Taking It With You: Platform Barriers to Entry and the Limits of Data Portability

Gabriel Nicholas

Abstract

Policymakers are faced with a vexing problem: how to increase competition in a tech sector dominated by a few giants. One answer proposed and adopted by regulators in the US and abroad is to require large platforms to allow consumers to move their data from one platform to another. Facebook, Google, Apple, and other major tech companies have enthusiastically supported this form of regulation, known as data portability, through their own technical and political initiatives. Today, data portability has taken hold as one of the go-to solutions to address the tech industry’s competition concerns.

This Article argues that despite the regulatory and industry alliance around data portability, today’s public and private data portability efforts are unlikely to meaningfully improve competition because they focus solely on mitigating switching costs, ignoring other barriers to entry that may preclude new platforms from entering the market. The technical implementations of data portability encouraged by existing regulation — namely one-off exports and API interoperability — address switching costs but not the barriers of network effects, unique data access, and economies of scale. This Article proposes a new approach to better alleviate these other barriers called collective portability, which would allow groups of users to coordinate to transfer data they share to a new platform, all at once. Although not a panacea, collective portability would provide a meaningful alternative to existing approaches and avoid the privacy/competitive utility trade off of one-off exports along with the hard-to-regulate power dynamics of API interoperability.

1 Gabriel Nicholas is a Joint Research Fellow at New York University School of Law Information Law Institute and the New York University Center for Cybersecurity, and a Fellow at the Engelberg Center on Innovation Law & Policy. Thanks to Katherine Strandburg, Aaron Shapiro, Mark Verstraete, Salome Viljoen, Randy Milch, Carrie Brown, and Lucas Daniel Cuatrecasas for their sage guidance. Special thanks to Michael Weinberg for introducing me to this topic.
Introduction

On March 30, 2019, Mark Zuckerberg released his op-ed, “Four Ideas to Regulate the Internet”.2 His first three ideas regarded speech, political advertisements, and privacy. The fourth addressed concerns that Facebook and its ilk had grown too large for any other platform to compete:3

“Finally, regulation should guarantee the principle of data portability. If you share data with one service, you should be able to move it to another. This gives people choice and enables developers to innovate and compete.”4

Zuckerberg was signaling a strategic shift in Facebook’s public policy strategy. Instead of fighting to keep Facebook data within the walls of the platform,5 Facebook would allow for data portability, the ability for users to move their data from one service to another. It was an instrument of self-regulation that responded directly to mounting calls for antitrust action.6 In the same way that the free movement of capital allows for competitive markets, so too would the free movement of data allow for competitive platforms, at least in theory.7 Facebook was not alone in adopting this approach — Apple, Alphabet, Microsoft, and Twitter all had data portability as part of their competition policy efforts.8 Data portability was also a central tenet of the

---

2 See generally Mark Zuckerberg, The Internet Needs New Rules. Let’s Start in These Four Areas, Wash. Post (Mar. 29, 2019), https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html.
3 See generally, e.g., Lina Khan, Amazon’s Antitrust Paradox, 126 Yale L.J. 710 (2017); Robert Reich, Break up Facebook (and while we’re at it, Google, Apple and Amazon), Guardian (Nov. 20, 2018), https://www.theguardian.com/commentisfree/2018/nov/20/facebook-google-antitrust-laws-gilded-age; Joe Nocera, Easiest Fix for Facebook: Break It Up, Bloomberg: Bloomberg Opinion (Nov. 21, 2018), https://www.bloomberg.com/opinion/articles/2018-11-21/fix-facebook-by-breaking-it-up-under-antitrust-regulation.
4 See Zuckerberg, supra note 2.
5 See generally Mark Andrejevic, Privacy, Exploitation, and the Digital Enclosure, 1 Amsterdam L.F. 47 (2009); Kevin Bankston, How We Can ‘Free’ Our Facebook Friends, New America (June 28, 2018), https://www.newamerica.org/weekly/edition-211/how-we-can-free-our-facebook-friends/.
6 Recent antitrust stirrings indicate that the government is interested in increasing competition between tech companies, particularly Alphabet, Amazon, Facebook, and Apple, who are all facing antitrust scrutiny by the Justice Department and other federal and state organizations. See “16 Ways Facebook, Google, Apple and Amazon Are in Government Cross Hairs”, N.Y. Times (Sept. 9, 2019), https://www.nytimes.com/interactive/2019/technology/tech-investigations.html.
7 See generally Victor Mayer-Schönberger & Thomas Ramge, Reinventing Capitalism in the Age of Big Data (2018); Peter Swire, Markets, Self-Regulation, and Government Enforcement in the Protection of Personal Information, in Privacy and Self-Regulation in the Information Age, U.S. Department of Commerce (1997), https://papers.ssrn.com/abstract=11477.
8 The Data Transfer Project is a joint framework intended to make it easier for companies to let users port their data between. It is an open source project being led by Apple, Facebook, Google, Microsoft, and Twitter. See generally Data Transfer Project Overview and Fundamentals, Data Transfer Project, (July 20, 2018), https://datatransferproject.dev/dtp-overview.pdf.
ACCESS Act,⁹ proposed by Senators Warner, Hawley, and Blumenthal that same year to improve competition in the tech sector.¹⁰

Data portability laws already exist in California, the European Union, and Singapore,¹¹ but they are not exclusively aimed at improving tech sector competition. Portability in the California Consumer Privacy Act (CCPA) for example is mostly meant to let consumers hold platforms accountable for how they track and sell personal data.¹² It says little about making data easily transferable to other platforms and does not require data to be sent in a machine-readable format.¹³

The right to data portability in Europe’s General Data Protection Regulation (GDPR)¹⁴ has the dual purposes of giving consumers greater control over their own data¹⁵ and encouraging competition between platforms.¹⁶ It requires data controllers to allow consumers to export their personal data in a machine-readable format, and where technically feasible, directly transmit data to the new controller.¹⁷ European legal scholars argue that this improves competition by lowering the cost for users to change platforms.¹⁸ In some cases though, the aims of data control and competition

---

⁹ Augmenting Compatibility and Competition by Enabling Service Switching Act of 2019, S. 2658, 116th Cong. (2019).
¹⁰ See also Consumer Online Privacy Rights Act, S. 2986, 116th Cong. § 105(a) (2019) (providing for a right to data portability stronger than that in the CCPA because it requires data to be made available in a computer-readable format, but weaker than the right provided in the GDPR because it does not require data controllers to transfer the data to a new entity upon request).
¹¹ Data portability laws have either passed or been proposed in other countries as well, including Australia, Kenya, and India.
¹² The CCPA gives consumers rights to access all personal data a platform has collected on them, delete that information, and access certain knowledge about how their personal data gets sold, including what data has been sold and what category of business has bought it. Users also have the right to prevent buyers from reselling it without permission. Cal. Civ. Code §§ 1798.105, .110, .115.
¹³ The CCPA only requires platforms to send twelve months worth of user data and has few restrictions about how data gets sent, explicitly permitting data to be shared through email or a toll-free number. See Cal. Civ. Code § 1798.130.
¹⁴ Regulation 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 Advancement of Such Data, and repealing Directive 95/46/EC, art. 20, 2016 O.J. L 119/1 [hereinafter GDPR].
¹⁵ See GDPR, supra note 14, at pmbl. 68 (“To further strengthen the control over his or her own data, where the processing of personal data is carried out by automated means, the data subject should also be allowed to receive personal data concerning him or her which he or she has provided to a controller...”).
¹⁶ “Indeed, the primary aim of data portability is to facilitate switching from one service provider to another, thus enhancing competition between services.” Article 29 Working Party, Guidelines on the Right to Data Portability, WP 242 at 4 (Dec. 13, 2016). See generally Helena Ursic, Unfolding the New-Born Right to Data Portability: Four Gateways to Data Subject Control, 15 SCRIPT-ed (2018).
¹⁷ GDPR, supra note 14, at art. 20.
¹⁸ See generally Inge Graef, Blurring Boundaries of Consumer Welfare: How to Create Synergies Between Competition, Consumer and Data Protection Law in Digital Markets, in Personal Data in Competition, Consumer Protection and Intellectual Property Law (Mor Bakhoum et al., eds.,...
come into direct conflict. For example, if a data subject tries to port data that overlaps with someone else’s, the GDPR says that it is more important to respect the privacy and erasure rights of the second data subject than portability rights of the first.

The Lei Geral de Proteção de Dados (LGPD), Brazil’s data protection law, is similar to the GDPR and also gives data subjects the right to port their data from one data controller to another. However, unlike the GDPR which requires the individual rights of the data subject to be considered before the “legitimate interests” of the data controller, the LGPD allows controllers to weigh the two considerations against one another. It is yet to be determined what this balance means for the competitive utility of Brazilian data portability since as of this writing, the law has not gone into effect.

None of these data portability laws attempt to address platform barriers to entry beyond switching costs. These barriers to entry, here meaning the structural economic disadvantages faced by some firms but not others, may still preclude new platforms from entering the market. A platform that can overcome an incumbent’s switching costs may still fail to overcome its network effects or economies of scale, for example.

In this paper, I argue that the technical approaches to data portability, as encouraged by existing regulation, fail to improve barriers to entry sufficiently to allow more new platforms to enter the market. As an alternative to the current approaches, namely one-off export portability and API interoperability, I recommend a new approach to better address these barriers called collective portability, which allows groups of users to coordinate in order to transfer data they share to a new platform. This not

2018); Damien Geradin & Monika Kuschewsky, Competition Law and Personal Data: Preliminary Thoughts on a Complex Issue (2013), https://ssrn.com/abstract=2216088.

19 See generally Tal Zarsky, The Privacy-Innovation Conundrum, 19 Lew & Clark L. Rev. 115 (2015).

20 See generally GDPR, supra note 14.

21 Abigayle Erickson, Note, Comparative Analysis of the EU’s GDPR and Brazil’s LGPD: Enforcement Challenges with the LGPD 44 Brooklyn J. Int’l L. 862, 880 (2019); Lei No. 13.709, de 14 de Agosto de 2018, Diário Oficial Da União [D.O.U.] de 15.8.2018 (Braz.), translated in Ronaldo Lemos et al., Pereira Neto, https://iapp.org/media/pdf/resource_center/Brazilian_General_Data_PROtection_Law.pdf (last visited June 7, 2020) [hereinafter LGPD].

22 See Steven Blickensderfer and Joseph W. Swanson, Brazil’s New Data Protection Law: An Overview and Four Key Takeaways for U.S. Companies, National Law Review (May 2, 2019), https://www.natlawreview.com/article/brazil-s-new-data-protection-law-overview-and-four-key-takeaways-us-companies; What is the LGPD? Brazil’s version of the GDPR, Proton Technologies AG, https://gdpr.eu/gdpr-vs-lgpd/.

23 See Peter Swire & Yianni Lagos, Why the Right to Data Portability Likely Reduces Consumer Welfare 72 Md. L. Rev. 338, 349 (2013).

24 See, e.g., Organisation for Economic Co-operation and Development, Glossary of Industrial Organisation Economics and Competition Law, 13–14 (1990), http://www.oecd.org/regreform/sectors/2376087.pdf; Yafit Lev-Aretz and Katherine J. Strandburg, Regulation and Innovation: Approaching Market Failure from Both Sides 4 (Pub. Law & Legal Theory Research Paper Series, Working Paper No. 19-48, 2019).
only helps new market entrants overcome switching costs, but also the barriers of unique data access, economies of scale, and network effects.

In addition to evaluating and improving upon data portability’s viability as a tool of competition regulation, this paper seeks to contribute to two broader bodies of literature. The first is the literature on data as an economic asset. Scholars often discuss the economic advantages and normative drawbacks of data-rich technical systems, but less often discuss the contexts in which data-based advantages fail to be realized. To the extent this conversation is had, it is mostly through analogies, such as debating whether or not data is the new oil. Through my analysis of data portability, I demonstrate that data is not a fungible asset, with value in isolation. Rather, the value of data depends on its context and the transferability of that value depends on the capacity to maintain that context in the move.

Second, this paper seeks to contribute to the larger literature on regulating tech companies. As scholars and policymakers consider how to manage platform bigness, data portability will be an important tool in the regulatory tool belt. But it is not a panacea, and what works as a matter of policy may fail in technical implementation. This paper will look at various real world examples of data portability that are well-known in Silicon Valley but not often analyzed in the legal scholarship. They will hopefully serve as a good prism through which to interpret existing scholarship and bolster new scholarship.

I build my argument on three sources of evidence. The first is theoretical work done on barriers to entry from the fields of law, economics, and information systems. Second is anecdotes from real data portability regimes and analogous efforts, mostly from real tech companies, large and small. However, despite the insistence that data portability improves competitiveness in the tech sector, I could not find any products built off of data specifically from one-off exports (despite the fact that Facebook has

---

25 See Lauren Henry Scholz, *Big Data is Not Big Oil: The Role of Analogy in the Law of New Technologies*, 86.863 Tenn. L. Rev. (forthcoming 2020), https://ssrn.com/abstract=3252543. But see Jonathan Vanian, *Why Data Is the New Oil*, Fortune (July 11, 2016), https://fortune.com/2016/07/11/data-oil-brainstorm-tech/; Joris Toonders, *Data is the New Oil of the Digital Economy*, WIRED (July 2014), http://www.wired.com/2014/07/data-new-oil-digital-economy/.

26 It is technically possible for a market entrant to build a new product off of a one-off export since they are computer readable, but in practice they are difficult if not impossible for developers to integrate into new or existing products. Data senders tend to provide little to no documentation about the structure of data users get from an export, leaving recipients to reverse engineer it themselves. Of the seven data portability exports studied for this Article, six had no documentation. The seventh, Twitter, seems to use something similar to its own documented, external facing API, though this is not confirmed anywhere. Also, the Twitter data set was the only one in a cleanly formatted JSON. None of them had any versioning, which meant that these platforms could introduce breaking changes at any time. See Gabriel Nicholas & Michael Weinberg, *Data Portability and Platform Competition: Is User Data Exported From Facebook Actually Useful to Competitors?*, Engelberg Center on Innovation Law and Policy (November, 2019).
allowed users to download their data since 2010, and Google since 2011. Therefore, for this part of the analysis, I turn to analogous initiatives in the telecom and banking sectors. The third source of evidence is real ported data. I downloaded and analyzed my own personal data from seven different platforms — Facebook, Google, Apple’s iCloud, Instagram, Snapchat, Spotify, and Twitter—to get a sense of how platforms could and could not actually leverage it to compete. My analysis is sprinkled throughout, though much of it is focused in Part IV with my analysis of data from Facebook’s Download Your Information tool.

Part I of this paper defines “data portability” and delineates between two approaches, one-off exports and APIs. Part II assesses how effectively each of these approaches helps competitors overcome four different barriers to entry raised by big data: switching costs, network effects, economies of scale, and unique data access. Part III highlights additional limitations of API portability, based on the market behavior of users and platforms, and opportunities, based on its innovative potential. I then propose my new approach in Part IV, collective portability. I look at it through the lens of Facebook’s Download Your Information tool, and then explore questions, hypotheticals, and challenges to the approach. I conclude with recommendations for regulators.

I. Current approaches to data portability

The International Organization for Standardization defines “data portability” as the “ability to easily transfer data from one system to another without being required to re-enter data”. Notably, this definition encompasses any means of transfer that could be considered ‘easy’ and any type of data. Legal regimes that give users a right to portability get more specific on these fronts, particularly the latter. The GDPR focuses its definition of data portability on personal data — Article 20 gives Europeans the rights to obtain “personal data concerning [a data subject] which he or she has provided”. It does not include information a data controller has inferred from provided personal data or any non-personal data, since that data is not considered ‘provided’ by the data subject (Subsequently, this paper will largely focus on personal-information-based products and services, since these are the ones already subject to regulation.)

Data portability also does not give a data subject access to someone else’s personal data, even if it overlaps with their own. This is particularly relevant in the context of

27 Alexia Tsotsis, Facebook Now Allows You To ‘Download Your Information’, TechCrunch (Oct. 6, 2010), https://techcrunch.com/2010/10/06/facebook-now-allows-you-to-download-your-information/.
28 See Data Liberation Blog, The Data Liberation Front Delivers Google Takeout, (June 28, 2011), http://dataliberation.blogspot.com/2011/06/data-liberation-front-delivers-google.html.
29 Int’l Standards Org., ISO Standard No. 18101-1-2019:: Automation systems and integration — Oil and gas interoperability, 3.23 (2019), https://www.iso.org/obp/ui/#iso:std:iso:ts:18101:1-2019::en:term:3.23 [hereinafter ISO].
30 Forms of transfer that would not be considered easy include manually data entry, an external tool, or having a specialist do it. See infra Part II.A.
31 GDPR, supra, note 14, art. 20.
deletion laws, like the GDPR’s “right to be forgotten”. If Alice is tagged in a photo that Bob took and uploaded, for example, portability regimes under the GDPR may allow Bob to port that photo but not Alice, since the photo is Bob’s data and someday he may want to delete it, which he could not do if Alice had it.

By and large though, platforms are autonomous in determining which data they make portable. A helpful way to conceptualize this is by imagining a platform as a circle-and-line graph of entities and connections, or points and edges. A point can represent a person or any discrete piece of content, like a message or a video. An edge represents a relationship between two points, like between two people (e.g. friends on Facebook), a person and content (e.g. liking a photo) or two pieces of content (e.g. a comment on a photo.) A specific data portability regime can be seen as which edges and points an individual user is allowed to export/transfer, and in what context. A regime may not make all edges portable, particularly if it is unclear which user should have control over it, and data may fall through the gaps.

There are two main types of portability, which differ in how they allow data to move. The first is one-off export portability, whereby users can download a snapshot of the data they have on one platform in a form that can be uploaded to another. One-off exports do not require an ongoing relationship between the platform data is being transferred away from (the data sender) and the platform data is being transferred into (the data receiver). Often, the user acts as an intermediary between the two platforms, moving data from one to the other, meaning the data sender and receiver never have a direct relationship. This allows the user to delete their data from the sending platform, if they so desire. In practice, one-off exports often take a long time to be processed.

---

32 GDPR, supra note 14, at pmbl. 65.  
33 When exactly the “right to be forgotten” applies to social media data has somewhat yet to be determined. Many cases fall between the gaps of the extreme examples given by the European Commission. See generally European Data Commission, Do we always have to delete personal data if a person asks?, https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/dealing-citizens/do-we-always-have-delete-personal-data-if-person-asks_en. The European Data Protection Board has begun to fill in some of these gaps by differentiating between “delisting” (from search engines) and “full erasure”. See Guidelines 5/2019 on the criteria of the Right to be Forgotten in the search engines cases under the GDPR, European Data Protection Board (Dec. 2, 2019), https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_201905_rtbsearch_engines_forpublicconsultation.pdf.  
34 See generally Alan Gibbons, Algorithmic Graph Theory (1985).  
35 The terminology can be confusing, because some refer to one-off export portability as “data portability” and API portability as “interoperability”. See generally Oscar Borgogno & Giuseppe Colengelo, Data Sharing and Interoperability: Fostering Innovation and Competition Through APIs, 35 Computer L. & Security Rev. (2019); Paul De Hert et al., The Right to Data Portability in the GDPR: Towards User-Centric Interoperability of Digital Services, 34 Computer L. & Security Rev. 193 (2018). I use the terms “one-off export portability” and “API portability” to avoid confusion and remind the reader that both, by the ISO definition, are forms of data portability.  
36 Erin Egan, Data Portability and Privacy, Facebook Newsroom 11 (Sept. 2019), https://about.fb.com/news/2019/09/privacy-and-data-portability/ [hereinafter Facebook Portability White Paper]. Refers to this relationship as an “open transfer.”
for data controllers to prepare — minutes to hours on some platforms,37 days to weeks on others38 — because the data sender may have to move a lot of data at once that is not stored in an easily accessible way.39

The second type of data portability is interoperability, whereby users allow two or more platforms to exchange information directly with one another. The ongoing connection between systems lets live, up to date information be transferred, often through defined and documented protocols called APIs (application programming interface). To avoid confusion with other definitions of “interoperability,”40 I will refer to this approach as API portability.41

In contrast with one-off exports, APIs require the data sender and data receiver to maintain an ongoing relationship.42 This lets the receiver get data faster and more often. It also lets the data sender observe and control what, when and how the receiver gets data, usually through rate limiting (limiting in how frequently they can make API requests) and key registration (having the recipient register with the sender to access the API, thus letting the sender see how they are using the API and if necessary, cut off access.) There are reasons outside of competition for data senders to retain this kind of control, such as to improve security, prevent spam, and keep their costs down, but these mechanisms have also been used for anti-competitive purposes, as will be discussed later.43

To understand the difference between one-off exports and APIs, imagine that Carlo keeps track of his recipes in an application called CookUp but wants to move to a new recipe app called Tazte. If CookUp allows for one-off export portability, Carlo can download his recipes, upload them to Tazte, and, if he wants, delete that data from CookUp. The export is just a snapshot so if Carlo adds a recipe to CookUp it would not also appear on Tazte. However, if CookUp gives API access to Tazte, perhaps through a “Sign In With CookUp” button, the two systems can directly connect and exchange live data with each other. Thus, if Carlo adds or deletes a recipe to CookUp, the two lists could potentially stay in sync.

37 E.g., Facebook, Twitter.
38 E.g., Spotify, Snapchat.
39 The databases used to store this data are often optimized for some operations over others. Since all platforms explored here were originally developed after portability laws existed, it is unlikely that their databases are optimized for downloading all of a user’s data.
40 “Interoperability” is often seen as synonymous with “compatibility”. In this context, the ISO definition makes the most sense: the “ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged.” Int’l Standards Org., & Int’l Electrotechnical Comm’n., ISO/IEC Standard No. 19941:2017Information technology — Cloud computing — Interoperability and portability, International Standards Organization 3.1.1 (2017), https://www.iso.org/obp/ui/#iso:std:iso-iec:19941:ed-1:v1:en.
41 APIs can be used for more than moving data around. They can also allow applications to expose features of their system (e.g. database storage, algorithms.) In this paper, I only consider the strict portability use case.
42 Facebook Portability White Paper, supra note 36, at 12. Facebook calls this a "partnership transfer."
43 See infra Part IIIA.
One-off export and API portability are not strictly dichotomous. For example, a data sender and data recipient may be in direct communication for a brief period of time and then cut off from one another. Or a data sender may only let users transfer their data to a platform if certain conditions are met, such as only allowing health data to be sent to another HIPAA-compliant platform.44

II. Data portability and barriers to entry

Competition problems in the tech sector differ meaningfully from competition problems in other sectors because technology firms often have business models that depend on and leverage the advantages of big data.45 Scholars argue that disparities in data held by technology firms can create barriers to entry that make it difficult for non-dominant platforms to compete.46 In this section, I review four of these barriers: switching costs, unique data access, economies of scale, and network effects. I have put them in order from best to illest suited to being alleviated by current approaches to data portability. These four barriers will serve as an operational definition of competition47 and be used to evaluate each different approach to data portability discussed.

A. Switching costs

If a user has a lot of their data on one platform, there may be a significant cost for them to switch to another platform. For example, if a user has logged hundreds of books they have read on Amazon’s Goodreads platform, they may not be willing to switch to a new, more preferable platform if it meant manually re-entering every book they read. In this way, lock-in and switching costs can act as barriers to entry.48

One-off data exports are naturally suited to lower switching costs — data portability is by definition the “ability to easily transfer data,”49 the opposite of lock-in. A one-off export is able to lower switching costs in the Goodreads example because the user values their data (reading history) for its own sake, not just in the context of other data on the Goodreads platform that could not be exported (e.g. comments from other

---

44 Facebook Portability White Paper, supra note 39, at 12. Facebook calls this a “conditioned transfer.”
45 I define this term with the four V's of big data: volume, velocity, variety, veracity. In other words, “big data” sets are large, frequently updated, gathered from various sources, and reasonably accurate.
46 See e.g. IBM Big Data & Analytics Hub, The Four V's of Big Data, Int'l Bus. Mach. Corp. (2015), www.ibm.com/bigdatahub/infographic/four-vs-big-data. Different sources expand or contract this mnemonic, ranging between three and six V's. See generally Matt Griffin et al., A Case Study: Analyzing City Vitality with Four Pillars of Activity—Live, Work, Shop, and Play, 4 Big Data, https://www.liebertpub.com/doi/full/10.1089/big.2015.0043
47 Competition can be an elusive term so it is easier to evaluate by taking on a non-comprehensive but also non-recursive operationalization. See generally Paul McNulty, Economic Theory and the Meaning of Competition, 82 The Quarterly Journal of Economics 639 (1968).
48 Daniel Rubinfeld & Michal Gal, Access Barriers to Big Data, 59 Ariz. L. Rev. 339, 364 (2017).
49 ISO, supra note 30.
users.) One-off exports work similarly well for other types of data that are difficult to move and users value in and of themselves, like photos, videos, and blog posts.\footnote{Facebook has already taken this approach with its port to Google Photos project. \textit{See} Steve Satterfield, \textit{Driving Innovation in Data Portability With a New Photo Transfer Tool}, Facebook Newsroom (Dec. 2, 2019), \url{https://about.fb.com/news/2019/12/data-portability-photo-transfer-tool/}.}

There is historical precedent for one-off exports\footnote{See generally Minjung Park, \textit{The Economic Impact of Wireless Number Portability}, 59 Journal of Industrial Economics 714 (2011). Mobile phone number portability is an example of a one-off export, not interoperability. When transferring a number, the sender and receiver do need to communicate with one another once the relevant data, originally allocated by the North American Numbering Plan Administrator, is sent. The FCC also mandates this process happens in a timely manner, and it usually takes about two hours. \textit{See generally} Federal Communications Commissions, CC Docket No. 95-116, Memorandum Opinion and Order and Further Notice of Proposed Rulemaking In the Matter of Telephone Number Portability, FCC 03-284, (Nov. 10, 2003). This was actually only in the top 100 Metropolitan Statistical Areas. It only applied to all markets (despite a Petition for Forbearance filed by a number of telecoms) starting May 24, 2004. \textit{Id.} at 4.} working to improve competition in the telecom sector. On November 24, 2003, the Federal Communications Commission (FCC) mandated mobile phone number portability, requiring all wireless carriers to allow their customers to move their phone number to another carrier, if they so desired.\footnote{\textit{Id.} at 4.} Previously, if a customer wanted to change carriers, they would also have to abandon their existing phone number. By removing this switching cost, the FCC hoped to increase competition between carriers.\footnote{For example, unlimited anytime minute plans became more similarly priced between AT&T and Verizon. Park, \textit{supra} note 54, at 734-36.} Looking at cost as a proxy for competition, the mobile number portability can be seen as an across the board success. Small, medium, and large phone plans all became cheaper, and price dispersion between companies offering analogous plans decreased.\footnote{\textit{See infra} Part III.B.}

Well designed APIs can drive switching costs even lower than one-off exports. Streamlined design elements like the “Sign in with Facebook” button can make for a more seamless transfer experience than downloading and uploading an export file. APIs also allow users to change services without abandoning the original, further reducing switching costs but also reducing the cost of going back to the incumbent.\footnote{\textit{See infra} Part III.B.}

It is important to note though that portability does not inherently make data flow from dominant incumbents to competitors. The telecom example sidesteps this dynamic because it is an industry mostly made up of a few dominant players. Tech though has a few large players and many small ones, and the hope of any competition effort is to make it easier for the two groups to better compete with each other. Reducing switching costs might actually \textit{harm} smaller competitors by helping incumbent platforms with strong network effects to pull in users.

This dynamic plays out on customer relationship management platforms (CRM). A small startup may opt for a smaller, cheaper CRM, such as Pipedrive, that competes...
with the industry-standard CRM, Salesforce, on price.\textsuperscript{56} Once the startup grows, they will have more clients who use Salesforce and will want to better integrate with their systems. The startup is incentivized to switch to the strongly networked Salesforce as soon as they are able to. Currently though, that can take a while because startups need to either buy a third-party tool or pay a specialist to make the switch.\textsuperscript{57} High switching costs could cause startups to switch away from small CRMs later than they otherwise would, making it even more difficult for them to compete with the network effects of the dominant platform.

**B. Unique data access**

Critics who argue that big data does not act as a barrier to entry point out that data is a non-rivalrous, non-exclusive resource.\textsuperscript{58} However, there are cases where a platform may have unique, functionally exclusive access to certain data. An IoT thermometer company may have unique access to data concerning a patient’s body temperature, or an oil company may have unique access to geothermal data, for example.

Data portability to some extent can mitigate this advantage. If an IoT thermometer company makes its data portable, a health tracking app can use that data to give its users a better sense of their overall health. However, the data sender will likely have more context about how the data was collected since they have more knowledge of their own internal systems. This may translate into knowledge about which data is more trustworthy. Additionally, the data sender may withhold some of the data it collects. They may not want to overload the recipient with excessive data, or they may want it for their own advantage, among other possible reasons. The IoT thermometer company may choose not to expose its internal diagnostic data because it is unnecessary, or because it undermines the confidence of the temperature taken, or because it undermines the user’s privacy.

The UK Department for Business Innovation & Skills is attempting to lower the barriers of uniqueness in the financial tech sector through a portability initiative called midata, which makes it easier for consumers to access information about their transactions across various industries. The goal is for consumers to be able to compare the prices of different products using their existing data.\textsuperscript{59} The initiative first focused on the banking industry — specifically, the government wanted to let consumers download their historical bank transaction data in a standardized format so they could use price comparison tools to see whether they could save by moving to another bank. By opening up this information that banks might otherwise have

\textsuperscript{56} Interview with Berk Birand, CEO of Fero Labs, in N.Y. City, N.Y. (Nov. 21, 2019) [hereinafter Birand Interview].
\textsuperscript{57} Id.; See generally \textit{Import2 Wizard}, \url{https://www.import2.com/}.
\textsuperscript{58} See generally Anja Lambrecht & Catherine Tucker, \textit{Can Big Data Protect a Firm from Competition?} (2015), \url{https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2705530}; Darren Tucker & Hill Wellford, \textit{Big Mistakes Regarding Big Data}, American Bar Association: Antitrust Source (Dec. 2014), \url{https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2549044}.
\textsuperscript{59} See generally U.K. Dep’t for Bus. Innovation & Skills, \textit{Review of the Midata Voluntary Programme}, (2014), \url{https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/327845/bis-14-941-review-of-the-midata-voluntary-programme-revision-1.pdf}. 
unique access to, the government hoped that customers would switch bank accounts more frequently, the way they do credit cards, thus improving competition.60

The midata initiative, which rolled out in 2014 and became compulsory in 2015, so far seems to be neither a significant success nor failure. The only major company61 that offers the hoped-for price comparison service is GoCompare, and their tool is not without criticism.62 The UK Department for Business Innovation & Skills has not released data on whether bank account switching has increased.

However, by standardizing the format the data must be made accessible in though, the initiative leaves itself open to future success. Previously, most UK banks let users download their data in some form or another, but they varied greatly in how far back they went in the transaction history (sometimes months, sometimes years), and often released the data in a non-machine readable format, like PDF. midata also let the government set unified privacy standards for banking data.63 And even if the data is underutilized now, standardization allows this data to potentially be used by new products in the future. The UK briefly ran the midata Innovation Lab to explore some of these potential uses for open data in banking and other industries.64

Sometimes, platforms can create a barrier to entry by having uniquely quick access to data.65 Stucke and Grunes discuss this in terms of “nowcasting”,66 the ability for a company to use the velocity of data they receive to identify trends before others. For example, Lyft’s dynamic pricing works by increasing the price when it notices certain patterns of orders occurring in an area (for example, people trickling out of a concert early before a massive rush.)67 The combination of timely data and broad historical data allows Lyft to more quickly and accurately adjust their price to the customer’s willingness to pay than a rival app with less users. Twitter likely uses a similar process to identify trending topics faster than other social media sites, thus further perpetuating itself as the place to go for the most up to date conversation.

60 See generally Id.
61 Numerous business leaders have praised the midata initiative (e.g. Martin Lewis from MoneySavingExpert, Alistair Crane from Grapple) but have yet to create their own comparison tools, so far as I can tell. See Press Release, U.K. Dep't for Bus., Innovation, & Skills, Businesses Get Creative with Consumer Data at the 'Midata' Innovation Lab Launch, (July 10, 2013), https://www.gov.uk/government/news/businesses-get-creative-with-consumer-data-at-the-midata-innovation-lab-launch.
62 See Midata: Which? First look, Which? (Mar. 26, 2015), https://www.which.co.uk/news/2015/03/midata-which-first-look-399235/. It also is unclear how much the tool is actually used, and since its initial release GoCompare has reduced the prominence of the midata tool on their website.
63 ATM transactions are unredacted, debit card transactions and fees redact card number data, and other transactions redact all descriptive data. See data for Personal Current Accounts, http://www.pcamidata.co.uk/ (last visited June 11, 2020).
64 See generally Midata Innovation Lab (2015), http://www.midatalab.org.uk/.
65 See Maurice E. Stucke & Allen P. Grunes, Big Data and Competition Policy 162-86(2016); Rubinfeld & Gal, supra note 48, at 353.
66 See generally Stucke & Grunes, supra note 66.
67 Adam Greenhall, Experimentation in a Ridesharing Marketplace - Part 2 of 3: Simulating a Rideshare Marketplace, Medium (Oct. 6, 2016), https://eng.lyft.com/https-medium-com-adamgreenhall-simulating-a-ridesharing-marketplace-36007a8a31f2.
One-off exports cannot address this kind of barrier to entry but APIs can. As mentioned before, one-off exports can take a long time to prepare all of a user's data, anywhere from minutes to months. APIs though allow for smaller chunks of information to be accessed more quickly (often in a fraction of a second) and more frequently.

C. Economies of scale

Economies of scale occur when a firm increases in efficiency as it increases in size. Usually, these gains in purchasing, managerial, and marketing power are eventually curbed by diseconomies of scale. Platforms that improve and grow from increased data collection though can experience unbounded economies of scale.68 Stucke and Grunes describe how the dynamic plays out:

"[T]he more people actively or passively contribute data, the more the company can improve the quality of its product, the more attractive the product is to other users, the more data the company has to further improve its product, which becomes more attractive to prospective users." 69

Growth and informational economies of scale are paramount virtues in the culture of Silicon Valley.70 Many platforms only become profitable once they reach a certain scale or market share, so they depend on venture capital, often for years, to offset their high upfront costs.

Many scholars, namely those in the Chicago School of competition law, argue that economies of scale do not themselves constitute barriers to entry.71 Others argue that the drive to reach and maintain economies of scale incentivize certain barrier creating behaviors. Khan describes two of these behaviors in the context of Amazon: predatory pricing and vertical integration.72 Firms may even go so far as to take data from other platforms, as Google Maps did when it scraped images from Yelp to fill out its information on places.73

Data portability can only help competitors overcome information economies of scale if the regime can bring a lot of data and users away from the data sender and to the data receiver. There is not much of a difference between APIs and one-off exports on this front. Though APIs may help receivers save on storage costs and give them access

---

68 See generally Robert Wilson, Informational Economies of Scale, 6 Bell Journal of Economics 184 (1975), https://www.jstor.org/stable/3003221.
69 Stucke & Grunes, supra note 66, at 170.
70 See generally, e.g., Peter Thiel & Blake Masters, Zero to One: Notes on Startups, or How to Build the Future, Crown Business (2014); Ben Horowitz, The Hard Thing About Hard Things: Building a Business When There Are No Easy Answers, Harper Business (2014).
71 Marc A. Eisner, Antitrust and the Triumph of Economics: Institutions, Expertise, and Policy Change 104 (reprint ed. 1991).
72 Khan, supra note 3, at 722-37. But see Daniel Crane, The Tempting of Antitrust: Robert Bork and the Goals of Antitrust Policy, 79 Antitrust L. J. 835, 852 (2014).
73 See Natasha Tiku, Yelp Claims Google Broke Promise to Antitrust Regulators, WIRED (Sept. 12, 2017), https://www.wired.com/story/yelp-claims-google-broke-promise-to-antitrust-regulators/.
to more up to date, targeted information, the relative economies of scale between the receiver and the sender may not change as much because the data continues to exist on the incumbent platform. Even if many users want to move away from an incumbent platform, informational economies of scale may mean that competitors are vastly inferior to the incumbent until those users actually move, creating a coordinated action chicken-and-egg problem.

Platforms have successfully used data portability to bootstrap their growth, but often at the expense of other normative values. LinkedIn adopted a legally dubious approach to data portability called “growth hacking” shortly after it was founded in 2002. Unlike Friendster, an open social network and the most popular at the time, LinkedIn was meant to connect professionals who knew each other in real life, allowing users to better trust recommendations made on the platform. This gated system made connections between users more meaningful, but it also meant less of them, which translated into a sparser network. LinkedIn’s monetization streams (job listings, subscriptions for improved communications, advertisements) all depended on the platform having many users, so they needed to find a way to grow.

To bootstrap their growth, LinkedIn ported contact lists from email clients, a novel strategy at the time. They built a plugin for the Microsoft Outlook email client that mined a user’s contacts and sent out invites to users who had not already signed up for LinkedIn. In 2011, they allowed users to upload their email contact lists from other clients with their “Add Connections” feature, and soon, up to 7% of users were uploading their address books, increasing the number of invitations to join LinkedIn by more than 30%. LinkedIn increased its access to contact information with the acquisition of Rapportive, which they used to build Sales Navigator that gave users rich information about their contacts inside of Gmail.

LinkedIn eventually was sending out so many invitations that they ran into legal trouble. When someone who was invited to LinkedIn through the Add Connections feature did not respond, LinkedIn would follow up two more times, without getting explicit permission from the inviter. If a person who was not a member of LinkedIn appeared in multiple uploaded contact lists, that could mean a huge volume of email. LinkedIn faced a lawsuit over this activity in 2015 and paid out $13 million in a

---

74 See Morgan Brown, Growth Studies - LinkedIn Growth Engine: The Never Ending Viral Loop, GrowthHackers (2014), [https://growthhackers.com/growth-studies/linkedin](https://growthhackers.com/growth-studies/linkedin).
75 See id.
76 See Ellen Lee, LinkedIn’s Startup Story: Connecting the Business World, CNN Money (June 2, 2009, 10:24 AM), [https://money.cnn.com/2009/06/02/smallbusiness/linkedin_startup_story.smb/index.htm](https://money.cnn.com/2009/06/02/smallbusiness/linkedin_startup_story.smb/index.htm) [https://web.archive.org/web/20180327012620/https://money.cnn.com/2009/06/02/smallbusiness/linkedin_startup_story.smb/index.htm](https://web.archive.org/web/20180327012620/https://money.cnn.com/2009/06/02/smallbusiness/linkedin_startup_story.smb/index.htm).
77 See id.
78 Brown, supra note 75.
79 Anthony Ha, Rapportive Announces Acquisition By LinkedIn, (Basically) Confirms $15M Price, TechCrunch (Feb. 22, 2012), [https://techcrunch.com/2012/02/22/rapportive-linkedin-acquisition/](https://techcrunch.com/2012/02/22/rapportive-linkedin-acquisition/).
80 See Jacob Kastreakes, LinkedIn Agrees to Settle Unwanted Email Lawsuit, Verge (Oct. 2, 2015), [https://www.theverge.com/2015/10/2/9444067/linkedin-email-lawsuit-settlement-add-connections](https://www.theverge.com/2015/10/2/9444067/linkedin-email-lawsuit-settlement-add-connections).
At that point though, they had already leveraged the scale of email to create their own scale. They had grown consistently, from 37 million members in 2009 to over 400 million members in 2016 when Microsoft acquired the company for $26.2 billion.

The growth story of LinkedIn shows how portability, especially when done shrewdly and through backchannels, can improve competition at the expense of end users. The “growth hacking” approach raises serious questions about privacy, security, and spam. Ideally, a portability regime should be able to allow for competition to thrive without impinging the rights and control of users over their data.

D. Network effects

Network effects occur when a product or service increases in value as more people use it. If one platform grows far larger than its competitors, that means network effects may entrench their advantage, creating a barrier to entry for others. There are two types of positive network effects: direct and indirect. Direct network effects occur when economic agents value a platform more when similar economic agents use it, like how telephone users gained value when others had telephones since it meant more people they could call. Similarly, Twitter benefits from positive network effects, since more active users on Twitter means more relevant and timely tweets.

Indirect positive network effects occur when an economic agent benefits from a different kind of agent using the platform. For example, Lyft drivers benefit from having more riders on the app and visa versa. More riders means drivers can more quickly find passengers, and more drivers means passengers have shorter pick up times. A failure to capture network effects can lead to a platform’s demise — a new ridesharing app will have a hard time attracting drivers without riders and visa versa; a rival to Twitter will have difficulty establishing itself as a dominant public forum if no one is there.

Evans outlines some strategies in their catalyst framework for new platforms entering the market to build up their own network effects. The value of multisided platforms, they argue, is that they help bring together distinct economic agents, reducing transaction costs for them to find one another. The difficult part for platforms is finding the initial catalyst to get those economic agents onto the platform in the first place.

This challenge can take on many forms. Sometimes, a platform needs to attract economic agents to only one side of the market first. For example, many social media applications match users with advertisers but only need users to initially get off the

---

81 Id.
82 See Ingrid Lunden, Microsoft Officially Closes Its $26.2B Acquisition of LinkedIn, TechCrunch (Dec. 8, 2016), https://techcrunch.com/2016/12/08/microsoft-officially-closes-its-26-2b-acquisition-of-linkedin/.
83 See Stucke & Grunes, supra note 66, at 162.
84 Michael Katz & Carl Shapiro, Systems Competition and Network Effects, 8 J. of Econ. Persps. 93, 96 (1994).
85 David Evans et al., Platform Economics: Essays on Multi-Sided Businesses 58-62 (2011).
ground (so long as they have funding to run their service.) In other cases, both sides of the market need to be catalyzed at once. A dating app that targets heterosexuals for example would need to have both straight men and straight women on the app to get either to join. And some platforms need one side to make a large, disproportionately risky initial investment. Video game systems for example must convince developers to make a big investment in their system by building games for it, often even before the system has reached the market and has any users.

Evans’ proposals for addressing catalytic problems involve coordinating groups of economic agents to move onto the platform all at once. They suggest contingent contracts, which gets one side to join after certain conditions have been met, a zig-zag strategy, where the platform alternates onboarding chunks of users onto each side, or a marquee strategy, where a platform tries to attract a prestige member who will bring a following with them.

Data portability could potentially be a part of any of these strategies. However, if individuals are only able to export their data one user at a time, it may be too difficult to organize enough users for the platform to reach catalyzation. Rather, groups of users would need to be able to move their data together in coordinated movements. As of today, no major platform offers a data portability tool that lets groups users migrate en masse to a new platform.

One might argue that platforms could use one-off exports to address network effects by allowing users to download their social graphs. Facebook allows something like this in their one-off export tool Download Your Information — users can download a list of their friends that allows them to match with any friend who uploads their own list. Without access to external contact information though, the exported social

---

86 Id. at 60.
87 Microsoft employed this strategy with their game streaming when they paid Ninja, a popular streamer of the most streamed game Fortnite, to move exclusively to their platform from Amazon’s platform, Twitch. See Tramel Raggs, Fortnite star Ninja leaves Twitch, to stream on Microsoft’s Mixer, Wash. Post (Aug. 1, 2019), https://www.washingtonpost.com/sports/2019/08/01/fortnite-star-ninja-leaves-twitch-stream-microsofts-mixer/.
88 See infra Part IV.
89 See Josh Constine, Facebook Shouldn’t Block You from Finding Friends on Competitors, TechCrunch (Apr. 13, 2018), https://techcrunch.com/2018/04/13/free-the-social-graph/; see generally Facebook Portability White Paper, supra note 36.
90 A brief technical note: when users download a list of their friends, in a file called friends.json, they get their name and the timestamp of their friendship. This combined with the name of the user who has uploaded their friends list could more or less work to connect two users if they both uploaded their friends list, without giving any information about mutual friends who did not consent. This is called an edge export. See Facebook Portability White Paper, supra note 36. However, the identifier is not strictly unique — if two John Mills each became friends with two Li Weis at the exact same second, the identifier would not be unique. See infra note 121.
91 I looked at my own friends.json file and found that it listed an email address or phone number for 1.5% of friends in that list. Facebook does not document what data a user should expect to see in their DYI export, so I do not know why it is in some and not others, but it may have to do with their privacy settings. See infra note 121.
graph cannot fuel growth. It only becomes useful on a wide scale once a platform has a large number of users, at which point it will benefit from its own network effects anyways.

APIs only address network effects better than one-off exports to the extent that data senders are comfortable sending more data knowing that they can cut off access.\(^{92}\) This data is limited in usefulness unless it also includes external contact data, which in the wake of Cambridge Analytica and other privacy scandals happens less often. In general though, APIs do not inherently make it easier for users to coordinate their movements or for platforms to overcome network effects.

### III. Additional considerations for API portability

The previous section outlined some of the advantages that APIs have over one-off exports in helping platforms overcome barriers to competition. However, APIs come with additional factors regulators should consider when making design choices about a legal portability. The main difference between one-off exports and APIs is the ongoing relationship between the data sender and receiver. One-off export regimes do not need to consider this much, since the user acts as an intermediary between the two. API regimes must consider this factor heavily, since it will likely be up to the regulator to balance the power dynamics between them.

This section describes three additional factors outside the specific competitive barriers to entry mentioned before that should be considered. First, competitors may be hesitant to build a product that depends on incumbent APIs out of fear that the incumbent will change them or cut off access. Second, APIs may not lead to products that are very differentiated from the incumbent, and users either may not switch to competitors in the first place or may backslide to the incumbent. Third, APIs may better serve complementary platforms rather than analogous direct competitors.

#### A. Hesitance to depend on incumbent APIs

API portability potentially lets competitors receive more frequent, up to date data from incumbents, but it also creates a complicated power dynamic problem for regulators to manage: from a technical standpoint, data senders have complete control over who can receive data and what data they can receive, and historically, platforms have used this to withhold data from threatening competitors. Until recently, Facebook explicitly disallowed applications that "replicate core functionality that Facebook already provides" in the platform policy for its social Graph API.\(^{93}\) They used this justification to cut off access to multiple social media

---

\(^{92}\) E.g. Facebook exhibits this behavior with data they share about a user’s friends. In their one-off export tool (Download Your Information), users can only download information that includes each of their friends’ names and the time they became friends. Their API (the Graph API) also gives a unique ID, though only limits the data to friends who have also installed the application. See Graph API Reference v7.0: User Friends - Documentation, https://developers.facebook.com/docs/graph-api/reference/v7.0/user/friends.

\(^{93}\) Josh Constine, Facebook ends platform policy banning apps that copy its features, TechCrunch (Dec. 4, 2018), https://techcrunch.com/2018/12/04/facebook-allows-competitors/.
competitors, including Vine, Yandex Wonder, and Voxer. Facebook did something similar to Prisma, a popular image filter app from 2016 that used a machine learning technique called "style transfer." In November 2016, Prisma allowed users to add their filters on Facebook Live streams. Within the same month, Facebook revoked Prisma's ability to use the Facebook Live API and released its own style transfer filter.

Other cases of Facebook restricting API access fall in the gray space between user protection and anti-competitiveness. Zynga used multiple Facebook APIs to grow a massive gaming platform on top of Facebook itself. The two companies had a symbiotic relationship, with Zynga at one point in 2011 accounting for 19 percent of Facebook's revenue and Facebook accounting for 93 percent of Zynga's. However, their relationship was contentious — Zynga was upset when Facebook started to take 30% of their revenue through their "Facebook Credits" system; Facebook was upset with Zynga for spamming its users with game invites. As Zuckerberg himself said in an interview:

“A lot of users like playing games, but a lot of users just hate games, and that made it a big challenge, because people who like playing games wanted to post updates about their farm or frontier or whatever to their stream. But people who don’t care about games want no updates. So we did some rebalancing so that if you aren’t a game player you’re getting less updates.”

94 According to emails released by the UK Parliament as part of the Six4Three case, this was personally approved by Mark Zuckerberg the day Vine launched. See generally Note by Damian Collins MP, Chair of the Digital, Culture, Media, & Sport Committee, https://www.parliament.uk/documents/commons-committees/culture-media-and-sport/Note-by-Chair-and-selected-documents-ordered-from-Six4Three.pdf.

95 See Josh Constine, Facebook Blocks Yandex's New Social Search App from Accessing Its Data Just Three Hours After Launch, TechCrunch (Jan. 24, 2013), https://techcrunch.com/2013/01/24/facebook-blocks-yandex-wonder/.

96 See Josh Constine, Facebook Is Cutting off Find Friends Data to "Competing" Apps That Don’t Share Much Back, Starting with Voxer, TechCrunch (Jan. 18, 2013), https://techcrunch.com/2013/01/18/facebook-data-voxer/.

97 See generally Vlad Savov, Prisma Will Make You Fall in Love With Photo Filters All Over Again, Verge (July 19, 2016), https://www.theverge.com/2016/7/19/12222112/prisma-art-photo-app.

98 Natasha Lomas, Facebook Has Cut Off Prisma’s Live Video Access, TechCrunch (Nov. 30, 2016), https://techcrunch.com/2016/11/30/facebook-has-cut-off-prismas-live-video-access/.

99 Tomio Geron, Facebook’s Dependence On Zynga Drops, Zynga’s Revenue To Facebook Flat, Forbes (July 31, 2012), https://www.forbes.com/sites/tomiogeron/2012/07/31/facebook-dependence-on-zynga-drops-zyngas-revenue-to-facebook-flat/.

100 Zynga Inc., Quarterly Report (Form 10-Q) (Oct. 26, 2012).

101 Zynga also threatened to create their own platform. Michael Arrington, Zynga Gunning Up (and Lawyering Up) for War Against Facebook with Zynga Live, TechCrunch (May 7, 2010), https://techcrunch.com/2010/05/07/zynga-gunning-up-and-lawyer-up-for-war-against-facebook-with-zynga-live/.

102 Exclusive: Discussing the Future of Facebook and the Facebook Ecosystem with CEO Mark Zuckerberg, InsideFacebook (2010), https://dc.uwm.edu/zuckerberg_files_transcripts/55.
In 2012, Facebook ended its special relationship with Zynga and introduced API changes that harmed Zynga’s business model.\textsuperscript{103} Zynga was no longer able to have users send out mass game invites to their friends, and shortly thereafter prevented users from inviting their friends altogether. Around this time, Zynga’s valuation dropped from $15 billion to $3 billion in six months.

The stories of Zynga and the cut off Facebook competitors serve as a warning to startups and the venture capitalists who fund them — a platform that depends on a third-party API for growth may fail if it changes. Were regulators to try to improve competition by enforcing API portability, they would have to figure out how to strike the balance preventing incumbents from breaking competing platforms that depend on their API and letting those incumbents innovate and update their service. The ideal balance may change dramatically between different types of products, making scalable policy difficult.

B. Backslide and undifferentiated products
If an incumbent offers an API or other means of interoperability, competitors that hope to build a similar product or service to that incumbent may have difficulty differentiating their offering, especially once price has reached an equilibrium.\textsuperscript{104} With undifferentiated products, even low switching costs may be too high for users to consider it worth it to change to a new product. Something like this happened in when the Federal Communications Commission (FCC) opened up competition between local exchange carriers (LECs, predominantly the seven Regional Bell Operating Companies, or "Baby Bells") and the long distance companies (e.g. AT&T, MCI, Sprint) in the Telecommunications Act of 1996.\textsuperscript{105}

After the divestiture from AT&T in 1984, LECs were regulated as natural monopolies and long distance companies were left to do more or less as they pleased. When the FCC passed the Telecommunications Act of 1996, they wanted the two to compete in each others markets. In order to give the long distance companies a chance at competing, the Act required LECs to provide them with "nondiscriminatory access to network elements on an unbundled basis"\textsuperscript{106} (a data sharing access scheme similar to an open API) at wholesale price. Only once state authorities, the FCC, and the Justice Department all decided that the LECs had adequately given access to their systems\textsuperscript{107} would they then in turn be allowed to go into the long distance market. This ended up giving the long distance companies a four year head start in being able to sell packages that bundled local and long distance service.

\textsuperscript{103} Tomio Geron, Zynga and Facebook Revamp Deal, Zynga Shares Drop, Forbes (Nov. 29, 2012), \url{https://www.forbes.com/sites/tomiogeron/2012/11/29/zynga-and-facebook-revamp-deal-zynga-shares-drop/}.

\textsuperscript{104} See, e.g., Luuk de Klein & Lisha Zhou, French, Australian agency chiefs share ideas on privacy, competition interplay, in PARR, ABA Antitrust Spring Meeting 18 (2019), \url{https://www.acuris.com/assets/PaRRABAreport2019_1.pdf}.

\textsuperscript{105} See Peter W. Huber et al., Federal Telecommunications Law, (2nd ed. 2011).

\textsuperscript{106} 47 U.S.C. § 251(c)(3) (LEXIS through Pub. L. No. 116-141).

\textsuperscript{107} See generally 47 U.S.C. § 271 (LEXIS through Pub. L. No. 116-141).
Despite allowing for competition on a more even footing, the 1996 Act did not create long term sustainable competition through this process. LECs and long distance services ended up offering essentially the same product, with similar costs and revenue, only adding the minor convenience of giving customers one bill instead of two. Very few customers actually switched their providers, and those switched when telecoms started paying customers to do so, often switched back when other telecoms made the same offer. Wireless eventually destroyed the profits of the long distance industry, and the long distance companies had not diversified enough to save their business model. Most of them were acquired by LECs for cheap, further concentrating the telecommunications industry.

The 1996 Act was a success in many ways, but the data portability aspect of it failed because customers were not willing to bear even a low switching cost of going to a competing product since it was so similar. Those who found it worthwhile to switch could also backslide. The competitors, here the long distance companies, could only differentiate their products so much because it entirely depended on a similar one from the incumbents, the LECs.

C. Opportunities for creative destruction

API portability is not without its potential in the competitive landscape. To understand it, it is useful to look at different notions of competition: neoclassical and Schumpeterian. Neoclassical economics argues that competition is defined by its desired outcome of improved product and lower price. This approach to competition is easy to model and leads more to the creation of directly competing products than wholesale innovation. Meanwhile, Schumpeter saw competition as a process of “creative destruction,” whereby innovative newcomers disrupt and dislodge incumbents. It is a process of explosive, unpredictable change that usually leads to a larger paradigm shift towards complementary products rather than directly competing ones.

The portability examples of telecoms and midata were built to encourage neoclassical competition. However, APIs may work better for encouraging Schumpeterian competition. The creation of complementary products avoids some of the problems raised in the Telecommunications Act of 199 — it does not let users to backslide to the incumbent nor does it create the problem of not being able to overcome low switching costs for small advantages. It also gives the complementary products a first mover advantage, even if the incumbent can follow suit later. This dynamic has played out in dating apps — Tinder used the Facebook API to make it easier for users to build their profiles with ported Facebook data. When Facebook released its own dating app later on (again, it took longer to build a complementary product), Tinder had built up enough of a brand name and network effect that it wasn’t a threat.

108 Birand Interview.
109 See generally Herbert Hovencamp, United States Competition Policy in Crisis: 1890-1955, 94 Minn. L. Rev. 311 (2010).
110 Joseph A. Schumpeter, History of Economic Analysis 835-38 (Elizabeth B. Schumpeter ed.; reprt. ed., rev. ed. 1984).
Some of the potential for Schumpeterian style competition through APIs can be seen with open banking in the financial tech sector. Companies like Mint, Yodlee, and Quicken leverage banking data, often from multiple sources, to give users a comprehensive picture of their finances. Another fintech company, Cushion, reads a user’s bank transaction history and negotiates lower fees on their behalf. Plaid, acquired by Visa in 2020 for $5.3 billion, has created APIs that lets third party developers interface with banks to further allow this industry to grow.\footnote{See Press Release, Visa, Visa To Acquire Plaid, (Jan. 13, 2020), https://usa.visa.com/about-visa/newsroom/press-releases.releaseId.16856.html.}

All of this growth has happened even without banks and tech companies cooperating. Open banking is an example of interoperability without explicitly permitted APIs — fintech companies often use unsanctioned logins and screen scraping to collect financial data\footnote{For example, Plaid using link injections to get bank credentials. See Ben Ellis, privacy/security concerns #68 Github Plaid/Link (Feb. 11, 2016), https://web.archive.org/web/20190415103059/https://github.com/plaid/link/issues/68; Whockey, Comment to Plaid Deletes GitHub Issue Exposing Imitation of Bank Login UIs, HackerNews (June 8, 2019), https://news.ycombinator.com/item?id=20133806.} while banks in turn improve their anti-scraping technology.\footnote{See Nathan DiCamilo, In Data Dispute with Capital One, Plaid Stands Alone, American Banker (July 17, 2018), https://www.americanbanker.com/news/in-data-dispute-with-capital-one-plaid-stands-alone.} At the same time, banks sue tech companies for “unauthorized access” under the Computer Fraud and Abuse Act\footnote{See generally 18 U.S.C. § 1030(a)(1) (LEXIS through Pub. L. No. 116-141).} and protection of “nonpublic personal information” under the Gramm-Leach-Bliley Act,\footnote{See generally 15 U.S.C. §§ 6801–6809 (LEXIS through Pub. L. No. 116-141).} while tech companies insist that the Dodd-Frank Act protects their use of consumer data.\footnote{See generally 12 U.S.C. § 5481 (LEXIS through Pub. L. No. 116-141).} All of this is to say that given how much innovation in the fintech sector with costly, jerry-rigged interoperability, it may only represent a fraction of the innovation that would be possible with officially sanctioned APIs.

Schumpeterian and neoclassical competition require different policies to induce.\footnote{See generally Tim Wu, Taking Innovation Seriously: Antitrust Enforcement If Innovation Mattered Most, 78 Antitrust L.J. 313 (2012).} Both have positives and negatives. Neoclassical competition is easier to regulate but is unlikely to bring about massive innovation. Schumpeterian creates a lot of value for firms and consumers but is by its nature difficult to predict. It is hard to prove that a data set has innovative potential either in the positive or the negative so it is difficult to design policy to encourage it. I will set aside further interrogation of this problem for a future paper. For now, I will turn to a new approach to data portability that is designed to maximally encourage neoclassical competition.

**IV. Collective portability — a new approach**

The previous two sections looked at approaches to data portability that focus on giving individual users access to their data. This reflects the way laws tend to regulate portability and platforms implement it. The individual approach works to alleviate
switching costs and some unique data access barriers but fails to address network effects and informational economies of scale. Previous analysis of existing theories and examples of portability suggests these barriers can be better addressed by groups coordinating to move their data all at once.

To this end, I propose a new approach to portability called collective portability. Collective portability is the ability for users to easily transfer their data to another platform in the same unit they originally used it in and alongside other users they used it with. If data originally shared and accessed within a group, the original platform makes it easy for groups to coordinate to move their data, and if they choose to, facilitate transfer and deletion. Like API portability, collective portability requires the data sender and receiver to be in direct communication with one another. However, since the relationship does not continue after the transfer is complete, backslide and balancing the power between senders and receivers is not a concern. It also means that users can delete their data from the data sender if they so choose.

Collective portability works better than one-off exports at improving neoclassical competition. Not only does it better address network effects and informational economies of scale, but it also allows data to be transferred that otherwise could not without raising privacy concerns. One-off exports only allow users to transfer data that is clearly their own. If data is shared by multiple people, it may end up in none of their exports. Collective portability includes that data in the transfer, making it more than the sum of everyone’s individual user exports without harming anyone’s privacy.

This can be seen in the data that Facebook makes available in its Download Your Information (DYI) tool. DYI allows users to download a computer-readable archive of all information they have ever entered into Facebook that it still has stored. The information includes such things as posts, events attended, comments, advertisements clicked on, and photos uploaded. It does not include data provided by other users, like tagged photos or likes from friends.118

I downloaded my personal DYI archive and here is an anonymized example of the data representing my comment on a photo that my friend (here named “Fannee Doolee”) posted to a group we are in:119

118 The information also does not include any inferences Facebook may have made about a user. For example, if Facebook has determined that a user supports gun owner rights and uses that information to show ads for guns, DYI does not give a user who downloads their information access to that insight. See Facebook Portability White Paper, supra note 36, at 13.

119 For this analysis, I downloaded my own Facebook information, which goes from 2006 to 2019. I downloaded my data in the machine-readable JSON format as opposed to the more human-readable HTML format, since this is likely what actual competitors would use to integrate the data. My data does not represent the entire range of possible data to download because I have not used every feature on Facebook. For example, I have never used Facebook's “check in” feature, so my data download has an empty folder where that data would be. I also may have had certain privacy settings enabled that prevented certain types of information from being downloaded. People who I have interacted with on Facebook may also have these settings. All of this is speculative because Facebook provides no documentation about what data structures a user should expect from their DYI export. I reached out to Facebook to get basic information on the rules underlying DYI or even the basic structure of data I could expect, but after six months of back and forth, they did not provide any, nor did they give any
Here, I get the text of the comment, the time I posted it, they type of item I commented on (a photo), the name of the group I am posting to (Little Wanderers), and the name of the person who owns it (Fannee Doolee). I do not get access to the photo itself, or any unique identifiers for the group, person, or photo. If Fanee Doolee downloaded all of her photos, she would similarly not have access to my comment. Therefore, even if Fanee Doolee and I both uploaded our data to the same competing photo-posting-and-commenting platform, it would not have a way to connect my comment to her photo.

It is unlikely that Facebook made this system design choice solely for the purpose of being anti-competitive — it also addresses legitimate privacy concerns. If Fanee Doolee does not want to upload her photo to the platform that I upload my data to, she would not expect me to be able to bring it there without her permission. And if a number of people who commented on that photo move to the new platform, she may not want that conversation moved somewhere without her permission. However, these privacy protections directly reduce the ability of competitors to recreate features of Facebook. A similar tradeoff occurs with other ported data, such as events, statuses, and likes.

Collective portability could address this issue at least as it pertains to the Facebook group. If Little Wanderers chose to move to another platform and Fanee Doolee and I both consented, the new platform would be able to show the relationship between

---

See generally Gabriel Nicholas, Portability Project, GitHub (Sept. 2019) https://github.com/gajeam/Portability-Project/ [hereinafter Portability Project GitHub].

120 Note that though there are two timestamps, they are identical. See generally Id.

121 If I were friends with multiple people named "Fannee Doolee", this would be a problem or in multiple groups called "Little Wanderers", this would be a problem. See generally Id.

122 This could be considered an "inappropriate information flow". See generally Helen Nissenbaum, Privacy in Context: Technology, Policy, and the Integrity of Social Life ( 2009).

123 Notably, it does not occur in data about private messages — users get both sides of the conversation and all reactions to messages, meaning that a competitor could in theory use this information to build a new messaging platform. See generally Portability Project GitHub supra note 119.
her photo and my comment, reducing switching costs for us and increasing network
effects for the new platform.124

Designing a regime for collective portability is not as straightforward as it is for one-
off exports, even in this relatively simple Facebook example. Below, I outline five
questions that platforms and policymakers would have to consider to implement and
regulate collective portability, and I apply them to different hypothetical applications.
Then, I look at the ways collective portability addresses the barriers to entry
discussed earlier. Finally, I point out new challenges collective portability raises.

A. Five questions for a collective data portability regime
With any portability regime, platforms must make design decisions about which
data they make available to users and how they make it available. Collective portability
raises additional design questions for a platform because it involves multiple users,
with many potential relationships between each other and their data. Five of these
questions are discussed below:

1. Who is in in the collective? For many platforms, groups will have clear boundaries
   — it usually is clear who is and is not on a group messaging thread, for example. Some
   platforms may have more informal boundaries around their groups and
   communities, or no boundaries at all. Nothing strictly delineates which users are
   and are not part of “legal Twitter.”125 Twitter therefore it may not be as good a fit
   for collective portability.

2. Which data is part of the collective? Again, this is a relatively easy question for
   smaller and simpler groups like message threads, but more difficult for larger,
   more fluid groups. For example, a community platform like Reddit would have to
decide whether or not it allowed groups to port data cross-posted from another
   community.126

3. How does the collective come to a decision to port their data? Collective portability
   requires the data sender to implement a governance mechanism to allow users in
   the group to decide to move to another platform. There is no one-size-fits-all
   design for this mechanism, especially for larger groups. If the data sender allows
   for a vote, how many users have to vote for the move to occur? How are inactive
   group members handled? How does the migration handle data from users who
   voted against it?

4. How do special user roles affect the transfer decision process? Some platforms allow
   certain users in a group to have more privileges than others. This makes some
   portability decisions trivial — if an employer wants to switch their team

124 The profile_information.json file included among other things, my birthday, email address,
current city, home town, family members, education and work experiences, and pages I had
liked ten years ago. See generally Id.
125 But see Sarah Mui, Best legal Twitter accounts of 2018, ABA Journal Web 100 (Dec. 1, 2018),
http://www.abajournal.com/magazine/article/best_law_twitter_2018.
126 u/EuroGeologist, How to crosspost on Reddit in 2018?, Reddit (May 18, 2018, 4:54 AM),
https://www.reddit.com/r/help/comments/8kbptp/how_to_crosspost_on_reddit_in_2018/.

Electronic copy available at: https://ssrn.com/abstract=3550870
communication tool, they can make a unilateral decision to transfer group data, even if they did not provide it all. Other questions may arise in non-professional communities, where administrators and moderators have less unilateral power.

5. How does the migration work? Once decisions have been reached about which users and which data ought to be transferred, it is up to the data sender to pass the information along to the data receiver. For groups with missing or incomplete data from users that opted out of the migration, the data receiver will have to make design decisions about how to show that to users. Similarly, the data receiver’s schema for representing data may not map perfectly onto the the sender’s, especially if the platforms do not share all the same features. Receivers need to decide whether to discard this data or handle it another way.

B. Hypothetical collective portability regimes

In this section, I apply the five questions outlined above to designing hypothetical collective portability regimes for Spotify, iMessage threads, and Facebook groups. The hypotheticals increase in complexity and are meant to show that designing these regimes may be difficult, but is not impossible.

1. Spotify

Imagine I use Spotify to stream music but I want to switch to YouTube Music because they offer a similar product at a lower cost. Spotify has years worth of data on my playlists and albums saved to my library. They also have my listening history, which means they can give me high quality recommendations for songs I might like.

I am the only user in this group, since I largely interact with the Spotify by myself, and therefore I can make unilateral decisions about my data. The data in this group are my playlists, music library, and listening history. Even if friends follow some of my playlists, most of the boundaries between my data and theirs are clear. Network effects, at least in terms of interacting within small groups, are less of a concern because the primary value of Spotify comes not from the network of other users but interaction with the data itself (i.e. listening to music as opposed to interacting with friends.)

If Spotify were to adopt a collective portability regime, they would facilitate me sending all three of these data sources to a competitor, like YouTube Music, and in turn, allow me to cancel, pause, or delete my Spotify account. YouTube Music may have to do some work on its end to map Spotify’s representation of songs onto its own, and likely they would only implement this if they believed the effort could actually get users away from Spotify.

2. iMessage

Imagine I am in an iMessage thread with four close friends. We all agree that we want to move our thread to WhatsApp because I am about to switch from an iPhone to an Android and WhatsApp offers better inter-operating system messaging.

The people in this group are me and my four friends, and the data is our thread, with every message, image, and reaction sent. With so few people in the thread, iMessage
may require unanimous consent to move. The voting mechanism could be built into iMessage itself and users would be able to pass along their WhatsApp credentials to iMessage to allow it to make the transfer. Some information that does not translate between the two systems, such as message reactions (iMessage has them, WhatsApp does not) gets lost.

3. Meme groups

Imagine I am in a ten-thousand person Facebook group called “Tech Competition Memes for Neo-Brandesian Teens” where dozens of irreverent, antitrust-related image macros get posted everyday. The admins of the group fear that Facebook is preventing their content from showing up on people’s newsfeeds, and suspiciously, new users seem not to be able to join. They want to move the group to a new platform MemeOn!, built specifically for anti-authoritarian meme groups.

The people in this group are the group members and the data is the posts in the group, including images and comments. The governance mechanism may be a direct or liquid democracy. Facebook would have to decide how to handle data from people who did not vote or voted against the move, though this issue would be easier for groups started after Facebook begins collective portability because they could set the options from the beginning. Administrators may have a special status, potentially deciding to hold the vote or choosing how many people would need to approve the vote for the group to move. Migration would raise many design questions, and either every migrating user would need to connect to their MemeOn! account (this could be tied in with voting) or the old data would only need to be there for archival purposes.

C. Collective portability and barriers to entry

In this section, I compare how well collective portability alleviates each of the previously established four barriers to entry compared to the existing approaches, one-off exports and APIs. I order them from least to most improvement.

1. Switching costs

Collective portability does not lower switching costs by being a simpler, more intuitive process for users to engage in. For large groups, governance mechanisms and third-party authentication flows may actually increase switching costs, especially if poorly designed. However, it reduces social switching costs by bringing other users from the network from the original to the new one. This is particularly helpful for nascent social networks. Often, users pioneering new social networks know few other people there, even if on social networks experiencing viral growth. If users do not have to leave behind their social connections on the original social network, they will experience lower switching costs.

---

127 See generally Steve Hardt & Lia C.R. Lopes, Google Votes: A Liquid Democracy Experiment on a Corporate Social Network, in 79 Defensive Publication Series, Technical Disclosure Commons, (June 5, 2015), https://www.tdcommons.org/dpubs_series/79/.

128 E.g., Wikitribune, Ello, Vero. All are social media platforms that surged in popularity in response to different Facebook scandals. They all had waitlists when they first launched because they could not handle the scale of user interest, but all failed to maintain that level of growth, potentially because too few people were on them.
2. Unique data access

Collective portability can open up unique data sources that entire platforms can be built around. By moving users en masse and allowing them to bring their data with, a new entrant can target a specific large group and create a product directly aimed at its needs in ways that a large multi-use platform cannot. For example, Facebook groups have been used by political activists to organize and patient groups to support one another, but the platform fails to meet the privacy needs of both groups. Both originally started on Facebook to reach the largest number of people. Now that they have reached those people, they do not have a way to coordinate movement to another, and therefore competitors are not incentivized to build products specifically aimed at them, even if they could build a better tailored product.

Collective portability cannot address temporal, “nowcasting” advantages because it does not maintain an ongoing connection between the sender and the receiver, so timely data cannot be sent.

3. Economies of scale

Collective portability better addresses economies of scale than other approaches to data portability because it allows for more data and more users to move over at once. This can help competing platforms reach a same or similar scale to incumbents more quickly. It also allows for scale without threatening privacy, the way LinkedIn’s contact scraping did, but also without compromising utility to competitors, like Facebook's DIY tool. Also, since groups agree to move their data together, less data will fall between the cracks as it would with users moving with individual one-off exports.

4. Network effects

Collective portability aligns closely with the strategies Evans recommends for overcoming the catalyst issues posed by network effects. If platforms create value by bringing distinct economic agents together and reducing the transaction costs of them finding one another, then new entrants to the market need to kickstart this growth by onboarding large chunks of users at once. Collective portability does exactly that. And for some platforms, such as messaging platforms, a few strong

---

129 See generally Madelyn Rose Sanfilippo & Katherine J. Strandburg, Privacy governing knowledge in public Facebook groups for political activism, 22 Information, Communication & Society 26, 2019, https://www.tandfonline.com/doi/full/10.1080/1369118X.2019.1668458; United States v. Facebook, Inc., No. 1:19-CV-2184 (D.D.C. July 24, 2019) (stipulated order), https://www.ftc.gov/system/files/documents/cases/182_3109_facebook_order Filed_7-24-19.pdf.

130 It may also not take having the same amount of data to compete. There are eventually diminishing returns on the amount of data and the features of the algorithm itself superceded. See generally Xinran He et al., Practical Lessons from Predicting Clicks on Ads at Facebook, AdKDD’14: Proceedings of the Eighth International Workshop on Data Mining for Online Advertising 1 (Aug. 24, 2014).

131 See generally Evans, supra note 86.
network connections may create network effects as powerful as many weaker connections.\textsuperscript{132}

D. Challenges to collective portability

Collective portability has some distinct advantages over current one-user-at-a-time portability approaches, as discussed previously. However, regulation and implementation of collective portability, beyond being difficult, threatens to exacerbate other problems faced by the tech industry. Platforms may also resist adoption of collective portability for competition and non-competition related reasons. Some of the challenges this approach may face are discussed below.\textsuperscript{133}

1. Implementation complexity

Platforms may have a difficult time implementing collective portability which in turn may lead to a complex design that users do not readily adopt. Some platforms may have difficulty determining which users or what data constitute a collective. Users may have different gradations of control over certain data (e.g. a Google Doc can have owners, editors, commenters, and viewers) and ownership is not always exactly clear (e.g. if a Twitter user quotes another user’s public tweet, is it their data?)\textsuperscript{134}

Governance mechanisms can also be difficult to design, especially for large communities where the choice to move platforms will most likely not be unanimous and users have different power and permissions.\textsuperscript{135} Additional normative and technical questions arise for groups and platforms that were built before implementing collective portability. Platforms that built it in from the beginning will have an easier time with these, allowing for rules to be set at the time of users signing up for the group or platform. If the rules determining which data gets sent are too complicated or not transparent enough, competitors may be hesitant to invest in integrating with it and users will not adopt it.

2. Incumbent resistance

Existing platforms are likely uninterested in implementing collective portability. If firms believed it would encourage users to move away from their platforms, they may not want to support it because they see it as sabotaging their own business. Collective portability is in some sense a zero sum game, with platforms vying for the same users,

\textsuperscript{132} E.g., I personally have this with Letterboxd, the movie reviewing app. I only have a few close friends on it, but we all write reviews for each other and all read them. A questions and answer service like Quora though may depend on many people, since they want a range of expertise.

\textsuperscript{133} See Mike Masnick, \textit{Protocols, Not Platforms: A Technological Approach to Free Speech}, Knight First Amendment Institute at Colum. U. (Aug. 21, 2019), https://knightcolumbia.org/content/protocols-not-platforms-a-technological-approach-to-free-speech, These problems are the same as the ones in Masnik’s approach to platform free speech and competition. The two proposals are very different, but both involve making it easier to move data around.

\textsuperscript{134} Twitter says, yes. When I downloaded my Twitter data, I got other people’s retweets, including their pictures.

\textsuperscript{135} StackOverflow gives users new permissions if their questions and answers are voted on as helpful by the community.
data, and attention. Large platforms with more to lose may be particularly against the initiative, meaning that Platforms may also not want to take responsibility for answering the hairy questions about how to send data laid out in the previous section since those answers could upset certain users and take a lot of resources. Platforms would be especially hesitant to invest in collective portability (and perhaps already are with their current portability regimes) since the whole process of sending data would not benefit their own platform. Platform resistance would put the onus on regulators to design competition-inducing collective portability standards, and they might not have the technical knowhow to do so effectively.

Finally, collective portability may run counter to the way that large technology firms conceptualize their platforms and design policy for their users. For example, research on how Facebook, Inc. sets rules for its users shows by following signals from individual actors (users and organizations). Collective portability would require a paradigm shift in Facebook’s approach to public policy to addressing the needs of pluralities rather than individuals.

3. Content moderation and filter bubbles

If users can easily organize and move their data from one platform to another en masse, de-platforming could become more difficult. It could make it easier for hate speech and other normatively undesirable behavior to more regrow on a new platform when stomped out on another. This dynamic has played out on a smaller scale between Reddit, a news aggregator and social network, and Voat, an almost identical service. In 2015, Reddit closed down five of its communities (or “subreddits”) for harassment, including r/FatPeopleHate with 150,000 members. A controversy over free speech on the platform broke out and some moved to Voat, then run by a two person team, where they restarted some of the closed communities. Voat’s FatPeopleHate group at one point had 18,700 members. The r/Incels subreddit, known for inciting violence against women including a mass shooting in Isla Vista, was also banned from Reddit and partially moved to Voat.

136 It is too speculative to say that collective portability would increase the number of users on the Internet or the amount of time they spend on multi-sided platforms.
137 See generally Matthias C. Kettemann & Wolfgang Schulz, Setting Rules for 2.7 Billion. A (First) Look into Facebook’s Norm-Making Systems: Results of a Pilot Study (Hans-Brewdow-Institut, Working Paper No. 1413-9413 2020).
138 See generally Emma Woollacott, Users Flock to Voat as Reddit Shuts Harassing Groups, Forbes (June 11, 2015), https://www.forbes.com/sites/emmawoollacott/2015/06/11/users-flock-to-voat-as-reddit-shuts-harassing-groups/.
139 See Adi Robertson, Welcome to Voat: Reddit Killer, Troll Haven, and the Strange Face of Internet Free Speech, Verge (July 10, 2015, 9:11 AM), https://www.theverge.com/2015/7/10/8924415/voat-reddit-competitor-free-speech.
140 Id.
141 See Chris Hauser, Reddit Bans ‘Incels’ Group for Inciting Violence Against Women, N.Y. Times (Nov. 9, 2017), https://www.nytimes.com/2017/11/09/technology/incels-reddit-banned.html; Jay Hathaway, Why Reddit Finally Banned One of Its Most Misogynistic Forums, Daily Dot (Nov. 11, 2017), https://www.dailydot.com/unclick/reddit-incels-ban/.
If Reddit had an option for collective portability, it could have made the migration of r/FatePeopleHate or r/Incels to Voat easier, potentially allowing them to keep the historical content and migrate users. Furthermore, if Reddit knew that hateful communities could easily reconstitute on new platforms, it may have been less incentivized to remove these communities in the first place under the logic of, if we don’t host them someone else will. Frictionless ideology-based migration could also increase filter bubbles,\textsuperscript{142} by making it easier groups with shared ideologies to flock to the same platform and isolate themselves from other ideas.

The capacity to deplatform is normatively ambivalent. Forums like 8chan, a forum linked to multiple mass shootings known for harboring white supremacism and antisemitism, should obviously be able to be deplatformed, as should child pornography and terrorist activity. However, deplatforming raises free speech concerns,\textsuperscript{143} as came up when Alex Jones was banned from YouTube and Facebook, seriously diminishing traffic to his show “Infowars”.\textsuperscript{144} The extent to which collective portability allows for certain kinds of deplatforming depends on specific choices in implementation.

\textbf{Conclusion}

Data portability as it exists today has yet to prove that it can improve competition in the tech sector. The GDPR has required platforms to allow users to download their personal data since 2018, yet no meaningful competitors have emerged from this effort. Successful predecessors to platform data portability such as mobile number portability only had to alleviate switching costs; new multi-sided platforms hoping to enter the market face additional barriers, and since data on an incumbent platform is mostly useful in the context of other users, one person porting their data in isolation is often insufficient to bootstrap competitors.

Collective portability addresses an incumbent’s multiple barriers to entry beyond switching costs. It allows users to coordinate to transfer their data with others and maintain the context it was used in, assuring competitors that they can use it for the same purpose. This may be an appealing tool for regulators hoping to improve platform competition, even if it comes with its own challenges around governance, content moderation, and industry buy-in.

What might good collective portability regulation look like? First, it would be entirely separate from data portability regulation aimed at giving users better control over their data. Singapore has a precedent for this in forthcoming amendments to their Personal Data Protection Act (PDPA), which separate out the competition and data control aims of data portability into two separate rights, the Data Portability

\textsuperscript{142}See generally Eli Pariser, The Filter Bubble: What The Internet Is Hiding From You, (1st ed. 2011).
\textsuperscript{143}See Jonah Goldberg, Deplatform on Social Media and Free Speech, Nat’l Rev. (May 13, 2019, 5:19 PM), https://www.nationalreview.com/corner/deplatforming-on-social-media-and-free-speech/.
\textsuperscript{144}Jack Nicas, Alex Jones Said Bans Would Strengthen Him. He Was Wrong., N.Y Times (Sept. 4, 2018), https://www.nytimes.com/2018/09/04/technology/alex-jones-infowars-bans-traffic.html.
Obligation¹⁴⁵ and the Access Obligation respectively.¹⁴⁶ Second, it would differentiate between small, new market entrants and large incumbents. This would prevent users and data from further concentrating on large platforms with strong network effects while also relieving under-resourced startups of a technical burden. Third, the regulation would create explicit carve outs for when portable user data overlaps. The GDPR says that if personal data overlaps between two or more data subjects, privacy and deletion rights supersede portability rights, leaving none of the data subjects able to port it. A collective portability version of this policy would say that if the data subjects were part of some closed off group, the platform should make it possible for all relevant data subjects to consent to a transfer or export.

¹⁴⁵ Singapore Personal Data Protection Commission, *Response to Feedback on the Public Consultation on Proposed Data Portability and Data Innovation Provisions*, 5 (Jan. 2020), https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Legislation-and-Guidelines/Response-to-Feedback-for-3rd-Public-Consultation-on-Data-Portability-Innovation-200120.pdf. The Data Portability Obligation is intended “to facilitate movement of consumer data from one service provider to another, so that consumers are better empowered to try out or move to new or competing service offerings.” It requires data controllers to transmit data in a machine-readable format directly to another controller if so requested.

¹⁴⁶ Id. The Access Obligation “is intended to allow individuals to access and verify their personal data in an organisation’s possession or under its control, and how their personal data has been used by the organisation.”