Harvard Journal of Law & Technology
Volume 36, Number 2 Spring 2023

THE PARADOX OF DATA PORTABILITY AND LOCK-IN EFFECTS

*Jiawei Zhang**

TABLE OF CONTENTS

I. INTRODUCTION 658

II. WHY DATA PORTABILITY? 663
A. Data Autonomy of Individuals 664
B. Competition in Digital Markets 665
C. Solving Lock-In Effects as a Bridge in Between 666

III. LOCK-IN EFFECTS 667
A. User Lock-In and Switching Costs 668
1. Market-Inherent Switching Costs ("MISCs") 670
2. Artificially Raised Switching Costs ("ARSCs") 671
B. Data Lock-In and Essential Facility Doctrine 673

IV. THE MISALIGNMENT OF DATA PORTABILITY AND LOCK-IN EFFECTS 675
A. Misalignments 675
1. Data Portability & Non-Data-Based Features 676
2. Data Portability & Switching Costs 677
3. Data Portability & Data-Based Concerns 679
B. Recalibration 680

V. CONCLUSION 682

* MPhil Candidate in Information Technology Law at the University of Oxford; LL.M. graduate from U.C. Berkeley Law School. All comments are welcomed via jiawei.zhang@law.ox.ac.uk. This piece won first place in the *Harvard Journal of Law & Technology* Note Competition. The author would like to extend his gratitude to Robert Burrell, Lothar Determann, Christopher Hockett, Yuan Hao, Prasad Krishnamurthy, Viktor Mayer-Schönberger, Joel Sanders, Paul M. Schwartz, Chuanhui Wang, as well as his friends, Shuang Liu, Ruiyang Ma, Xiangyu Ma, Amirhossein Salehi, Ruikai Yan, and Chunting Guo for their invaluable suggestions and considerate support at different stages of drafting this Note. He also expresses special thanks to Jillian Paffenbarger and other editors of the *Harvard Journal of Law & Technology* for their excellent editorial suggestions and to Dr. Liu Chak Wan for his generous financial support for the author's MPhil program at Oxford.

I. INTRODUCTION

Data, considered the new oil,[1] is arguably the most valuable resource in the digital age. Data has also been compared to the plankton that nourishes ocean life in a Darwinian Sea, as free access to data is indispensable for companies to continuously innovate and engage in effective competition.[2] Accordingly, intensive discussions in the last decade have taken place concerning the role of data in the digital economy. A keenly debated topic is the most appropriate regulatory approach toward the tech giants that control unprecedented volumes of data. In these debates, support for data portability rules [3] has

1. *See The World's Most Valuable Resource Is No Longer Oil, but Data*, ECONOMIST (London) (May 6, 2017), https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data [https://perma.cc/U3D2-GWGQ]; Janos Barberis, *From Fintech to Techfin: Data Is the New Oil*, ASIAN BANKER (May 16, 2016), https://www.theasianbanker.com/updates-and-articles/from-fintech-to-techfin:-data-is-the-new-oil [https://perma.cc/9TSF-X32J]. *But see* VIKTOR MAYER-SCHÖNBERGER & THOMAS RAMGE, ACCESS RULES: FREEING DATA FROM BIG TECH FOR A BETTER FUTURE 23–25 (2022) (arguing that data is not like oil because "data doesn't burn when used" and "its value increases precisely through diverse, repeated, and combined use"); MICHAEL MANDEL, PROGRESSIVE POL'Y INST., THE ECONOMIC IMPACT OF DATA: WHY DATA IS NOT LIKE OIL P2 (July 12, 2017), https://progressivepolicy.org/wp-content/uploads/2017/07//PowerofData-Report_2017.pdf [https://perma.cc/7BD4-KF8F] (arguing that "[d]ata should not be compared to oil" because "it is not a scarce commodity, is nonrival, and cannot be monopolized").

2. *See* MAYER-SCHÖNBERGER & RAMGE, *supra* note 1, at 75–79; *see also* MAURICE E. STUCKE & ALLEN P. GRUNES, BIG DATA AND COMPETITION POLICY ¶¶ 4.01–.13 (2016) (illustrating the strategies that some companies use to adapt to big data implications with empirical evidence).

3. For support on data portability, see, for example, Whitney Nixdorf, *Planting in a Walled Garden: Data Portability Policies to Inform Consumers How Much (if Any) of the Harvest Is Their Share*, 29 TRANSNAT'L L. & CONTEMP. PROBS. 135, 138, 148–49 (2020) (explaining that the benefits of data portability include increasing transparency and choice, enhancing users' control over their information, decreasing lock-in, and encouraging innovation and competition); Paul De Hert, Vagelis Papakonstantinou, Gianclaudio Malgieri, Laurent Beslay & Ignacio Sanchez, *The Right to Data Portability in the GDPR: Towards User-Centric Interoperability of Digital Services*, 34 COMPUT. L. & SEC. REV. 193, 203 (2018) (proposing a wide interpretation of the scale of data portability to further strengthen users' control over data and foster innovation); Barbara Van der Auwermeulen, *How To Attribute the Right to Data Portability in Europe: A Comparative Analysis of Legislations*, 33 COMPUT. L. & SEC. REV. 57 (2017) (proposing greater data portability rights for similar reasons); Helena Ursic, *Unfolding the New-Born Right to Data Portability: Four Gateways to Data Subject Control*, 15 SCRIPTED 42, 45 (2018) (arguing that "data portability could increase transparency of data processing and could allow data subjects to better control their online identities"); Eva Fialová, *Data Portability and Informational Self-Determination*, 8 MASARYK U. J.L. & TECH. 45, 54 (2014) (arguing that data portability "would ensure the control of the data subject with respect to his/her data, and herewith to guarantee the informational self-determination"); Gabriela Zanfir, *The Right to Data Portability in the Context of the EU Data Protection Reform*, 2 INT'L DATA PRIV. L. 149, 162 (2012) (arguing that "data portability is vital for the global development of cloud computing" and can act as a "catalyst[] of firmer, clearer, and more coherent policies for international data transfers"); Joaquín Almunia, V.P. Eur. Comm'n for Competition Pol'y, Eur. Comm'n, Address at the European Commission Privacy Platform Event: Competition and Personal Data Protection (Nov. 26, 2012), https://ec.europa.eu/

overwhelmingly outweighed opposition.[4] One of the most widely cited arguments for data portability is that it is pivotal to competition and innovation due to its ability to solve lock-in problems in digital markets.[5]

Lock-in effects[6] are ubiquitous, especially in digital markets. Users will be locked into a system "[w]hen the costs of switching from one brand of technology to another are substantial."[7] Some argue that platform users face substantial lock-in effects because they "can only

---

commission/presscorner/detail/en/SPEECH_12_860 [https://perma.cc/TTG3-CJ9R] (noting that "effective competition requires that customers can switch by taking their own data with them").

4. For doubts on data portability's impacts on fair competition, see, for example, Peter Swire & Yianni Lagos, *Why the Right to Data Portability Likely Reduces Consumer Welfare: Antitrust and Privacy Critique*, 72 MD. L. REV. 335, 338–39 (2013) (arguing that per se requirements of data portability will cause false positives in market competition analysis because data portability rules apply to both start-up companies and monopolists). For doubts on data portability's impacts on consumer welfare, see, for example, Jan Krämer, *Personal Data Portability in the Platform Economy: Economic Implications and Policy Recommendations*, 17 J. COMPETITION L. & ECON. 263, 273 (2021) (mentioning that, "as data can be easily ported to the entrant, the new provider has less [sic] incentives to economize on data use and increases the amount of data collected," which can consequently result in a reduction in consumer welfare); Jan Krämer & Nadine Stüdlein, *Data Portability, Data Disclosure and Data-Induced Switching Costs: Some Unintended Consequences of the General Data Protection Regulation*, 181 ECON. LETTERS 99, 99 (2019) (arguing that because of the GDPR right to port data, an incumbent content provider "has less [sic] incentives to preserve users' privacy"); Michael Wohlfarth, *Data Portability on the Internet*, 61 BUS. & INFO. SYS. ENG'G 551, 552 (2019) (arguing that "data portability is not necessarily beneficial for users because [content providers] entering the market have an incentive to increase the amount of data users have to reveal"). For doubts on data portability's impacts on innovation, see, for example, Barbara Engels, *Data Portability Among Online Platforms*, 5 INTERNET POL'Y REV., no. 2, 2016, at 1, 7–8 (cautioning that extending the right to data portability without proper nuance could hamper innovation); Thomas M. Lenard, *If Data Portability Is the Solution, What's the Problem?*, 43.1 REGULATION 10, 10 (2020), https://www.cato.org/sites/cato.org/files/2020-03/regv43n1-5.pdf [https://perma.cc/QHU7-544T] ("[Data portability] would also reduce potential returns for winners and therefore the incentive to invest and innovate.").

5. *See, e.g.*, Christopher S. Yoo, *When Antitrust Met Facebook*, 19 GEO. MASON L. REV. 1147, 1154 (2012) (noting the argument made by some scholars that "the inability to move data from one social networking site to another can create a form of lock-in"); Aaron Perzanowski & Jason Schultz, *Digital Exhaustion*, 58 UCLA L. REV. 889, 900 n.50 (2011) ("Switching costs would be reduced further if consumers were assured data portability between platforms."); Nixdorf, *supra* note 3, at 148 ("[T]he most often cited benefit of a right to data portability is decreased consumer 'lock-in' caused by high switching costs and network effects."); Aysem Diker Vanberg & Mehmet Bilal Ünver, *The Right to Data Portability in the GDPR and EU Competition Law: Odd Couple or Dynamic Duo?*, 8 EUR. J.L. & TECH., no. 1, 2017, at 1, 6 ("Data portability will indeed have a significant impact on avoiding consumer lock-in and switching costs."); Jay P. Kesan, Carol M. Hayes & Masooda N. Bashir, *Information Privacy and Data Control in Cloud Computing: Consumers, Privacy Preferences, and Market Efficiency*, 70 WASH. & LEE L. REV. 341, 470 (2013) ("[D]ata mobility in the cloud would facilitate consumer participation and reduce transaction costs for consumers when moving from one provider to another.").

6. Because lock-in effects are not per se anticompetitive, this Note uses this concept as well as the verb "lock[ed]" in a neutral way.

7. CARL SHAPIRO & HAL R. VARIAN, INFORMATION RULES: A STRATEGIC GUIDE TO THE NETWORK ECONOMY 104 (1999).

change to another platform at the cost of leaving their data."[8] Tech companies have significant incentives to entrench their dominant positions by guarding their exclusive access to the data they collect from users and keeping their systems closed.[9] In this situation, users might be less likely to switch to an alternative system and will eventually be locked into a single system. Hence, data portability rules have frequently been identified as an appropriate tool to mitigate lock-in effects and ensure a more competitive market.[10]

The European Union ("EU") has led the way in accepting data portability rules through the General Data Protection Regulation ("GDPR").[11] It has subsequently sought to expand the scope of data portability in the Digital Markets Act ("DMA").[12] Some U.S. scholars and policymakers also advocate that the U.S. federal government follow the EU's path and apply data portability rules more widely.[13] Those advocating this approach clearly assume that data portability provides an effective solution to user lock-in effects.[14] However, there is cause to question whether this is necessarily true.

Most studies have approached this issue from the perspective of the competitive relationships among digital platforms. Relevant

---

8. Engels, *supra* note 4, at 2.

9. *See id.*; Nathan Newman, *Search, Antitrust, and the Economics of the Control of User Data*, 31 YALE J. ON REGUL. 401, 425–26 (2014) (arguing that Google maintains its dominance by excluding potential competitors' access to users' data); *see also* Inge Graef, Sih Yuliana Wahyuningtyas & Peggy Valcke, *Assessing Data Access Issues in Online Platforms*, 39 TELECOMMS. POL'Y 375, 379–83 (2015) (describing the American and European judicial approaches to data access in the context of antitrust allegations against social media platforms).

10. *See supra* note 5 and accompanying text.

11. *See* Regulation (EU) 2016/679 of the European Parliament and of the Council 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), art. 20(1), 2016 O.J. (L 119) 1, 45 [hereinafter GDPR]. Although GDPR's main purpose is to enhance data subjects' control over their personal data, the requirement of providing a "machine-readable" format implies some competition concerns. *See* Article 29 Data Protection Working Party, Guidelines on the Right to Data Portability, WP 242, at 4 (Apr. 5, 2017) [hereinafter *WP29 Data Portability Guidelines*].

12. *See* Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on Contestable and Fair Markets in the Digital Sector and Amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), art. 6(9), 2022 O.J. (L 265) 1, 36 [hereinafter DMA].

13. *See* Augmenting Compatibility and Competition by Enabling Service Switching Act of 2021, H.R. 3849, 117th Cong. § 3 (2021) (proposing a data portability duty to "promote competition, lower entry barriers, and reduce switching costs for consumers and businesses online"); Data To Go: An FTC Workshop on Data Portability, FED. TRADE COMM'N (Sept. 22, 2020), https://www.ftc.gov/news-events/events/2020/09/data-go-ftc-workshop-data-portability [https://perma.cc/UTV4-GYW9] (arguing that "data portability may benefit competition by allowing new entrants to access data they otherwise would not have so that they can grow competing platforms and services"); *see also* Nixdorf, *supra* note 3, at 160–62 (proposing that the Federal Trade Commission promulgate data portability rules to solve "lock-in" problems).

14. *See* Nixdorf, *supra* note 3, at 165.

discussions include, but are not limited to, whether big data is a valuable resource for companies;[15] whether big data has created entry barriers and entrenched the market dominance of the tech giants;[16] whether data portability can mitigate the anticompetitive effects of data monopolies;[17] and whether data portability will create free-riding problems and harm platform incentives to collect and analyze data.[18] However, data portability involves more than the relationships among digital entities; users also play an indispensable role in the process of data porting.

This Note approaches data portability from the user perspective and discusses how the current preference for data portability overestimates its capacity to solve platforms' lock-in problems. Through a

---

15. *See, e.g.*, Andrew McAfee & Erik Brynjolfsson, *Big Data: The Management Revolution*, 90 HARV. BUS. REV. 60, 68 (2012) (acknowledging barriers to big data usage for companies, but arguing that the benefits are worth it); Enric Junqué de Fortuny, David Martens & Foster Provost, *Predictive Modeling With Big Data: Is Bigger Really Better?*, 1 BIG DATA 215, 223 (2013) (demonstrating that "institutions with larger data assets . . . potentially can obtain substantial competitive advantage"). *But see* Anja Lambrecht & Catherine E. Tucker, *Can Big Data Protect a Firm from Competition?*, 1 CPI ANTITRUST CHRON., Jan. 2017, at 1, 5, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2705530 [https://perma.cc/K4DP-BA7U] (concluding that "by itself, big data is unlikely to be valuable").

16. *See, e.g.*, Newman, *supra* note 9, at 425–41 (explaining Google's strategies to protect its competitive dominance by expanding its control over users' data); European Data Protection Supervisor Press Release EDPS/2014/06, Privacy and Competitiveness in the Age of Big Data (Mar. 26, 2014) https://ec.europa.eu/commission/presscorner/detail/en/EDPS_14_6 [https://perma.cc/8F77-AQQA] ("The collection and control of massive amounts of personal data are a source of market power for the biggest players in the global market for internet services . . . ."); Daniel L. Rubinfeld & Michal S. Gal, *Access Barriers to Big Data*, 59 ARIZ. L. REV. 339, 345 (2017) (analyzing "the different types of barriers that limit entry into the different links of the data-value chain"). *But see* Lambrecht & Tucker, *supra* note 15, at 8 (emphasizing that "the simple act of amassing big data by itself does not confer a long-term competitive advantage"); Darren S. Tucker & Hill B. Wellford, *Big Mistakes Regarding Big Data*, THE ANTITRUST SOURCE, Dec. 2014, at 1, 4, 12 (arguing that partly due to the "ubiquity, low cost, wide availability, and fleeting value" of data, "big data creates no durable barriers to entry or any other significant competitive threat"); Andres V. Lerner, *The Role of "Big Data" in Online Platform Competition*, TELECOMMS. & REGULATED INDUS. J. 1, 62 (Aug. 2014) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2482780 [https://perma.cc/4FTW-AB8S] ("[E]ven if there were significant economies of scale from the collection of user data in online markets, the mere existence of such scale economies do not establish that large providers have monopoly power or that there is harm to consumers and competition, and do not justify more aggressive antitrust intervention.").

17. *See supra* note 5 and accompanying text; *cf.* Peter Swire, *The Portability and Other Required Transfers Impact Assessment (PORT-IA): Assessing Competition, Privacy, Cybersecurity, and Other Considerations*, 6 GEO. L. TECH. REV. 57, 115–18 (2022) (questioning whether data portability "reduce[s] [the] lock-in effect and facilitate[s] switching to competing providers" and concluding that "the prominence of lock-in effects in all the case studies suggests the importance of identifying the cause of lock-in effects early in consideration of a [data portability] initiative").

18. *See* Krämer, *supra* note 4, at 273 (arguing that, under a data portability framework, "the new provider has less [sic] incentives to economize on data use and increases the amount of data collected"); Krämer & Stüdlein, *supra* note 4, at 99 (demonstrating that "customers of the new [content provider] are worse off, while customers of the incumbent [content provider] are better off"); *see also* Engels, *supra* note 4, at 13 (arguing that data portability could "also hamper innovation by making data too available").

discussion of the concrete example of the web browser market, this Note highlights that users may not have sufficient motivation to port inferred and derived data that are functionally essential to digital market competition. Instead, users will only port data that would otherwise raise their switching costs but is less important for enhancing competitiveness.[19] This Note concludes that such limited data porting is insufficient to remedy the digital marketplace's lack of competition and innovation.

Parts II and III lay the theoretical foundations for the following discussions. Specifically, Part II analyzes two proposed objectives for data portability, which are first, to enhance the data autonomy of individuals and, second, to reinvigorate competition in digital markets. While the first, individual-oriented objective requires a relatively small applicable scope of data portability, the second ambition to resurrect digital markets entails much broader portability of derived or inferred data. These two objectives are logically bridged by the expectation of solving lock-in effects.

Part III, which uses the web browser market as an example, illustrates user lock-in and data lock-in problems. This analysis is split into two Sections. The first categorizes different sources of user lock-in effects, such as lock-in by non-data-based features and data-based features. It also explains how market-inherent switching costs ("MISCs") and artificially raised switching costs ("ARSCs") create user lock-in problems. The second Section then focuses on data lock-in problems and argues that the essential facility doctrine is not applicable to all kinds of data, because not all data are functionally essential to digital market competition.

Part IV comprises the main contribution of this Note by highlighting the paradox of data portability and lock-in effects. It first illustrates three ways in which the goals of data portability and its actual functions are misaligned. First, data portability is not capable of solving lock-in effects by non-data-based features. Second, data portability is less competent to mitigate the lock-in effects generated by MISCs. Third, although data portability can lower ARSCs, it is far from sufficient to handle data-based concerns. Part IV then recalibrates the relationship between the real functions of data portability and the goal of

19. In this sense, it is comparable to the "privacy paradox" discussion in academia — privacy policymakers found that while users highly value their privacy, their real-world behavior did not reflect this. *See generally* Nina Gerber, Paul Gerber & Melanie Volkamer, *Explaining the Privacy Paradox: A Systematic Review of Literature Investigating Privacy Attitude and Behavior*, 77 COMPUTS. & SEC. 226, 226 (2018) (exploring the privacy paradox that, while "privacy of their personal data is an important issue for online users worldwide, most users rarely make an effort to protect this data actively and often even give it away voluntarily"); Alessandro Acquisti, *Privacy in Electronic Commerce and the Economics of Immediate Gratification*, 2004 PROC. 5TH ACM CONF. ON ELEC. COM. 21, 23–24 (explaining the dichotomies between privacy attitudes and behavior by analyzing the individual decision-making process).

eliminating detrimental lock-in effects. It argues that data portability rules governed by individualistic goals have already been sufficient to address potential market-related concerns — that is, to reduce ARSCs. Expanding the applicable scope of data portability to address further market goals is therefore unnecessary and meaningless.

Finally, Part V advises policymakers not to expect data portability to have the same functionalities as mandatory data sharing.

II. WHY DATA PORTABILITY?

The concept of data portability discussed in the following Parts should be clarified upfront. It includes both "one-off export" portability and interoperability as two types of portability.[20] One-off export portability enables users to "download a snapshot of the data they have on one platform in a form that can be uploaded to another,"[21] while interoperability "allow[s] two or more platforms to exchange information directly with one another."[22] However, other researchers define data portability more narrowly and use it in parallel with the concept of interoperability.[23] This Note chooses to conduct its analysis using a narrower conceptualization and focuses only on so-called "one-off export" portability.

Goals for data portability can be primarily distilled into two clusters: the first cluster relates to enhancing data subjects' autonomy over their data, and the second cluster is aimed at invigorating competition and innovation in digital markets.[24] Different policy goals require different scopes of data subject to the portability rules. For example, in legislative documents, the GDPR primarily prioritizes the individual side[25] and, as such, data portability applies only to the data that a data subject "has *provided* to a controller"[26] ("provided data"). In contrast, the DMA aims to complement the GDPR by accommodating market objectives with a greater range of data portability.[27] Thus, under the

---

20. *See, e.g.*, Gabriel Nicholas, *Taking It with You: Platform Barriers to Entry and the Limits of Data Portability*, 27 MICH. TECH. L. REV. 263, 270–71 (2021).

21. *Id.* at 270.

22. *Id.* at 271; *see also infra* note 115.

23. *See, e.g.*, Engels, *supra* note 4, at 4 (differentiating the concepts of data portability and interoperability using examples from Facebook); Nixdorf, *supra* note 3, at 147 (explaining that interoperability entails technical compatibility of systems while data portability requires transferability of personal data). Similarly, DMA follows this narrower conception in its provisions. *See, e.g.*, DMA, *supra* note 12, art. 2(29), 6(7), 6(9).

24. *See* Salomé Viljoen, *A Relational Theory of Data Governance*, 131 YALE L.J. 573, 621 (2021) ("Data portability combines elements of the data control claim and the market efficiency claim to enhance competitive opportunity via individuals' market actions.").

25. GDPR, *supra* note 11, recital 68, at 13 (emphasis added) ("*To further strengthen the control over his or her own data* . . . the data subject should also be allowed to receive personal data concerning him or her which he or she has provided to a controller . . . .").

26. *Id.* art. 20(1), at 45 (emphasis added).

27. *See* DMA, *supra* note 12, recital 59, at 15.

DMA, apart from provided data, any data "*generated* through the activity of the end user in the context of the use of the relevant core platform service" is also included.[28]

A. Data Autonomy of Individuals

The primary objective of data portability is to enhance individuals' control over their personal data. For instance, the right to data portability under the GDPR mainly aims to "empower data subjects regarding their own personal data," although it may also objectively facilitate service switching and enhance competition.[29] Similarly, the data portability arrangements in the California Consumer Privacy Act ("CCPA") and the California Privacy Rights Act ("CPRA") elevate autonomy considerations of data subjects over competition considerations of the digital market because, under the CCPA, it is not mandated that ported data be in a machine-readable format.[30]

The rationale behind this goal is that data portability can enable data subjects to establish control over the transfer and reuse of their data and better facilitate equality[31] and the "free development of personality."[32] This goal represents individuals' data autonomy or informational self-determination — to "determine for themselves when, how, and to what extent information about them is communicated to others."[33] Moreover, by enhancing data subjects' control over their data, data portability can also improve the transparency of data processing and bridge the information asymmetry between tech giants and internet users.[34]

Based on this individualistic objective, the scope of data that is subject to data portability rules should be properly tailored, not only to protect data subjects' autonomy and self-determination, but also to avoid disproportionate disclosure. For example, despite some conceptual uncertainties,[35] under the GDPR's rules, according to the WP29 Data Portability Guidelines, the provided data not only includes the

---

28. *Id.* art. 6(9), at 36 (emphasis added).

29. *WP29 Data Portability Guidelines*, *supra* note 11, at 4.

30. *See* CAL. CIV. CODE §§ 1798.105, .110, .115. Unlike Article 20 of the GDPR, which mandates machine-readable format for ported data, CCPA and CPRA do not require such provisions. This implies that CCPA and CPRA do not prioritize the interests in data transfer between digital platforms. *See* Nicholas, *supra* note 20, at 266 (arguing that data portability rules under CCPA "offer[] little utility to new market entrants since [they] do[] not even require platforms to make their data available in a machine-readable format").

31. *See* Ursic, *supra* note 3, at 58–67.

32. *Id.* at 59.

33. *See* Viljoen, *supra* note 24, at 599 (quoting ALAN F. WESTIN, PRIVACY AND FREEDOM 7 (1967)).

34. *See* Kesan et al., *supra* note 5, at 372–73.

35. *See* Krämer, *supra* note 4, 266–67 (distinguishing between "volunteered data," "observed data," and "inferred data").

data *provided* by users, such as mailing address, username, and age, but also the data *observed* from their activities, such as their search history, traffic data, and location data.[36] However, inferred and derived data, which are created by the data controller using the data provided by users, are excluded from the GDPR's portability rules.[37]

B. Competition in Digital Markets

On the market side, data portability is often treated as a panacea to rejuvenate competition and innovation in digital markets.[38] Specifically, data portability is expected to achieve this ambition in two main ways. First, some commentators believe that data portability has the potential to solve lock-in effects, enrich consumers' choices, remove market entry barriers, and enable competitors to acquire a larger user base by lowering users' switching costs (*user-based concerns*).[39] Phone number portability provides a suitable example of user-based concerns. This is because it mitigates the need for users to spend a large amount of time informing their contacts of their new number, liberating them from one system and allowing them to freely choose their telecommunications service providers.[40]

Second, data portability is also expected to enable potential competitors to access more users' data and eventually achieve the free flow of data by mandating machine-readable data transfer and free data sharing (*data-based concerns*).[41] The necessity of data portability derives from the underlying belief that high volumes of data generate an overwhelmingly competitive advantage and increase market concentration. This is because consumers' information will be locked into a single system and market entrants will not be able to obtain a sufficient user base to generate enough revenues and gain access to the necessary

36. *See WP29 Data Portability Guidelines*, *supra* note 11, at 9–10.

37. *Id.* at 10–11.

38. For instance, one of the objectives of data portability under the DMA is "[t]o ensure that gatekeepers do not undermine the contestability of core platform services, or the innovation potential of the dynamic digital sector, by restricting switching or multi-homing . . . ." DMA, *supra* note 12, recital 59, at 15.

39. *See, e.g.*, Nixdorf, *supra* note 3, at 138 (noting that "the purpose of the regulation is to provide consumers with greater choice, control, and empowerment, and to 're-balance' the relationship between consumers and data controllers"); Van der Auwermeulen, *supra* note 3, at 68 (explaining that one "objective of data portability is to reduce the lock-in of the consumers," which would require online providers to share data and, in turn, "reduce monopoly power and therefore improve competition in the market").

40. *See, e.g.*, 47 C.F.R. § 52.23 (2023) (mandating that local exchange carriers allow users to port their phone numbers to different carriers when switching services).

41. *See, e.g.*, Rubinfeld & Gal, *supra* note 16, at 349–51 (arguing that big data serves as a market entry barrier and that portability is an important way to access data); Diker Vanberg & Ünver, *supra* note 5, at 14 (arguing that mandatory data portability is necessary where "data owned by an incumbent is necessary for the appearance of a new product or service").

amount of data to provide high-quality services.[42] Thus, it seems that data portability rules are necessary to stop a vicious circle from becoming embedded.

These market-based objectives might entail a broader scope of data portability. With such objectives, secondary data, like inferred data and derived data, should also be subject to data portability.[43] This is because provided data (including observed data) is not as valuable as secondary data and thus is insufficient to address the dearth of competition and innovation in digital markets.[44] These inferred and derived data include information relating to users' political leanings, demographic status, and shopping interests.[45]

For example, compared to the GDPR, the DMA's data portability rules use arguably broader wording to accommodate the act's market-based objectives. For instance, data "provided by the end user or *generated* through the activity of the end user" are within the scope of data portability.[46] The fact that the DMA does not follow the previous concept (i.e., observed data) adopted in the WP29 Data Portability Guidelines, but instead uses another, indicates that implied and derived data are also included. Some scholars point to Recital 59 of the DMA to support this view.[47]

C. Solving Lock-In Effects as a Bridge in Between

Whether addressing user-based or data-based concerns, two logical chains start with data portability enhancing users' control over their data and end with a more competitive environment and greater

---

42 . *See infra* note 51 and accompanying text; *see also* JACQUES CRÉMER, YVES-ALEXANDRE DE MONTJOYE & HEIKE SCHWEITZER, EUR. COMM'N, COMPETITION POLICY FOR THE DIGITAL ERA 98–100 (2019), https://data.europa.eu/doi/10.2763/407537 [https://perma.cc/H98D-Y2TX] (explaining why competition law should require mandatory data access). *But see* Emmanuel Syrmoudis, Stefan Mager, Sophie Kuebler-Wachendorff, Paul Pizzinini, Jens Grossklags & Johann Kranz, *Data Portability Between Online Services: An Empirical Analysis on the Effectiveness of GDPR Art. 20*, PROC. ON PRIV. ENHANCING TECHS. 351, 366 (2021) (showing that "incumbents are most generous in exporting data and offering opportunities to import data").

43. *See, e.g.*, Wing Man Wynne Lam & Xingyi Liu, *Does Data Portability Facilitate Entry?*, 69 INT'L J. INDUS. ORG., Mar. 2020, at 1, 3, 13.

44. *See, e.g.*, Swire, *supra* note 17, at 73 (noting the comparative value of secondary data, because "inferred and derived data can provide powerful economic advantages to the companies that hold the largest and most nuanced databases about individuals").

45. *See id.*

46. DMA, *supra* note 12, art. 6(9), at 36 (emphasis added).

47. *See, e.g.*, Damien Geradin, Konstantina Bania & Theano Karanikioti, The Interplay Between the Digital Markets Act and the General Data Protection Regulation 5 (Aug. 29, 2022) (unpublished manuscript) (on file with SSRN), https://papers.ssrn.com/abstract=4203907 [https://perma.cc/GQ5T-2DKK] ("[T]hat the DMA may cover inferred and derived data finds support in Recital (59), which lays down that the data portability obligation it establishes is '[t]o ensure that gatekeepers do not undermine [. . .] the innovation potential of the dynamic digital sector.'").

innovation in digital markets.[48] In other words, the two clusters of objectives, the individualistic objectives and market-based objectives, are intrinsically interlinked with each other. However, although the objectives of data portability appear ambitious and the foregoing logical chains seem reasonable,[49] a more nuanced observation finds that reality might be different from what many people expect.

This is because the issue of addressing lock-in effects serves as a crucial bridge between the initial point of departure (i.e., data autonomy of individuals) and the ultimate goal (i.e., competition in digital markets). Users hold the key to this bridge, as they have the freedom to decide whether or not to utilize data portability to move their data. If users are not actually locked in by data, do not initiate data porting to trigger the downstream outcomes, or have no incentive to do so, both logical chains will collapse and the market targets will become unattainable — user-based and data-based hindrance will still exist and incumbents will still dominate the markets. In this sense, to eliminate lock-in effects with data portability, the issue of user lock-in is primary while the issue of data lock-in is ancillary.

III. LOCK-IN EFFECTS

The user-based and data-based barriers that data portability aspires to address are intrinsically linked and generate a feedback loop — collection and use of huge amounts of data by the first mover in the digital market can lead to better services and performance, which in turn can attract more users, enlarge network effects, and further entrench the first mover's dominant position with even more collected data to refine its performance.[50] However, this data feedback loop is subject to great controversy.[51] This Note argues that even if this feedback is present, the capacity for data portability to solve it is highly questionable.

---

48. For the logical chain of user-based concerns, see supra note 39 and accompanying text; for the logical chain of data-based concerns, see supra note 41 and accompanying text.

49. Specifically, for user-based concerns, data portability is expected to enable users to exercise their data autonomy to switch platforms with lower switching costs so that competitors can accordingly acquire a larger user base. For data-based concerns, data portability is assumed to be capable of unlocking users and their data from a dominant platform and thus achieving free sharing of large volumes of valuable data.

50. *See* MARC BOURREAU, ALEXANDRE DE STREEL & INGE GRAEF, CERRE, BIG DATA AND COMPETITION POLICY: MARKET POWER, PERSONALISED PRICING AND ADVERTISING 35–37 (2017) (providing a detailed explanation of the data feedback loop); Van der Auwermeulen, *supra* note 3, at 58 (explaining that the first mover in the digital markets, on the one hand, benefits from the network effects and large volumes of data collected from users, and, on the other hand, hinders users from switching to its competitors and raises barriers for potential market entrants).

51. *See* Patrick Bajari, Victor Chernozhukov, Ali Hortaçsu & Junichi Suzuki, *The Impact of Big Data on Firm Performance: An Empirical Investigation* 39–40 (Nat'l Bureau of Econ. Rsch., Working Paper No. 24334, 2018) ("[W]e do not see evidence for a version of the 'data

As web browser services contain both data-based and non-data-based features, and data portability among different browsers has occurred, this Note chooses the web browser market as an exemplar to explain the lock-in effects and their relationship with data portability (see Table 1).

A. User Lock-In and Switching Costs

Tech giants routinely contend that they face intense competitive pressure because users can switch to a competitors' platform easily or use homogeneous services offered by different companies — so-called "multihoming."[52] However, they actually have great incentives to devise various features to prevent multihoming and to lock users into their own systems.[53] The features that potentially have lock-in effects are either data-based or non-data-based.[54] For example, in terms of data-based features of a web browser, Google Chrome ("Chrome") collects data from users' browsing activities and uses algorithms to produce personalized content tailored to users' interests.[55] Chrome also stores

---

feedback loop' theory, wherein adding new products leads to an indirect network effect with more accurate forecasts leading to more customers/sales leading in turn to more accurate forecasts."); Lerner, *supra* note 16, at 6 ("The fact that cross-platform network effects are essentially one-sided fundamentally weakens or eliminates the possibility of a feedback loop that locks users and advertisers to a dominant platform."); D. Daniel Sokol & Roisin Comerford, *Antitrust and Regulating Big Data*, 23 GEO. MASON L. REV. 1129, 1148 (2016) (concluding that, "[i]n reality, the strength of the feedback loop may be grossly overstated").

52. For example, Larry Page, Google's co-founder, said in his 2012 open letter that "when our products don't work or we make mistakes, it's easy for users to go elsewhere because our competition is only a click away." Larry Page, *2012 Update from the CEO*, ALPHABET INV. RELS., https://abc.xyz/investor/founders-letters/2012/ [https://perma.cc/S69B-USAP]; *see also* Adam Kovacevich, *Google's Approach to Competition*, GOOGLE PUB. POL'Y BLOG (May 8, 2009), https://publicpolicy.googleblog.com/2009/05/googles-approach-to-competition.html [https://perma.cc/ZQM6-WZTZ] ("Competition is just one click away."). "Multihoming" originally referred to "purchasing connections from multiple providers and routing traffic among them in real time." Stanley Besen, Paul Milgrom, Bridger Mitchell & Padmanabhan Srinagesh, *Advances in Routing Technologies and Internet Peering Agreements*, 91 AM. ECON. REV. 292, 292 (2001). Currently, "multihoming" is mainly used to indicate "the ability for an individual to use multiple platforms to access similar services." Kenneth A. Bamberger & Orly Lobel, *Platform Market Power*, 32 BERKELEY TECH. L.J. 1051, 1067 (2017).

53. For example, proposed strategies to create lock-in effects include releasing new functionalities, promoting a broader level of openness to grow user network size, adjusting pricing, and denying interoperability. *See* Kalina Staykova & Jan Damsgaard, *How Digital Platforms Compete Against Diverse Rivals*, 20 MIS Q. EXEC. 275, 292 (2021).

54. *See, e.g.*, Lerner, *supra* note 16, at 30 (describing how the quality of platform services is dependent not only on data, but also on other inputs, including "engineering resources, innovation, and quality testing").

55. For an overview of relevant features, see Google Chrome, *Meet the Features that Set Chrome Apart*, https://google.com/chrome/browser-features/ [https://perma.cc/8WWY-97TP] [hereinafter *Features*]; *see also* Kate O'Flaherty, *It's Time to Ditch Chrome*, WIRED UK (June 6, 2021, 6:00 AM), https://www.wired.co.uk/article/google-chrome-browser-data [https://perma.cc/A99T-G2S] (explaining that "Google's Chrome app can collect data

recommended passwords that users accept and automatically fills them when users log in to their accounts on websites.[56] In terms of non-data-based functions, Chrome's design of its tabs, address bar, and sync function, as well as the appearance of different modes, may make users feel comfortable and give them a sense of efficiency.[57] These features can generate efficiencies, network effects, and anticompetitive switching costs that may either independently or collectively lock users into a single system.

Theoretically, users' lock-in effects derive from high switching costs.[58] Some scholars broadly interpret switching costs as "costs that are incurred when switching from one supplier of a particular good or service to another supplier, including money costs and the value of users' time."[59] On this view, among these costs are compatibility costs (such as network effects), contractual costs, and transaction costs.[60] However, some scholars treat the concept of switching costs separately from that of network effects, and both concepts are regarded as two of the main sources of lock-in problems.[61] This Note follows the former broader conception of switching costs and divides switching costs into two subsets, "market-inherent switching costs" and "artificially raised switching costs."[62] The Note will examine whether and to what extent data portability can address these costs.[63]

---

including [users'] location, search and browsing history, user identifiers and product interaction data for 'personalisation' purposes").

56. *See Generate a Password*, GOOGLE CHROME HELP, https://support.google.com/chrome/answer/7570435 [https://perma.cc/LVK4-L63P] [hereinafter *Generate a Password*]; *see also* Sanam Ghorbani Lyastani, Michael Schilling, Sascha Fahl, Michael Backes & Sven Bugiel, *Better Managed than Memorized? Studying the Impact of Managers on Password Strength and Reuse*, 2018 PROC. 27TH USENIX SEC. SYMP. 203, 204, 207–08; Fahad Alodhyani, George Theodorakopoulos & Philipp Reinecke, *Password Managers — It's All About Trust and Transparency*, 12 FUTURE INTERNET 189 (2020). However, the existing literature focuses heavily on security issues, rather than the potential competitive effects that this Note subsequently discusses.

57. *See Features*, *supra* note 55.

58. *See* SHAPIRO & VARIAN, *supra* note 7.

59. Aaron S. Edlin & Robert G. Harris, *The Role of Switching Costs in Antitrust Analysis: A Comparison of Microsoft and Google*, 15 YALE J.L. & TECH. 169, 176 (2013).

60. *Id.* at 193–96.

61. *See* Joseph Farrell & Paul Klemperer, *Coordination and Lock-In: Competition with Switching Costs and Network Effects*, *in* 3 HANDBOOK OF INDUS. ORG. 1967, 1970 (Mark Armstrong & Robert Porter eds., 2007) ("Switching costs and network effects bind customers to vendors if products are incompatible, locking customers or even markets in to early choices."); Catherine Tucker, *Digital Data, Platforms and the Usual [Antitrust] Suspects: Network Effects, Switching Costs, Essential Facility*, 54 REV. INDUS. ORG. 683, 683–84 (2019) (considering network effects and switching costs as two sources of market power).

62. *See* discussion *infra* Sections III.A.1–.2.

63. *See* discussion *infra* Part IV.

1. Market-Inherent Switching Costs ("MISCs")

Market-inherent switching costs ("MISCs"), also called endogenous switching costs or natural switching costs, are the costs "that arise from the nature of the product(s) or their market."[64] Although MISCs have raised antitrust concerns,[65] they are not per se detrimental and anticompetitive. Instead, they are neutral or even beneficial to consumers and should not be attributed to the entrenched parties.[66] For instance, direct network effects exist when "the utility that a user derives from consumption of the good increases with the number of other agents consuming the good."[67] Similarly, indirect network effects are generated by the nature of the multisided market, with "the value delivered to each user in one user group (say, consumers) increas[ing] as the number of users in another, interdependent user group (producers) grows."[68] Other switching costs such as the costs to search, purchase, and learn a new system also belong to MISCs, because they are naturally generated and inevitable in daily economic activities.

First, web browser users could plausibly be attracted by, for example, the highly personalized content or targeted advertisements certain browsers provide based on their user profiling.[69] This advertising is enabled by the data collected from other users with similar browsing

64. Edlin & Harris, *supra* note 59, at 176.

65. *See, e.g.*, First Amended Complaint for Injunctive and Other Equitable Relief at 53, Fed. Trade Comm'n v. Facebook, Inc., No. 20-cv-03590 (D.D.C. Aug. 19, 2021) ("Facebook's monopoly power is durable due to significant entry barriers, including direct network effects and high switching costs."); *see also* Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CALIF. L. REV. 479, 522 (1998) ("[N]etwork effects may foreclose competition entirely or limit effective competition to that occurring between members of the same network."); William E. Cohen, *Competition and Foreclosure in the Context of Installed Base and Compatibility Effects*, 64 ANTITRUST L.J. 535, 540–41 (1996) (arguing that network effects will raise efficiency concerns and "can be significant sources of first-mover advantage").

66. *See* William J. Kolasky, *Network Effects: A Contrarian View*, 7 GEO. MASON L. REV. 577, 585–86 (1999) (explaining that network effects are ubiquitous and arguing that they "can be either positive or negative").

67. Michael L. Katz & Carl Shapiro, *Network Externalities, Competition, and Compatibility*, 75 AM. ECON. REV. 424, 424 (1985).

68. ALEX MOAZED & NICHOLAS L. JOHNSON, MODERN MONOPOLIES: WHAT IT TAKES TO DOMINATE THE 21ST CENTURY ECONOMY 168 (2016).

69. Despite some controversies, some empirical evidence indicates that consumers prefer targeted advertisements to random ones. *See Consumers Say They Prefer Targeted to Random Online Ads*, MKTG. CHARTS (Apr. 19, 2013), https://marketingcharts.com/digital-28825 [https://perma.cc/P594-UMTQ] (showing that 40.5% of survey respondents chose to see targeted ads while only 16.1% preferred random ads). Some studies also demonstrated that consumers are more "willing[] to share their data in return for a more personalized and targeted shopping experience." *See* Grace Nasri, *Why Consumers Are Increasingly Willing To Trade Data for Personalization*, DIGIT. TRENDS (Dec. 10, 2012), https://www.digitaltrends.com/social-media/why-consumers-are-increasingly-willing-to-trade-data-for-personalization/#ixzz2x28Ahiyj [https://perma.cc/F263-QSGQ]. For theories explaining why targeted advertisements do not harm consumers and competitive process, see Lerner, *supra* note 16, at 12–19.

interests and histories.[70] In essence, such lock-in effects are derived from MISCs, and more specifically, from direct network effects. Second, users may continue to use a particular browser due to its indirect network effects. For example, some who use Chrome may be incentivized to continue using that browser due to its compatibility with Gmail, Google Docs, and other extensions. Third, users may also be reluctant to leave the browser of their choice because of other non-data-based features, such as its user interface or tab management system. These features are efficient and user friendly, so they raise relatively few antitrust concerns, despite some objective MISCs and lock-in effects.[71]

2. Artificially Raised Switching Costs ("ARSCs")

Unlike MISCs, artificially raised switching costs ("ARSCs"), as the term implies, are costs generated exogenously, and they involve strategic human participation.[72] Compared to MISCs, ARSCs are more problematic in the eyes of antitrust enforcers, because raising or maintaining switching costs might be held to be illegal and directly attributed to the market participant who intentionally undertakes this action.[73] The most common approach used to raise switching costs is the signing of long-term exclusive-dealing contracts.[74]

In the web browser market, a noteworthy example of ARSCs is recommended passwords. Specifically, when a user sets up an account using Chrome or Apple's Safari web browser, the browser automatically recommends a randomly selected password that combines a long

70. *See* Newman, *supra* note 9, at 431 (explaining that, for Google's services, "data feeds the accumulating profile that Google has not only on the user as an individual, but on aggregated profiles of people like them").

71. Under the rule of reason framework in antitrust law, potential anticompetitive effects should be weighed against the efficiencies. *See generally* Standard Oil Co. v. United States, 221 U.S. 1 (1911) (establishing the rule of reason principle); United States v. Microsoft Corp., 253 F.3d 34, 84 (D.C. Cir. 2001) (applying the rule of reason framework to Microsoft's tying conduct and holding that both anticompetitive nature and justifications should be considered and carefully balanced).

72. *See, e.g.*, Edlin & Harris, *supra* note 59, at 176 (comparing inherent switching costs to strategic switching costs and explaining that "[s]trategic switching costs reflect choices made by firms designed to create switching costs or increase them above their inherent level").

73. *See, e.g.*, *id.* at 185–88 (explaining that a court found Microsoft's "browser wars" to be anticompetitive because they were "an effort to maintain high switching costs among operating systems by maintaining the applications barrier to entry").

74. *See, e.g.*, ZF Meritor, LLC v. Eaton Corp., 696 F.3d 254, 289 (3d Cir. 2012) (holding that a manufacturer's alleged use of long-term agreements "functioned as unlawful exclusive dealing agreements," because it "unlawfully foreclosed a substantial share of the HD transmissions market" and harmed competition); *cf.* Omega Env't, Inc. v. Gilbarco, Inc., 127 F.3d 1157, 1163 (9th Cir. 1997) (footnote omitted) (ruling that "the short duration and easy terminability of these [one-year] agreements negate substantially their potential to foreclose competition").

chain of symbols, numbers, and letters.[75] Users do not need to memorize or otherwise record these passwords because the browser stores them and automatically fills in the password for users the next time they log in to the relevant account. This is the so-called "autofill" feature.[76] For users, these passwords are complex and difficult to memorize.

On one hand, recommended passwords and autofill features are convenient and efficient. However, if password portability is not available, users might be locked into a single browser because switching to another browser would mean that a large number of passwords could be lost unless a user spends a significant amount of time resetting them. In this sense, recommended passwords could act like long-term exclusive-dealing contracts that bind online services providers, users, and web browsers and further entrench the browser's dominant position.[77] Moreover, multihoming would become difficult because, compared to the passwords given by entrenched platforms, such as Chrome and Safari, users are less likely to accept and store their passwords in a system that they doubt they will use for a long time.

That does not mean that recommended passwords and autofill features are not user friendly. However, in the absence of appropriate remedies, their advantages can be eclipsed by their anticompetitive effects. Specifically, the features of recommended passwords will raise users' switching costs, restrict users' multihoming choices, and negatively affect browser market competition in the long run. This may draw the attention of antitrust regulators in calling for ex ante interference, such as imposing data portability requirements.

It is noteworthy that the differentiation between MISCs and ARSCs is not clear-cut. In fact, switching costs exist on a spectrum, and in many cases, are interwoven. Sometimes, a product or feature simultaneously has both MISCs and ARSCs. The phone number system is a typical example — it connects all people in that phone system and meanwhile hinders them from switching among alternative carriers if there is no portability in place.[78] Similarly, in the online browser market, the feature of recommended passwords bridges two sides of the market by connecting users to online services, but it could also stealthily raise the costs that users incur to switch to other browsers. In particular, users who are considering changing to another browser may

---

75. *See Generate a Password*, *supra* note 56; *see also* Safari User Guide, *Autofill Your Username and Password in Safari on Mac*, https://support.apple.com/guide/safari/autofill-user-name-and-password-info-ibrwf71ba236/14.0/mac/11.0 [https://perma.cc/GG43-CMS7].

76. *See, e.g.*, *id.*

77. The indirect network effects generated by recommended passwords and autofill features can reinforce the switching costs, which in turn will generate larger network effects. *See* Bamberger & Lobel, *supra* note 52, at 1068–69 ("Network effects and switching costs reinforce each other to create lock-in . . . .").

78. Colleen Bryan, *Number Portability for Consumers: Taking Your Wireless Number with You*, 16 LOYOLA CONSUMER L. REV. 267, 269 (2004).

encounter a high temporal cost to change all their passwords in the new platform or turn to the original browser for the password records. Similarly, the bookmark feature common across browsers raises comparable concerns but is less problematic than recommended passwords.[79]

B. Data Lock-In and Essential Facility Doctrine

Data lock-in is, to some extent, a derivative of user lock-in problems. This is because data lock-in is the result of the fact that most users are locked into the dominant system due to MISCs and/or ARSCs and, as such, market entrants with relatively small user bases struggle to gain sufficient data and generate economies of scale to effectively compete with the incumbents.[80] To solve data lock-in and to avoid markets that always favor incumbents, some researchers propose that data should be treated as "essential facilities" for competition and that digital entrants should be eligible to access these "essential facilities."[81]

However, such a generalized statement fails to stand up to nuanced examination. To constitute an essential facility, data would need to fulfill several constituent elements.[82] First and foremost, data would need to be an essential input such that "competition must fail without access to [it]."[83] However, data is not necessarily as valuable as people expect. The value of data is not merely decided by its volume, but also by its quality and type.[84] This means that data is not infallible; instead, it is

---

79. Hypothetically, this may be because users rely more heavily on passwords to access their desired websites and services than they rely on bookmarks. Consequently, features such as recommended passwords and autofill would be more closely aligned with the ARSCs, while the bookmark feature might be closer to the MISCs end of the spectrum.

80. *See, e.g.*, Zachary Abrahamson, *Essential Data*, 124 YALE L.J. 867, 871 (2014) (explaining that, as argued in recent legal disputes between online platforms, "[c]ompetitors could not duplicate the data because of network effects: each user who used the monopolist's platform made that platform more valuable to every other user").

81. *See, e.g.*, *id.* at 867 (making the claim that "the essential facilities doctrine sometimes should require open access to data").

82. Robert Pitofsky, Donna Patterson & Jonathan Hooks, *The Essential Facilities Doctrine Under U.S. Antitrust Law*, 70 ANTITRUST L.J. 443, 448 (2002) (footnotes omitted) ("[T]o establish antitrust liability under the essential facilities doctrine, a party must prove four factors: '(1) control of the essential facility by a monopolist; (2) a competitor's inability practically or reasonably to duplicate the essential facility; (3) the denial of the use of the facility to a competitor; and (4) the feasibility of providing the facility to competitors.'" (quoting MCI Commc'ns Corp. v. AT&T Co., 708 F.2d 1081, 1132–33 (7th Cir. 1983))).

83. *See* Tucker, *supra* note 61, at 690.

84. *See, e.g.*, *id.*, at 690–91 (discussing studies showing the limitations of the predictive power of data); *see also* Commission Decision (EC) No. COMP/M.4731 of 11 Mar. 2008, art. 8, 2008 O.J. (C 927) 71 ("Competition based on the quality of collected data thus is not only decided by virtue of the sheer size of the respective databases, but also determined by the different types of data the competitors have access to and the question [sic] which type eventually will prove to be the most useful for internet advertising purposes."); Geoffrey A. Manne & R. Ben Sperry, *The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework*, CPI ANTITRUST CHRON., May 2015, at 1, 9 ("Information is

inherently limited to specific times[85] and functions.[86] Put simply, data is not always essential to digital market competition.

For example, web browsers may lock users' data collected from their activities, such as browsing history or acceptance of specific functions like recommended passwords. However, not all these data are valuable in terms of browser market competition. For instance, recommended passwords belong to "provided data" but are not functionally valuable for other browser competitors because access to these passwords will not generate additional competitive advantages in the essential functioning of a web browser. In comparison, browsing history data might be useful for competitors to analyze users' interests and correspondingly produce tailored content and targeted advertisements.[87] Furthermore, those inferred or derived data, such as user profiling, would be comparatively more valuable in platform competition.[88]

important to companies because of the value that can be drawn from it, not for the inherent value of the data itself.").

85. *See, e.g.*, Tucker, *supra* note 61, at 686 (In the context of search engines, "more data on what consumers did in the past is not useful."); Margrethe Vestager, Competition Comm'r, Eur. Comm'n, Competition in a Big Data World, Digital Life Design Conference 2016 (Jan. 17, 2016), https://ec.europa.eu/commission/presscorner/detail/ov/speech_16_5224 [https://perma.cc/P9A8-D9NJ] ("It might not be easy to build a strong market position using data that quickly goes out of date. So we need to look at the type of data, to see if it stays valuable.").

86. *See, e.g.*, Bajari et al., *supra* note 51, at 5 (suggesting that increasing the data volume of different products has relatively little impact on forecasting performance); Lerner, *supra* note 16, at 27 ("Some types of user data may be useful for some purposes, but not for others.").

87. *See, e.g.*, David Martens, Foster Provost, Jessica Clark & Enric Junqué de Fortuny, *Mining Massive Fine-Grained Behavior Data To Improve Predictive Analytics*, 40 MIS Q. 869, 869 (2016) (finding that "when using fine-grained behavior data, there continues to be substantial value to increasing the data size across the entire range of the analyses"). *But see* Tucker, *supra* note 61, at 691 (explaining with empirical evidence that "browsing data alone is often not predictive about the kind of person who is doing the browsing").

88. *See, e.g.*, Swire, *supra* note 17, at 73 ("This inferred and derived data can provide powerful economic advantages to the companies that hold the largest and most nuanced databases about individuals.").

Table 1: Data Portability Objectives and Lock-In Effects

Objectives	Sources of Lock-In Effects		Browser Features	Competitive Effects
User-based Concerns (Section III.A)	Non-data-based features (III.A)		*e.g.*, appearance, tabs, address bar, sync function	Efficient
	Data-based features	MISCs (*e.g.*, Network Effects) (III.A.1)	*e.g.*, personalized content based on user profiling, compatibility with other extensions	Neutral, but likely problematic
		ARSCs (III.A.2)	*e.g.*, recommended passwords	Potentially anticompetitive
Data-based Concerns (Section III.B)	Essential facility (III.B)		N/A	Depends on the different types and functions of data

IV. THE MISALIGNMENT OF DATA PORTABILITY AND LOCK-IN EFFECTS

As described previously, the proponents of data portability have the ambition to solve lock-in effects to promote the free flow of data and eventually reinvigorate competition and innovation in the digital markets.[89] However, users become locked into a single system not only by data-based features but also by non-data-based features.[90] For users who are locked into a given system by data-based features, it remains questionable whether or not they have enough incentives to exercise their right to data portability when switching from one platform to another. This Part of the Note will unpack this paradox of data portability and lock-in effects and indicate how the objectives of data portability are misaligned with its actual functions.

A. Misalignments

In straightforward terms, the paradox is that users only have incentives to port those data that would otherwise be used to raise their switching costs but that are not functionally valuable to digital market competition. However, they are less likely to exercise the right to data portability to trigger the free flow of data that is arguably more essential to the competition of core services. In other words, the features that have the potential to lock users in are not largely driven by data that is functionally valuable for digital market competition. The result is that

89. *See* discussion *supra* Section II.B.
90. *See* discussion *supra* Section III.A.

data portability has a very limited role to play in solving lock-in problems and tackling the lack of digital market competition. This Note argues that the incapability of data portability in this regard originates from three misalignments (see Table 2).

1. Data Portability & Non-Data-Based Features

First, data portability is not capable of solving lock-in effects created by non-data-based features. In some circumstances, data does not function very effectively, and non-data-driven features play a key role in an application's performance and user experience.[91] Users may get locked into a system by features that are code-based and eligible for patent protection. In this sense, the innovation of the non-data-based features, instead of the data per se, is decisive in the quality of a system's services and users' switching motivation. Given this, data portability would naturally have little effect on solving such lock-in problems. In other words, users would not be locked in a system by data even without data portability.

Empirical research indicates that, in most cases, users do not seriously suffer from lock-in problems because multihoming is very common and effortless in users' daily online activities.[92] Moreover, the fact that most users are unfamiliar with data portability may also imply that users are actually not strongly locked in by data, and hence they are

91. For empirical research of data having little effect on users' experience, see Lesley Chiou & Catherine Tucker, *Search Engines and Data Retention: Implications for Privacy and Antitrust* 19 (Nat'l Bureau of Econ. Rsch., Working Paper No. 23815, 2017) (The result that "reducing the length of storage of past search engine searches [did not] affect[] the accuracy of search . . . suggest[s] that the possession of historical data confers less of an advantage to firms who own the data than is sometimes supposed."); *see also* Bajari et al., *supra* note 51, at 39–40 ("[W]e do not see evidence for a version of the 'data feedback loop' theory, wherein adding new products leads to an indirect network effect with more accurate forecasts leading to more customers/sales leading in turn to more accurate forecasts."); Nico Neumann, Catherine E. Tucker & Timothy Whitfield, *Frontiers: How Effective Is Third-Party Consumer Profiling? Evidence from Field Studies*, 38 MKTG. SCI. 918, 920–22 (2019) (finding in a study that brokers with larger data profiles do not necessarily produce more accurate results). The potential explanations for these results are, first, "personalization and customization undermine the potential for economies of scale and scope in data" and, second, "what may matter is the quality of the algorithm and the underlying engineering team," not the data. Tucker, *supra* note 61, at 686–87.

92. *See, e.g.*, Lerner, *supra* note 16, at 5 (arguing that "users can, and often do, utilize multiple online services, even for the same type of task" because no exclusive contracts are binding on users); Edlin & Harris, *supra* note 59, at 204 ("[T]he ability of consumers to use a combination of general and vertical search engines to find information is not hindered by switching or 'multi-homing' costs."). For empirical evidence, see, for example, Ryen W. White & Susan T. Dumais, *Characterizing and Predicting Search Engine Switching Behavior*, 2009 PROC. 18TH ACM CONF. ON INFO. & KNOWLEDGE MGMT. 87, 89 ("Of the 14.2 million users in our log sample, . . . 7.1 million (50.0%) switched engines within a search session at least once, and 9.6 million (67.6%) used different engines for different sessions (i.e., engaged in between-session switching).").

indifferent to their right to data portability.[93] In this case, "competition is just one click away"[94] even though no data portability is equipped.

For instance, if users are locked in by a specific web browser due to its tabs, address bar, or other non-data-based features, data portability would become ineffective and unnecessary. This is because users would switch to other browsers that offer more desirable counterpart features. In this scenario, data itself is relatively less important and has minimal impact on raising users' switching costs.

2. Data Portability & Switching Costs

Second, data portability is rarely suitable to mitigate the lock-in effects generated by MISCs. Users might be locked in by MISCs generated by data-based features,[95] but that does not mean that data portability is necessarily suitable for and capable of dealing with MISCs. MISCs are intrinsically neutral and not per se unlawful. Some features that ostensibly impose MISCs are actually efficient and user friendly.[96] Users locked in by these features may not have strong incentives to leave and port their data. For example, users may become attached to their browser's bookmark feature and stay with their well-categorized bookmarking system.[97]

Additionally, some features may gain a competitive advantage through the feedback loop between high volumes of collected data and the already established network effects, raising antitrust concerns.[98] However, users are less likely to port these data. These data are generally inferred or derived from users' daily activities and thus they are invisible to users and have very limited capabilities to lock users into

93. *See* Sophie Kuebler-Wachendorff et al., *The Right to Data Portability: Conception, Status Quo, and Future Directions*, 44 INFORMATIK SPEKTRUM 264, 266 (2021) (showing that the right to data portability "is the least known right of the GDPR, with less than a third of participants indicating that they have heard of it").

94. *See* Kovacevich, *supra* note 52.

95. *See* discussion *supra* Section III.A.1.

96. *See* Kolasky, *supra* note 66 and accompanying text.

97. This is currently not a significant concern as many major web browsers are equipped with a portability function that enables users to import or export their bookmarks. *See, e.g.*, *Import Bookmarks and Settings*, GOOGLE CHROME HELP, https://support.google.com/chrome/answer/96816?hl=en-GB [https://perma.cc/A3SL-B2QW]; *Import Bookmarks, History and Passwords in Safari on Mac*, SAFARI USER GUIDE, https://support.apple.com/en-gb/guide/safari/import-bookmarks-and-passwords-ibrw1015/16.1/mac/13.0 [https://perma.cc/C983-JLM6] [hereinafter *Safari Import*]; *Export Firefox Bookmarks to an HTML File to Back Up or Transfer Bookmarks*, MOZILLA SUPPORT, https://support.mozilla.org/en-US/kb/export-firefox-bookmarks-to-backup-or-transfer [https://perma.cc/PX82-5XZZ]; *What's Imported to Microsoft Edge*, MICROSOFT EDGE SUPPORT, https://support.microsoft.com/en-us/microsoft-edge/what-s-imported-to-microsoft-edge-ab7d9fa1-4586-23ce-8116-e46f44987ac2 [https://perma.cc/A3UB-NGCB] [hereinafter *Edge Import*].

98. *See* BOURREAU ET AL., *supra* note 50.

that system.[99] And, although data portability enables users to transfer data freely, it remains powerless to mitigate network effects.[100] Users have to act collectively and simultaneously to port their data and switch their services, but this rarely occurs.[101] In this situation, the crux of the present issue lies in network effects, rather than data.

Therefore, it is reasonable to predict that users will not switch to a new platform with fewer databases and lower performance. In fact, a digital company in the competitive market would have great incentives to invest in data-driven innovation based on its large database to improve its performance and prevent user turnover.[102] This is relatively harmless to users and does not undermine market competition. For another example, if web browser users appreciate highly targeted content, they will not turn to an alternative platform that possesses fewer users and less data and expect it to produce more accurate predictive results. In essence, network effects alone or network effects in combination with data, but not the data alone, discourage users from switching.

However, data portability is well suited to lowering data-based ARSCs. Without data portability, some data may eclipse other features that are relatively more valuable for competition and hinder users from switching to another system. But we should not expect to see this hindrance. Recommended passwords exemplify the data-based ARSCs. They are data intended to provide convenience and enhance security, yet this feature is not related to data analytics and thus is not significant to the competition of browsers' main functions. Analogous to customers being locked into a carrier system by their phone numbers, users accepting a large number of recommended passwords have to forgo consideration of other more preferable features and keep using the platform in which their passwords are stored. However, if users can port their complicated recommended passwords to another platform without any hindrance, competition for other functions will return. In fact, password portability features have been put in place on many browsers to

---

99. *See* SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER 186 (2019) (arguing that the inferred or derived data is "hidden from our view" for the benefit of "surveillance capitalists").

100. *See* Nicholas, *supra* note 20, at 287 (arguing that current data portability "fails to address network effects and informational economies of scale"); Krämer, *supra* note 4, at 274 (arguing that data portability does not alleviate the lock-in generated by user-side network effects).

101. *See* MAYER-SCHÖNBERGER & RAMGE, *supra* note 1, at 120.

102. *See* Esmeralda Florez Ramos & Knut Blind, *Data Portability Effects on Data-Driven Innovation of Online Platforms: Analyzing Spotify*, 44 TELECOMMS. POL'Y, Oct. 2020, at 1, 13 (discussing incentives for innovation in data analysis); *see also* Syrmoudis et al., *supra* note 42, at 366 ("Incumbents . . . know better how to use the [right to data privacy] for defending their positions by building enhanced trust with consumers, which in turn can lead to them providing more data.").

lower and even eliminate ARSCs.[103] This case is also corroborated by economic models.[104]

3. Data Portability & Data-Based Concerns

Third, although data portability can lower ARSCs, it is far from sufficient to handle data-based concerns. The nature of data portability should be re-emphasized here. Data portability is different from mandatory business-to-business data sharing. Instead, data portability is essentially "business-to-consumer/user-to-business" — users act as a bridge in between.[105] As illustrated before, users only have the incentive to port their data to remove ARSCs that otherwise create lock-in effects, whereas they have fewer reasons to port data when they are locked in by non-data-based features or data-based features with MISCs. However, those data with ARSCs which are subject to portability are quite limited in scope. This means that the derivative data-based concerns will not be solved because even if users exercise their right to portability, the ported data are insufficient to support effective competition and innovation. For example, password portability only solves user lock-in problems but leads to no direct increase in the volume of valuable data for the new browsers.[106]

However, it is possible that market entrants can make the most of data portability by inducing users to transfer as much data as possible.[107] The strong incentives for free-riding may lead to an increasing amount of collected data as well as a decrease in the innovation of other features, which will ultimately render consumers worse off.[108] Moreover, large-scale data porting may seriously raise users' concerns about privacy and security.[109]

To conclude, users only have strong incentives to port data that might otherwise create ARSCs. In contrast, users who are locked in by non-data-based features or by data with MISCs might show lower interest in porting their data to another platform. Such deficiencies mean

---

103. *See, e.g.*, *Manage Passwords*, GOOGLE CHROME HELP, https://support.google.com/chrome/answer/95606 [https://perma.cc/Z62W-AN8U]; *Safari Import*, *supra* note 97; *Export Login Data from Firefox*, MOZILLA SUPPORT, https://support.mozilla.org/en-US/kb/export-login-data-firefox [https://perma.cc/5V76-8YRT]; *Edge Import*, *supra* note 97.

104. *See, e.g.*, Wynne Lam & Liu, *supra* note 43, at 6–7 (vindicating that without big data analytics, greater data portability is negatively correlated with switching costs).

105. *See* discussion *supra* Section II.C.

106. For competitive value disparity of different data, see discussion *supra* Section III.B.

107. *Cf.* Wohlfarth, *supra* note 4, at 552 ("[D]ata portability is not necessarily beneficial for users because [content providers] entering the market have an incentive to increase the amount of data users have to reveal.").

108. *See supra* note 18.

109. *See, e.g.*, James Grimmelmann, *Saving Facebook*, 94 IOWA L. REV. 1137, 1194–95 (2009) ("As social-network-site data becomes more portable, it also becomes less secure — and thus less private."); *see also* Diker Vanberg & Ünver, *supra* note 5, at 6.

that data portability will not effectively solve data-based concerns in digital markets.

Table 2: Misalignment of Objectives and Real Functions of Data Portability

Objectives	Sources of Lock-In Effects		Data Subject to Portability	Results	Competition Value
User-Based Concerns (Section III.A)	Non-data-based features (III.A)		N/A	Not applicable (IV.A.1)	Depends
	Data-based features	MISCs (*e.g.*, Network Effects) (III.A.1)	Secondary Data (Inferred/ Derived Data)	Rarely applicable (IV.A.2)	High
		ARSCs (III.A.2)	Provided Data (incl. Observed Data)	Applicable (IV.A.2)	Low
Data-Based Concerns (Section III.B)	Essential facility (III.B)		All	Insufficient (IV.A.3)	Limited

B. Recalibration

The objective of data portability should be recalibrated to its real functions. As discussed previously, data portability has two clusters of objectives: individualistic objectives and market-based objectives.[110] Under the market-based objectives, since users lack sufficient motivation to initiate large-scale data porting, the main function of data portability lies in removing data-based ARSCs.[111] However, this function can be objectively achieved under the individualistic goals of data portability. This is primarily because data provided by users during the use of digital services will in turn create ARSCs and lock users into that system when data portability is unavailable. After export and import features are enabled, users will be able to transfer the provided data when switching to another system.

However, users are less likely to be locked in by other data, such as inferred and derived data. They are therefore less likely to port those data when they switch platforms. In other words, data portability under individualistic goals is sufficient to achieve possible market-related goals — that is, to reduce ARSCs. For instance, browser users can exercise their data autonomy to port stored passwords and bookmarks that they have created to a new browser platform. Simultaneously, this

110. *See* discussion *supra* Sections II.A, II.B.
111. *See* discussion *supra* Section IV.A.

process also removes the user lock-in effects. This means that broadening the applicable scale of data portability to inferred and derived data is meaningless — a more extensive portability rule will not achieve a marginally greater outcome since, as illustrated in the first Section of this Part, users have less incentive to port secondary data.[112]

But that does not mean that these secondary data should not be publicly accessible. Researchers have paid more attention to the positive effects of free data sharing and have found that it has important implications for public welfare.[113] Given that data portability is not an appropriate tool to execute the large-scale mandating of data sharing, these secondary data, under conditions of privacy and security, might be compulsorily shared and accessed in other data governance regimes.[114] Hence, data portability rules should be appropriately tailored to coordinate with other policy arrangements, such as interoperability and pooling,[115] to tackle the lock-in problems more holistically and further enhance consumer welfare. The core difference between data pooling and data portability is that data pooling binds the relationship only among businesses, without the direct involvement of individual users.[116] When markets operate effectively with the aid of interoperability and pooling, data sharing will be achieved autonomously by voluntary agreement.[117]

---

112. *See* discussion *supra* Section IV.A.

113. *See, e.g.*, Lothar Determann, *California Privacy Law Vectors for Data Disclosures* 13–14 (U.C. Hastings Research Paper Forthcoming, 2022) (manuscript at 13–14) (on file with author) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4146903 [https://perma.cc/ZHS7-P2SY] ("Data collection, usage and sharing will increase, in fact: must increase, to better research and cure diseases; treat patients with personalized, precision medicine; develop artificial intelligence; enable autonomous cars to recognize and protect people; support global communications; create reliable block-chains; and protect national and international security."); *see also* MAYER-SCHÖNBERGER & RAMGE, *supra* note 1, at 125 ("Such a mandate for data sharing, as previously set out, would not only break up the power asymmetries between Big Tech monopolists and users freely handing over their data. It could also help an even grander vision turn into reality: open data.").

114. As for the secondary data that enhances the network effects, data portability can be incentivized but not mandated. *See* Alexander Macgillivray, *Summary of Comments Received Regarding Data Portability*, THE WHITE HOUSE: PRESIDENT BARACK OBAMA (Jan. 10, 2017, 9:19 AM), https://obamawhitehouse.archives.gov/blog/2017/01/10/summary-comments-received-regarding-data-portability [https://perma.cc/Z7YE-63HE] (stating that many commenters prefer incentivizing portability over mandating it).

115. "Interoperability" refers to a situation where "technology or systems of multiple firms are linked in a way that permits users to process instructions for all of them simultaneously," and "pooling" refers to compulsory data sharing. Compared to data portability, "interoperability" and "pooling" directly regulate the business-to-business relationship. Herbert Hovenkamp, *Antitrust and Platform Monopoly*, 130 YALE L.J. 1952, 2032–37 (2021); *see also* Wynne Lam & Liu, *supra* note 43, at 13 (footnote omitted) (arguing that "forced data sharing may be helpful in allowing potential entrants to have sufficient access to data to train their algorithm . . . instead of relying on consumers' motive to switch or port data").

116. *See* Hovenkamp, *supra* note 115, at 2035 ("[P]ooling as an antitrust remedy would place the data into a common database equally accessible by all participating search firms, subject to user rights to withhold.").

117. *See id.* at 2033.

V. CONCLUSION

There has been a rising clamor for legislation to expand data portability. In addition to the objective of enhancing platform users' data autonomy, many argue that data portability would rejuvenate competition and innovation in digital markets by liberating users and their data from being locked into single systems. Based on this ambitious belief, there is a growing propensity to attempt to enlarge the scope of data subject to data portability.

However, as this Note has demonstrated, the market functions of data portability have been overestimated. Those who are adamant that the scope of data subject to data portability should be broadened to fulfill its market objective seemingly fail to realize that it is users, not legislators, who have the final say on the volume and the categories of data that are ported. If users do not have strong incentives to port data that are functionally essential to the digital market competition, the logical chain to invigorate the competitive environment in digital markets by enhancing users' data autonomy will naturally break down. Therefore, this Note has put the academic dispute about data and platforms to one side and instead has sought to shed light on this issue from the perspective of users.

It is also important to recognize that users are not, in fact, locked in by all categories of data. Instead, there is a sliding scale — the less that users are locked in by data, the lower the likelihood they will port their data. Therefore, compared to inferred and derived data, users have stronger incentives to port provided data that will otherwise cause ARSCs. The paradox is that the data which are ported are relatively small in volume and are less essential to digital market competition. In contrast, the inferred and derived data that users are less likely to port are arguably more significant to realizing the market goals.

This Note has illustrated this paradox by analyzing three sources of user lock-in effects. First, data portability is incapable of solving lock-in effects by non-data-based features. Second, data portability is rarely suitable to mitigate the lock-in effects generated by MISCs. Third, data portability has the potential to remove the data-based ARSCs. Thus, data portability can only solve highly limited data lock-in problems. This limited function can be achieved objectively by the individualistic goals of data portability. A broader scope of data portability is unnecessary and meaningless because relying on data portability to reinvigorate competition is merely the wishful thinking of legislators. In reality, it will not lead to any major changes.

This Note is both predictive and heuristic. It provides a warning to policymakers that they should not expect data portability to have the same functionalities as mandatory data sharing. Given the fact that the EU is broadening the applicable scope of data portability, comparative

empirical research should be conducted between the EU, the United States, and other jurisdictions. While this Note looks forward to the emergence of more statistics to support its key arguments, it welcomes any normative criticisms and opposing empirical observations.