WORKING PAPERS

N° 1520

March 2024

"The Economics of the Cloud"

Gary Biglaiser, Jacques Crémer and Andrea Mantovani

Toulouse School of Economics

The Economics of the Cloud[1]

Gary Biglaiser
University of North Carolina at Chapel Hill
Jacques Crémer
Toulouse School of Economics
Andrea Mantovani
TBS Business School

March 29, 2024

[1] This report has been written with the support of a research grant from Microsoft Corporation to the Toulouse School of Economics. The authors are the only parties responsible for the opinions expressed therein, which in no way implicate Microsoft or TSE.

In the preparation of this report, we have spoken to numerous people and presented the ideas it discusses to numerous audiences. We are very grateful for the feedback which we have received.

Abstract

The aim of this report is to present the main facets of the development of cloud services, its economics and the related policy issues. We begin by surveying the sector, its growth and the significant increase in concentration in recent years. We then discuss the tools that economics gives us to study these phenomena before turning to a critical analysis of some of the most prominent policy reports which have been produced on the topic. We finally turn to a more detailed look at the (meagre) economic literature on the industry and of the economic theories which could be used for deeper analysis.

JEL codes: K21, L13, L51, L86, O33

Contents

Background

Cloud computing is changing the face of the digital industries. While firms, governments, and non-profit organizations used to run most of their computing in-house, first through computers and then through servers, in recent years, more and more of the world's computations and data storage are carried out "in the cloud", and its reach extends beyond mere computation. Cloud computing can indeed improve access to advanced IT solutions and boost innovation and productivity. Cap Gemini, a major provider of information technology and consulting, predicts that, thanks to 5G, cloud computing and telecommunications will merge to create an interconnected universe, which will generate colossal amounts of data.[1] Revenue for cloud services between 2021 and 2022 reached \$191 billion, and the cloud market is "projected to be worth \$376.36 billion by 2029".[2]

We are witnessing the birth of a new industry: the rental of computation capacity, and this raises important analytical and policy questions. Yet, little research has been done on the economics of the cloud; most of the discussions of its recent development are descriptive, with little economic analysis. The aim of the report is threefold: first, to describe the state of the art in the economics of cloud computing; second, to indicate possible pathways to produce research of high quality; third, to develop the interest for the topic in the economic profession — the issues related to cloud computing are not only important from a policy viewpoint, they are also intellectually exciting. In so doing, we will try to understand how standard economic analysis can help to understand the cloud.

As cloud computing has gained momentum, it recently attracted a lot of attention from policy makers as there are concerns in terms of market concentration and controversial practices adopted by dominant cloud providers that some believe to have increased switching costs and lower interoperability. Notwithstanding the recent entry of new players, the cloud sector remains concentrated. Indeed, the aggregate market share of the three largest cloud computing platforms - Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) - has significantly increased in the last few years, from 50% in 2017 to 66% in 2023.[3] It is therefore natural to ask

[1] `https://www.capgemini.com/fr-fr/perspectives/blog/cloud-5g-agents-indissociables-revolution-de-donnee/`.

[2] `https://aag-it.com/the-latest-cloud-computing-statistics/`.

[3] We follow estimates from Synergy Research Group as shown in this chart: `https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/`. The market shares are imprecise and

how many cloud providers should there be "at equilibrium". Given the technology of the cloud, with substantial fixed costs, one expects a high degree of concentration. Still, many questions remain unanswered. In particular, *how much* competition is possible in the cloud industry? Does concentration create systemic risks? How should the toolkit of competition policy be applied to this industry? If the level of competition is considered to be too limited, how can policy makers intervene in order to increase it? And what would be the consequences of these interventions on innovation?

Cloud computing therefore poses delicate challenges to regulators and policy makers. We believe this is just the beginning, as we reckon there will be in future years increasing policy interest and calls for regulation of the industry. It is therefore paramount that these debates be informed by the best possible evidence, and analyzed by the most appropriate economic tools. At the present time, there is justifiably much talk about the way in which Artificial Intelligence (AI) is reshaping the world. However, not sufficient attention has been devoted to this silent revolution that is transforming multiple business activities, while simultaneously propelling the growth of AI and the Internet of Things (IoT). The protagonist of this revolution is cloud computing, and in this report we want to further explore this phenomenon.

In Section 1, we briefly introduce cloud computing and its major actors. We describe its technical characteristics, as well as its service and deployment models. We also outline its impressive growth. Many reports and market studies have been written on the cloud, and we will not try to be comprehensive, but rather to highlight the elements that are important for the rest of this report.

In Section 2, we discuss the concepts of economics which can illuminate cloud computing: switching costs, pricing of cloud service, interoperability, network effects, and standards.[4] We keep this section not very technical, as in Section 4 we go in more detail in what the economic literature teaches us.

The growth of the cloud, its importance for the digital sector and hence for the economy as a whole, and the concentration of the industry have, not surprisingly, attracted the attention of regulators. Section 3 is therefore devoted to a discussion of the most important policy initiatives of governments and policy makers. We discuss the relevant regulations, in particular Europe's Data Act, as well as influential reports such as those produced by the

depend on the product definition as well as the statistical sources. However, different sources agree on the order of magnitude and the trend.

[4] Other factors, such as scale and scope economies, are also important for understanding the economics of the cloud. We will not explicitly discussed them, as their presence in the cloud sector does not necessitate novel insights through the lens of economic analysis.

Dutch (ACM, 2022) and the French (FCA, 2023) competition authorities.

In Section 4, we provide a more technical exploration of the economic literature that deals with issues relevant to cloud computing. We first briefly review the scant academic literature which is directly aimed at understanding the cloud sector. We then turn to a review of what we can learn, and what we cannot learn, from the literature which explores the consequences of switching costs, network effects, data, and pricing.

In the concluding section, we advise regulators and policymakers to proceed with caution when taking decisions related to stakeholders in cloud computing, a relatively unexplored sector. Furthermore, we advocate for conducting rigorous research using the economic toolkit at the disposal of our colleagues to better inform such decisions.

1 Cloud computing

Cloud computing is the provision of computing services offered over the Internet ("the cloud"). The most important services are the use of servers, storage, databases, networking, software, and data analytics. The following definition, from the UK's Information Commissioner's Office, is useful to introduce some of the economic concepts that will be used in this report:

> "cloud services are digital services that enable access to a scalable and elastic pool of shareable computing resources."[5]

The terms "scalable" and "elastic" refer to computing resources that are flexibly allocated by the cloud provider to accommodate demand fluctuations. The term "shareable" specifies that such computing resources are provided to multiple users who share a common access to the service, but that the processing is carried out separately for each user.

Another important definition was proposed in 2011 by the US National Institute of Standards and Technology (2011):

> "cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.".[6]

[5] https://ico.org.uk/for-organisations/the-guide-to-nis/digital-service-providers/?q=DSP.

[6] https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf.

Cloud computing, therefore, provides on-demand services. In particular, the customer buys access to computing resources such as hardware and software that are available on a scalable and elastic basis, without requiring human interaction with each service provider. These resources are provided via a network and accessed through public internet or private connection. Given its characteristics, cloud computing is a key driver for digital transformation, as it can improve access to advanced IT solutions, boosting innovation and productivity across a wide range of sectors.

From an economic point of view, this is a further step in the "platformization" of the digital economy; as we will see, the cloud service providers behave in many ways as platforms, and this change requires the use of the most recent tools of economic analysis to be studied.

The users of the cloud are firms — there is little direct sale to individuals. Among these firms, it is sometimes useful to distinguish two types of users, who differ in terms of when they adopted the cloud: digital natives and migration businesses. Digital natives are companies that from their inception used the cloud for their computing needs and, therefore, do not have to, or did not have to, adapt their IT system to the cloud. By contrast, migration businesses decided to use cloud services at a later stage, after relying on in-house servers or offline services. However, this distinction can become blurred in practice, as many customers with existing IT infrastructure also possess significant digital-native workloads that were developed in the cloud and never resided on-premises. Furthermore, when existing workloads are migrated to the cloud, they must often undergo significant restructuring to optimize their performance and take full advantage of the capabilities of the cloud.

1.1 Service models

Cloud business or service models are all built on top of cloud computing, and they are traditionally classified according to three standard definitions:

1. **Software as a service (SaaS).** This is the most outsourced model, as it provides a complete software solution that can be purchased on a pay-as-you-go basis. Customers rent the provider's applications that run on a cloud infrastructure which can be accessed through the Internet, most often through a mobile app, sometimes through web browser. For instance, a Customer Relationship Management (CRM) software such as Salesforce is a SaaS, but so are many consumer-oriented SaaS solutions, such as Netflix, Spotify, and banking apps.

2. **Platform as a service (PaaS).** The customer does not control the cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications; as Microsoft puts it: "You manage the applications and services you develop, and the cloud service provider typically manages everything else.".[7]

   In recent years, cloud service providers have offered versions of PaaS that allow their customers to abstract more and more from the management of the infrastructure. Serverless and Function as a Service (FaaS) architectures enable users to have granular, flexible, and increasingly elastic access to computing resources.[8] For our purposes, the precise definition of these models is not important. However, as we discuss below, they may have consequences for the cost and ease of switching providers.

3. **Infrastructure as a service (IaaS).** This is the least outsourced model, and it involves computing resources being supplied by a cloud service provider. The user does not manage or control the underlying cloud infrastructure, but has control over operating systems, storage, and deployed applications.

The exact boundaries between each of these three service models are not clear-cut; they are evolving, and some services do not fully fit into these models. Our discussion below applies to IaaS and PaaS and is less relevant to SaaS.

1.2 Deployment models

Cloud services are also distinguished by their "deployment models".

1. **Private Cloud** is usually adopted by a single business or organization. It mainly consists of computing resources tailored to the specific needs and preferences of individual customers or group of customers. The choice of this model can be dictated by the fact that the customer wants more control, security, and customization of the resources than public cloud services can provide. There can be regulatory reasons why some organizations must use private clouds.

[7] `https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-paas.`

[8] See `https://www.bmc.com/blogs/serverless-faas/.` for an introduction to these concepts.

2. **Public Cloud** is mainly operated by third-party managed platform and it is the most common cloud deployment model. Cloud services are open to all customers willing to pay and computing resources are shared between them. Customers normally use public cloud services for less-sensitive applications and for storing data that do not require frequent access. Moreover, their demands may vary over time. Some public cloud resources are available for free, while others require either a subscription or are on a pay-as-you-go system.

3. **Hybrid Cloud** is a combination of the previous two typologies which involves a mix of on-premises, private cloud and third-party public cloud services. Workload is typically shared between the on-premises data center and the public cloud. This can give the customer greater flexibility and more data deployment options, but it requires work to ensure compatibility between the different environments.

4. **Multi-cloud** defines the situation in which customers purchase cloud services from more than one supplier. In the terminology that we will use in this report, the main difference between hybrid and multi-clouds resides in their architecture: a hybrid cloud comprises a mixture of private and public cloud services, whereas a multi-cloud model is based on the combination of two or more public cloud services (some authors include what we call hybrid cloud in multi-cloud). Customers using multiple public clouds may easily access their preferred services and gain some bargaining power against their cloud providers. We understand that most organizations of any significant size use a multi-cloud model.

1.3 Not only about computing

On 15 October 2020, all the EU member states signed a declaration on "Building the next generation cloud for businesses and the public sector in the EU",[9] which stated[10]

> "Cloud computing enables data-driven innovation and emerging technologies, such as 5G/6G, artificial intelligence and Internet of Things. It allows European businesses and the public sector to run and store their data safely, according to European rules and standards."

[9] `https://ec.europa.eu/newsroom/dae/redirection/document/70089`.
[10] `https://digital-strategy.ec.europa.eu/de/node/362/printable/pdf`.

The cloud is therefore not only about storage capacity and computing power. Cloud providers also offer database and innovative services such as artificial intelligence (AI) and machine learning tools. The cloud has indeed the potential to create an array of benefits for both companies and the public sector that can propel innovation and generate economic growth.

Cloud computing profoundly influences strategic sectors such as telecoms and broadcasting. Telecom providers experience substantial advantages through the automation and seamless implementation of updates provided by public cloud services. Broadcasters benefit from the flexibility to operate across diverse delivery networks, formats, and viewing devices. They also benefit from the extensive and scalable data storage and processing capabilities offered by the cloud. In the public sector, cloud services play an increasingly pivotal role, not just in meeting data storage needs but also in enhancing the operational agility of national and local governments. For instance, in May 2021, the French government announced[11] a "National Strategy for the Cloud".[12] One of its three pillars was "The digital transformation of public services: a new «Cloud at the center» policy", which was described as follows:

> "With the adoption of the «Cloud at the center» doctrine, the French government makes the Cloud a prerequisite for any new digital project within the public sector, so as to accelerate the transformation of the public sector for the benefit of users ...
>
> The Cloud now becomes the default hosting method for the public sector digital services, for any new digital product, and for products undergoing a substantial evolution."

The opportunity to save money by reducing their expenditure on in-house data centers and resourcing also has beneficial consequences in terms of competition. Renting servers and computing capacity lowers entry costs. Startups especially benefit from the improved scalability and flexibility provided by cloud computing, which allows to access new technologies that were only available to larger companies. This, in turn, can lead to new products and innovations, from which the entire economy may benefit, as explained by Etro (2009). According to Impink (2022), startups that outsource IT to a cloud platform develop superior data analytics capabilities, which are linked

[11] The documents pertaining to the "Stratégie Nationale pour le Cloud" are in French. The translations are ours.

[12] `https://www.numerique.gouv.fr/uploads/Strategie-nationale-pour-le-cloud.pdf`.

to greater product differentiation and startup growth.[13] Furthermore, additional jobs are required and new competences can be developed, creating the conditions for sustained economic growth. DeStefano, Kneller and Timmis (2020) find that the adoption of cloud services leads to more employment and revenue, especially for young firms. Regarding its environmental impact, Park, Han and Lee (2023) offer empirical evidence of the positive influence of cloud computing adoption on energy efficiency.

Cloud computing is also essential for the development of the Internet of Things (IoT), which broadly refers to the interconnectedness of various devices or objects over the Internet. On the consumer side, examples of IoT applications are smartwatches, smartphones, smart home devices or self-driving cars. It has also many industrial applications, such as machine to machine communication; as a consequence, there is much European interest for IoT as it will be an important leverage for the development of "industry 4.0". In all of these fields, cloud computing provides the tools and services needed to create and maintain applications, but provides also the required connectivity. Cloud computing is therefore not only important in order to store IoT data, but also to enable systems to be automated in a cost-effective way.

Finally, the development of cloud computing has become inextricably intertwined with that of Artificial Intelligence, particularly generative AI or foundation models. For instance, without the computation power of the public cloud, OpenAI would not have been able to develop its foundation models and resulting applications, including ChatGPT. We do plan to explore the issues that this raises in future work, but will,essentially, not discuss it in the current report.

1.4 The growth of the cloud and its most important players

It is beyond the scope of this document to present a complete history of the development of the cloud industry. However, to understand the analytical and policy conundrums that it creates, it is important to be aware of both its very rapid expansion in the less than 20 years of its existence, and of its increased concentration in recent years. We comment briefly on these aspects at the end of this subsection and at different points in this document.

The term cloud computing appears for the first time in an internal Compaq document in 1996. However, it is generally accepted that the modern-day cloud was invented by Amazon in the years 2002 to 2006, as it realized

[13] Chen, Guo and Shangguan (2022) demonstrate the positive effect of cloud computing on firm performance.

that during much of the year it had unused computing capacity that could profitably be rented to other firms. Amazon Web Services (AWS), the subsidiary of Amazon which provides on-demand cloud computing services, was launched in 2006. There was no real competition for many years,[14] and AWS remains the largest cloud service provider worldwide.

Microsoft's cloud computing platform, Azure, is AWS's closest competitor and has significantly grown its market share since entering the market in 2010.[15] The variety of services it offers has increased substantially over the years, and today Azure is a major contributor to Microsoft's profits, much in the same way than AWS is a major contributor to Amazon's profits. Overall, AWS seems to be favored by small businesses and start-ups, whereas Azure is preferred by larger, more installed businesses.

Google Cloud Platform (GCP) is the main challenger to AWS and Azure. Google announced App Engine in 2008, marking its entry into cloud computing services. App Engine provided a platform for developing and hosting web applications in Google-managed data centers. The service became generally available in November 2011. Although its market share has grown in recent years, GCP remains significantly smaller than the two market leaders.

According to estimates from Synergy Research Group,[16] as of the second quarter of 2023, AWS has a 32% market share of worldwide cloud infrastructure while the next two competitors, Azure and GCP, have respectively 22% and 11%. The total market share of the three "hyperscalers", as they are often, and not entirely affectionately, called, was therefore 65%. (In the rest of the report we will follow the general usage and use that term.) This was the same aggregate market share as in 2022, but a substantial increase over the middle term: their aggregate market share was 61% in 2021 and 50% in 2017.[17]

Alongside the three large firms, there exist smaller cloud providers. Some of them are large firms, such as Alibaba, IBM, and Oracle, which have a relatively small footprint in the cloud service industry.[18] Others are spe-

[14] In an interview at the Economic Club of Washington in 2018, Jeff Bezos said that they "faced no like-minded competition for seven years. It's unbelievable." https://www.youtube.com/watch?v=zN1PyNwjHpc&t=2975s.

[15] It was officially launched as Windows Azure in February 2010 and renamed Microsoft Azure on March 25, 2014.

[16] https://www.srgresearch.com/articles/quarterly-cloud-market-once-again-grows-by-10-billion-from-2022-meanwhile-little-change-at-the-top.

[17] https://www.srgresearch.com/articles/q3-cloud-spending-up-over-11-billion-from-2021-despite-major-headwinds-google-increases-its-market-share.

[18] Alibaba currently holds the fourth position with a market share of 4%, while

cialized smaller players. A leading example is OVHcloud, the largest European provider of cloud services, with 1.4 million users worldwide.[19] Another important European cloud provider is IONOS, headquartered in Germany. Created in 1988, first specialized in web hosting and associated services, it now offers cloud computing services and has become an important European player.

These European firms are relatively small when compared to the hyperscalers, and their market share, even within Europe, is decreasing.[20] The EU is contemplating whether it would be more appropriate to support a "national champion" or a consortium of suppliers. Both solutions are under consideration, although more concrete initiatives have been taken in support of the former. The European Investment Bank, for example, recently supported the growth of European cloud leader OVHcloud with a €200 million loan for investments in Europe.[21] This loan is part of a series of efforts by EU members to strengthen Europe's cloud infrastructure, and it will assist OVHcloud in building 15 new data centers by the end of 2024, opening up possibilities for sustainable development. As per the latter solution, it is noteworthy to cite the European Alliance for Industrial Data, Edge and Cloud, which aims to "strengthen the position of EU industry on cloud and edge technologies. It aims to serve the needs of EU businesses and public administrations that process sensitive categories of data, and has the objective to increase Europe's leadership position on industrial data".[22] Compared to other initiatives, this alliance is initiated by a public authority, which gives it a certain legitimacy, even though its role has so far been rather limited.

Although we have not seen it discussed much in the literature, there does exist a third category of cloud service providers: specialized clouds. CoreWeave[23] specializes on running massive computations on NVIDIA chips and seem to be faster and cheaper than the hyperscalers for such task as

IBM and Oracle occupy the sixth and seventh positions, respectively, each with a market share of approximately 2%, according to estimates from Synergy Research Group: https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/.

[19]The company was founded in 1999 to provide dedicated servers and other web services. In 2019 OVH adopted OVHcloud as its public brand name.

[20]Figure 1 on page 19 of European Alliance for Industrial Data and Cloud (2023), drawn from Synergy Research Group data, shows the local market share of European Cloud providers dropping from 27% to 13% between Q1 2017 and Q2 2022.

[21]https://www.eib.org/en/press/all/2022-504-france-la-bei-soutient-le-developpement-du-leader-europeen-du-cloud-ovhcloud-avec-un-pret-de-200-millions-d-euros-destine-a-ses-investissements-en-europe.

[22]https://digital-strategy.ec.europa.eu/en/policies/cloud-alliance.

[23]https://www.coreweave.com/.

estimation of AI models and video rendering. This is clearly a marginal phenomenon at this point in time, but offering specialized services could be an entry point for future, more generalist, competitors.

1.5 The cloud ecosystem

In this report, we stress the economics of the core of cloud services: the provision of computing power to its customers. However, it is important to note that this core is actually a platform on which a number of auxiliary activities have sprung.

At the outset of the cloud, firms rented virtual servers on the cloud, and used on these servers software which they could also have run in-house. More and more, they purchase "managed services", *i.e.* computing services which have been designed to work directly on the cloud, and for which they do not have to manage or install the underlying software. This can range from services which help them manage their data, such as Snowflake or Databricks, to more specialized services such as database management tools, Internet of Things technology, AI compute, and many, many others. These services are provided by independent firms, but also by the cloud providers, often in competition with each other. These tools can be bought either directly from the vendors or through online stores hosted by the cloud providers. There are also suppliers of services which help their clients to manage their relationship with the cloud service providers, both from a technical point of view and for cost reduction; Datadog is a leading example.

2 The economist toolkit for studying the economics of the cloud

In the previous section we discussed the technical characteristics of the functioning and deployment of cloud computing services. In this section, we analyze the characteristics of the cloud that are relevant in the light of the economic analysis that we suggest in this report. We also want to specify the role of cloud computing within the economics of digital industries, which are characterized by innovation, increasing returns to scale, intellectual property rights, switching costs, data, network effects, and two-sidedness, among others. The economics of cloud computing is characterized by similar features, but their application to this sector requires some adjustments. Moreover, there are business practices that are specific to cloud computing, such as egress fees, that have to be properly evaluated.

2.1 Switching costs

Switching costs are the costs incurred by customers when changing the supplier of a service or product. For individuals, these costs are not limited to monetary costs but also include psychological costs, time loss, learning, and effort-based costs, and more. For firms, which are more relevant in the case of the cloud, in addition to monetary costs, there can be costs of adapting production processes to the new inputs and of retraining employees. It seems intuitive that switching costs are an impediment to competition: they limit the incentives of customers to purchase from competitors in response to an increase in price by their current suppliers. For this reason, competition authorities and regulators are concerned about switching costs. We will see later, in 4.2.1, that the economics of switching costs is actually quite complex, but in this section we will simply describe these costs and highlight their particularities in the case of the cloud.

Switching from one cloud service provider to another can be expensive and difficult, (although Jin, Peng and Wang (2023) provide, to the best of our knowledge, the only attempt to measure these costs in the economic literature — we discuss this paper in more detail in 4.1.3). One reason is that different cloud providers have distinct interfaces and techniques, which implies that some of the customer's software needs to be modified. Another reason is that few engineers are proficient in multiple clouds, which makes the modification process hard and may require extensive retraining. Moreover, learning new features may take time. For instance, Netflix has relied on AWS since 2006 for most of its computing and storage needs, such as databases, analytics, recommendation engines, and more.[24] AWS enables Netflix to deliver billions of hours of content every month to customers worldwide. Software engineers at Netflix have been trained and have become experts in AWS. Moving to a different cloud provider would involve significant retraining costs, which are hard to estimate but can be safely assumed to be quite high.

All of this would be quite standard and could be analyzed through the standard, well-developed, economic theory of switching costs. However, a new phenomenon is prominent in the case of the cloud industry, which, to the best of our knowledge, has not been studied in the economic literature: the switching costs are dependent, in part, on the choices made by the users.[25] All cloud providers offer a range of products, ranging from open-

[24] `https://aws.amazon.com/solutions/case-studies/innovators/netflix/`.

[25] A similar phenomenon certainly arises in other industries where switching costs are important. However, we are not aware of any formal analysis of this phenomenon in the

source software to proprietary solutions which are specific to its platform. This has important consequences for the cost of switching. For instance, a user who mostly uses IaaS services, installs virtual Linux servers and uses open-source databases, will find it easier to change suppliers, especially if it deploys solutions such as Kubernetes, which are designed to make software more portable. Indeed, Argonaut, a start-up which helps its customers manage their workflow in the cloud, provides the following advice to start-ups (cloud credits are free access to cloud computing resources): "Use open-source solutions where possible and third-party data stores so you can step away from the cloud provider and leverage another year or two of free credits from another cloud provider."[26,27]

On the other hand, a user which makes intensive use of proprietary services, such as serverless computing, will find the cost of migration higher. For the purposes of this document, we will therefore distinguish between two types of switching costs:

- *Exogenous switching costs*: those switching costs which are intrinsically due to difference of technologies of the cloud providers and are technologically given.
- *Endogenous switching costs*: those costs which are influenced by the technological choices specifically of the users.

As we will discuss later, endogenous switching costs create externalities between users: the choice of technologies with low switching costs by some users may induce cloud providers to lower their prices, benefiting other users.

Of course, the cloud service providers themselves also affect switching costs, in two ways. They can develop technologies that are more or less compatible or adhere more or less to public standards. Introducing new technologies can also have a side consequence of making switching more difficult for the customers who choose to adopt them. In the other direction, cloud providers offer tools and software to ease migration to their own platform, and to decrease its cost: AWS does so with "AWS Application Migration Service", and Azure with "Azure Migrate". As these strategies, and the responses of the users, can be analyzed with existing tools in economic theory, we will spend little time discussing them.

economic literature. A quick look at the managerial literature on switching costs indicates that it concentrates on providing advice to increase customer retention.

[26] `https://www.argonaut.dev/blog/cloud-credits-startup-guide`.

[27] The Central Digital and Data Service of the UK government publishes an useful webpage which provides recommendations to cloud users on the management of switching costs. `https://www.gov.uk/guidance/managing-technical-lock-in-in-the-cloud`.

The issues of interoperability and portability, which will be discussed in 2.3, are also highly intertwined with switching costs, which can be substantially lowered when cloud services can seamlessly integrate with other products or systems of rival cloud providers, and when customers can easily migrate workloads, applications and data between cloud providers at relatively low costs. Large companies often use multiple clouds for different services, and therefore switching to select the best tools becomes easier for them if such services are both interoperable and portable.

2.2 Pricing and egress fees

The pricing of cloud services is quite peculiar, extremely complicated, and not easily comparable to that of other similar services. The price list for cloud services is usually published on each provider's website. Contrary to traditional IT services, which require a one-time upfront payment for licenses, cloud services are priced according to a pay-as-you-go model. Users have a lot of choices available, and must manage a complicated array of possibilities. All of this would certainly deserve a deep dive to understand it better. We will not pursue this path, but rather concentrate on the "egress fees", which have been the focus of much attention from the regulators.[28]

Egress fees refer to the price that users must pay to move their data out of the cloud. Traditionally, users did not have to pay anything to upload data to the cloud, but cloud providers imposed exit or egress fees to move data out of the cloud.[29] This has been deemed anti-competitive by some observers and some competition authorities, and the industry is evolving on this point, both on its own volition and pushed by regulators.[30] However, we still find it worthwhile to understand better this practice, as it could provide important lessons about the economics of the industry.

Transferring data from a cloud provider's network to another network entails a bandwidth cost. However, cloud providers have typically included the possibility to egress a certain amount of data for free (usually the first 100GB). Above this threshold, a fee is applied, which increases for low volumes of data transfer, and then decreases for higher volumes.[31] Moreover,

[28]Smaller cloud provider services have also complained about the "cloud credits" offered by the hyperscalers. We discuss them in 3.1.2.

[29]The interested reader will find the pricing of Azure for "bandwidth" at `https://azure.microsoft.com/en-us/pricing/details/bandwidth`.

[30]For instance, as we discuss in 3.1.1, the European Data Act restricts the use and size of egress fees.

[31]See in particular Figure 5.4 on page 121 of Ofcom (2023), together with the discussion in para. 5.122 on page 120.

certain customers have negotiated private discounts with cloud providers, adding complexity to discussions about the magnitude of egress fees. For all these reasons, the presence of such fees appears closely linked to the issue of switchability, especially with the creation of endogenous switching costs, that we introduced in 2.1. Relatedly, it is not even clear what are the distortions created by egress fees, and what could be the effect of removing them. In 3.1.1 we discuss whether egress fees are deemed as anti-competitive or not, and the decisions that regulators and policy-makers have taken.

2.3 Interoperability and portability

Interoperability and portability are crucial for organizations that seek flexibility and look for the best fit solutions in the dynamic landscape of cloud computing. However, it is important to keep the distinction between the two clear. Interoperability involves repeated communication between services of different providers, whereas portability is related to a one-time transfer of data, applications and workloads for a user who changes supplier.

These two concepts play an important role in the analysis of multi-cloud strategies and of switching, as we introduced in 2.1. Other things being equal, in an open digital ecosystem, users can switch to services of the highest quality or most competitively priced services seamlessly. This creates a sustainable form of competition that contributes to competitive prices, quality and innovation. To achieve this situation, improving interoperability and data portability is of great importance. (However, mandating interoperability for new products can, of course, preclude innovation.)

On the one hand, regarding portability, an interesting definition can be found in the Gartner Report (2023): “A cloud-based application is portable if its full life cycle can be moved from one cloud provider’s environment to a different provider’s environment. [...] In other words, you are not simply porting the application. You are porting the application’s entire technical and organizational environment, including all associated tools, which may be quite complex and multifaceted, as in the case of the organization’s value stream delivery platform”.[32] Portability therefore makes it easier for customers to switch among cloud providers to find the best solution for their business. This also reduces the risk of lock-in, where a customer becomes heavily dependent on a single cloud provider.

On the other hand, as multi-cloud becomes more diffused, with companies adopting two, three, or more clouds, demand for interoperability is

[32]See page 4 of this report.

growing, resulting in some strategic partnerships.[33] For instance, Microsoft and Oracle announced in 2022 a partnership to facilitate interoperability across their respective cloud services. Poor interoperability, on the contrary, would reinforce lock-in effects at the tasks level. Users would then be required to use the same provider or a third-party service running on the same cloud infrastructure, making it more difficult to migrate an individual task to a competing cloud. This would essentially turn migration into a "whole-or-nothing" proposition, presumably making it more difficult. Kubernetes, an open-source container orchestration system for automating deployment, scaling, and management of containerized applications, plays an important role in enabling interoperability and multi-cloud.

2.4 Network effects

Direct network effects occur when the value of a service to a particular user increases when more users join the service.[34] There are two types of direct network effects. One-sided network effects arise when users are all of the same "type": for instance, in a telephone network or in a social network, every user will potentially connect with every other user. There are two-sided network effects when the users can be allocated into two categories (*i.e.,* sides) and the users of one category (*i.e.,* side) connect with the users of the other: this is the case in marketplaces, for instance, where consumers form one side and do not interact with each other, but interact with the merchants which form another side.

It is also customary to distinguish direct and indirect network effects. The examples in the previous paragraph are representative of direct market effects. Indirect market effects arise when the influence of the presence of a user on the other user is not caused directly by the presence of that user. There are often indirect network effects in two-sided networks. A buyer on a marketplace does not directly benefit from the presence of more buyers, but benefits indirectly if these additional buyers attract more sellers, thus providing for a more diversified offer of products for sale.

Of course, as the economic literature has extensively analysed, network effects of all types favor the creation of large firms and furthermore protect incumbents as users find it difficult to coordinate migration to potentially "better" networks. We do not know of any evidence on the size of network

[33] According to Flexera (2023), 87% of companies embrace a multi-cloud strategy (see in particular page 18); `https://info.flexera.com/CM-REPORT-State-of-the-Cloud`.

[34] Throughout, we assume that all network effects are positive, while in some markets there maybe negative network effects.

externalities, and it should be a priority of empirical work to try to find ways to measure them. We will therefore just offer the following very preliminary thoughts.

Ofcom (2023, para. 6.77, p. 185) identifies some direct network effects: some customers asserted that they prefer to use cloud providers that are "more popular amongst other users in their stakeholder group or supply chain". However, the evidence they provide stems from one firm which explained that the reason was "they would not be paying egress fees when exchanging data/content with them". For analytical purposes, we feel it is quite important to reserve the terms network externalities for fundamental feature of the industry, not for benefits which stem from the tariffs charged by the suppliers. There may be other reasons why direct network effects occur, which align with our more "fundamental" definition. First, for some application, the speed of transfer of data between two users might be crucial, and this speed could be improved if the two users belong to the same cloud.[35] It could also be that custom software would more easily be shared by clients of the same cloud service providers.

There are clearly some two-sided network effects. For instance, the service provider with a larger number of users will have more third-party services in its marketplace than a provider with relatively fewer users. According to different studies, AWS and Azure have both twice as many third-party products in their marketplace than GCP.[36] This obviously attracts more users, which in turn attracts even more services from third parties and therefore more revenues. The size of these effects is very difficult to estimate.

Many authors and industry participants point to a form of indirect network effects. The market for engineering competence is "thicker" for clouds with large installed bases. This effect is particularly pronounced as it seems that few engineers are competent in the use of several clouds.

Choosing cloud providers that are popular within an industry may be important for start-ups, especially those considering future acquisition. Potential purchasers may view favourably companies using the same cloud provider in order to minimize post-acquisition integration costs. However, it is worth noting that much of the software may need to be rewritten after an acquisition, mitigating this aspect to some extent.

Finally, we should mention that services such as Snowflake and Rescale[37]

[35] The only evidence we have for this claim is anecdotal. A supplier of spatial images told us that the use of its images was easier and more efficient for clients hosted by the same cloud.

[36] See, for example, page 48 of ACM (2022).

[37] Snowflake partners with AWS, Azure and Google Cloud `https://www.snowflake.`

add a layer of abstraction "on top of" the cloud service providers. Because these services function on top of the three hyperscalers, they presumably provide network externalities that favor them.

We should stress once again that everything discussed in this section is very preliminary, and it is crucial that empirical research be conducted to disentangle these effects.

2.5 Standards and open source

Cloud services usually follow certifications in which they meet specific security and compliance standards, especially in terms of security. The Cloud Security Alliance[38] (CSA) promotes the use of best practices for providing security assurance within cloud computing, and its security standards are usually adopted by large cloud providers. However, there exist standards in cloud services without certification, as in the case of some software/technical standards. Open source standards are not usually regulated by a specific entity.

There is a very close relationship between collaboration among cloud computing providers and open source. For instance, Amazon, Google and Microsoft are large contributors to the technology of Kubernetes, which we discussed in 2.3. On the one hand, this can promote standardization and compatibility, but on the other hand, it could theoretically be an instrument for collusion, although we have seen no indication that this is the case.

The issue of standardized cloud technologies is also strongly associated with interoperability, and the lack of standards could hinder switching and multi-cloud technologies. A support for existing industry standards and open-source software may not help improve their availability and quality but also improve interoperability and portability in cloud services. Setting technology standards has led to increased interoperability and portability in various sectors, for example in the telecommunications industry. However, the economics of the adoption of standards is quite complex, with many trade-offs.[39] European regulations encourage the development and the use of standards.

An additional key issue involves compatibility, in particular whether it is feasible for entrants to achieve one-way compatibility with an established

`com/`; Rescale adds Oracle Cloud Platform to this list `https://rescale.com/`.

[38] `https://cloudsecurityalliance.org/`.

[39] Although it seems intuitive that standards improve compatibility and therefore are efficiency-inducing, it is easy to develop models where they also reduce competition and, as a consequence, welfare. There is also some empirical evidence that this *can* be the case.

standard.[40] With the network advantage enjoyed by the incumbents, entry may not be possible without compatibility with one of the incumbent platforms. This strategy is especially relevant in the cloud ecosystem because compatibility is easier to achieve and less costly.

3 Policy initiatives in the cloud sector

An evaluation of the level of competition in the cloud sector is a difficult task, given the nature of the sector, and the very rapid changes it has incurred. On the one hand, there is potentially a lot of competition, at least if we consider the number of players involved. On the other hand, the degree of concentration is very high, as seen in 1.4. Several factors contribute to explaining the current situation, including substantial fixed costs required to enter the market, increasing returns to scale, the industry's high level of technicality, and switching costs. Large players may also benefit from network effects, as we discussed in 2.4. These factors exogenously reduce market contestability and are probably the main explanation for the oligopolisitic nature of the industry. On the other hand, competition authorities and regulators have, as they are supposed to, examined the strategies of the hyperscalers to identify any anti-competitive element which would have reinforced their market power, for instance by raising switching costs or diminishing portability and interoperability.

Of particular interest for the purpose of this section are the market studies conducted by the Netherlands Authority for Consumers and Markets (ACM, 2022), the French Competition Authority (FCA, 2023), the Japan Fair Trade Commission (JFTC, 2022), and the UK's Office of Communications (Ofcom, 2023).[41]

The first three market studies were issued by national authorities, whereas the fourth is a market study into the supply of public cloud infrastructure services in the UK prepared by Ofcom, the regulator and competition authority for the UK communications industries. Ofcom has concurrent functions with

[40] One-way compatibility means that the software written for the incumbent technology can be used on the entrant's technology.

[41] In the US, the Federal Trade Commission (FTC) issued a request for information (RFI) on March 2023, focusing on the competitive dynamics of cloud computing and on data security. The RFI was completed on June 2023 and at the time of writing this report the FTC was still evaluating the response received. The interested reader can find more information here: `https://www.ftc.gov/news-events/news/press-releases/2023/03/ftc-seeks-comment-business-practices-cloud-computing-providers-could-impact-competition-data`.

the Competition and Market Authority (CMA), the competition regulator in the United Kingdom, to whom it can refer to carry out a proper market investigation.[42]

The fifth report was written by Professor Frédéric Jenny for CISPE, (CISPE, 2021), which defines itself as "a non-profit trade association for infrastructure as a service (IaaS) cloud providers in Europe" which was created to "promote data security and compliance within the context of cloud infrastructure services".[43] Its membership include about two dozen small and medium-size European firms active in the cloud sphere as well as AWS.

Finally, the last report that we include was written by Professor Daniel Schnurr for the Centre on Regulation in Europe (CERRE, 2022), whose membership is very diversified: it includes regulatory authorities, firms active in different industries, including all the "GAMAM", as well as the Toulouse School of Economics! This report is part of a larger project entitled "Improving the Data Act".[44]

These reports are in general well-detailed and represent important progress for the understanding of the cloud sector. However, they focus on explaining why some practices *could* be anti-competitive but do not provide a complete economic analysis, either of these practices or of the industry as a whole.

Two recent European regulations bear on the cloud industry. The Digital Markets Act, which came into force on 1 November 2022, lists cloud computing services as "core platform services" which are subject to the act. However, for procedural reasons which are not germane to our discussion here, no firm was designated as a gatekeeper for that industry. We understand this is due to legal considerations. More important from our viewpoint is the recently adopted Data Act, adopted by the European Council on 27 November 2023, which imposes a number of obligations on cloud services. In this report, we will only consider the rules aimed at lowering the cost of switching — that is the rules concerning egress fees and interoperability.

3.1 Anti-competitive concerns

In this section we discuss the practices of cloud providers that competition authorities and other commentators have highlighted as being problematic.

[42] Ofcom can commission market studies related to commercial activities connected with communications if they believe such activities may be detrimental to consumers.

[43] `https://aws.amazon.com/blogs/security/aws-announces-cispe-membership-and-compliance-with-first-ever-code-of-conduct-for-data-protection-in-the-cloud/`.

[44] More information can be found here: `https://cerre.eu/publications/data-act-towards-a-balanced-eu-data-regulation/`.

Title			
Market Study Cloud Services	Netherlands Authority for Consumers and Markets	ACM (2022)	Competition problems caused by barriers to switching and poor interoperability
Report on Trade Practices in Cloud Service sector	Japan Fair Trade Commission	JFTC (2022)	Survey of trade practices and competition in the cloud sector; focus on conducts that might restrict competition.
Market Study on Competition in the Cloud Computing Sector	French Competition Authority	FCA (2023)	Analysis of practices that could restrict competition; focus on egress fees and cloud credits.
Cloud Services Market Study: final report	UK's Office of Communications	Ofcom (2023)	The degree of competition in cloud infrastructure services; focus on main barriers to switching: egress fees, technical barriers, and committed spend discounts.
Report on Cloud Infrastructure Services	Cloud Infrastructure Services Providers in Europe	CISPE (2021)	Potentially anti-competitive practices of dominant cloud providers, such as bundling, self-preferencing, and licensing practices that increase switching costs.
Switching and interoperability between data processing services in the proposed Data Act	Centre of Regulation in Europe	CERRE (2022)	Analysis of the third part of the Data Act, which contains provisions to facilitate switching and interoperability between data processing services and data spaces.

Table 1: Most relevant market studies and reports.

It is definitely **not** our aim to adjudicate competition law debates. On the other hand, given that much of the policy discussion has centered over these debates, we feel it is important to discuss them. Our focus is entirely on the analytical issues that they raise and to point out the research directions which could help their resolution.

In what follows, we first identify which are these controversial practices and then evaluate the policy interventions that have been suggested. Our initial focus is on egress fees, likely the most prominent example, followed by an examination of other practices, including committed spend discounts, bundling, and license limitations.

3.1.1 Egrees fees

We introduced egress fees in 2.2. They have attracted the attention of regulators, who fear that they are anti-competitive, because they discourage switching between cloud service providers and multi-cloud strategies. In this section, we review these concerns.[45]

Before discussing how regulators have analyzed egress fees, it may be worthwhile to point out that European policy makers have jumped the gun and already started controlling them. Indeed, the European Data Act has a complex and layered approach to egress fees. After a three-year transition period, cloud service providers will not be able to impose "switching charges" when one of their customers migrates to a new service. In that context, egress fees are also forbidden (see Article 29, along with the definition of switching charges in Article 2(36)). On the other hand, when a cloud service is used "in parallel with" another cloud service, egress fees can be charged, as long as they do not exceed the cost of providing the service. We assume that this applies to circumstances where, for instance, in the normal course of business, a user transfer data between, let us say, two parts of its accounting system that are housed on different clouds. Finally, the Act does not regulate egress fees when they apply to the transfer of data to a third party or to one's own servers, as long as this is not part of a switching strategy.

However, the passage of the Data Act does not mark the end of the debate on egress fees. First, because there are many jurisdictions in the world which it does not affect; second, because the Data Act could be amended at some

[45] It is also the case that all the large cloud providers provide a free tier of service for new clients (see, for instance the "12 months free" offer of AWS, `https://aws.amazon.com/free/?all-free-tier`), and the smaller providers complain that they cannot compete in this dimension with the hyperscalers. As this issue has been relatively little debated, we will not discuss it.

time; third, because thinking about the form that egress fees took before the intervention of the regulator can help us understand the dynamics of competition in the cloud.

According to Ofcom (2023)[46] and the Dutch Competition Authority ACM (2022), there are indications that egress fees have exceeded the incremental cost of providing data transfer; Gans, Hervé and Masri (2023) confirm this from proprietary data from AWS. Moreover, Ofcom (2023) indicates that AWS, Azure and GCP charge egress fees which are 5-10 times higher than smaller rivals, such as OVHcloud and Oracle.[47] This was also supported by ACM (2022), according to which the fees for the three major providers are between 0.05 USD and 0.09 USD per GB, depending on the volume transferred,[48] whereas in the case of Oracle and OVHcloud, transporting data from the cloud costs 0.0085 USD per GB and 0.01 EUR per GB, respectively, regardless of the volume. Both Ofcom (2023) and ACM (2022) conclude that egress fees have a significant impact on the use of multi-cloud solutions and on the ability to switch between cloud providers. For what it's worth, the UK communications regulator has also launched a customer research study and found, not surprisingly, that most of the respondents were in favor of reducing or removing egress fees, even though only a few of them identified egress fees as the main barrier to switching.[49]

The French Competition Authority FCA (2023) also identifies egress fees as a major concern for the industry, as customers are often unable to anticipate their future needs in terms of data traffic and bandwidth usage. Given the way these fees are structured, they can increase the risk of customers being locked-in by making it more difficult for cloud users to leave their primary provider or to use several providers at once in a multi-cloud environment.

All in all, the common view that emerges from all these reports is that policy makers are concerned that egress fees substantially reduce the ability of customers to switch to a different cloud provider or to engage in multi-cloud strategies, and thereby induce inefficiencies. However, the lack of economic analysis in support of these claims is surprising, especially given the number of pages dedicated to this subject. A point which we want to reit-

[46]The answer to Ofcom consultation by the Centre for Competition Policy at the University of East Anglia (Ennis, Evans and Mariuzzo, 2023) provides an analysis of Ofcom (2023) which is very much in agreement with ours.

[47]See in particular Figure 5.4 on page 121.

[48]See Table 2 on page 58.

[49]The interested reader can find more information in Ofcom (2023): in particular, while 78% of respondents believed that egress fees should be reduced or removed (see para. 1.20 on page 7), only 6% identified these fees as the most relevant obstacle when deciding whether to switch or not (see para. 5.153 on page 133).

erate is that a serious analysis of the effect of egress fees is necessary before concluding that they are necessarily anti-competitive. Moreover, CERRE (2022) warns that eliminating egress fees for all cloud providers may disproportionately burden smaller players. Hyperscalers' portfolio activities may indeed enable them to better absorb foregone revenues resulting from the withdrawal of these fees.

We already mentioned that there is a substantial lack of research on the economic effect of egress fees. A partial exception is Gans et al. (2023), who investigates the effect of eliminating egress fees. They conclude that other fees can possibly increase as a result of this policy change, such as storage costs, thus penalizing customers. Moreover, the full elimination of all egress fess could lead to an inefficient transfer of data out of the cloud provider. Their analysis represents therefore an additional warning sign against the undesirable effects of the full elimination of egress fees. However, they rely on the very strong assumption, which we believe not warranted in practice, that the cloud service market is competitive and that, absent regulatory intervention, every element of the service is priced at marginal cost.

In the more general literature, recent papers theoretically investigate how to regulate platform fees, broadly suggesting to impose fee caps (Gomes and Mantovani, 2024; Wang and Wright, 2022; Bisceglia and Tirole, 2023). These studies can be taken into consideration when studying the socially optimal level of egress fees, as well as those of other tariffs in the cloud industry.

3.1.2 Committed spend discounts and cloud credits

Committed spend discounts define the situation in which a large customer agrees to spend a predetermined amount with a cloud provider in return for a percentage discount. If the customer does not reach the set amount, it pays the difference — for economists, this is an example of a very standard pricing strategy, two-parts tariff, which has been extensively studied. Committed spend discounts may benefit the cloud provider by helping to forecast future demand, thus enabling an appropriate investment in infrastructure. Interestingly, these discounts may also provide advantages for large customers as they may enable them to better negotiate prices with the hyperscalers. Ofcom (2023) reports that committed spend discounts are widely adopted by the three hyperscalers.

Cloud credits are the units of virtual currency required to perform certain tasks on the cloud, such as running a simulation in a cloud environment. They can be used to pay for resources such as storage, compute, and bandwidth, providing customers with an opportunity to explore and experiment the wide range of products and services available in the cloud. AWS credits, for instance, are automatically applied to bills to help cover the costs associated to AWS. Cloud credits are of real use and add value for many companies, especially startups, who can avoid substantial investments that could hamper their development, but also for cloud providers, who use them to spread and encourage adoption of their technology.

These instruments are therefore directed towards different types of customers. Committed spend discounts target large established customers with sophisticated purchasing habits and a deep understanding of their needs. Conversely, cloud credits are mainly designed for new customers, or to facilitate experimentation and trial by existing ones. In general, though both instruments may provide benefits in terms of lower prices and ease of payment, they can encourage some customers to use a single cloud provider, thus limiting the ability of smaller players to gain scale. Ultimately, this could restrict competition in the cloud computing industry, and for this reason they raised anti-competitive concerns.

The UK's communications regulator expressed particular concern about the negative effects of committed spend discounts for competition, especially for smaller cloud providers that do not offer the full range of products available from the hyperscalers.[50] The use of these discounts may indeed further raise barriers to entry for these players. However, in this case as in the case of egress fees, we notice a lack of economic analysis, and some of the responses gathered by Ofcom (2023) were ambiguous. For some customers, "the level of growth in commitment [for committed discounts] is not problematic as they are growing their cloud usage fast, or because they have bargaining power, they can use to mitigate the pressure."[51] For other customers the situation is different, and these discounts may create a barrier to multi-cloud. It would be appropriate to conduct an analysis of the costs and benefits for the different groups of customers.

The French Competition Authority (FCA, 2023) is more concerned about the use of cloud credits, which are considered together with egress fees as possible reinforcements of lock-in effects. Also in this case, however, the market study admits that there are circumstances where these credits may have a pro-competitive role, as they may help companies, especially startups, to enter the market. Indeed, by using such credits, companies may avoid bearing substantial investments costs. It is therefore necessary that, in order

[50]The interested reader will find the discussion on pages 142-152 of Ofcom (2023) particularly relevant.

[51]See para. 5.230 on page 151 of Ofcom (2023).

to benefit from these cloud credits, "as efficient competing cloud providers are able to offer them profitably".[52] It is not specified, however, how this should be implemented.

3.1.3 Bundling and licensing practices

Competition policy has long grappled with the issue of dominant market players leveraging their position to gain advantages in complementary goods markets. Indeed, a multi-product dominant firm can use bundling to foreclose access of a single product rival to one of the markets it serves (see, among others, Whinston, 1990; Carlton and Waldman, 2002; Nalebuff, 2004). This concern is echoed for the cloud market in reports by the UK's communications regulator (Ofcom, 2023) and by Frédéric Jenny (CISPE, 2021). Their analyses center on large cloud providers, particularly Microsoft, which, according to them, having established dominance in server software pre-cloud era, is now using potentially anti-competitive bundling and licensing practices. Oracle, with its strong position in database management software, but much lower market share in the cloud, faces less scrutiny in the cloud market, although some argue that its market strategies raise similar concerns (CISPE, 2021).[53]

The French Competition Authority (FCA, 2023) also points to the presence of clauses that can limit the options for customers to change provider by increasing migration costs, but it remains rather generic.

3.1.4 Technical restrictions on interoperability

While a degree of complexity is inevitable given the nature of cloud services, some technical restrictions may be engineered by leading cloud providers

[52] See FCA (2023), page 7.

[53] This report is not the place to discuss and even less evaluate these very complex licensing issues, all the more that much of the information is not generally available (for instance, a large proportion of the sources in Ofcom (2023) is redacted). However, it may interest the reader to have an illustration of the complexity of the issues. To the best of our understanding, traditionally many users bought licenses for on premises use; that is, they could run the software on their own servers or on "dedicated hosted cloud services", that is on servers rented to them by outside parties. For quite some time, Azure allowed the use of that software on cloud service providers. In 2019, Microsoft announced that this practice would no longer be permitted when the cloud service provider was Alibaba, Amazon, Google or Microsoft (unless the user had purchased a license that allowed for migration to other hardware). Transfer to smaller cloud service providers were unaffected. We will let the readers decide whether they think of this as positive discrimination with respect to small (mainly European) competitors, or as unfair to the largest competitors. Of course, this is but one of the allegations of anti-competitive practices in CISPE (2021).

to prevent some of their functions to effectively work with those of rival providers. Ofcom (2023) claims that such technical barriers are more significant than they should be and identifies a number of AWS and Microsoft cloud services with potential interoperability limits. Microsoft (Ofcom, 2023, page 104) responded that limited interoperability may, in some cases, be the natural result of innovation, given that cloud providers are continuously upgrading their services with the latest security enhancements and innovations. CISPE (2021) identifies different forms under which software providers create limits to interoperability. Amongst others, the most relevant are specificities of a technical nature, whereby software providers may employ operating proprietary language to reduce the ease of interaction between systems.[54]

Articles 30(1) and 35 of the European Data Act[55] set requirements for the technical aspects of switching between data processing services. Cloud providers must guarantee "functional equivalence" after switching. Recital 86, on page 69, states

> "Functional equivalence means re-establishing, on the basis of the customer's exportable data and digital assets, a minimum level of functionality in the environment of a new data processing service of the same service type after switching, where the destination data processing service delivers a materially comparable outcome in response to the same input for shared features supplied to the customer under the contract."

Indeed, one of the aims of the Data Act is to allow customers to easily switch their data and other digital assets between competing providers of cloud and other data processing services. It also aims to boost interoperability and provides safeguards on international data transfers. However, the report by Ennis et al. (2023) on the Data Act proposal of 2022 warned that the functional equivalence criterion might be challenging to operationalize in practice.[56] What the report questioned is whether the destination provider should guarantee the same performance as the original provider. This, however, may depend on multiple factors, not all attributable to the original provider, which should only be held responsible for making its best effort to ensure the maintenance of a minimum level of functionality in the desti-

[54]See Subsection 3.2, especially 3.2.2.2 on page 44; `https://cispe.cloud/new-study-links-unfair-software-licences-to-distortion-of-competition-in-cloud-infrastructure-market/`.

[55]The interested reader will find a much deeper dive into the portability and interoperability requirements of the Data Act in Ennis and Evans (2023)

[56]Similar concerns are also raised by Ennis and Evans (2023).

nation provider. The final text of the Data Act, adopted on November 27, 2023, has only partially embraced this perspective.

3.2 An overall view and other issues

The economics of the cloud are specific enough that we believe the practices and strategies covered in this section must be studied in the specific context of that industry. It seems to us that a priority is switching costs and egress fees. Committed spend discounts and cloud credits are probably less crucial factors in reinforcing lock-in effects.

The European Data Act contains various obligations on providers of cloud services to promote competition in the sector. The reports that we consulted and cited are more negative, and point to the existence of market features and excessive concentration caused by hyperscalers' decisions. Ofcom (2023) considers egress fees, restrictions on interoperability and committed spend discounts as barriers that are difficult to overcome, as they make unpractical for customers to change provider or use multiple suppliers. This may result in significant price surge and increasing difficulty for customers to access the best quality products. ACM (2022) highlights that discount structures may render even a partial switch to a competitor unattractive, and egress fees reinforce lock-in effects. FCA (2023) insists on reducing switching fees, imposing restrictions on use of cloud credits, and introducing obligations on cloud services providers to promote interoperability. Finally, JFTC (2022) includes an interesting survey that revealed that, following an hypothetical price increase of 5 to 10%, respondents would neither switch from a cloud service to on-premise service, nor to another cloud provider. For the Japanese Competition Regulator, it is then clear that a lock-in effect is present. However, as already pointed out in different parts of this report, we believe that more economic analysis is needed to substantiate these affirmations.

There are many other issues discussed in the policy reports which took center stage in our discussion above. We conclude by listing some of them, without comments or analysis, although they would deserve longer developments.

Turning to market concentration, Ofcom (2023) pointed to the high levels of profitability for the hyperscalers coupled with a gradual increase in market concentration. This reveals the presence of limits to the overall level of competition. JFTC (2022) also highlighted that the degree of market concentration is increasing. Apart from the "usual suspects" identified above (limits to interoperability and contractual restrictions, including egress fees), this report also indicates that the presence of economies of scale and scope, as well as strong indirect network effects due to an increase in related businesses, are important explanatory factors. Similar issues are evoked by FCA (2023), which concludes that hyperscalers benefit from economies of scale and product ranges offered in their ecosystems. They also have access to a preexisting and consolidated customer base that may enable them to take advantage of significant network effects.

In our analysis we did not consider the possibility that mergers and acquisitions increase the level of market concentration.[57] ACM (2022) warns that vertical mergers in cloud-related markets may increase market concentration. FCA (2023) suggests to competition authorities to be particularly vigilant when it comes to merger deals involving cloud providers, such as IBM's acquisition of software provider Red Hat in 2019.[58] An important point is that the acquisitions of companies with a relatively small turnover may not be a subject to merger oversight. The DMA, however, requires gatekeepers to inform the European Commission about these mergers, and this may intensify scrutiny.

4 Literature review

In this section, we first delve into the relatively restricted research conducted on the economics of cloud computing. Subsequently, we explore pertinent economic literature that aids in contemplating the economic challenges that emerge in the cloud market. Emphasis will be placed on elucidating potential interconnections among various economic concepts.

4.1 Cloud literature

The work specifically on the cloud can be divided into three areas: pricing, infrastructure, and switching costs.

[57] Dohan and Mariuzzo (2023) empirically investigate the effect of mergers on innovation in the cloud computing market. They initially find that leading firms tend to acquire young startups, whereas non-leading firms tend to acquire more established firms to gain market share. They then conduct an ex-post evaluation of how mergers in this market affect the innovation output – measured by patents – and show a positive impact of mergers on innovation.

[58] https://www.ibm.com/investor/articles/ibm-completes-acquisition-of-red-hat.

4.1.1 Pricing

As stated above, discussions among regulators and policy analysts have extensively covered the pricing structure of cloud services, particularly emphasizing the egress fees imposed by major cloud service providers. The prevailing view suggests that these fees act as a deterrent for users considering a switch between cloud service providers and are strategically employed by prominent incumbents to discourage migration to smaller competitors. However, it's noteworthy that there is a lack of comprehensive analysis regarding the specific levels of these fees, and no thorough attempt has been made to construct a model for determining their socially optimal level.

A partial departure from this general perspective is found in the work of Gans et al. (2023), who explore the potential ramifications of removing egress fees. Although they don't present a fully developed model, their analysis appears to operate under the assumption of a competitive market for cloud services. In such a scenario, compelling cloud providers to set egress fees below their marginal costs could introduce distortions in resource allocation and trigger increases in other fees. Furthermore, the complete elimination of all egress fees might result in an inefficient transfer of data out of the cloud provider. (As discussed in 3.1.1, we harbor skepticism regarding the assumption that egress fees are aligned with marginal costs; if they were, it would become crucial to comprehend why ingress fees are set below marginal costs and the distortions thus created).

Hummel and Schwarz (2022) explore regional disparities in cloud service pricing. Their study involves a model featuring a monopolistic provider operating in regions with diverse characteristics and customers facing uncertain demands for cloud services. The findings indicate that the monopolist will strategically set prices to generate proportionally less excess capacity in larger markets compared to smaller ones. Additionally, the monopolist will opt for lower prices in the larger locations. Intriguingly, the monopolist aims to attract more users to the larger locations, despite having less expected excess capacity. This strategy is driven by the desire for a more predictable demand at larger locations, akin to the logic of the law of large numbers, enabling better fine-tuning of capacity.

4.1.2 Cloud infrastructure

A few researchers have delved into the realm of cloud infrastructure. In Leka (2022), the functioning of the cloud is presented, along with a brief discussion of certain practices employed by cloud providers. In a dissertation at the Toulouse School of Economics, Lam (2015) conceptualizes cloud computing as a resource for firms facing capacity constraints. In her model, two user firms initially determine their individual capacity levels, followed by the cloud computing provider deciding on its capacity and pricing. Subsequently, demands are realized. The investigation focuses on how the correlation between users' demands and the cost of cloud computing infrastructure influences the investment incentives of both users and cloud providers. In a monopoly cloud provider scenario, users opt for an inefficiently high level of investment to alleviate the holdup problem associated with procuring from a high-priced monopoly provider during peak demand. Conversely, the cloud provider tends to select insufficient capacity relative to the social optimum. The cloud provider's capacity increases with demand correlation only if costs are sufficiently low, as the downside of excess capacity is minor compared to the substantial gains reinforced by demand correlation. In the case of two cloud providers, the holdup problem is partially alleviated. Notably, simulations demonstrate that investments by cloud providers increase with demand correlation, even when the costs of the cloud provider are significant.[59]

Impink (2022) examines the effects of different cloud infrastructures on the product differentiation among start-up firms. The cloud lowers the cost of launching a start-up by reducing the capital requirements. This encourages venture capital firms to fund more start-ups, as they face less risk if the start-up fails. However, relying on the cloud for most of the development reduces the flexibility of the start-ups and may lead them to produce more similar products, as they share the same cloud resources. This can also limit the startup's ability to switch to another cloud provider.

Impink also studies the consequences of the use of the cloud for the diversification of the software developed by start-ups. He finds that the product development tools, proposed by cloud service providers in order to facilitate app creation, increase the technical homogeneity of their users. On the other hand, data analytic tools, which enable firms to analyze data, make their functionalities more differentiated. The reason for this is that

[59] Lam assumes that the cloud service provider uses peak load pricing in order to allocate capacity efficiently across time. Although we have seen no direct evidence of peak load pricing, it does exist under a different form under the term of "post pricing", which is designed for "interruptible" workloads. Users quote a price and when the price of CPU falls below this amount, their computations start. On the other hand, if the price of CPU increases, the workload might be interrupted (it seems that for an extra charge users can purchase an option where load will not be interrupted). There could be some interesting economics in all of this. See `https://azure.microsoft.com/en-us/products/virtual-machines/spot/` (our understanding is that this type of service is offered by all major cloud service providers).

these “analytic technologies are more modular, making the fit with the IT platform (cloud) and compatibility of these technologies with each other less important to producing needed data resources”.

4.1.3 Switching costs

As discussed in 2.1, we know only of one paper which treats of switching specifically in the context of the cloud, Jin et al. (2023). This is an empirical paper in which the three authors estimate switching costs, focusing on IaaS services. They find that, as expected, low price elasticity for most products, but they also observe that there are substantial costs of adopting new products even from the same provider. They demonstrate very high benefits from the use of the cloud. Interestingly, they find that suppressing user inertia would increase their surplus by a large amount - 62% - but also provider's revenue by a similar order of magnitude as it would increase cloud usage.

4.2 General Literature

We now explore segments of economic literature conducted in diverse contexts and tailored for different markets. This exploration will provide insights that can enhance our understanding of the cloud market, encompassing research on switching costs, compatibility, network effects, and pricing.

4.2.1 Switching costs

Switching costs encompass the financial and resource expenditures a user faces when transitioning from one supplier to another. These costs involve various elements, such as adapting code to meet the specifications of a new provider, but do not include any contractual payments owed to the original service provider. For instance, when a company shifts from one cloud service provider to another, its engineers must invest time in adjusting their code to fit the specifics of the new provider, and the salaries of these engineers contribute to the overall switching costs. It is important to note that egress fees, which are part of the contractual agreements between the provider and its clients, are not, according to our definition, part of switching costs.

The concept of switching costs was introduced in the economic literature by Paul Klemperer; early surveys of the theoretical literature can be found in Klemperer (1995) and Farrell and Klemperer (2007); a discussion of policy implications can be found in National Economic Research Associates (2003), particularly Annex C. The early literature commonly assumed that all users faced the same prices, regardless of the length of their relationship with the supplier. It was posited that the supplier lacked the ability to offer special incentives to attract new customers. This led to a trade-off between "harvesting" and "investing". The supplier could harvest existing clientele by imposing high prices that current users would tolerate rather than switching to a lower-priced competitor. Alternatively, the supplier could invest by adopting lower prices to attract new users, creating an opportunity for long-term harvesting.

Beginning with Chen (1997), numerous authors have investigated "subscription models", wherein suppliers adjust prices based on users' purchasing histories. This approach severs the link between harvesting and investing incentives, although many findings applicable to the early models remain relevant in this scenario.

To comprehend Klemperer's fundamental insight, let us consider the scenario where a market for a new service emerges, attracting several firms vying to offer the product. Due to the existence of switching costs, customer acquisition becomes valuable, prompting firms to initially offer very low prices to entice customers (akin to the credits offered in the cloud industry). The elevated prices charged once customers commit to a supplier are offset, at least partially, by the special conditions that competitors are willing to offer to induce them to switch. The extent of this offset depends on various factors, primarily the level of competition in the industry. This trade-off remains consistent across all types of models.

To better understand the importance of customer acquisition and the influence of future competition in the regulation of suppliers, we employ the framework used by Biglaiser, Crémer and Dobos (2013). Let us begin with a straightforward one-period model. Initially, there exists a single incumbent supplier that all numerous (identical) users have patronized in preceding periods. Additionally, there are multiple new entrants.[60] The cost of providing service to one user is c and the cost of switching from the incumbent to one of the entrants is σ. It is clear that the entrants will not want to offer a price lower than c, and that the incumbent can charge $c + \sigma$,[61] and makes a profit of σ per user.

To delve into the intricacies of user acquisition value, we expand the model of the preceding paragraph by introducing a second period, once again featuring numerous entrants. If users opt for a different supplier in the second period, they face the same switching cost σ. All entities, including suppliers

[60] Much of the literature has focused on duopoly models, which are in some sense more relevant for the cloud sector, but make the intuition of the model less clear.

[61] For simplicity, we assume that when the users are indifferent they purchase from the incumbent.

and users, share a common discount factor denoted as $\delta \leq 1$.

In the second period, entrants set prices at cost, and all incumbents (that is all firms which sold to at least one customer in the first period) choose a price of $c+\sigma$ — ensuring that no customer switches to a different supplier. The profit generated by a firm from each user it retains at the start of the second period (those who purchased in the first period) is σ. From the perspective of the first period, this value is $\delta \times \sigma$. Consequently, in the initial period, entrants are willing to price down to $c - (\delta \times \sigma)$ to attract users. To retain its customer base, the incumbent, in the first period, must set a price equal to or less than $(c - \delta\sigma) + \sigma$. It will choose the upper bound, and make a per user profit of $(1 - \delta)\sigma$ in the first period. Thus, from the vantage point of the beginning of the first period, the present discounted value of the incumbent's per-user profit over the two periods is

$$\underbrace{(1-\delta)\sigma}_{1^{\text{st}} \text{ period}} + \underbrace{\delta \times \sigma}_{2^{\text{nd}} \text{ period}} = \sigma,$$

which is the same as the profit in the one period model! In the first period, the entrants compete away the second period profit of the incumbent. It is a mistake to think that the incumbents can charge a price equal to their cost plus the level of switching cost in every period.

This argument can be extended to any number of periods, and even, subject to some technicalities, to an infinite number of periods.[62] It is a general result, that, with identical users, an incumbent can only generate the static profits due to switching costs from its current user base even if the market interactions extend through time.

A significant portion of the literature assumes that users share identical switching costs. However, in many real-world scenarios, if not the majority, users have different switching costs. Several models have attempted to explore the implications of this variability. In one version (see, for instance, Taylor, 2003), switching costs are identically and independently distributed over time and users. At the start of each period, each user draws its switching cost for that period from a distribution that remains constant across all users and periods. In this scenario, all users hold the same future value for the supplier that successfully attracts them, resulting in dynamics very

[62]Papers that study infinite horizon switching cost models include Beggs and Klemperer (1992) and Padilla (1995). The standard setting is for there to be two firms and, as in the model, which we have just presented, homogeneous user switching costs; the focus is on the evolution of market shares and on the effect of switching costs on prices (see also Klemperer (1987)).

similar to those with homogeneous switching costs, except that users with low migration costs change suppliers.

In reality, one would think that the disparities in switching costs among users are more structural. A user with a low switching cost in 2023 is likely to have a similarly low switching cost in 2024. This complicates the strategies of entrants: when they attract some but not all users from a rival firm, the users they draw are typically those with the lowest switching costs. In subsequent periods, these users are more inclined to leave the firm in response to higher prices, yielding limited future profits. This diminishes the incentive for entrants to aggressively price below costs to attract new users, benefiting incumbent firms which can charge relatively high prices and retain the most valuable users. Biglaiser et al. (2013) demonstrate that the presence of users with low or zero switching costs enhances the profits of an incumbent firm, even if it doesn't retain those users who migrate to other firms.[63]

We have extensively examined the literature on switching costs for several reasons: a) they hold significance in the cloud industry, b) many observers contend that they pose a substantial barrier to competition, and c) we assert that further analysis is warranted, given that certain aspects of switching costs in the cloud remain unexplored in the existing literature. Regarding the first point, as we have previously argued in this report, there is little doubt about that switching costs are non-negligeable in the cloud (we also discuss this below). The second point becomes evident when reviewing the reports outlined in Section 3. The discussion above, in our view, demonstrates that while the presence of switching costs undeniably favors incumbents, it also heightens competition for users and may be less detrimental to competition than a first intuition would suggest.[64] (It is also the case that for multi-homing users there can be intense competition for the "next workload", *i.e.,* the next tasks which has not been allocated to one cloud or the other.)

Regarding the third point, a critical aspect that has been entirely over-looked in the literature is the users' ability, at a certain cost, to determine

[63] More precisely: adding low switching costs customers to a population of high switching costs customers increases the profits of the incumbent. On the other hand, the incumbent is always better off if a low switching cost customer is replaced by a high switching cost customer.

[64] We want to stress the "may" in this sentence. More research is clearly needed. For instance, a number of the reports reviewed in Section 3 stress the fact that the smaller European competitors find it financially difficult to meet the rebates given to new users by the "hyperscalers". This would be represented in the very simple model which we presented above by a lower discount factor δ for these firms. We do not believe that the consequences of such a change in the standard models have been studied. From a policy viewpoint, it might be important to think through the different ways to remedy this distortion.

the level of switching cost. This occurs when users opt for software with functionalities across multiple clouds instead of tailoring it to a specific supplier—such as utilizing open-source solutions instead of, presumably more efficient, proprietary solutions from their current cloud service provider. More broadly, users can influence switching costs by choosing Infrastructure as a Service (IaaS) over Platform as a Service (PaaS), among other considerations.[65] Hence, the user must carefully weigh the advantages of reducing its switching costs, thereby enhancing its ability to switch providers in response to price changes (or less favorable changes compared to competitors), and its potential to capitalize on an appealing offer from an alternative provider. This must be balanced against the benefits of optimizing its system for the current provider. From an economic analysis standpoint, the intriguing aspect arises from the creation of "externalities" among users: when one user lowers its switching costs, competition intensifies, benefiting other users. It is crucial to thoroughly investigate the implications of this phenomenon, both to comprehend the dynamics of the cloud service market and to assess the repercussions of regulations aimed at facilitating transitions between cloud service providers.

4.2.2 History based pricing

The value that a user attaches to using the cloud is contingent on the type and intensity of how utilize it, information that their current cloud provider can observe and which provides insights into the user's willingness to pay for the service. As prices are often individually negotiated, the cloud provider may seek to tailor the charges—potentially through special offers or cloud credits—based on the user's switching costs and unique preferences for specific cloud services. The work of Hagiu and Wright (2023) offers valuable insights for delving deeper into this subject.

Economics has a longstanding tradition of studying firms employing behavioral, history-dependent pricing and investigating the consequences of such practices. In scenarios with switching costs, as discussed in Chen

[65] Multi-homing, where a user utilizes different clouds for various purposes, could serve as a means to diminish switching costs for two reasons. First, it implies that the user will possess in-house expertise on multiple clouds. Second, it facilitates partial migration, which is less risky than a complete change of suppliers. Additionally, multi-homing may help alleviate one of the barriers to migration: the complexity of understanding the pricing of cloud services. With experience across several clouds, users can gain a better understanding of the financial implications of their choices.

Each of the considerations outlined in the preceding paragraph presents an opportunity for insightful and theoretically profound analysis.

(1997) and Taylor (2003), excessive switching of users and lower welfare, compared to the efficient amount, may occur with sufficient competition between firms. Fudenberg and Tirole (2000) examine a setting where buyers are sorted through short-term and long-term contracts, finding that whether users switch providers at too low or high a level depends on the feasibility of long-term contracts and whether users' preferences regarding firms are fixed or change over time.

Cloud providers are able to offer prices which depend both on the timing of users' usage and its past purchasing history. This is evident, for instance, when new users receive substantial discounts, including zero prices, when they initially join a cloud. This pricing strategy is closely related to the use of egress fees, which lacks a clear counterpart in other markets. The closest analogy is found in contract theory with breach penalties if a buyer fails to purchase from a firm in the future, as explored in Aghion and Bolton (1987).[66] However, there are at least two distinctions between egress fees and breach penalties. First, many cloud users seek to multi-home on different cloud providers to leverage different efficiencies. Second, even when a user is transferring data off the cloud, it does not necessarily imply a move to another provider; they may be utilizing the data for internal purposes.

4.2.3 Compatibility

Connected to the economic impact of switching costs is the ability of users to transfer their data from one cloud provider to another while maintaining the same functions on the new provider. The Data Act places obligations on "data processing services", with Article 30 mandating full interoperability for Infrastructure as a Service (IaaS) and requiring most Platform as a Service (PaaS) to create open interfaces facilitating both switching and multi-cloud usage.

The existing literature on compatibility often explores settings where users construct a system comprised of multiple complementary products. Matutes and Regibeau (1988) initiated this literature in a static setting. In a more recent contribution, Jeon, Menicucci and Nasr (2023) examined a two-period duopoly model with switching costs and firms employing behavioral pricing: prices in the second period are influenced by purchasing histories. They discovered that in a symmetric setting, firms opt for product incompatibility, when switching cost are relatively high, in order to soften future price competition. Requiring data portability, which in their model is

[66] Fudenberg and Tirole (2000) analyze a two period model with changing preferences and breach penalties.

equivalent to lowering switching costs, encourages firms to choose compatible products more frequently, but this benefits users only if a non-negative pricing constraint is binding, which is not applicable in cloud computing. Clearly, more work is needed to explore the robustness of these results. (We discuss further this paper in 4.2.4.)

An older work by Farrell and Saloner (1992) features the incorporation of interoperability most relevant to the cloud setting. Users can adopt one of two technologies, and there are network effects if they use the same technology. Users are horizontally differentiated, and have the option to purchase a "converter" which allows them to benefit from an imperfect level of network externalities from the users of the alternative technology. Farrell and Saloner explore various market structures, including perfect competition in both platform prices and the converter, a monopolist in all three markets, and a duopoly in the platform markets with a competitive converter market. A central question is whether a converter increases the likelihood of equilibria where users are on different platforms or makes it more likely that one platform dominates. The findings indicate that a monopolist uses a converter to enhance its ability to price discriminate between users, and that, in equilibrium, the use of converters is more prevalent under duopoly than under monopoly. Notably, interoperability differs from the purchase of a converter because both the cloud provider and the user can influence the degree of interoperability. As users design their networks to be adaptable to multiple cloud providers, this leads to a higher degree of interoperability.

4.2.4 Data

The competitiveness of the cloud market is enhanced by the ability of users to switch from one provider to another, in particular thanks to the portability of data. (The competitiveness is also influenced by other factors such as the competition for the next workload.) The European Data Act aims at facilitating the transfer of data between cloud providers in order to increase the competitiveness of the cloud services market. The rationale is that enabling users to move their data benefits both the user, who can enhance their activities with the new provider, and the competitive dynamics in the market.

The existing economics literature contains relatively few articles that study data-related issues. Current research often focuses on the use of data to either better match a consumer's tastes with the products of a supplier or to allow firms to extract more surplus from each consumer. Lam and Liu (2020) make a distinction between two types of data: data directly provided by the user to the provider (considered portable) and data generated by the provider by inferring information from the user's "raw" data (not portable). Their analysis, within a two-period model where users transition from an incumbent to an entrant, highlights two effects of data portability. While entry is facilitated for a given level of data provision, data portability encourages users to provide more data in period 1. By supplying more data in period 1, the incumbent can offer enhanced services in period 2, making it less likely for users to switch to an entrant. Network effects among users also contribute to reducing the likelihood of switching, not due to the coordination effect but because users tend to underestimate the positive externality of providing data, leading to insufficient data provision. Data portability mitigates this problem by encouraging users to provide more data, potentially making switching less likely if the enhanced services provided by the incumbent are significant. Thus, data portability may inadvertently bolster the incumbent's advantage while simultaneously improving efficiency.

In a broader context, firms having better data about a user can either increase or reduce the utility offered to the users. Increased data allows a firm to customize its offerings, enhancing the surplus in the user-firm interaction. Conversely, having more data about a user may enable a firm to extract more surplus. When firms compete for users, whether data is pro or anti-competitive depends on how increasing data affects user offerings and the reactions by rivals. De Cornière and Taylor (2023) present a framework to analyze the effects of data competition when firms compete in "utility space".

Finally, in 4.2.3 we discussed the article by Jeon et al. (2023) who introduced switching costs in a model of complementary products. Actually, as the title of the article indicates and as we discussed earlier, they model data portability as a reduction of the switching costs for the users of cloud services. This is an example of the strategy that economists have typically used when studying data: they assume the consequences of access to data for individual users, and study how changing these affect the equilibrium of competition between firms. Often, the formal model could be interpreted in other ways. For instance, the model of Section 7.3 of Jeon et al. (2023) could also be interpreted as a model of the consequences of a greater similarity between the programming interfaces of different cloud services.

4.2.5 Network effects

As discussed in 2.4, network effects occur when the value that a user derives from a platform increases with the presence of others on the same platform, either directly or indirectly. Direct effects emerge as more firms use the same cloud service, for instance by reducing the time needed for users to exchange data. Indirect effects arise when a larger user base incentivizes cloud providers and software-designing firms to optimize physical infrastructure and software for the cloud. Even if there are no switching costs at the individual level, network effects can favor the incumbents, creating something akin to what is sometimes called "social" switching costs: each user hesitates to migrate to another platform, even if it definitely think it would be a better choice, for fear that the others do not follow.

As in the case of switching costs, when determining prices and contract terms, a cloud provider considers the benefits of higher profits from current users who find it difficult to migrate due to network effects versus the potential loss caused by the lower attractiveness for new users. Another effect arises, specific to the network effect setup: more new users increase the value of the platform for all users. Cabral (2011) presents a dynamic duopoly model with network effects and showed that the larger firm tends to attract new users when the sizes of firms are not too different. On the other hand, when the firms are of very different sizes, the larger firm harvests its current clients while the small firm prices aggressively to expand its network. This is clearly not the case in the cloud services industry, and it might be of interest to speculate why. One reason could be that the industry as a whole is still expanding. The second is the fact that the three bigger firms are of not too different sizes, and therefore are aggressively competing against each other to build their network. It could be interesting to expand Cabral's model to more than two firms to see which insights could be drawn for the cloud industry.

Because users favor the presence of other users rather than the intrinsic quality of the platform, network effects often result in multiple "equilibria", where the configuration depends on the, often history-dependent, users' beliefs about which provider others will choose. Various "equilibrium selection" criteria, that is ways of predicting which equilibrium will obtain, have been proposed. They often emphasize the importance of incumbency in shaping user expectations.

In a static framework, Caillaud and Jullien (2003) and most subsequent works assume that, out of equilibrium, agents coordinate on the equilibrium less favorable to the entrant. Hagiu (2006) and Hałaburda and Yehezkel (2013) adopt similar strategies. Jullien (2011) extends the favorable expectations concept to multi-sided platforms. In a dynamic setting, Halaburda, Jullien and Yehezkel (2020) build an infinite-horizon duopoly model where incumbency, modeled through a "belief approach", plays a crucial role.

Unlike previous papers that assume that on each side of the market all users are identical, Biglaiser and Crémer (2020) analyze a single-sided, infinite-horizon multi-platform model with user heterogeneity. To select an equilibrium, they conduct a thought experiment akin to fictitious play, where users assess the utility of moving to any platform assuming no other user migrates. If no one gains, users remain on their respective platforms; if some users can gain, a small number switch until no user can gain. The equilibrium may result in single-platform dominance, but not necessarily. An incumbent may prefer high tariffs for users with significant network benefits while allowing other types of users to move to a rival platform.

5 Conclusion

Cloud computing has profoundly transformed the way businesses and organizations run their IT systems. By leveraging distributed networks of servers and software, cloud systems have enabled the provision of faster, more reliable, and scalable computing services. They also make possible many other innovative services and applications.

In this report, we have reviewed the main features and challenges of cloud computing, such as the different types of cloud models, the role of cloud infrastructure service providers, the increasing concentration of the industry, and the main concerns of policy makers. We have also analyzed the economic aspects of cloud computing, such as the market structure and competition, the pricing strategies including egress fees, cloud credits and committed spend discounts.

We have two parallel messages for regulators and for our fellow academics. For regulators, our main message is to exercise some caution. We understand their concerns with some of the practices in the industry and sympathize with their desire to intervene early to increase the degree of competition in the industry. However, there is much that is not understood, and this leads us to the message to our fellow academics, especially economists. In the past forty or so years, we have developed an impressive set of tools to understand the consequences of platforms - we have done so by taking the technology as given in our analysis. The emergence of the cloud enables us to use these tools to go deeper into the way the technological infrastructure of the IT industry works. We should seize this opportunity. The problems are intellectually fascinating; we have the basic tools to study them; and the policy implications are significant.

References

ACM: Netherlands Authority for Consumers and Markets (2022) "Market Study Cloud Services," URL: https://www.acm.nl/en/publications/market-study-cloud-services.

Aghion, Philippe and Patrick Bolton (1987) "Contracts as a Barrier to Entry," *The American Economic Review*, pp. 388–401.

Beggs, Alan and Paul Klemperer (1992) "Multi-period competition with switching costs," *Econometrica: Journal of the Econometric Society*, pp. 651–666.

Biglaiser, Gary and Jacques Crémer (2020) "The value of incumbency when platforms face heterogeneous customers," *American Economic Journal: Microeconomics*, Vol. 12, No. 4, pp. 229–269.

Biglaiser, Gary, Jacques Crémer, and Gergely Dobos (2013) "The value of switching costs," *Journal of Economic Theory*, Vol. 148, No. 3, pp. 935–952.

Bisceglia, Michele and Jean Tirole (2023) "Fair gatekeeping in digital ecosystems," TSE Working Paper.

Cabral, Luis (2011) "Dynamic price competition with network effects," *The Review of Economic Studies*, Vol. 78, No. 1, pp. 83–111.

Caillaud, Bernard and Bruno Jullien (2003) "Chicken & egg: Competition among intermediation service providers," *RAND journal of Economics*, pp. 309–328.

Carlton, Dennis W. and Michael Waldman (2002) "The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries," *The RAND Journal of Economics*, Vol. 33, No. 2, pp. 194–220, URL: http://www.jstor.org/stable/3087430.

CERRE (2022) "Switching and interoperability between data processing services in the proposed Data Act," by D. Schnurr, URL: https://cerre.eu/wp-content/uploads/2022/12/Data_Act_Cloud_Switching.pdf.

Chen, Xi, Ming Guo, and Wuyue Shangguan (2022) "Estimating the impact of cloud computing on firm performance: An empirical investigation of listed firms," *Information & Management*, Vol. 59, No. 3, p. 103603.

Chen, Yongmin (1997) "Paying customers to switch," *Journal of Economics & Management Strategy*, Vol. 6, No. 4, pp. 877–897.

CISPE (2021) "Report on Cloud Infrastructure Services," by F. Jenny, URL: `https://cispe.cloud/new-study-links-unfair-software-licences-to-distortion-of-competition-in-cloud-infrastructure-market/`.

de Cornière, Alexandre and Greg Taylor (2023) "Data and competition: A simple framework," Toulouse School of Economics Working Paper No. 1404.

DeStefano, Timothy, Richard Kneller, and Jonathan Timmis (2020) "Cloud computing and firm growth," CESifo Working Paper No. 8306.

Dohan, Thanh and Franco Mariuzzo (2023) "Digital platform mergers and innovation: Evidence from the cloud computing market," CCP Working Paper 23-05.

Ennis, Sean and Ben Evans (2023) "Cloud Portability and Interoperability under the EU Data Act: Dynamism versus Equivalence," `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4395183`.

Ennis, Sean, Ben Evans, and Franco Mariuzzo (2023) "CCP Response to the Ofcom Consultation: Cloud Services Market Study (Interim Report)," Center for Competition Policy Response.

Etro, Federico (2009) "The economic impact of cloud computing on business creation, employment and output in Europe," *Review of Business and Economics*, Vol. 54, No. 2, pp. 179–208.

Farrell, Joseph and Paul Klemperer (2007) "Coordination and lock-in: Competition with switching costs and network effects," *Handbook of industrial organization*, Vol. 3, pp. 1967–2072.

Farrell, Joseph and Garth Saloner (1992) "Converters, compatibility, and the control of interfaces," *The journal of Industrial Economics*, pp. 9–35.

FCA: French Competition Authority (2023) "Market Study on Competition in the Cloud Computing Sector," Summary Opinion, URL: `https://www.autoritedelaconcurrence.fr/en/opinion/competition-cloud-sector`.

Flexera (2023) "2023 State of the Cloud Report," URL: `https://info.flexera.com/CM-REPORT-State-of-the-Cloud`.

Fudenberg, Drew and Jean Tirole (2000) "Customer poaching and brand switching," *RAND Journal of Economics*, pp. 634–657.

Gans, Joshua, Mikaël Hervé, and Muath Masri (2023) "Economic analysis of proposed regulations of cloud services in Europe," *European Competition Journal*, pp. 1–47.

Gartner Report (2023) "Cloud Governance Best Practices: Managing Vendor Lock-In Risks in Public Cloud IaaS and PaaS."

Gomes, Renato and Andrea Mantovani (2024) "Regulating Platform Fees Under Price Parity," forthcoming, *Journal of the European Economic Association*, URL: `https://doi.org/10.1093/jeea/jvae014`.

Hagiu, Andrei (2006) "Pricing and commitment by two-sided platforms," *The RAND Journal of Economics*, Vol. 37, No. 3, pp. 720–737.

Hagiu, Andrei and Julian Wright (2023) "Data-enabled learning, network effects and competitive advantage," *RAND Journal of Economics*, Vol. 54, No. 4, pp. 638–667.

Halaburda, Hanna, Bruno Jullien, and Yaron Yehezkel (2020) "Dynamic competition with network externalities: how history matters," *The RAND Journal of Economics*, Vol. 51, No. 1, pp. 3–31.

Hałaburda, Hanna and Yaron Yehezkel (2013) "Platform competition under asymmetric information," *American Economic Journal: Microeconomics*, Vol. 5, No. 3, pp. 22–68.

Hummel, Patrick and Michael Schwarz (2022) "Efficient Capacity Provisioning for Firms with Multiple Locations: The Case of the Public Cloud," in *Proceedings of the 23rd ACM Conference on Economics and Computation*, pp. 1018–1039.

Impink, Stephen Michael (2022) "Outsourcing IT and Technological Differentiation: Evidence from Digital Startups," Available at SSRN 4200675.

European Alliance for Industrial Data, Edge and Cloud (2023) "European Industrial Technology Roadmap for the Next-Generation Cloud-Edge," URL: `https://digital-strategy.ec.europa.eu/en/news/european-alliance-industrial-data-edge-and-cloud-presents-its-first-deliverables`.

Jeon, Doh-Shin, Domenico Menicucci, and Nikrooz Nasr (2023) "Compatibility Choices, Switching Costs, and Data Portability," *American Economic Journal: Microeconomics*, Vol. 15, No. 1, pp. 30–73.

JFTC: Japan Fair Trade Commission (2022) "Report on Trade Practices in Cloud Services Sector," URL: https://www.jftc.go.jp/en/pressreleases/yearly-2022/June/220722_2EN.pdf.

Jin, Chuqing, Sida Peng, and Peichun Wang (2023) "Sticky consumers and cloud welfare," Available at SSRN 4573326.

Jullien, Bruno (2011) "Competition in multi-sided markets: Divide and conquer," *American Economic Journal: Microeconomics*, Vol. 3, No. 4, pp. 186–219.

Klemperer, Paul (1987) "The competitiveness of markets with switching costs," *The RAND Journal of Economics*, pp. 138–150.

——— (1995) "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *The Review of Economic Studies*, Vol. 62, No. 4, pp. 515–539.

Lam, Wing Man Wynne (2015) "Competition in the Market for Flexible Resources: an Application to Cloud Computing," LCII Policy Brief 2015/2, Liège Competition and Innovation Institute.

Lam, Wing Man Wynne and Xingyi Liu (2020) "Does data portability facilitate entry?" *International Journal of Industrial Organization*, Vol. 69.

Leka, Enxhi (2022) "What Should Economists Know About the Cloud? A Literature Review on Digital Economics," Calgary: International Telecommunications Society (ITS).

Matutes, Carmen and Pierre Regibeau (1988) "" Mix and match": product compatibility without network externalities," *The RAND Journal of Economics*, pp. 221–234.

Nalebuff, Barry (2004) "Bundling as an entry barrier," *The Quarterly Journal of Economics*, Vol. 119, No. 1, pp. 159–187.

National Economic Research Associates (2003) "Switching costs," Report prepared for the Office of Fair Trading and the Department of Trade, UK.

Ofcom (2023) "Cloud Services Market Study: final report," URL: `https://www.ofcom.org.uk/consultations-and-statements/category-2/cloud-services-market-study`.

Padilla, A Jorge (1995) "Revisiting dynamic duopoly with consumer switching costs," *Journal of Economic Theory*, Vol. 67, No. 2, pp. 520–530.

Park, Jiyong, Kunsoo Han, and Byungtae Lee (2023) "Green cloud? An empirical analysis of cloud computing and energy efficiency," *Management Science*, Vol. 69, No. 3, pp. 1639–1664.

Taylor, Curtis R (2003) "Supplier surfing: Competition and consumer behavior in subscription markets," *RAND Journal of Economics*, pp. 223–246.

Wang, Chengsi and Julian Wright (2022) "Regulating platform fees," Unpublished manuscript, Natl. Univ. Singapore.

Whinston, Michael D (1990) "Tying, Foreclosure, and Exclusion," *The American Economic Review*, Vol. 80, No. 4, pp. 837–959.