’ associated with one platform to run on another” assuming a two-sided market (Pollock, 2009, p.155). Thus, Pollock considers interoperability being determined by indirect network effects. Additionally, the impact of the ability to control the mode of interoperability itself is investigated. Thus, the author allows the platform to directly control the costs of flow of information, i.e., the costs for interoperability. Nevertheless, although interoperability plays a pivotal role in online markets, the mentioned studies do not depict the concept of data portability for several reasons. In general, interoperability should not be confounded with the portability of data (c.f., Graef, 2015). Additionally, next to several technical dimensions, the central economic distinction can be seen in (i) the role of network externalities, which are not necessarily relevant in the context of data portability as a user’s lock-in in data-driven markets is crucially influenced by the (amount of) data revealed at a certain online service and not solely by network externalities (c.f., examples provided in Section 1), and (ii) the scope of influence in controlling the flow of data: since the regulation is binding for all services alike, online services are left with no possibility to strategically set the amount of data that can be ported, i.e., online services are unable to control the costs for portability.

Our proposed game-theoretic model, which will be outlined in the following Section, captures the trade-offs for the involved parties and considers the specific aspects of data-driven business models. We use this model to answer the following five research questions:

*RQ 1* How does a right to data portability affect the amount of data online services collect?

- RQ 2 What is the effect of a right to data portability on the incumbent's and the entrant's profit?
- RQ 3 How does a right to data portability affect service variety?
- RQ 4 How does a right to data portability affect the consumers?
- RQ 5 Which regime (data portability or no data portability) is more efficient with regard to total welfare?

### 3. Outline of the Economic Model

To highlight the competitive effects of data portability and to capture the effects on market entry and innovation, a market with two competing, differentiated CPs ( $i = A, B$ ) offering substitutable services, and heterogeneous users having a natural preference for one of the CPs is considered. Thus, a unit mass of users, uniformly distributed between zero and one (Hotelling, 1929) and CPs located at either end of the users' preference spectrum (d'Aspermont et al., 1979; Gehrig and Stenbacka, 2004) differentiated by some primary attribute (Irmén and Thisse, 1998; Iyer and Soberman, 2000) is assumed, i.e., CP  $A$  is located at  $x = 0$  and CP  $B$  is located at  $x = 1$ . Hence, both CPs are in competition for the same users. Moreover, to investigate the competitive effects on the amount of required data, market entry and innovation, we assume two time periods ( $t = 1, 2$ ) and assume that CP  $A$  is active in  $t = 1$  and  $t = 2$ , whereas CP  $B$  enters in  $t = 2$ . Thus, CP  $A$  might be classified as an incumbent content provider, whereas CP  $B$  is an entrant.

In period  $t = 1$ , a user located at  $x$  choosing to become active at CP  $A$  derives a utility of  $U_A^{t=1}(x) = v_A - \tau x - r_A$ , where  $v_A$  is the base-utility a user derives from being active at CP  $A$  (e.g., determined by the service's functionalities, the ease-of-use, etc.),  $\tau$  is a parameter specifying the competitive intensity in the market (i.e., if  $\tau$  is high, the user's inherent preference for one of the CP gets relatively more important), and  $r_A$  is the disutility a user derives from entering (personal) data at CP  $A$  (see introductory examples stated above; we refer to  $r_i$  as data consumption or amount of data required or collected at CP  $i$ ). As commonly specified, we assume users to be myopic, i.e., they only consider their utility in the first period, when they decide to be active at CP  $A$  in  $t = 1$  (i.e., if  $U_A^{t=1}(x) \geq 0$ ). Thus, the indifferent user between choosing CP  $A$  and not choosing any CP at all is located at  $x^{*,t=1} = \frac{v_A - r_A}{\tau}$  and users located at  $x > x^{*,t=1}$  are not active at CP  $A$  in  $t = 1$ .

In period  $t = 2$ , CP  $B$  enters the market. Users now have the choice between two competing CPs and choose the one from which they derive the higher utility. A user located at location  $x$  and active at CP  $A$  derives in the second period a utility of

$$U_A^{t=2}(x) = \begin{cases} v_A - \tau x - \theta \cdot r_A & , \text{ if } U_A^{t=1}(x) \geq 0 \\ v_A - \tau x - \theta \cdot r_A - r_A & , \text{ else} \end{cases}.$$

Note that  $\theta$  is the strategic variable of CP  $A$  in  $t = 2$  and  $\theta \cdot r_A$  equals the amount of *additional* data required by CP  $A$  in  $t = 2$  for users being active at CP  $A$ . This utility function captures the fact that, in the second period, users do not experience a further disutility from data already revealed in  $t = 1$  if the same CP is used again. For example, if a user entered (personal) data (e.g., its name, address, date of birth, interests, or uploaded photos and documents), it does not have to re-enter, re-validate or re-upload this information. Conversely, users who were *not* active in  $t = 1$  have to enter *all* required data if they decide to become active in the second period. Thus, these users need to enter data with an amount of  $\theta \cdot r_A + r_A$  (if they decide to become active at CP  $A$  in  $t = 2$ ) or  $r_B$  (if they decide to become active at CP  $B$ , respectively). For these users, the corresponding disutility is thus independent from the ability to port data, because they have not revealed any data previously. A user located at  $x$  who becomes active at CP  $B$  in  $t = 2$  derives a utility of

$$U_B^{d,t=2}(x) = \begin{cases} v_B - \tau(1-x) - r_B + r_A, & \text{if } U_A^{t=1}(x) \geq 0 \text{ with data portability } (d = P) \\ v_B - \tau(1-x) - r_B, & \text{else } (d = NP) \end{cases}.$$

The utility function  $U_B^{d,t=2}(x)$  captures the effect that users becoming active at CP  $B$  need to enter all required data (i.e.,  $r_B$ ) either if users have not been active in  $t = 1$ , or if there is no possibility to port already revealed data ( $d = NP$ ). Moreover, the function accounts for the effects of data portability if users switch from CP  $A$  to CP  $B$ : if users have been active at CP  $A$  in the first period, i.e.,  $U_A^{t=1}(x) \geq 0$ , and are able to port the already entered data to the new CP ( $d = P$ ), they only need to reveal the *net amount of required data* at CP  $B$  in  $t = 2$ , i.e.,  $r_B - r_A$ . Thus, if we assume the amount of information required at CP  $B$  to be at least as high as the amount of information required at CP  $A$ , i.e.,  $r_B \geq r_A$ , users only have to reveal the *additional information* required by CP  $B$ .<sup>2</sup>

By solving  $U_A^{t=2}(x) = U_B^{d,t=2}(x)$  and assuming the market to be fully covered, the indifferent user in  $t = 2$  is thus located at

$$x^{*,d,t=2} = \begin{cases} -\frac{\theta \cdot r_A + r_A - r_B - \tau - v_A + v_B}{2\tau}, & \text{with data portability } (d = P) \\ -\frac{\theta \cdot r_A - r_B - \tau - v_A + v_B}{2\tau}, & \text{without data portability } (d = NP) \end{cases}.$$

Based on the market shares given by the location of the indifferent user, the profits of the CPs can be specified. We assume that CPs with data-driven business models benefit from data entered in one period also in later periods as the obtained information is still valuable (e.g., in terms of the ability to target ads, or tailor offerings). As we neither consider fixed, nor variable costs, the total profit of CP  $A$  for both periods is thus given by

---

<sup>2</sup> If CP  $B$  provides a relatively higher base-utility than CP  $A$  ( $v_A < v_B$ , see Section 4 for further details), CP  $B$  always requires more data in equilibrium, i.e., switching users need to enter *additional* data. If we drop this assumption, the interpretation of the “net amount of data” gets less intuitive. Note that the assumption is unnecessary from an analytical perspective.



$$\pi_A^d = \underbrace{x^{*,t=1} \cdot r_A}_{\pi_A^{d,t=1}} + \underbrace{x^{*,d,t=2} \cdot (r_A + \theta \cdot r_A)}_{\pi_A^{d,t=2}}.$$

CP  $B$ , which is only active in  $t = 2$ , makes a total profit of

$$\begin{aligned}\pi_B^d &= (1 - x^{*,t=1}) \cdot r_B + (x^{*,t=1} - x^{*,d,t=2}) \cdot ((r_B - r_A) + r_A), \\ \pi_B^d &= (1 - x^{*,d,t=2}) \cdot r_B.\end{aligned}$$

Note that we implicitly assumed that CPs cannot discriminate between old, new or switching users, i.e., the amount of data a CP collects in  $t = 2$  is independent from a user's decision in  $t = 1$ . Thus, all users active at a CP need to reveal the same amount of data (we refer to the limitations in Section 6.2 for a discussion of occurring effects if this assumption is relaxed).

Figure 1 illustrates the assumed market setting. Herein, squares above the user depict the (net) amount of data (illustrated by symbols) different users ( $j = 1, 2$ ) would have to reveal in the considered period for becoming active at the respective CP. In contrast, circles underneath the CPs indicate the amount of data a CP requires. In this illustration, user 1 is active in period one, whereas user 2 becomes active only in period two. *Without* data portability, user 1 has to re-enter the data already revealed to CP  $A$  at CP  $B$ , if it wants to switch to CP  $B$  in the second period (thus, he needs to re-enter: star, moon and heart, and additionally needs to enter: thunderbolt). In contrast, *with* data portability, user 1 has the ability to port its already entered data and thus only has to enter the net amount of required data (here: thunderbolt), if it wants to switch to CP  $B$ . For user 2, who has not been active in the first period, both cases are identical, i.e., user 2 has to enter all of the CP's required data independent from the considered regime (i.e., star, moon, heart and sun to become active at CP  $A$  or star, moon, heart and thunderbolt to become active at CP  $B$ ). Note that Figure 1 only illustrates the (net) amount of data that is required by the CPs and needs to be entered by users in the respective period. The user's actual decision which CP to patronize is not illustrated since it depends (inter alia) on the base-utilities.

In summary, the considered two-stage game proceeds as follows:

- Stage 1** The incumbent CP  $A$  sets the amount of required data  $r_A$  for period  $t = 1$  anticipating CP  $B$ 's action in period  $t = 2$ . Then, users decide whether to become active at CP  $A$  (if  $U_A^{t=1}(x) \geq 0$ ).
- Stage 2** Both CPs simultaneously set the amount of required data for  $t = 2$ , i.e., CP  $A$  sets  $\theta$  and CP  $B$  sets  $r_B$ . Then, users again decide at which CP they choose to become active. Under the full market coverage assumption, users in  $t = 2$  are active at exactly one CP. If  $U_A^{t=2}(x) > U_B^{d,t=2}(x)$ , users are active at CP  $A$  and vice versa.

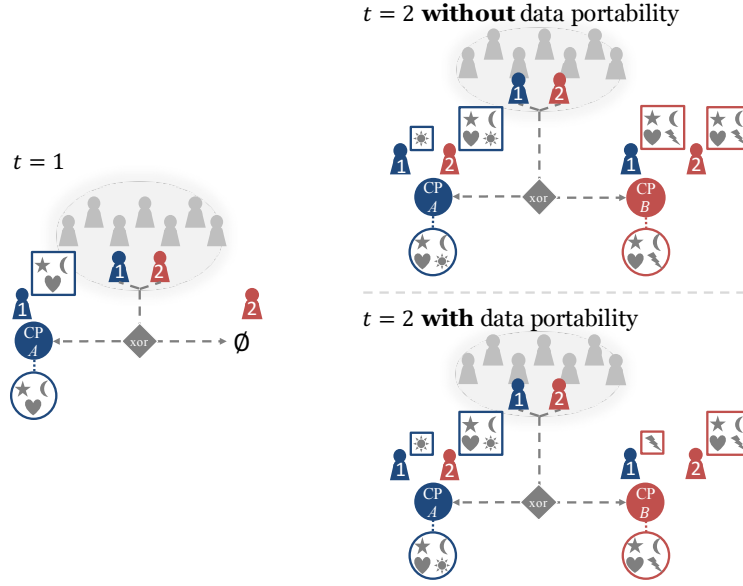


Figure 1. Illustration of the assumed market setting and effects of data portability.

#### 4. Results and Discussion

To identify the effects of a right to data portability, we compare the case *with data portability* ( $d = P$ ) to the case *without data portability* ( $d = NP$ ). In doing so, we solve for the subgame perfect equilibrium through backward induction beginning in *Stage 2* to deduce the equilibrium amounts of required data. The results are successively used to analyze the effects on CPs' profits, consumer's surplus and total surplus.

In *Stage 2* both CPs compete for users and revenues. Consequently, a CP's decision is affected by the decision of its competitor and the corresponding actions of users, i.e., the CPs take into account the amount of data required by the competing CP. Analytically, these effects are captured by simultaneously solving  $\frac{\partial \pi_A^{d,t=2}}{\partial \theta} = 0$  and  $\frac{\partial \pi_B^d}{\partial r_B} = 0$ , which yields the CP's equilibrium amount of required data for period  $t = 2$  (in the following, a  $*$  in the superscript refers to an equilibrium result).

In *Stage 1* CP  $A$  serves the market as monopolist. Nevertheless, it anticipates the effects on second-period profits in its decision how much data to collect. Analytically, we use the equilibrium results of *Stage 2* (i.e.,  $r_B^{*,d}$  and  $\theta^{*,d}$ ) to specify CP  $A$ 's profit over two periods ( $\pi_A^d$ ) and then solve  $\frac{\partial \pi_A^d}{\partial r_A} = 0$  to obtain the optimal amount of required data for CP  $A$  in period  $t = 1$  (i.e.,  $r_A^{*,d}$ ).

#### 4.1. Amount of Required Data by the CPs

A comparison of equilibrium amounts of required data shows that CP *A* requires a higher amount of data under the regime *without* data portability ( $d = NP$ ).<sup>3</sup> Interestingly, the period one amount of data consumption of CP *A* *without* data portability is even higher than the monopoly consumption  $r_{Monopoly}^*$ , i.e., the amount CP *A* would require without the entry of CP *B*:

$$r_A^{*,NP} = \frac{3\tau + 10v_A - v_B}{17} > \frac{v_A}{2} = r_A^{*,P} = r_{Monopoly}^*.$$

This highlights the effect of anticipated entry: Intuitively, CP *A* requires a high amount of data to generate (higher) switching costs to weaken competition in later periods. The effect of weakened competition even dominates the (negative effect of) reduced period one market shares and, compared to a regular one-period monopoly, reduced profits. The observation that CP *A* requires an even higher amount of data than in monopoly is, at first sight, in contrast to the traditional switching cost literature. Here, anticipated entry results in price wars lowering early-period prices to gain market shares, which can be harvested in later periods (c.f., Klemperer 1989; Klemperer, 1995). But, within our considered setting depicting a data-driven market environment, lock-ins are not generated by participation alone (e.g., network externalities or the functionalities of a service, c.f., Section 5 for a relaxation of this assumption), which could be stimulated by low prices, but by a user's invested effort to enter or reveal (personal) data. Thus, lock-in effects do also play a pivotal role for CPs in these market environments, although the underlying rationales differ compared to traditional market environments since (i) data required by a CP (i.e., “prices” set) in early periods are directly relevant to CPs' profits in later periods, and (ii) the incumbent's “price setting” is (additionally) constrained by entrants in later periods. *With* data portability ( $d = P$ ), the incumbent CP requires the monopoly amount of data. Because lock-in effects vanish through the users' ability to port data to the competing CP in following periods, the incumbent CP is constrained since it cannot benefit from establishing lock-ins anymore. Consequently, CP *A* maximizes its profit in the first period by requiring the same amount of data it would require in a one-period game, where it acts as monopolistic CP.

With respect to the amount of required data in the second period, this restricting effect of increased competition is also observable: the incumbent CP always requires less data if users are able to port their data, i.e.,

$$\theta^{*,NP} = \frac{15\tau - v_A - 5v_B}{3\tau + 10v_A - v_B} > \frac{6\tau - v_A - 5v_B}{3v_A} = \theta^{*,P}.$$

---

<sup>3</sup> The amount of required data *with* data portability can only be higher if  $\tau < -\frac{1}{2}v_A + \frac{1}{3}v_B$ . This induces that  $x^{*,d,t} < 0$ , which is not feasible. In successive analyses, we additionally require that  $x^{*,d,t} \leq 1$  (c.f., Appendix 7.1 for details).

Conversely, evaluating optimal data collection of CP  $B$  reveals that the required amount of data *with* data portability is always higher than in the case *without* data portability:

$$r_B^{*,NP} = \frac{16\tau - 9v_A + 6v_B}{17} < \tau - \frac{v_A - v_B}{3} = r_B^{*,P}.$$

Intuitively, CP  $B$  requires more data *with* data portability because users that switch from CP  $A$  only have to reveal the *net amount of required data* due to the possibility to port the already entered data. Consequently, these users experience less disutility in the case of switching, because the net amount of required data is lower ( $r_B - r_A \leq r_B$ ), all else being equal, leading to higher market shares and profits for the entrant under this regime.

Next, we analyze the factors influencing the equilibrium amount of required data by conducting comparative statics. First, we find that CP  $A$ 's first-period data consumption increases in its base-utility  $v_A$ , whereas its second-period data consumption decreases in  $v_A$ , i.e.,  $\frac{\partial r_A^{*,d}}{\partial v_A} > 0$  and  $\frac{\partial \theta^{*,d}}{\partial v_A} < 0$ . The negative effect on the second-period amount of required data by CP  $A$  can be explained by the incumbent's rationale to protect its market share in a competitive environment: Through an increased base-utility, CP  $A$  is able to require a large(r) amount of data in period one. Protecting this market share in period two (through a comparably low amount of required data in this period) dominates the potential positive effects arising from requiring more data in the second period. On the contrary, if its base-utility is decreasing, protecting market shares does not dominate the positive effects of requiring additional data in period two. Second, an increase in CP  $B$ 's base-utility  $v_B$  lowers CP  $A$ 's data collection: in period one to increase the share of users that are locked-in, in period two due to stronger competitive forces. Since lock-ins disappear *with* data portability, the first-period amount of required data is unaffected by  $v_B$  (in  $d = P$ ). In conclusion,  $\frac{\partial r_A^{*,NP}}{\partial v_B} < 0$ ,  $\frac{\partial \theta^{*,d}}{\partial v_B} < 0$  and  $\frac{\partial r_A^{*,P}}{\partial v_B} = 0$ . Third, the competitiveness of the market ( $\tau$ ) has an unambiguous effect on CP  $A$ 's data consumption: the lower the competitive intensity of the market (increase in  $\tau$ ), the higher the amount of required data, i.e.,  $\frac{\partial r_A^{*,NP}}{\partial \tau} > 0$  and  $\frac{\partial \theta^{*,d}}{\partial \tau} > 0$ . Finally, for CP  $B$ , comparative statics (also) show that an increase in the competitor's base-utility ( $v_A$ ) reduces the amount of required data. In contrast to the incumbent, an increase in the own base-utility ( $v_B$ ) unambiguously increases the amount of required data. The effect of the competitiveness of the market on CP  $B$ 's data consumption is similar to the effect on CP  $A$ 's data consumption, i.e., the lower the competitive intensity of the market, the higher the amount of required data.

**Proposition 1 (amount of required data)**

(a) Under a data portability regime, incumbents require less user data, whereas entrants unambiguously increase their amount of required data for users becoming active at their service.

(b) A (in terms of service quality) strong competitor or a higher competitive intensity of the market reduces the CP's amount of required data. If a CP increases its own quality, it requires more data in the first period being active.

#### 4.2. CPs' Profits

To analyze CPs' profits ( $\pi_i^d$ ), we evaluate optimal profits given the just derived equilibrium amounts of required data. In the feasible parameter range (c.f., Appendix 7.1), the incumbent always suffers from data portability ( $\pi_A^P \leq \pi_A^{NP}$ ), whereas the entrant always benefits from data portability ( $\pi_B^{NP} \leq \pi_B^P$ , see Appendix 7.2 for details). Thus, data portability leads to a higher service variety (or innovation) as entrants are more likely to enter the market due to higher profits. Thus, if the market is dominated by a single firm, data portability may be a suitable device to foster competition.

Comparative statics show that an increase in the CP's own (competitor's) base utility has always a positive (negative) effect on its profits (i.e.,  $\frac{\partial \pi_i^d}{\partial v_i} > 0$  and  $\frac{\partial \pi_i^d}{\partial v_{-i}} < 0$  for  $i = \{A, B\}$  and  $-i$  representing the competing CP  $i$ ). Interestingly, the effect of a less competitive market (i.e., an increase in  $\tau$ ) is ambiguous: with respect to  $\pi_A^P, \pi_A^{NP}$  and  $\pi_B^{NP}$ , a less competitive market is always beneficial only if the competing CP  $-i$  is strong in terms of its base-utility, i.e.,  $v_{-i} \gg v_i$ . Otherwise, the effect of  $\tau$  depends on the parameter constellation.<sup>4</sup> With regard to  $\pi_B^P$ , the effect is unambiguous: the less competitive the market, the higher the profit.

#### 4.3. Consumer's Surplus

To examine the effects on consumer's surplus ( $CS_i^d$ ), we compare the users' utility accounting for the different regimes. With respect to users active at CP A, consumer's surplus for both periods is:

$$CS_A^d = \underbrace{\int_0^{x^{*,t=1}} U_A^{t=1}(x) dx}_{\text{period } t=1} + \underbrace{\int_0^{x^{*,d,t=2}} U_A^{t=2}(x) dx}_{\text{period } t=2}.$$

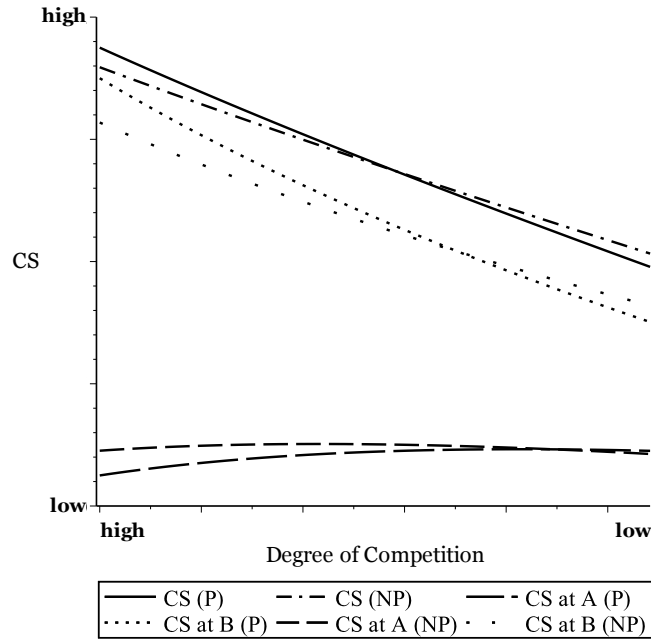
Note that users active at CP B differ with regard to their utility depending on the availability of a right to data portability. *With* data portability, users that have not been active in the first period derive a utility of  $U_B^{NP,t=2}(x)$ . If they have been active at CP A in the first period, switch to CP B and port their data, they derive a lower (net) disutility for a given amount of required data by CP B leading to an utility of  $U_B^{P,t=2}(x)$ . If data portability is *not* enforced, all users becoming active at CP B derive a utility of  $U_B^{NP,t=2}(x)$ . In total, consumer's surplus can be calculated by:

---

<sup>4</sup> Formally, the derivative changes its sign within the feasible parameter range. The effect of an increase in  $\tau$  on  $\pi_A^P$  is positive if  $\tau > \frac{1}{6}(\sqrt{22v_A^2 - 12v_Av_B + 4v_B^2})$  and an increase in  $\tau$  on  $\pi_A^{NP}$  is positive if  $\tau > \frac{1}{6}(\sqrt{26v_A^2 - 12v_Av_B + 4v_B^2})$ . The effect on  $\pi_B^{NP}$  is positive if  $\tau > \frac{6v_B - 9v_A}{16}$ .

$$CS_B^d = \begin{cases} \int_{x^{*,P,t=2}}^{x^{*,t=1}} U_B^{P,t=2}(x)dx + \int_{x^{*,t=1}}^1 U_B^{NP,t=2}(x)dx & , \text{ with data portability } (d = P) \\ \int_{x^{*,NP,t=2}}^1 U_B^{NP,t=2}(x)dx & , \text{ without data portability } (d = NP) \end{cases}.$$

Comparing consumers' surpluses in equilibrium, it can be seen that no data portability may leave users actually better off. Thus, the sum of consumers' surpluses at both CPs may even decrease under a data portability regime, i.e.,  $CS_A^P + CS_B^P < CS_A^{NP} + CS_B^{NP}$  (see Appendix 7.3 for details). Consequently, although data portability is most commonly justified by the potential benefits for end customers (c.f., Macgillivray & Shambaugh, 2016; European Commission, 2016b), this goal is not necessarily warranted. Figure 2 illustrates the possible negative effect on total consumers' surplus for a specific parameter constellation ( $v_A = 1$ ,  $v_B = 4$ ) by showing total consumers' surplus, as well as the consumer's surplus at each CP for different degrees of competition.



**Figure 2. Illustration of total consumer's surplus for  $v_A = 1$  and  $v_B = 4$  with data portability ( $d = P$ , solid line) and without data portability ( $d = NP$ , dash-dotted line).**

It can be shown that a (relatively) low degree of competition (high  $\tau$ ) may lead to users being worse off *with* data portability.<sup>5</sup> More precisely, as we have shown above, CPs require higher amounts of data if the degree of competition is low (because an *increase* in  $\tau$  increases the amount of required data:  $\frac{\partial r_i^{*,d}}{\partial \tau} > 0$  and  $\frac{\partial \theta^{*,d}}{\partial \tau} > 0$ , see above). This, in turn, increases the

<sup>5</sup> More precisely, if  $\tau > \frac{174v_B - 822v_A + 17\sqrt{6658v_A^2 - 752v_Av_B + 16v_B^2}}{726}$  users are better off without data portability (see Appendix 7.3).

disutility a user derives from being active at the considered service, which, consequently, reduces consumer's surplus (i.e.,  $\frac{\partial (CS_A^d + CS_B^d)}{\partial \tau} < 0$ ). However, if the CPs' base-utilities are relatively equal ( $v_B < \frac{447}{160} \cdot v_A$ ), consumers unambiguously benefit under a data portability regime. Additionally, higher base-utilities always affect the consumer's surplus positively ( $\frac{\partial (CS_A^d + CS_B^d)}{\partial v_i} > 0$ ).

#### 4.4. Total Surplus

Finally, the total surplus ( $TS^d$ ) being the sum of consumer's surplus and CPs' profits, i.e.,

$$TS^d = \sum_{i=A,B} (\pi_i^d + CS_i^d)$$

is examined (see Appendix 7.4 for details). Within the feasible parameter range, it can be concluded that the total surplus is unambiguously increasing with data portability, i.e.,  $TS^P > TS^{NP}$ . Thus, although consumers might be worse off in some cases and CP *A* always experiences lower profits under a regime with a right to data portability, the increased profits of CP *B* always outweigh these effects.

#### **Proposition 2 (welfare analysis)**

- (a) Data portability unambiguously increases an entrant's profit arguably increasing service variety. In contrast, an incumbent always suffers under a data portability regime.
- (b) The possibility to port data from one service to another has ambiguous effects on consumer's surplus. If both services offer a comparable quality for users, consumer's surplus always increases. Conversely, if the entrant offers a qualitatively better service, users may suffer under a data portability regime.
- (c) Total surplus always increases under a data portability regime.

### 5. The Role of Network Effects

As highlighted in the previous sections, network effects are not a precondition for online CPs to become successful and are not necessarily the (main) source of users to become locked-in. Nevertheless, the utility a user derives may additionally be affected by the number of other users active at that platform, i.e., positive network effects may exist influencing the user's decision, but also the CPs' strategies in setting the amount of required data. Intuitively, the presence of network effects may reduce the incentive to switch to an entrant CP because the utility a user derives from the already installed base may outweigh the potentially higher base utility from the joining CP – even though data already entered can be ported to that joining CP with a right to data portability. To investigate the role of network effects formally, we modify the

user's utility function and incorporate positive network effects, i.e., a positive effect of the number of active users at a certain CP. More precisely, we assume that the total number of active users at the considered CP has a positive effect on a user's utility, i.e.,  $U_A^{d,t,NWE}(x) = U_A^{d,t}(x) + \gamma \cdot x^{*,d,t}$  for CP A and  $U_B^{d,t,NWE}(x) = U_B^{d,t}(x) + \gamma \cdot (1 - x^{*,d,t})$  for CP B, respectively with  $\gamma > 0$  (please note the superscript *NWE* from now on referring to the scenario accounting for network effects). By changing the utility function, the location of the indifferent user also changes. The indifferent user in period  $t = 1$  is now located at  $x^{*,t=1,NWE} = \frac{v_A - r_A}{\tau - \gamma}$ , the indifferent user in period  $t = 2$  is now located at  $x^{*,P,t=2,NWE} = \frac{r_A \cdot \theta + \gamma + r_A - r_B - \tau - v_A + v_B}{2(\gamma - \tau)}$  (with data portability) and  $x^{*,NP,t=2,NWE} = \frac{r_A \cdot \theta + \gamma - r_B - \tau - v_A + v_B}{2(\gamma - \tau)}$  (without data portability).<sup>6</sup> The profit functions as well as our proposed two stage game remain qualitatively unchanged (we refer to Appendix 7.5 for details).

We solve for the subgame perfect equilibrium using backward induction and derive the period one and period two amount of required data as shown in Appendix 7.5. Compared to the case without network effects (c.f., Section 4), one can easily show that CPs never require more data. Nevertheless, it is worth noting that the incumbent CP does not change the amount of required data with data portability in  $t = 1$ , i.e.,  $r_A^{*,P} = r_A^{*,P,NWE}$ . All other results highlighted with respect to prices deduced in Section 4.1 still hold, i.e.,  $r_A^{*,NP,NWE} > r_A^{*,P,NWE}$ ,  $\theta^{*,NP,NWE} > \theta^{*,P,NWE}$  and  $r_B^{*,NP,NWE} < r_B^{*,P,NWE}$ . Consequently, the existence of positive direct network effects has a negative impact on the amount of data a CP requires, i.e.,  $\frac{\partial r_i}{\partial \gamma} \leq 0$  and  $\frac{\partial \theta}{\partial \gamma} < 0$ . Nevertheless, the incumbent can<sup>7</sup> benefit in terms of profits from the existence of network effects due to higher market shares in the first period. Conversely, the entrant always realizes lower profits. Unsurprisingly, consumers are unambiguously better off if positive direct network effects are considered.

Although the consideration of network effects does have ramifications on all actors within our considered market setting, our deduced effects of introducing a right to data portability qualitatively remain unchanged. More precisely, the incumbent always suffers from the existence of a right to data portability, the entrant is always better off, and total surplus always increases. The effect of data portability on users remains ambiguous: although consumer's surplus with a right to data portability is higher in more cases, i.e., the intersection of both

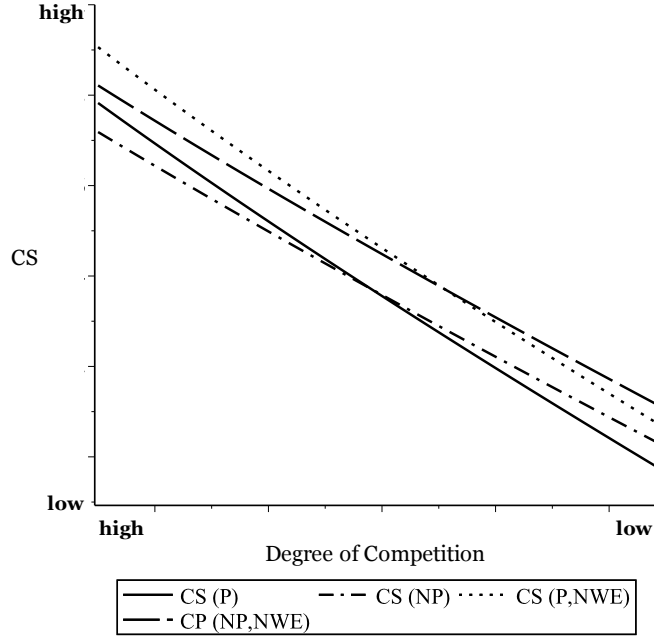
---

<sup>6</sup> We thereby rely on the concept of fulfilled-expectations, i.e., in equilibrium, the network size determined by the location of the indifferent user equals the expected one.

<sup>7</sup> The incumbent's profit functions with and without the existence of network effects intersect within the feasible parameter range. If the competitive intensity of the market is low, the incumbent realizes higher profits with network effects. Formally, if  $\tau < \frac{\gamma}{2} + \frac{1}{6}\sqrt{9\gamma^2 + 22v_A^2 - 8v_Av_B + 4v_B^2}$  (with data portability) and if  $\tau < \frac{\gamma}{2} + \frac{1}{6}\sqrt{9\gamma^2 + 26v_A^2 - 12v_Av_B + 4v_B^2}$  (without data portability), the incumbent realizes higher profits if network effects are considered.



functions is shifted to the edge of the feasible parameter range (c.f., Figure 3), users nevertheless may experience reduced surplus. Figure 3 illustrates total consumer's surplus with and without a right to data portability for a scenario accounting for network effects (dotted and dashed curves), and for a scenario without considering network effects (solid and dash-dotted curves).



**Figure 3. Total consumer's surplus with and without a right to data portability for scenarios with and without network effects for  $v_A = 1$ ,  $v_B = 4$  and  $\gamma = 0.05$ .**

## 6. Conclusion

Data portability allows users to transfer their data entered at a certain (online) service to another (online) service. Although some services have implemented such features voluntarily, a standardized and mandatory ability for users to port (personal) data is pursued by the European Commission for all online services available in the EU's member states through the General Data Protection Regulation (European Commission, 2016b). Additionally, this topic also gains momentum for non-European policy makers, as the U.S.-based request for information illustrates (c.f., Macgillivray and Shambaugh, 2016).

Despite the importance of this issue resulting from the far-reaching implications on business strategies of online services and so on the total economy, we are –to the best of our knowledge– the first to analyze the resulting competitive effects theoretically. In doing so, we not only shed light on current policy issues, but also highlight relevant implications on the interface of the IS, the technical and the economic realm to better understand and develop systems' value propositions and to provide insights for policy makers and managers, alike. For this purpose, we propose a game-theoretic model that captures firms' strategic incentives and

identify the feasible market outcomes together with the implications for relevant stakeholders within data-driven markets.

In conclusion, we find that on the one hand, the existence of a right to data portability fosters market entry, which arguable enhances service variety and innovation, but on the other hand, active services unambiguously suffer. Although the existence of network effects dampens these effects, they do not vanish. Whereas such an outcome might be desired by policy makers to alleviate concerns about dominant online services, we highlight that also end users may actually suffer from a right to data portability, because the joining service has an incentive to increase the amount of collected data compared to a regime without data portability. Nevertheless, as the total surplus unambiguously increases due to higher overall profits, a decision about a mandatory data portability rule invokes a complex assessment. In the following, we outline policy as well as managerial implications based on the obtained results and discuss avenues for further research.

### **6.1. Policy and Managerial Implications**

*From a policy perspective*, the rationales to introduce a (general) right to data portability are clearly focused on the protection of end users (see, e.g., European Commission, 2016b, Article 1). Consequently, our results stated above induce that data portability should *not* be applied to all online services as consumer's might actually be worse off. On the other hand, considering the total economy, overreaching goals such as the Digital Single Market Strategy (DSM strategy) within the European Union (c.f., European Commission, 2016a) or former-president Obama's executive order on competition from April, 2016 (c.f., Obama, 2016) highlight the importance of open, fair and non-discriminatory (data-driven) markets. As we show that the entrant's profit always increases under data portability, a right to data portability may attribute to these goals. As a result, our findings evoke the necessity for policy makers to carefully weigh whether they want to promote market entry to stimulate innovation and successively service variety, or purely focus on consumer's surplus.

If services should be incentivized to enter the market, data portability should be strictly enforced and exceptions should be scarce. To date, the concept of data portability proposed by the European Commission solely focuses on personal data revealed by the users themselves. Hence, data revealed by third persons (say, reviews for a private lift, or endorsements on professional networking sites) are excluded in the current version of the proposed regulation. Therefore, policy makers might think of extending the scope of data that can be ported.<sup>8</sup> In most cases, this strategy would be in line with the goal of enhancing consumer's surplus. Nevertheless, it has to be taken into consideration that (i) porting sensitive data (e.g., credit card

---

<sup>8</sup> As highlighted in the mid-term review on the implementation of the Digital Single Market strategy, the European Commission already "subject to impact assessment, prepare[s] a legislative proposal [...] which takes into account [...] the principle of porting non-personal data" (European Commission, 2017, p.11).

numbers, tax IDs, social security numbers, etc.) bears important privacy and security risks, although users entered these data voluntarily, and (ii) there are instances where users are actually worse off with a right to data portability, as we have shown above. Our results suggest that users might be worse off if the qualities of the considered services are asymmetric, e.g., the entrant has a superior value proposition providing the user a higher base-utility. Arguably, entry is then beneficial for the entrant even without a right to data portability. Consequently, one may hypothetically think of a concept where data portability is only granted to some services. Whereas this seems possible in theory, the likeliness of success of such an approach is questionable as (i) this concept would contradict popular “neutrality regimes”, which might get increasingly important on a service level (c.f., Easley et al., forthcoming), (ii) the current political view aims at giving end users back the control of their (personal) data – independent from the considered service (c.f., European Commission, 2016b), and (iii) the nature of the Internet with independent parties and hard-to-control dataflow make supervision costly.

*From a managerial perspective*, it has to be emphasized that incumbent services have an unambiguous incentive to inhibit the concept of data portability since their opportunity to soften competition vanishes, leading to reduced profits. In contrast, entrant services or start-ups should promote the concept of data portability since their flexibility in setting the amount of data that is required rises, leading to higher profits and thus, earlier profitability. If services have no possibility to influence the scope of data that can be ported (see above), incumbents should pursue a differentiation strategy if the entrant is superior in terms of base-utility. This reduces the competitiveness of the market and so benefits the incumbent. For this purpose, incumbents may try to change (aspects of) their service offering to escape the fierce competition with the joining service. In contrast, a strategy designed to imitate the competitor can be seen as an incumbent’s opportunity if the entrant is relatively equal in terms of base-utility. This might be achieved by matching all of the entrant’s value propositions to increase the competitiveness of the market and thus, profits (see effects of the market’s competitiveness on profits outlined in Section 4.1 and 4.2). The entrant always benefits from a less competitive market. Thus, it always needs to shape the market accordingly, i.e., differentiate as much as possible from the incumbent, e.g., by acting as the industry’s innovation leader.

## **6.2. Limitations and Avenues for Future Research**

Finally, we wish to conclude by highlighting possible model extensions and limitations. There are several interesting aspects of data-driven business models that should be taken into consideration and analyzed in future studies. First, the market environment could be changed to capture the effects of data portability on two existing services. In our terminology, CP *B* would then already be active in period one and data can be ported from CP *A* to CP *B* and vice versa. Arguably, as the CP’s flexibility in setting the amount of required data is reduced, CP’s unambiguously suffer under a regime that enforces data portability. Conversely, such a scenario

would be beneficial for end users. Secondly, the possibility to discriminate between new users and existing users might be seen as a possible model extension. However, it has to be noted that this extension is rather of theoretical nature since services arguably need to require the same amount of data from all users to reduce programming efforts and to provide a streamlined and consistent (service) portfolio. With data portability, the entrant would then collect a relatively high amount of data from new (i.e., not switching) users and additionally maintain flexibility for the share of users that may switch services, leading to reduced consumer's surplus. Thirdly, we assumed that data entered once has no effect on a user's utility in succeeding periods. Whereas we believe that this is a suitable benchmark, one may argue that the disutility of already entered data only diminishes over time, i.e., the effects of trust for a certain service might be included into the analysis. This could be achieved by assuming that there is a lower (or no) disutility if the same service is used again, whereas there is some disutility if the same data is ported to another service. Finally, a right to data portability arguably induces positive effects on other, unrelated CPs which are not modeled within this study focusing on the effects on competing CPs. Thus, the positive effect of data portability on service variety and innovation may be even stronger. Similar, some users might have an increase in utility due to the mere possibility to port (personal) data, which would increase consumer's as well as total surplus.

## 7. Appendices

### 7.1. Thresholds for the Feasible Parameter Range

In this paper, we build on Hotelling's model of horizontal differentiation (c.f., Hotelling, 1929) in order to identify the competitive effects of data portability. In doing so, we assume that a unit mass of users is uniformly distributed in  $[0,1]$ . By calculating market shares, which can directly be deduced from the location of the indifferent user (c.f., Section 3), we formally need to ensure that the indifferent user is in all cases located within the interval of  $[0,1]$ . Consequently, for the regime with data portability and for the regime without data portability, we require  $x^{*,t=1} \leq 1$ ,  $x^{*,d,t=2} \geq 0$  and  $x^{*,NP,t=2} \leq 1$ , i.e., the CPs' market shares are always positive and do not exceed 100%. As highlighted within our analysis, we assume full market coverage in  $t = 2$  for analytical tractability. Consequently, the entrant's base-utility needs to be large enough so that (at least) one user can potentially port its user data from CP  $A$  to CP  $B$ . Formally, we capture this requirement by ensuring that the indifferent user of CP  $B$  is located left from the period-one indifferent user of CP  $A$ , i.e.,  $x^{*,NP,t=2} < x^{*,t=1}$ . Please note that we only specified binding restrictions, i.e., all other restrictions are less restrictive (e.g.,  $x^{*,t=1} \geq 0$ ).

### 7.2. CPs' Profits ( $\pi_i^d$ )

With data portability ( $d = P$ ), the CP's profits are:

$$\pi_A^P = \frac{18\tau^2 + 12\tau(v_A - v_B) + 11v_A^2 - 4v_Av_B + 2v_B^2}{36\tau},$$

$$\pi_B^P = \frac{(3\tau - v_A + v_B)^2}{18\tau}.$$

Without data portability ( $d = NP$ ), the CP's profits are:

$$\pi_A^{NP} = \frac{18\tau^2 + \tau(18v_A - 12v_B) + 13v_A^2 - 6v_Av_B + 2v_B^2}{34\tau},$$

$$\pi_B^{NP} = \frac{(16\tau - 9v_A + 6v_B)^2}{578\tau}.$$

To determine whether CP's are better off under data portability, we calculate the intersection of the CP's profit functions under the different regimes (i.e.,  $\pi_i^P$  and  $\pi_i^{NP}$ ). Although the profit functions intersect two times, both intersections are outside the feasible parameter range given by the restrictions specified in Appendix 7.1. Consequently, the effect of data portability on the incumbent and entrant is unambiguous. It can easily be shown that the incumbent (entrant) always suffers (benefits) from data portability, i.e.,  $\pi_A^P \leq \pi_A^{NP}$  and  $\pi_B^{NP} \leq \pi_B^P$ .

### 7.3. Consumer's Surplus ( $CS_i^d$ )

With data portability ( $d = P$ ), the consumer's surplus equals:

$$CS_A^P = \frac{-45\tau^2 + \tau(24v_A + 30v_B) + 22v_A^2 - 8v_Av_B + 5v_B^2}{72\tau},$$

$$CS_B^P = \frac{-45\tau^2 + \tau(12v_A + 6v_B) + 7v_A^2 - 4v_Av_B + 7v_B^2}{72\tau}.$$

Without data portability ( $d = NP$ ), the consumer's surplus equals:

$$CS_A^{NP} = \frac{-1368\tau^2 + \tau(264v_A + 912v_B) + 763v_A^2 - 88v_Av_B + 152v_B^2}{2312\tau},$$

$$CS_B^{NP} = -\frac{(16\tau - 9v_A + 6v_B)(80\tau - 45v_A + 38v_B)}{2312\tau}.$$

To determine whether users are better off under data portability, we calculate  $CS_A^P + CS_B^P = CS_A^{NP} + CS_B^{NP}$  and reorder the result with respect to  $\tau$ . This leads to two solutions labeled by  $\tau_1$

and  $\tau_2$ . It can be shown that  $\tau_1 = \frac{174v_B - 822v_A + 17\sqrt{6658v_A^2 - 752v_Av_B + 16v_B^2}}{726}$  can be within the feasible

parameter range specified in Appendix 7.1, whereas  $\tau_2 = \frac{174v_B - 822v_A - 17\sqrt{6658v_A^2 - 752v_Av_B + 16v_B^2}}{726}$  is

always outside of that feasible parameter range. Consequently, the effect of data portability on consumer's surplus is ambiguous and users may suffer from a right to data portability. Whereas the effect is positive if  $\tau < \tau_1$ , the effect is negative if  $\tau > \tau_1$ . Please note that  $\tau_1$  is not always

within the feasible parameter range: if  $v_B < \frac{447}{160}v_A$ , the intersection is always outside the feasible parameter range.

#### 7.4. Total Surplus ( $TS^d$ )

With data portability ( $d = P$ ), the total surplus is:

$$TS^P = \frac{-18\tau^2 + 36\tau(v_A + v_B) + 55v_A^2 - 20v_Av_B + 10v_B^2}{72\tau}.$$

Without data portability ( $d = NP$ ), the total surplus is:

$$TS^{NP} = \frac{-200\tau^2 + \tau(888v_A + 496v_B) + 783v_A^2 - 500v_Av_B + 178v_B^2}{1156\tau}.$$

All intersections of the functions are outside the feasible parameter range specified by the restrictions given in Appendix 7.1. Consequently, the effect on total surplus is unambiguous. It can easily be shown that total surplus always increases with data portability, i.e.,  $TS^P > TS^{NP}$ . Please note that this result assumes that total surplus is the unweighted sum of producer's and consumer's surplus.

#### 7.5. Network Effects – The Amount of Required Data ( $r_i^{*,d,NWE}$ and $\theta^{*,d,NWE}$ ) and CPs' Profits ( $\pi_i^{*,d,NWE}$ )

##### The Amount of Required Data ( $r_i^{*,d,NWE}$ and $\theta^{*,d,NWE}$ )

As highlighted in Section 5.1, with network effects, a user's utility function changes. Since the location of the indifferent user changes, the corresponding profits change yielding different equilibrium amounts of required data. For CP A, the equilibrium amount of required data in  $t = 1$  equals:

$$r_A^{*,NP,NWE} = r_A^{*,NP} - \frac{3}{17}\gamma = \frac{3\tau + 10v_A - v_B - 3\gamma}{17},$$

$$r_A^{*,P,NWE} = r_A^{*,P} = \frac{1}{2} \cdot v_A,$$

and in  $t = 2$

$$\theta^{*,NP,NWE} = \frac{15\gamma - 15\tau + v_A + 5v_B}{3\gamma - 3\tau - 10v_A + v_B},$$

$$\theta^{*,P,NWE} = \frac{6\tau - v_A - 2v_B - 6\gamma}{3v_A}.$$

For CP B, the equilibrium amount of required data (in  $t = 2$ ) equals:

$$r_B^{*,NP,NWE} = r_B^{*,NP} - \frac{16}{17}\gamma = \frac{16\tau - 16\gamma - 9v_A + 6v_B}{17},$$

$$r_B^{*,P,NWE} = r_B^{*,P} - \gamma = \tau - \gamma - \frac{v_A - v_B}{3}.$$

### **CPS' Profits ( $\pi_i^{*,d,NWE}$ )**

The calculation of the CP's profits incorporating network effects qualitatively remains unchanged compared to the base model (c.f., Section 3 for details). Using the location of the indifferent users (c.f., Section 5) and the equilibrium amount of required data (c.f., Appendix 7.5), the CP's profits *with* data portability ( $d = P$ ) and *with* network effects yields:

$$\pi_A^{P,NWE} = \frac{18(\gamma^2 + \tau^2) + (-36\tau - 12v_A + 12v_B)\gamma + 12\tau(v_A - v_B) + 11v_A^2 - 4v_Av_B + 2v_B^2}{36\tau - 36\gamma},$$

$$\pi_B^{P,NWE} = \frac{(3(\tau - \gamma) - v_A + v_B)^2}{18\tau - 18\gamma}.$$

*Without* data portability ( $d = NP$ ) and *with* network effects:

$$\pi_A^{NP} = \frac{18(\gamma^2 + \tau^2) + (-36\tau - 18v_A + 12v_B)\gamma + \tau(18v_A - 12v_B) + 13v_A^2 - 6v_Av_B + 2v_B^2}{34\tau - 34\gamma},$$

$$\pi_B^{NP} = \frac{(16(\tau - \gamma) - 9v_A + 6v_B)^2}{578\tau - 578\gamma}.$$

## References

- Caminal, R., and Matutes, C. 1990. "Endogenous switching costs in a duopoly model," *International Journal of Industrial Organization* (8:3), pp. 353-373.
- Chen, P. Y., and Hitt, L. M. 2002. "Measuring switching costs and the determinants of customer retention in Internet-enabled businesses: A study of the online brokerage industry," *Information Systems Research* (13:3), pp. 255-274.
- d'Aspremont, C., Gabszewicz, J. J., and Thisse, J. F. 1979. "On Hotelling's "Stability in competition", " *Econometrica* (47:5), pp. 1145-1150.
- European Commission 2016a. "Online Platforms," accessed on 07.06.2017, available at <https://ec.europa.eu/digital-single-market/online-platforms-digital-single-market>.
- European Commission 2016b. "Regulation (EU) 2016/679 (...) of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)," accessed on 07.06.2017, available at [http://ec.europa.eu/justice/data-protection/reform/files/regulation\\_oj\\_en.pdf](http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf).
- European Commission 2017. "Communication from the commission (...) on the Mid-Term Review on the implementation of the Digital Single Market Strategy (SWD(2017) 155 final)," accessed on 07.06.2017, available at [http://ec.europa.eu/newsroom/document.cfm?doc\\_id=44527](http://ec.europa.eu/newsroom/document.cfm?doc_id=44527).
- Drozdiak, N., and Schechner, S. 2016. "EU Files Additional Formal Charges Against Google," accessed on 07.06.2017, available at <http://www.wsj.com/articles/google-set-to-face-more-eu-antitrust-charges-1468479516>.
- Easley, R., Guo, H., and Kraemer, J. forthcoming. "From Network Neutrality to Data Neutrality: A Techno-Economic Framework and Research Agenda," *Information Systems Research*.
- Facebook n.d. "Downloading Your Info," accessed on 07.06.2017, available at <https://www.facebook.com/help/131112897028467/>.
- Farrell, J., and Klemperer, P. 2007. "Coordination and lock-in: Competition with switching costs and network effects," in *Handbook of Industrial Organization*, M. Armstrong and R. H. Porter (eds.), Volume 3, Amsterdam: Elsevier, pp. 1967–2072.
- Farrell, J., and Saloner, G. 1992. "Converters, compatibility, and the control of interfaces," *The Journal of Industrial Economics* (40:1), pp. 9-35.
- Gehrig, T., and Stenbacka, R. 2004. "Differentiation-Induced Switching Costs and Poaching," *Journal of Economics & Management Strategy*, (13:4), pp. 635-655.
- Google n.d. "Download your data," accessed on 07.06.2017, available at <https://support.google.com/accounts/answer/3024190?hl=en>.
- Graef, we. 2015. "Mandating portability and interoperability in online social networks: Regulatory and competition law issues in the European Union," *Telecommunications Policy* (39:6), pp. 502-514.



- Hotelling, H. 1929. "Stability in competition," *The Economic Journal* (39:153), pp. 41–57.
- Irmen, A., and Thisse, J. F. 1998. "Competition in multi-characteristics spaces: Hotelling was almost right," *Journal of Economic Theory* (78:1), pp. 76-102.
- Iyer, G., and Soberman, D. 2000. "Markets for product modification information," *Marketing Science* (19:3), pp. 203-225.
- Katz, M. L. and Shapiro, C. 1994. "Systems competition and network effects," *The Journal of Economic Perspectives* (8:2), pp. 93-115.
- Klemperer, P. 1987a. "Markets with consumer switching costs," *The Quarterly Journal of Economics* (102:2), pp. 375-394.
- Klemperer, P. 1987b. "The competitiveness of markets with switching costs," *The RAND Journal of Economics* (18:1), pp. 138-150.
- Klemperer, P. 1995. "Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade," *The Review of Economic Studies* (62:4), pp. 515-539.
- Klemperer, P. 1989. "Price wars caused by switching costs," *The Review of Economic Studies* (53:3), pp. 405-420.
- Macgillivray, A., and Shambaugh, J. 2016. "Exploring Data Portability," accessed on 07.06.2017, available at <https://obamawhitehouse.archives.gov/blog/2016/09/30/exploring-data-portability>.
- Obama, B. 2016. "Executive Order -- Steps to Increase Competition and Better Inform Consumers and Workers to Support Continued Growth of the American Economy," accessed on 07.06.2017, available at <https://obamawhitehouse.archives.gov/the-press-office/2016/04/15/executive-order-steps-increase-competition-and-better-inform-consumers>.
- Petcu, D., and Vasilakos, A. V. 2014. "Portability in clouds: approaches and research opportunities," *Scalable Computing: Practice and Experience* (15:3), pp. 251-270.
- Pollock, R. 2009. "The control of porting in platform markets," *The Journal of Economic Asymmetries* (6:2), pp. 155-180.
- Ranabahu, A., and Sheth, A. 2010. "Semantics centric solutions for application and data portability in cloud computing," *Proceedings of the International Conference on Cloud Computing Technology and Science (CloudCom)*, pp. 234-241.
- Ray, S., Kim, S. S., and Morris, J. G. 2012. "Research Note-Online Users' Switching Costs: Their Nature and Formation," *Information Systems Research* (23:1), pp. 197-213.
- Swire, P., and Lagos, Y. 2013. "Why the right to data portability likely reduces consumer welfare: Antitrust and privacy critique," *Maryland Law Review* (72:2), pp. 335-380.
- Valero, J. 2015. "Tirole: Brussels must level the playing field for online platforms," accessed on 07.06.2017, available at <http://www.euractiv.com/section/digital/news/tirole-brussels-must-level-the-playing-field-for-online-platforms/>.