Cloud services market investigation

Competitive landscape working paper

23 May 2024

This is one of a series of consultative working papers which will be published during the course of the investigation. This paper should be read alongside the Issues Statement published on 17 October 2023 and other working papers published.

These papers do not form the inquiry group's provisional decision report. The group is carrying forward its information-gathering and analysis and will proceed to prepare its provisional decision report, which is currently scheduled for publication in September/October 2024, taking into consideration responses to the consultation on the Issues Statement and responses to the working papers as well as other submissions made to us. Parties wishing to comment on this paper should send their comments to CloudMI@cma.gov.uk by **13 June 2024.**

The Competition and Markets Authority has excluded from this published version of the working paper information which the inquiry group considers should be excluded having regard to the three considerations set out in section 244 of the Enterprise Act 2002 (specified information: considerations relevant to disclosure). The omissions are indicated by [✂]. Some numbers have been replaced by a range. These are shown in square brackets. Non-sensitive wording is also indicated in square brackets.

Contents

Tables

Figures

Appendix

1. Introduction to cloud services

1.1 Cloud services are increasingly important inputs to many businesses and organisations across the UK economy. They support most sectors, including communications, manufacturing, retail, hospitality and financial services as well as public and voluntary sector bodies. Without cloud services many digital businesses providing services to consumers would not be able to function in the way they do today.

1.2 Cloud services provide access to computing resources on demand, via a network. Customers buy access to the computing resources as a service and typically do not own the underlying hardware and software.

1.3 In this market investigation we are considering the supply of public cloud infrastructure services (cloud services) in the UK.[1] For these purposes:

(a) 'Cloud infrastructure services' means services that provide access to processing, storage, networking, and other raw computing resources (often referred to as infrastructure as a service, IaaS) as well as services that can be used to develop, test, run and manage applications in the cloud (often referred to as platform as a service, PaaS).[2] [3]

(b) 'Public cloud computing' means that cloud services are open to all customers willing to pay, with computing resources shared between them.

(c) 'Public cloud infrastructure services' are therefore cloud infrastructure services delivered via a public cloud model.

1.4 One way in which these service models are differentiated is by the level of control the customer has over the management and maintenance of the computing resources.[4] IaaS, PaaS and SaaS form a vertical 'cloud stack', where each layer is notionally built on top of the previous one(s).[5]

(a) IaaS are cloud services that provide access to raw computing resources (compute, storage, and network) for processing workloads[6] and storing data. The hardware associated with these computing resources take the form of servers and networking equipment owned and managed by the IaaS provider

[1] CMA, Public cloud infrastructure services, Issues statement (publishing.service.gov.uk) 17 October 2023. ('issues statement').
[2] Ofcom, Cloud services market study, final report, Terms of Reference (ofcom.org.uk), 5 October 2023 ('Terms of Reference').
[3] Some services may not 'fit' neatly into these service models and the lines between each of IaaS, PaaS and software as a service (SaaS) may be blurred (for example, see paragraph a)). However, we still consider them to be useful to inform our analysis in this market investigation. For an explanation of SaaS, please see paragraph 1.4 (c) below.
[4] Control refers to the involvement the customer has in the management and maintenance of the computing resources themselves, as opposed to the freedom it affords them to, for example, choose between providers.
[5] In practice, this vertical stack is not strictly applied. For example, SaaS may be built and deployed using IaaS only.
[6] An application, service, capability or a piece of work

(and typically held on racks in a remote data centre). To allow and manage that access, IaaS also includes some necessary software, including networking and virtualisation.[7] The IaaS service model provides the customer with the highest level of control over the cloud stack, including over the operating system, applications, and data. IaaS should be distinguished from bare metal[8] services.

(b) PaaS are cloud services that provide access to a virtual environment for customers to develop, test, deploy and run applications. They include application development computing platforms and pre-built application components and tools which customers can then use to build and manage full applications. The customer has less control over the cloud stack compared to IaaS – they still manage applications and data but not the PaaS computing platform (including its operating system) and the pre-built application components and tools.

(c) SaaS are complete applications hosted in the cloud. Like PaaS, they can be offered by the cloud provider that owns the underlying raw compute resources or by an independent software vendor (ISV). The service provider(s) manages all hardware and software.

1.5 In addition to public cloud computing, there are two other cloud deployment models:

(a) private cloud – a cloud deployment model in which computing resources (like the hardware) are used exclusively by one customer; and

(b) hybrid cloud – a cloud deployment model in which public and private clouds are combined and presented to the customer as a single unified cloud.

1.6 Cloud computing is distinct from traditional IT where assets (such as servers and networking hardware) are usually located on-premises and are not part of the cloud.

1.7 This working paper primarily focuses on public cloud infrastructure services, however other cloud deployment models and on-premises alternatives are discussed where relevant.

[7] Virtualisation is the process of using software to create an abstraction layer over servers that allows the hardware elements of a single server to be divided into multiple virtual servers, commonly called virtual machines. Each virtual machine runs its own operating system and behaves like an independent server, even though it is running on just a portion of the actual underlying server hardware. The software that creates, runs and manages virtual machines is called a hypervisor.

[8] Bare metal services offer access to dedicated servers (or 'hosts') with no or limited software installed (eg no operating system or virtualisation).

Structure and contents of this working paper

1.8 This working paper sets out our initial analysis of:

(a) the competitive landscape of cloud services including: (i) the nature of the customer base and trends in the usage of public cloud; (ii) the customer journey; (iii) customer preferences and parameters of competition between public cloud providers; and (iv) the main providers focusing on the vertically integrated suppliers of cloud services (which we refer to as 'cloud providers');

(b) how customers switch and multi-cloud,[9] including the prevalence of this;

(c) the relevant markets that we are considering as part of this investigation;

(d) shares of supply relating to the markets that we are considering as part of this investigation;

(e) market outcomes including profitability;

(f) barriers to entry and expansion in cloud services; and

(g) artificial intelligence and its potential impact on cloud services.

1.9 At the end of this working paper we also set out an overview of our emerging views based on the evidence to date.

1.10 As part of our initial analysis we have asked market participants, including cloud providers and customers, a range of questions. This has involved gathering quantitative evidence, for example to inform our estimates of market shares, and qualitative evidence, for example customers' views on different cloud providers. Where the evidence we have gathered is qualitative in nature, we have given a narrative summary of the key points that we consider emerge from the evidence.

1.11 We also commissioned qualitative customer research from Jigsaw Research. This research was intended to capture a wider range and a different set of customers from those we spoke to through direct channels. Evidence from this research, including evidence relevant to the initial analysis set out in this paper, will be provided in the Jigsaw report. We will consider the evidence from the Jigsaw report alongside the evidence outlined in this paper in our ongoing work.

[9] At its most basic, multi-cloud involves the placement by a customer of at least one workload on one provider's cloud, and at least one workload on an alternative provider's cloud. See also section 3.

2. Competitive Landscape

2.1 In a market investigation we need to understand how a market operates in practice as this allows us to apply the appropriate framework to our analysis. We collect and analyse evidence about the main characteristics of the reference market and use this to inform our assessment of whether any features may be harming competition in the market.[10]

2.2 In order to understand the main characteristics of the reference market, this section sets out evidence on:

(a) The nature of the customer base, such as the types of customer that are purchasing cloud services and trends in the usage of these services;

(b) the customer journey, including how products are purchased;

(c) customer preferences and the parameters of competition; and

(d) the main providers focusing on the vertically integrated suppliers of cloud services[11] (which we refer to as 'cloud providers') including their business models and strategies.

The demand for public cloud

Public cloud customers

2.3 Cloud services are increasingly important inputs to many businesses and organisations across the UK economy. They support most sectors, including communications, manufacturing, retail, hospitality and financial services as well as public and voluntary sector bodies. Without cloud services many digital businesses providing services to consumers would not be able to function in the way they do today.

2.4 Evidence we have seen from cloud providers shows that there is a large number of customers of cloud services in the UK. [12, 13] These customers are present in a range of different industries: for example, Amazon Web Services (AWS) listed 30 different industries in the data it provided to us.[14] Some of these industries have

[10] See CC3 (Revised), Guidelines for market investigations: Their role, procedures, assessment and remedies (publishing.service.gov.uk), ('CC3') paragraph 97.
[11] That is, vertically integrated suppliers of cloud services that operate their own cloud infrastructure, ie they own the underlying raw computing resources.
[12] [✂] response to the CMA's information request [✂].
[13] [✂] response to the CMA's information request [✂].
[14] AWS response to the CMA's information request [✂].

specialised use cases due to regulatory requirements (eg financial services) or prescribed procurement rules (as in the public sector).[15]

2.5 AWS, Microsoft and Google supplied customer level data which included revenue and industry categorisation. The categories provided by each provider were based on their own reporting and therefore do not align with each other – some providers supplied more granular categories or different groupings of industries to others.

2.6 Nevertheless, our analysis of revenue across categories allowed us to observe that customers in financial services, IT or software services and retail are important categories.

2.7 Customers also vary in terms of their size. Evidence from cloud providers shows that a small number of high-spend customers are responsible for a significant proportion of providers' UK revenue and a large number of low-spend customers are responsible for a small proportion of their revenue. In particular, the top 10% of customers account for a very large majority of revenues and the top 1% account for over half of revenues. [16, 17, 18]

Growth in demand for cloud services

2.8 AWS was the first provider to supply cloud services in 2006 using infrastructure that AWS developed initially to support its online retail business.[19] Microsoft made Microsoft Azure generally available in 2010,[20] and Google made Google Cloud generally available in 2011.[21]

2.9 Since AWS launched its IaaS offering, the cloud services landscape has expanded significantly with various providers operating at both the IaaS and PaaS level. This expansion has seen a proliferation of services and products offered by providers. For example, AWS' portfolio is now comprised of over 200 individual services and products, including compute, storage, databases, analytics, networking, mobile, developer tools, management tools, internet of things, security, and enterprise applications.[22]

2.10 To better understand how the cloud services markets may develop in the future, we have looked at trends in customer spending, cloud provider revenue and datacentre capacity.

---

[15] See section 7.
[16] [✂] response to the CMA's information request [✂].
[17] [✂] response to the CMA's information request [✂].
[18] [✂] response to the CMA's information request [✂].
[19] About AWS (archive.org), accessed 10 February 2024.
[20] Windows Azure Platform Now Generally Available in 21 Countries | Blog Azure | Microsoft Azure, accessed 11 February 2024.
[21] The History of Google Cloud Platform (pluralsight.com) , accessed 11 February 2024.
[22] Overview of Amazon Web Services - Overview of Amazon Web Services, accessed 13th March 2024

Customer spending

2.11 Gartner regularly forecasts growth in worldwide end-user spending[23] and we have seen evidence that providers use these forecasts for market monitoring. It forecasts that worldwide spending on all public cloud services will grow from $478 billion in 2022 to $679 billion in 2024.

2.12 This forecast covers products and services beyond public cloud infrastructure services in the UK as defined in this market investigation: Gartner forecasts that worldwide IaaS spending will increase from $120 billion in 2022 to $182 billion in 2024 and that PaaS spending will increase from $120 billion in 2022 to $176 billion in 2024.[24]

Cloud provider revenue

2.13 Global forecasts we have received from cloud providers also show that demand for cloud services will continue to grow. In particular, the internal analysis of one cloud provider shows that it expects year-on-year revenue growth from its cloud services to continue [✂].[25] [✂].[26]

2.14 Table 2.1 sets out the estimated UK revenue from IaaS and PaaS in the UK from 2019 to 2022 as set out in Ofcom's market study.[27] It was calculated from 11 companies'[28] UK IaaS and PaaS revenues between 2019 and 2022 and an estimate of the remainder of UK revenues associated with IaaS and PaaS using Synergy and IDC data.[29] We will update these figures, including adding data for 2023.

Table 2.1: UK IaaS and PaaS revenues, 2019-2022 (£bn)

	2019	*2020*	*2021*	*2022*	*Annual growth*
IaaS	[1.5-2.0]	[2.0-2.5]	[2.5-3.0]	[4.0-4.5]	30-35%
PaaS	[0.5-1.0]	[1.0-1.5]	[1.5-2.0]	[1.5-2.0]	40-45%
IaaS and PaaS	**[2.5-3.0]**	**[3.5-4.0]**	**[4.5-5.0]**	**[7.0-7.5]**	**35-40%**

Source: Table 4.8 of Cloud services market study final report (ofcom.org.uk)

2.15 Table 2.1 shows that in the UK the revenues for IaaS and PaaS grew from £[2.5-3.0] billion in 2019 to £[7.0-7.5] billion in 2022. This represented an annual growth rate of [35-40]% per year. IaaS revenues were higher in the UK in 2022 at £[4.0-4.5] billion compared to PaaS revenues which were at £[1.5-2.0] billion, but PaaS

[23] For example: Gartner Forecasts Worldwide Public Cloud End-User Spending to Grow 18% in 2021 and Gartner Forecasts Worldwide Public Cloud End-User Spending to Reach Nearly $500 Billion in 2022, accessed 10 February 2024
[24] Gartner Forecasts Worldwide Public Cloud End-User Spending to Reach $679 Billion in 2024, accessed 10 February 2024.
[25] [✂] response to the CMA's information request [✂].
[26] [✂] response to the CMA's information request [✂].
[27] Cloud services market study (final report) - Ofcom
[28] The 11 companies were: [✂].
[29] See paragraphs A1.24 to A1.27 of Annex 1 of Cloud services market study - Final report – Annexes (ofcom.org.uk)

revenues had a higher annual growth rate over that period of [40-45]% compared to [30-35]% for IaaS.

2.16 We have seen evidence of some new developments in public cloud services that may drive how the cloud is used in the future.

2.17 First, both Gartner and KPMG identified the rise of what they termed as 'industry cloud platforms' or 'vertical clouds' as a key trend driving cloud spending.[30] Industry cloud platforms combine a provider's IaaS, PaaS and SaaS capabilities into a whole-product offering which targets the industry-specific needs of customers.[31] For example, one provider offers industry-specific cloud solutions for the healthcare, financial services, non-profit, retail and government sectors, amongst others.[32]

2.18 Gartner predicts that 'by 2027, more than 70% of enterprises will use industry cloud platforms to accelerate their business initiatives, up from less than 15% in 2023'.[33] If this trend continues then it may influence competitive dynamics based on the strengths of each cloud provider's industry focused solutions. We will consider this further as part of our ongoing work to understand cloud providers' strategies, see paragraphs 2.66 to 2.149 below.

2.19 Second, there is evidence of an increased deployment of serverless computing and serverless architectures including function-as-a-service (FaaS) and backend-as-a-service (BaaS).

2.20 Since the serverless model abstracts away all the infrastructure and provisioning related to an application, customers only focus on developing the applications, rather than the scaling and provisioning of the underlying infrastructure, which is automatically scaled up or down based on the application traffic.

2.21 Gartner predicts that, due to the 'simplicity' and reduced overhead of lowered infrastructure maintenance', 'half of global enterprises will have deployed FaaS by 2025, up from only 20% in 2020'.[34]

2.22 Consistent with this, Microsoft service's (Azure Functions) UK customer base grew by 96% between 2019 and 2022.[35]

---

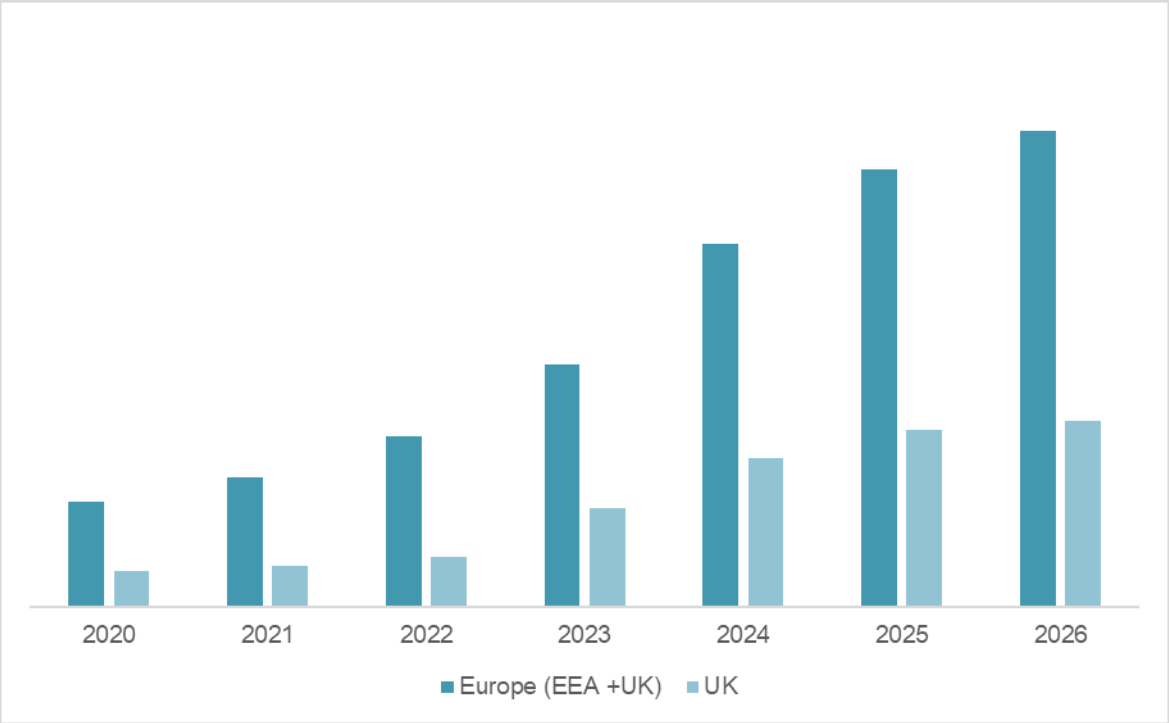
[30] The future of cloud is industry specific - KPMG Global and Gartner Forecasts Worldwide Public Cloud End-User Spending to Reach $679 Billion in 2024, accessed 10 February 2024.
[31] Gartner Forecasts Worldwide Public Cloud End-User Spending to Reach $679 Billion in 2024, accessed 10 February 2024.
[32] Microsoft Industry Clouds, accessed 10 February 2024.
[33] Gartner Forecasts Worldwide Public Cloud End-User Spending to Reach $679 Billion in 2024, accessed 10 February 2024.
[34] The CIO's Guide to Serverless Computing, Gartner. Also see The Rise of Serverless Computing, Communications of the ACM, Published December 2019, Volume 62 no.12, accessed 10 February 2024.
[35] [✂] response to Ofcom's information request [✂].

Datacentre capacity

2.23 We have considered cloud providers' current and forecast datacentre capacity because this would be consistent with increasing demand over time; if providers forecast increases in capacity, then this is consistent with an expectation among them that demand for cloud services will continue to increase.

2.24 Figure 2.1 shows the total data centre capacity for Europe (European Economic Area + UK) and the UK in megawatts (MW) for the period 2020 to 2026. This is based on MW as we understand that this is the industry standard metric and includes actual capacity for the period 2020 to 2023 and forecast capacity for 2024 to 2026.[36] This is based on data from AWS, Microsoft, Google, IBM and Oracle.

Figure 2.1: Total datacentre capacity for Europe and the UK in MW, 2020-2026



Source: CMA analysis of cloud provider data. [✂];

2.25 These data show that data centre capacity has increased significantly since 2020 and is forecasted to increase substantially in the near future. However, this expansion is projected to occur at a slower rate compared to previous years.

How customers purchase cloud services

2.26 The ways in which services are purchased can influence the nature of competition. In particular, where purchases are individually negotiated the details of these negotiations can influence the competitive process as customers may vary in their

[36] See Section 5 for a discussion of this data including relevant caveats.

needs and their ability to negotiate, and providers may tailor their negotiating strategies to particular customers. We describe below the channels through which customers purchase cloud services, and (where relevant) how contracts are negotiated.

Purchase channels

2.27 Customers purchase cloud services from all the main providers through competitive tenders, bilateral negotiations, from a provider's online portal or marketplace, or through suppliers of professional services, including authorised resellers.[37] Providers have submitted that different customers favour different ways of sourcing cloud services.[38]

2.28 A minority of customers carry out competitive tenders: these are particularly prevalent among public sector customers which often have specific requirements on how to procure IT services.[39] Tenders are more infrequent in the private sector, although one provider said that large companies will typically use competitive tenders to procure either all or some of their cloud services - especially when they are first sourcing them.[40]

2.29 Large enterprise customers – classified by one provider as those with an estimated spend of over £1 million per year[41] - generally procure cloud services through bilateral negotiations with providers.[42] One provider said that this allows a range of customers, including those with higher spend rates to secure bespoke contracts tailored to their needs.[43]

2.30 Smaller enterprise customers are more likely to purchase cloud services directly from providers through their online portals[44] and/or their marketplaces.[45,46] These customers generally pay for cloud services on a pay-as-you-go basis, paying the providers' listed prices.[47] In addition, one provider said that small and medium-

[37] Responses to Ofcom's information requests [✂]; Responses to the CMA's information requests [✂].
[38] The evidence from providers suggests that the purchase channel favoured by different types of customers is not affected by whether said customers are existing or new. Namely, a customer who chooses to procure cloud services through bilateral negotiations will most often renegotiate their contract when the latter comes to term. Similarly, providers submitted that public bodies tend to use tenders (existing and new customers) and that customers who buy cloud services from providers' online portals on a PAYG basis, just purchase the amount of cloud they need when they need it.
[39] Responses to Ofcom's information requests [✂].
[40] [✂] response to Ofcom's information request [✂]; [✂].
[41] [✂] response to the CMA's information request [✂].
[42] Responses to the CMA's information requests [✂].; [✂] response to Ofcom's information request [✂].
[43] [✂] response to the CMA's information request [✂].
[44] Online portals refer to the provider's website where customers can purchase cloud services at listed prices.
[45] Marketplaces are an online platform, where cloud providers and ISVs can offer services to customers, which run on the underlying infrastructure of the provider offering that marketplace. Both Google and Microsoft said customers can procure cloud services through their marketplaces.
[46] [✂] response to the CMA's information request [✂].
[47] [✂] response to the CMA's information request [✂].

sized customers often acquire cloud services through listed prices or bilateral negotiations.[48]

2.31 Some customers purchase cloud services through resellers: these are a network of partners authorised to resell cloud services from the main providers.[49]

(a) Resellers often independently negotiate prices with their customers, owning the billing and contractual relationship with them.[50] Customers who choose resellers tend to require additional technical services (eg IT consulting and support with cloud migration),[51] and may be in an early stage in defining their need for cloud services.[52]

(b) One provider said that small and medium sized business are more likely to engage with its partners for services, as opposed to purchasing cloud services directly,[53] which may reflect the fact that smaller businesses are less likely to have large in-house IT capabilities.

(c) Another provider also said that many customers obtain their cloud services through a combination of direct and partner relationships, depending on their circumstances and workloads.[54]

2.32 Some providers offer incentive programmes to resellers to promote sales of certain services; for example, one provider offers payments for hitting certain sales thresholds[55] and another provider offers discounts for resellers purchasing its cloud services.[56]

2.33 Evidence we have seen to date shows that:

(a) customers buy the large majority of cloud services directly from providers (ie using their online portals or through bilateral negotiations);

(b) there are some notable exceptions, such as one provider for whom a substantial proportion of sales come through sell partners or resellers; and

(c) for most providers, sales through marketplaces[57] constitute only a small percentage of their estimated UK sales by revenue.

---

[48] [✂] response to the CMA's information request [✂].
[49] Responses to Ofcom's information requests [✂].
[50] Responses to Ofcom's information requests [✂].
[51] [✂] response to Ofcom's information request [✂].
[52] [✂] response to Ofcom's information request [✂].
[53] [✂] response to the CMA's information request [✂].
[54] [✂] response to Ofcom's information request [✂].
[55] [✂] response to Ofcom's information request [✂].
[56] [✂] response to Ofcom's information request [✂].
[57] See footnote 45 for a definition of marketplaces.

Contract negotiations

2.34 We describe below the types of contracts customers enter into, and the process of contract negotiation by enterprise customers.

Types of contracts

2.35 There are two main types of contracts that customers can enter when purchasing cloud services: customer agreements and enterprise agreements.

(a) Customer agreements are the standard contracts cloud providers offer through their online marketplaces. When customers have a customer agreement, they typically assume the cloud provider's listed terms and prices. Several cloud providers said that the vast majority of customers are on this type of contract and so do not negotiate additional terms.[58]

(b) Enterprise agreements, which are individually negotiated, are generally reserved for larger customers with higher spending.[59] One provider submitted that they contain commonly requested terms such as invoicing and regulatory compliance commitments.[60]

2.36 Several cloud providers said that UK customers maintain a single contractual relationship with them for the cloud services they deploy.[61] There are exceptions to this though:

(a) One cloud provider said if a large customer has affiliates or subsidiaries then each may have separate agreements and billing accounts.[62]

(b) Another provider said that customers may, in limited instances, want separate contracts for different projects or for teams in different geographies who require isolated accounts for security purposes. This concerns mostly large customers in the banking and financial sector.[63]

2.37 Evidence from cloud providers shows that only a minority of customers negotiate their contracts.[64]

(a) For example, one cloud provider submitted a breakdown of the proportion of its UK customers who contracted its cloud services at listed prices vis-à-vis those who did so through competitive tenders or bilateral negotiations, and hence were more likely to negotiate bespoke contractual terms. This showed

[58] [✂] response to the CMA's information request [✂]; [✂] response to Ofcom's information request [✂].
[59] [✂] response to the CMA's information request [✂].
[60] [✂] response to the CMA's information request [✂].
[61] Responses to the CMA's information requests [✂].
[62] [✂] response to the CMA's information request [✂].
[63] [✂] response to the CMA's information request [✂].
[64] Responses to the CMA's information requests [✂]; [✂] response to Ofcom's information request [✂].

that less than [✂] of the provider's customers in 2022 negotiated contracts bilaterally or through competitive tenders. However, these customers accounted for over [✂] of the provider's UK revenue in 2022.[65]

Emerging view on types of contract

2.38 Based on the evidence we have seen to date, our emerging view is that most customers have standard contracts that have been agreed without negotiation. Some larger customers either engage in bilateral negotiations or tenders and are able to negotiate terms that depart from standard contracts.

How contracts are renegotiated

2.39 One provider said that contracts are often renegotiated at the end of a customer's contract period,[66] although another provider said that renegotiations can also occur during the contract term if the customer's circumstances or requirements change or if there is a change in the regulatory environment.[67] One of these providers said that either the customer will approach it to discuss a renewal or it would reach out to the customer,[68] whereas others said they would typically renegotiate at the customer's request.[69]

2.40 Most providers said that they do not differentiate between new and existing customers, and that the renegotiation process for ongoing contracts is broadly similar to that for new contracts.[70] One provider said that both existing and new customers are eligible for the same discounts, and customers who renegotiate the same overall terms upon the expiry of their original contract are not charged higher prices.[71]

2.41 Two providers said that renewing customers may get particular benefits:

(a) One provider said that, as renewing customers have spent more time using its cloud service, the commercial terms will likely better reflect their requirements.[72]

(b) Another provider said that, while it prefers to keep the same terms, further discounts based on the usage of its cloud services may be offered upon renewal of an existing contract.[73]

---

[65] [✂] response to the CMA's information request [✂].
[66] [✂] response to Ofcom's information request [✂].
[67] [✂] response to Ofcom's information request [✂].
[68] [✂] response to Ofcom's information request [✂].
[69] [✂] response to the CMA's information request [✂] and [✂] response to Ofcom's information request [✂].
[70] Responses to the CMA's information requests [✂]; Responses to Ofcom's information requests [✂].
[71] [✂] response to Ofcom's information request [✂].
[72] [✂] response to the CMA's information request [✂].
[73] [✂] response to the CMA's information request [✂].

2.42 We asked large customers[74] of cloud services about whether they had ever renegotiated their contracts with public cloud providers and whether the terms of the contract improved, worsened or remained broadly the same as part of any renegotiation.

2.43 In response, many of the customers we spoke to said they had renegotiated their contracts for public cloud services in the past, many of them in the previous three years.[75] Their responses indicate that the outcome of these renegotiations often depends on whether customers can commit to higher spending. In particular:

(a) Many customers renegotiated improved terms with at least one of the providers they source cloud services from.[76] This included:

(i) Improved terms consisted primarily of more sizeable discounts, driven by increases in customers' spending commitments.[77] For example one customer said that – based on its own experience dealing with AWS – if customers are unable to commit to increased levels in spend, there is a real risk that their provider would not offer a like-for-like agreement, particularly in terms of the discount available.[78]

(ii) An increased discount even if other terms worsened after renegotiation[79] [80] or more favourable payment terms with a provider even if they were still worse than those offered by other suppliers and third-party vendors.[81]

(iii) Improved non-price commercial terms (eg features, services, liabilities), from leveraging the potential use of new technology (eg generative AI)[82] or innovations in the cloud industry. [83]

(b) In a few cases customers said that some of their terms had worsened upon renegotiation.[84] For example, one customer cited price list increases and reduced discounts, while another customer said its cloud provider had made enterprise support conditional on purchasing a more expensive support package the customer did not need as the cloud provider was no longer offering the support package that it previously provided.

---

[74] We classified large customers as those with over £5m revenue. To the extent that a provider did not have customers generating over £5m worth of revenue, we used their top 10 customers
[75] [✂].
[76] [✂].
[77] [✂].
[78] [✂] response to the CMA's information request [✂].
[79] For example, one customer said that '[✂].
[80] Responses to the CMA's information requests [✂].
[81] [✂] response to the CMA's information request [✂].
[82] [✂] response to the CMA's information request [✂].
[83] [✂] response to the CMA's information request [✂].
[84] Responses to the CMA's information requests [✂].

(c) One customer said that on renewal its provider tried to raise costs, but it was able to push back against this price increase.[85]

2.44 In fewer cases customers said that their contract terms had remained broadly the same after renegotiation.[86] For example, one customer said its discount level was reduced after the renegotiation, and a minimum spend commitment was introduced. However, its provider compensated it by offering additional cloud credits linked directly to targets for the number of workloads the customer migrated to the provider's platform within a specified timeframe.[87]

Parameters of competition

2.45 We refer to the ways in which providers flex their offerings to meet customer preferences as the parameters of competition.

2.46 We present below evidence we have seen to date from both providers and customers on the relative importance of various parameters. We then describe the extent to which parameters are negotiated by enterprise customers.

Evidence from cloud providers

2.47 Cloud providers have identified a number of parameters of competition and we set these out below.

2.48 Most cloud providers identified price as a key parameter of competition for cloud services.[88] Providers compete on this parameter by adjusting the prices customers face through:

(a) Discounts: including committed spend agreements which are predominantly available to larger customers (see paragraph 2.31) and publicly listed discounts available to all customers. In relation to the latter, one provider said that it offers public discounting options which 'enable customers who can pay upfront fees and/or commit to use a service for a specific period' to optimise their costs.[89]

(b) Credits: most providers offer credits to discount customers' use of cloud services, encouraging them to trial a provider's cloud capabilities.[90] These credits can be offered through programmes targeted at specific customer groups (eg, [✂] and [✂] have credit programs for new start-up customers[91]).

[85] [✂] response to the CMA's information request [✂].
[86] [✂].
[87] [✂] response to the CMA's information request [✂].
[88] Responses to Ofcom's information requests [✂]; Responses to the CMA's information requests.
[89] [✂] response to the CMA's information request [✂].
[90] Responses to Ofcom's information requests [✂].
[91] Responses to Ofcom's information requests [✂].

(c) Free tiers: free tiers function similarly to credits in that they allow customers to trial some of a providers' services at no cost. Providers can impose limitations on these free tiers. For example, one provider said it enables customers to benefit from certain of its services free of charge for a certain amount of time or usage.[92] Similarly, another provider allows customers to take advantage of its free tier to use a selection of its products, provided that they remain within specified monthly usage limits.[93] On the other hand, another provider offers a free tier for many services with duration and capacity stipulated by service.[94]

2.49 Aside from price, providers also identified a number of additional parameters of competition in the supply of cloud services.

(a) Several providers said ease of migration is a factor considered by consumers.[95]

(i) For example, one provider said that it seeks to offer technical solutions that make it easier for customers to transfer all or part of their workloads to and from them.[96]

(ii) Another provider submitted that it offers dedicated technical support, professional services and guidance to ensure prospective customers can smoothly migrate data to its cloud services.[97]

(iii) One provider said that, as a challenger, developing products and features that facilitate multi-cloud strategies is critical to its ability to win workloads from the two incumbents.[98]

(b) Several providers said security and data protection are important factors in driving customer acquisition and retention.[99]

(i) One provider said that it anticipates data privacy and cybersecurity will become increasingly important drivers of customer choice in the future, as more customers migrate their workloads to cloud, and regulatory requirements intensify.[100]

[92] [✂] response to Ofcom's information request [✂].
[93] [✂] response to Ofcom's information request [✂].
[94] [✂] response to Ofcom's information request [✂].
[95] Responses to Ofcom's information requests [✂]; Responses to the CMA's information requests [✂].
[96] [✂] response to the CMA's information request [✂].
[97] [✂] response to Ofcom's information request [✂].
[98] [✂] response to the CMA's information request [✂].
[99] Responses to the CMA's information requests [✂]; Responses to Ofcom's information requests [✂].
[100] [✂] response to Ofcom's information request [✂].

(c) Two providers said innovation is a parameter of competition which is continuously striven for by providers.[101]

(i) One of them said that it competes on this parameter by continuously striving to develop products that are innovative and meet evolving demands.[102]

(ii) The other said that the dynamism of the sector for IT services makes iterative innovation necessary to attract and retain customers.[103]

(d) Another provider said that the availability of advanced cloud features was also an important factor influencing customer choice.[104]

(e) Several providers said that reliability[105] of a platform is an important consideration for customers.[106] One of the providers said that it works hard to ensure its cloud platform has lower down time than those of its main competitors.[107]

(f) Two providers highlighted that customers want flexibility to deploy cloud services in combination with their traditional IT infrastructure (hybrid deployment).[108]

(g) One provider said customers care about the elasticity of its services, which allows them to provision the resources that they actually need, knowing they can instantly scale up or down along with the needs of their business.[109]

(h) One provider said that customers' historical choices remain a leading factor in influencing their choice of cloud provider today. The provider said that the vast majority of companies choose to remain with their original provider (often either [✂] or [✂]).[110]

(i) Another provider said that brand trust is an important factor diving consumer choice.[111]

(j) One provider said that technical support was an important customer criterion for choosing a cloud provider but did not clarify what this support would

---

[101][✂] response to Ofcom's information request [✂]; [✂] response to the CMA's information request [✂].
[102] [✂] response to Ofcom's information request [✂].
[103] [✂] response to the CMA's information request [✂].
[104] [✂] response to the CMA's information request [✂].
[105] Reliability is often captured by the uptime: a measure of the amount of time that a system or service is available and operational without any planned downtime.
[106] [✂] response to the CMA's information request [✂]; [✂] response to Ofcom's information request [✂]; [✂] response to the CMA's information request [✂].
[107] [✂] response to Ofcom's information request [✂];
[108] [✂] response to the CMA's information request [✂]; [✂] response to the CMA's information request [✂].
[109] [✂] response to the CMA's information request [✂].
[110] [✂] response to Ofcom's information request [✂].
[111] [✂] response to the CMA's information request [✂].

entail.[112] Another provider also said that its customer service is one of its key strengths.[113]

2.50 Two providers said that different customer groups prioritise different factors when choosing a cloud provider.[114]

(a) One of these providers clarified that it does not perceive these differences to be driven by customer size or where customers are based.[115] It said that:

(i) Customers in regulated sectors (eg, financial services, healthcare or telecommunications) tend to prioritise infrastructure security, resilience and data privacy in order to fulfil their regulatory obligations and the needs of their customers. This point was reiterated by another provider.[116]

(ii) Similarly, public sector customers may prioritise considerations around data sovereignty and national security, as well as the costs of the services procured.

(iii) Customers in the technology sector and developers may prioritise advanced and innovative features, functionality and technical flexibility, and customisation.

(iv) Customers active in media, gaming and streaming will prioritise data latency, favouring proximity to the data centre to minimise the former.

(b) The other provider said that in industries such as retail, there may be a greater focus on cost-competitiveness, modernised workloads, availability and reputation.[117]

2.51 One provider said that the increasingly heterogeneous nature of customers' needs has 'opened opportunities for existing and new cloud providers to differentiate themselves in different industry and workload verticals, without the need for hyperscaler scale'.[118]

Evidence from customers

2.52 We asked large customers to rate the importance of a list of factors their organisation considers when choosing their main public cloud. They were asked to

[112] [✂] response to the CMA's information request [✂].
[113] [✂] response to the CMA's information request [✂].
[114] Responses to the CMA's information requests [✂].
[115] [✂] response to the CMA's information request [✂].
[116] [✂] response to the CMA's information request [✂].
[117] [✂] response to the CMA's information request [✂].
[118] [✂] response to the CMA's information request [✂].

rate these factors from one to five, where one is not important at all and five is very important. Not all customers ranked every factor.

2.53 The following factors were identified as the most important by the large customers we heard from when choosing their main public cloud provider, with an average rating that rounds to either five or four.

(a) Service quality**:** almost all customers considered this to be important or very important. Although one customer said this is largely not a differentiator between the main providers as they have similar offerings.[119]

(b) Price, including discounts or cloud credits: most customers considered this to be important or very important.

(c) Data sovereignty requirements: the majority of customers considered compliance with data localisation, privacy and protection regulations to be very important and one considered it to be important. Some customers said that almost all providers can support customers' usual requirements around data protection and sovereignty.[120]

(d) Range of cloud infrastructure services: the majority of customers considered this to be important or very important. One customer added that the main cloud providers have broadly similar capabilities in this regard, with AWS having a slightly superior offering.[121]

(e) Number and location of data centres: the majority of customers considered this to be important or very important. Some customers said this was important as they must comply with data sovereignty regulations.[122] Other customers cited resiliency concerns. One customer said that it is important for services to be distributed across multiple locations to minimise the risk of disruptions, but added that, provided there are sufficient locations, the absolute number of locations between providers is not the ultimate determining factor when choosing a provider.[123]

2.54 The following factors were identified as the next most important factors by customers when choosing their main public cloud provider, with an average rating that rounds to three.

(a) Cost and ability to use software licences: we will consider software licensing in more detail in a separate working paper.

---

[119] Responses to the CMA's information requests [✂].
[120] Responses to the CMA's information requests [✂].
[121] [✂] response to the CMA's information request [✂].
[122] Responses to the CMA's information requests [✂].
[123] Responses to the CMA's information requests [✂].

(b) Cloud-specific skills of employees: one customer who identified it as very important, said that cross-skilling its employees between different cloud providers can prove costly in terms of productivity loss.[124] On the other hand, another customer who gave this factor the lowest rating, said that it can just match the skills of its employees to the platform of choice, rather than the other way around.[125]

(c) Existing relationship with the cloud provider: a few customers said that switching providers is both time-consuming and costly.[126]

(d) Ease of integration with existing technology: there were multiple cases where customers considered this to be important or very important. A few of these customers said that this factor was important, as they needed their cloud solutions to integrate with their on-premises environments.[127] A handful of other customers did not assign much importance to this factor, as they either considered most providers to have similar capabilities in this respect or they used third-party solutions to facilitate integration.[128] Additionally, a few customers said that this factor is of decreasing importance to them, as they are moving all their workloads to the cloud.[129]

2.55 The following were considered as of moderate importance by customers, with an average rating that rounds to two:

(a) Range of cloud infrastructure services offered by ISVs: some customers considered this to be important. A few customers said that most ISVs are usable on all providers' clouds, so this is not a differentiating factor between providers.[130] One customer said that the extent to which a provider offers services via ISVs can be taken as a signal of the quality of its cloud offering, as ISVs must have had faith in its cloud services to have built their software stack on them.[131]

(b) Ease of integration with other public clouds: there were cases where customers considered this to be important, with a few of them citing their current multi-cloud strategy as the rationale behind their ranking.[132, 133] Some of the customers who gave it a low importance rating said that integration with other public clouds can happen through internet facing APIs or intermediaries, so a provider's capabilities in this area are not a major

[124] [✂] response to the CMA's information request [✂].
[125] [✂] response to the CMA's information request [✂].
[126] Responses to the CMA's information requests [✂].
[127] Responses to the CMA's information requests [✂].
[128] Responses to the CMA's information requests [✂].
[129] Responses to the CMA's information requests [✂].
[130] Responses to the CMA's information requests [✂].
[131] [✂] response to the CMA's information request [✂].
[132] See below section 3, types of multi-cloud.
[133] Responses to the CMA's information requests [✂].

concern when choosing a cloud provider.[134] Others of these customers said that most providers have similar capabilities in this respect, with one adding that no provider offers seamless integration with their competitors.[135]

(c) AI capabilities: in a handful of cases customers considered this to be important or very important, with many giving it a rating of two. However, there were also multiple cases where customers said that providers' AI capabilities are becoming an increasingly important consideration.[136] For example, one customer said that it increasingly sees AI capabilities as a differentiator between providers.[137]

Negotiations

2.56 We also asked providers and customers about the terms they negotiate on as this can indicate the terms and thus parameters of competition that cloud providers are flexing in practice.

2.57 Cloud providers told us that large customers typically negotiate the price of service and many of them choose to enter into CSDs.[138] This is particularly the case for customers requiring continued cloud services and who can broadly predict their minimum spend across an extended timeframe.[139]

2.58 We asked large customers about the key considerations or parameters when negotiating terms with public cloud providers. Most customers we spoke to said that price was a key consideration when negotiating terms with providers. For example, there were cases where customers mentioned minimum committed spend discounts[140] and in some of those cases the customer[141] said that important factors that they negotiated on included the timeframe over which said commitment needs to be achieved and the ability to roll over some of the commitment within a grace period.

2.59 We will consider committed spend agreements and discounts in more detail in a separate working paper.

2.60 Cloud providers reported a variety of levels of willingness to negotiate non-price terms. In particular:

(a) One cloud provider said that it enables its customers to include amendments to their [✂] Agreements to facilitate their meeting of their regulatory needs,

[134] Responses to the CMA's information requests [✂].
[135] Responses to the CMA's information requests [✂].
[136] Responses to the CMA's information requests [✂].
[137] [✂] response to the CMA's information request [✂].
[138] Responses to the CMA's information requests [✂]; Responses to Ofcom's information requests [✂].
[139] [✂] response to the CMA's information request [✂]; [✂] response to Ofcom's information request [✂].
[140] Responses to the CMA's information requests [✂].
[141] Responses to the CMA's information requests [✂].

additional security commitments, and modified liability terms. However, the cloud provider said it does not negotiate with customers on the technical characteristics of its services.[142]

(b) Another cloud provider said that, for its strategically important customers, contracts tend to include specific terms and services where certain product, legal, financial – and, where applicable, pricing – provisions are negotiated on a more bespoke basis. It also said it regularly concludes amendments to its standard template contract if there is a willingness from customers to further negotiate commercial and/or legal conditions.[143]

(c) A third cloud provider said that it can only occasionally adjust the terms and conditions available to customers, as its cloud services agreement has standard terms that reflect the underlying specification and characteristics of the service. It said that it will typically avoid amending these terms unless the service is bespoke or there is a substantial and critical reason to do so. It said that its willingness to amend the technical characteristics of services ultimately depends on what service is being procured. If it is something that can be adjusted or is bespoke, then it said it will consider exploring this.[144]

(d) A fourth cloud provider offers similar terms to all its customers, aside from those in highly regulated industries who might receive additional terms based on their specific regulatory requirements. It also stated that it rarely negotiates with customers on technical characteristics because to do so goes against the nature of a standardized service such as cloud.[145]

2.61 A few large customers said that it is challenging to negotiate bespoke non-price terms with providers, as cloud providers often offer “off the shelf” cloud services with standardised contract dimensions around service level agreements, liability provisions and technical specifications.[146]

2.62 The non-price parameters that customers we spoke to mentioned more often were:[147]

(a) Service performance, availability, and reliability. Many customers considered the availability and reliability of cloud services to be a key parameter in negotiations with providers. Many of these customers said they negotiate

[142] [✂] response to Ofcom’s information request [✂].
[143] [✂] response to the CMA’s information request [✂].
[144] [✂] response to Ofcom’s information request [✂].
[145] [✂] response to Ofcom’s information request [✂].
[146] Responses to the CMA’s information requests [✂].
[147] In addition to these points, three customers [✂, ✂, ✂] said that the scalability and flexibility of a provider’s cloud services is a key considerations and three other customers [✂, ✂, ✂] said they considered parameters associated with the exit terms of their contracts.

enhanced Service Level Agreements with providers[148, 149] or more generally negotiated to ensure reasonable performance levels.[150]

(b) Security and data protection. There were cases where customers said that data protection and security provisions are key considerations in their negotiations with providers. For example, a few customers said that where data is hosted relates to data protection considerations.[151]

(c) Commercial terms and contract length. Many customers said that they would seek to negotiate commercial terms[152] while others said the term of the contract is an important consideration.

(d) Support from providers. There were cases where customers also said that the level of support offered by providers was an important parameter of negotiations. Some of these customers referred to support with the migration to cloud, as well as with the initial deployment process, some referred to providers offering training courses or credits and some referred to the provision of professional services.[153]

Main providers

2.63 In this section, we provide an overview of the main providers of cloud services in the UK. We also include independent software vendors (ISVs) which supply PaaS and do not own any of the underlying raw computing services.[154]

Cloud providers in the UK

2.64 Based on evidence we have gathered to date, AWS, Microsoft, Google, IBM and Oracle are, in that order, the largest vertically integrated cloud providers in the UK by revenue (see Section 5 where we set out our analysis of shares of supply).

2.65 As well as setting out relevant information on each of these providers, we also consider evidence on their business strategies, including their pricing strategies and their customer base. We have also considered evidence on their business strategies from analyst reports and customers' views on each provider.

2.66 This provides context to our wider analysis: in particular, understanding the history, business strategies and relative strengths and weaknesses of different

[148] Responses to the CMA's information requests [✂].
[149] SLAs guarantee an agreed minimum level of uptime, speed, reliability, and availability of cloud services, and set out the compensation and remedies customers will receive in the event of service failure or disruption. [✂]
[150] Responses to the CMA's information requests [✂].
[151] Responses to the CMA's information requests [✂].
[152] Examples provided included terms of use, enterprise agreement terms or the level of investment a provider makes into a customer account.
[153] Responses to the CMA's information requests [✂].
[154] The cloud providers and ISVs also offer SaaS services, which are not included in the reference made by Ofcom.

providers allows us to properly understand competitive dynamics and as such feeds into our assessment of the extent of any market power as well as into the four theories of harm outlined in our issues statement.

2.67 The section below is based on evidence collected and analysed to date. We will continue to consider the business strategies of these cloud providers. This will include:

(a) analysing internal documents received from these providers on their strategy and business plans including around product development decisions.

(b) analysing evidence received from these providers on the relative strengths and weaknesses of each of the main suppliers of IaaS and PaaS.

Amazon Web Services (AWS)

2.68 AWS is a subsidiary of Amazon and started providing public cloud services in 2006.[155]

2.69 AWS' first service was 'Amazon S3(TM)' a storage service[156] and it has subsequently expanded to offer a large number of IaaS, PaaS and SaaS services.

2.70 AWS describes itself as 'the world's most comprehensive and broadly adopted cloud, offering over 200 fully featured services from data centers globally.'[157]

2.71 AWS lists its top product categories as compute, storage, database, networking and content delivery, analytics, machine learning and security, identity and compliance.[158] Of these, [✂] accounted for [✂] public cloud revenue in both the UK and globally in 2022. The next largest category is [✂] with [✂] in both the UK and globally in 2022.[159]

2.72 AWS serves 245 countries and territories across the world and its operations are organised within 33 geographic regions; it has announced plans to expand to four more 'AWS regions'.[160]

2.73 Amazon, the parent company of AWS, is also active in online retail (eg amazon.co.uk), entertainment (eg Amazon Prime Video, Audible, Amazon Music), devices and services (eg Alexa, Fire Tablets) and delivery and logistics.[161] AWS provides solutions to some of these business such as Amazon Prime Video.[162]

---

[155] About AWS (archive.org), accessed 10 February 2024.
[156] Amazon-Web-Services-Launches (aboutamazon.com), accessed 10 February 2024.
[157] What is AWS? - Cloud Computing with AWS - Amazon Web Services, accessed 10 February 2024.
[158] Cloud Products (amazon.com), accessed 10 February 2024.
[159] AWS response to the CMA's information request [✂].
[160] Global Infrastructure (amazon.com), accessed 10 February 2024.
[161] Amazon: What We Do (aboutamazon.com), accessed 10 February 2024.
[162] Prime Video Case Study (amazon.com), accessed 10 February 2024.

2.74 Amazon categorises its operations into North America, International and AWS.[163] Amazon's total global revenue was $575 billion in 2023 with AWS accounting for 16% of that revenue ($91 billion).[164] This figure has increased from 10% in 2017 with AWS revenue growing at a compound average growth rate of 32% per year over the same period.[165]

Public cloud strategy and views of customers

Evidence from AWS

2.75 AWS said that its 'overarching business strategy' is to 'work backwards' from 'what would be most attractive to the customer'.[166] AWS also said that the competition it faces provides it with strong incentives to innovate and that, given its overarching business strategy, this innovation is customer focused. In particular:

(a) AWS said that, given the competition it faces, it 'needs to keep innovating to attract and retain customers'[167] and that it had 'regularly introduced new services and thousands of new features in existing services each year' as well as continuously improving 'the quality and security of its offerings.'[168]

(b) AWS said that, due to the 'speed at which things change', 'the guiding principle behind the development of AWS' offer is iterative innovation.' AWS said that this allows it 'to continuously consider what AWS' customers may want or need and drives it to constantly innovate and develop new products and services.'[169]

(c) AWS said that it 'remains close to customers and focuses on elements it knows they will value over the long-term (eg, performance, security, breadth and depth of features and functionality, and cost performance of AWS' cloud services).'[170]

2.76 AWS said that its services are typically priced on a pay as you go basis.[171] AWS said that its pricing strategy is to 'sets its prices to ensure its offerings compete

---

[163] The AWS segment consists of amounts earned from global sales of compute, storage, database, and other services for start-ups, enterprises, government agencies, and academic institutions. These first two categories primarily consist of amounts earned from retail sales and advertising and subscription services. amzn-20231231 (sec.gov), accessed 10 February 2024.
[164] Amazon 2023 Form 10-K, page 24.
[165] CMA analysis of Amazon's reported accounts: Amazon 2023 Form 10-K, page 24; Amazon 2017 Form 10-K, page 25.
[166] AWS' response to the CMA's information request [✂].
[167] AWS' response to Ofcom's information request [✂].
[168] AWS' response to the CMA's information request [✂].
[169] AWS also said that it 'focuses on customer needs and prioritises efficient utilisation of resources and time to market. What AWS builds is driven by what customers tell AWS. Some novel offerings were created in response to a single customer's needs, but once developed, can be used to also benefit other customers.' AWS response to Ofcom's information request [✂].
[170] AWS' response to Ofcom's information request [✂].
[171] AWS' response to Ofcom's information request [✂].

effectively with alternatives available to customers'.[172] AWS said it seeks to 'continually lower customers' costs where possible through price reductions over the long run to stay competitive […] and to provide a cost-effective proposition.'[173]

2.77 AWS also offers a free tier, like other providers, and AWS said that this 'enables many customers to benefit from certain AWS services free of charge for a certain amount of time or usage.'[174]

2.78 AWS said that its customers comprise 'a variety of types' such as 'start-ups, mid-market companies, established enterprises, governments, and academic organisations' and 'span a wide range of industries'.[175]

2.79 We asked each of the main cloud providers to list their understanding of the factors UK customers value the most when choosing them as a provider. In response, AWS said that there are a number of reasons why customers may decide to use it including:[176]

(a) Innovation – AWS said that customers can focus their efforts on innovating and developing applications that differentiate their business and transform customer experiences instead of focusing on the underlying infrastructure.

(b) Resources – AWS said customers can quickly increase resources as they need them.

(c) Elasticity – AWS said that customers can instantly scale up or down along with the needs of their business, which also reduces cost and improves the customer's ability to meet their user's demands.

(d) Flexibility – AWS said that customers can use AWS' services in combination with their on-premises IT infrastructure (eg, hybrid) and with other cloud and IT services providers (eg, multi-cloud).

(e) Global reach – AWS said that customers can use AWS to deploy globally quickly.

(f) Cost savings – AWS said that customers can trade capital expense for variable expense, and only pay for IT as they consume it.

[172] AWS' response to Ofcom's information request [✂].
[173] AWS' response to the CMA's information request [✂].
[174] AWS' response to Ofcom's information request [✂].
[175] AWS' response to Ofcom's information request [✂].
[176] AWS' response to the CMA's information request [✂].

Evidence from analyst reports and customers

2.80 AWS is identified in analyst reports as being the leading cloud provider. One analyst said that not only is AWS the leader, but it is managing to keep ahead of the pack and is differentiating by investing in hardware.[177]

2.81 More generally based on our review of analyst reports:[178]

(a) AWS is identified as having strengths in areas such as: the breadth and depth of its capabilities; its strong ecosystem; its high level of innovation; its high brand awareness with customers; and its investments in hardware that has boosted compute capabilities.

(b) AWS is identified as having weaknesses such as: eroding customer relationships by focusing on the short term especially around contract renewal; a relatively weak strategy to support customers seeking multi-cloud[179] and 'sovereign solutions'; its usability (eg lack of quality service documentation) and costs, such as egress fees; its connection with Amazon means it has lost some retail business; and its open source offering which is less of a priority for AWS compared to other providers.

2.82 We asked large customers to rate the suitability of a list of public cloud providers[180] as alternatives to their main public cloud provider[181] based on their perception or any direct experience. They were asked to rate these providers from one to five, where one was not effective as an alternative and five was fully effective as an alternative. We note that in asking about whether a cloud provider is a suitable alternative to the customer's main provider this does not capture that some cloud providers may not be suitable as a main provider, but be a suitable alternative for some workloads. In some cases cloud providers may exert at least some competitive constraint based on an offering targeted at certain market segments.

2.83 AWS was the main provider for many of the customers we spoke to and, among other customers, most identified AWS as an effective or fully effective alternative with its average rating rounding to 4. These customers generally identified AWS as being effective or very effective as an alternative due to reasons such as its wide range of products and services.[182] One customer did not consider AWS as an alternative due to [✂].[183]

---

[177] [✂].
[178] [✂].
[179] The analyst report notes that 'AWS shows little incentive or interest in pursuing meaningful multi-cloud strategies on behalf of its customers'. [✂]
[180] The list provided was: AWS, Microsoft, Google, Oracle, IBM, OVHcloud, Alibaba and 'other'.
[181] This was self-identified based on revenue spend.
[182] Eg [✂].
[183] [✂] response to the CMA's information request [✂].

Microsoft

2.84 Microsoft started providing cloud services in 2008 through Windows Azure.[184] This was originally designed for developers to deploy apps in the cloud and was thus a PaaS product.[185]

2.85 Microsoft made Windows Azure more widely available in 2010[186] and subsequently expanded to offer a large number of IaaS, PaaS and SaaS services. Microsoft's website describes Azure as a cloud platform with 'more than 200 products and cloud services'.[187]

2.86 In 2022, Azure's top product category was compute which accounted for around [✂] of its revenue both in the UK and globally. The next largest category was storage which accounted for around [✂] of its revenue both in the UK and globally.[188]

2.87 Microsoft's website lists over 70 regions where Azure is either available or coming soon.[189]

2.88 Microsoft is also active in a range of other products and services and organises its services and products into three operating segments: Productivity and Business Processes (eg Office 365, LinkedIn), Intelligent Cloud (eg cloud services, enterprise services) and More Personal Computing (eg Windows operating system, search and news advertising).[190]

2.89 Microsoft's total global revenue was $212 billion in its 2023 financial year[191] with Microsoft's Intelligent Cloud segment, which includes Azure, accounting for 41% of that revenue ($88 billion).[192] This figure has increased from 28% in 2017 with Microsoft's Intelligent Cloud segment revenue growing at a compound average growth rate of 21% per year during that period.[193]

---

[184] About Microsoft - Stories, accessed 11 February 2024.
[185] Microsoft launches Windows Azure - Stories, accessed 11 February 2024.
[186] Windows Azure Platform Now Generally Available in 21 Countries | Blog Azure | Microsoft Azure, accessed 11 February 2024.
[187] What is Azure—Microsoft Cloud Services | Microsoft Azure, accessed 11 February 2024.
[188] [✂] response to the CMA's information request [✂].
[189] Choose the Right Azure Region for You | Microsoft Azure, accessed 11 February 2024.
[190] Segment Information - Microsoft Investor Relations, accessed 11 February 2024.
[191] 1 July 2022 to 30 June 2023.
[192] CMA analysis of Microsoft FY23 Form 10-K, page 45.
[193] CMA analysis of Microsoft FY23 Form 10-K, page 45; Microsoft FY18 Form 10-K, page 35.

Public cloud strategy and views of customers

Evidence from Microsoft

2.90 Microsoft said that it discusses customers' business requirements and illustrates 'how Microsoft's solutions could align with their operations.'[194] Microsoft also said that it has a unit 'dedicated to facilitating customers with their project rollouts' and 'committed to maximising customer utilisation of the platform and optimising customer experiences and platform usage.'[195] Further, Microsoft said that it 'invests heavily in assisting customers transition to the cloud.' Microsoft said this included through 'technical support, information tools and discount programs to aid migration.'[196]

2.91 Microsoft said that a large number of customers pay for its services on a pay-as-you-go-basis. Microsoft said that these 'prices are "metered" differently, depending on the nature of the service.'[197] Microsoft said that its pricing strategy is 'to ensure that its prices are competitive with the comparable services offered by key competitors, including for example AWS, GCP and ISVs (eg Snowflake).'[198]

2.92 In relation to IaaS, Microsoft said that it 'considers itself [✂] because prices for these services are typically set by reference to [✂]'.[199] In relation to PaaS, Microsoft said that it 'takes a "market based" approach which considers the incremental value the service provides to customers as well as the price charged by competitors for similar services.'[200]

2.93 Microsoft also said that its approach is dependent on whether the product/service is comparable to the products/services of other providers, where it will refer to the public prices of comparable products/services,[201] or an innovative product/service on the 'rare' occasions where there is no directly comparable product/service.

2.94 On the latter Microsoft said that '[✂].'[202]

2.95 Despite these differences, Microsoft said that it '[✂],' as its pricing policy is either constrained by competitor pricing for directly comparable cloud services or by a long run incentive to develop the market for certain nascent cloud services.'[203]

---

[194] Microsoft response to the CMA's information request [✂].
[195] Microsoft response to the CMA's information request [✂].
[196] Microsoft response to the CMA's information request [✂].
[197] Microsoft gave the examples of 'Azure Blob Storage is priced on a per gigabyte basis while Azure Compute is priced on a per hour or per month basis depending on the compute power of the virtual machine.' Microsoft response to Ofcom's information request [✂].
[198] Microsoft response to the CMA's information request [✂].
[199] Microsoft response to the CMA's information request [✂].
[200] Microsoft response to the CMA's information request [✂].
[201] Microsoft response to Ofcom's information request [✂].
[202] Microsoft response to Ofcom's information request [✂].
[203] Microsoft response to Ofcom's information request [✂].

2.96 Microsoft also said that, like others, it 'offers credits to encourage customers to try Azure.' Microsoft said that this is because it 'knows that customers have a lot of choices when it comes to cloud computing services' and 'Azure credits provide customers with a no cost path to explore and test Microsoft's service.'[204]

2.97 Microsoft said that it targets all customers and it has a range of customers based on type (eg public vs private sector[205]), size (eg large enterprise customers to start-ups)[206] and in terms of industries.[207]

2.98 When asked to list its understanding of the factors that UK customers value the most when choosing Microsoft over other cloud providers, Microsoft said that 'there is no standard scenario when customers choose between cloud providers' so no uniform/standard set of factors.[208]

2.99 Microsoft said that it had identified the following non-exhaustive factors as important in driving new customers and customer retention: security, brand trust, compliance (eg with regulatory / technical requirements), cost, hybrid scenarios, ease of migration, efficiency and up time, and availability of advanced cloud features such as analytics and AI.' [209]

Evidence from analyst reports and customers

2.100 Microsoft is identified by analyst reports as being the second leading provider overall behind AWS. One analyst said that Microsoft is closing the gap on AWS both globally and particularly in Europe.[210]

2.101 More generally based on our review of analyst reports:[211]

(a) Microsoft is identified as having strengths in areas such as: the breadth of its capabilities and its complementary ecosystems partners; its early differentiated position in broad segments such as telecom, healthcare, manufacturing, retail and financial services; its hybrid and multi-cloud offering; its powerful and diverse platform with a distinct advantage due to Microsoft 365 and Windows products; its success in attracting developers; being well placed to attract further customers/workloads due to the large number of Windows workloads still on premises; and being well placed to leverage long standing customer relationships.

---

[204] Microsoft said that [✂] Microsoft response to Ofcom's information request [✂].
[205] Microsoft response to the CMA's information request [✂].
[206] Microsoft response to the CMA's information request [✂]; Microsoft response to Ofcom's information request [✂].
[207] Microsoft response to the CMA's information request [✂].
[208] For example, Microsoft said '[✂]' whereas Microsoft said '[✂]' Microsoft response to the CMA's information request [✂].
[209] Microsoft response to the CMA's information request [✂].
[210] [✂].
[211] [✂].

(b) Microsoft is identified as having weaknesses such as: past security issues and a lack of innovation; opaque costs – eg due to poor cost management capabilities; 'punitive licensing' – eg when products not used on Azure; high profile outages; needing more wireless communications capabilities and compute instance types; and lack of visibility into Azure's capabilities.

2.102 We asked large customers to rate the suitability of a list of public cloud providers as alternatives to their main public cloud provider (see paragraph 2.82 above).

2.103 Microsoft was the main provider for many customers we spoke to and among other customers most identified Microsoft as an effective or fully effective alternative with its average rating rounding to 4. These customers generally considered Microsoft's offering to be similar to or equally as effective as that of AWS (who was the main provider for most of the customers),[212] although a few said Microsoft's offering was slightly less advanced/extensive when compared to AWS' offering.[213]

2.104 Scores lower than 4 were given due to a weaker offering in specific sectors [✂, ✂],[214] lack of direct engagement and complex solution integrations in some cases[215] and a more limited PaaS offering and poorer experience with account management.[216]

Google

2.105 Google started providing cloud services in 2008 when it previewed the Google App Engine, a platform enabling businesses to develop applications (PaaS), and then launched Google Storage for Developers in 2010.[217] Initially only being available to developers, it became more widely available in 2011[218] and expanded to offer a large number of IaaS, PaaS and SaaS services. Google Cloud's website identifies it as having over 150 products.[219]

2.106 Google Cloud is available in over 200 countries and territories based on 40 regions. It has announced plans to continue expanding in eight regions.[220]

[212] Eg [✂].
[213] Eg [✂].
[214] Responses to the CMA's information requests [✂].
[215] [✂] response to the CMA's information request [✂].
[216] [✂] response to the CMA's information request [✂].
[217] Google App Engine Blog: Introducing Google App Engine + our new blog, accessed 11 February 2024 Google Storage for Developers: A Preview - The official Google Code blog, accessed 11 February 2024.
[218] The History of Google Cloud Platform (pluralsight.com), accessed 11 February 2024.
[219] Cloud Computing Services | Google Cloud, accessed 11 February 2024.
[220] Global Locations - Regions & Zones | Google Cloud, accessed 11 February 2024.

2.107 Google is also active in a range of other products and services including digital advertising and products and services relating to mobile ecosystems.[221] Google runs some of its own products and services (eg Google Search) on the same infrastructure as Google Cloud.[222]

2.108 Alphabet, the parent company of Google, organises its operations into three segments: Google Services (eg advertising, Google Maps, YouTube), Google Cloud (eg Google Cloud Platform, Google Workspace collaboration tools) and Other Bets[223] (combination of all other services).[224]

2.109 Alphabet's total global revenue was $307 billion in 2023 with Google Cloud, which includes Google Cloud Platform, accounting for 11% of that revenue ($33 billion).[225] This figure has increased from 4% in 2017 with Alphabet's Google Cloud segment revenue growing at a compound average growth rate of 42% during that period.[226]

Public cloud strategy and views of customers

Evidence from Google

2.110 Google has said that it is a 'challenger' as its market share (globally and in the UK) is closer to smaller players (eg Oracle, IBM) than to AWS and Microsoft.[227] Google said that this affects how it competes and its strategy. For example, Google said that it is often 'seeking to compete to be the second or even third providers' as most customers already use either 'AWS (particularly for digital native) or Microsoft (particularly for traditional enterprise customers).'[228] Google said that this means it is particularly important for Google to show it adds value to customers and meets their needs and to make switching and multi-clouding of workloads as easy as possible for potential new and existing customers.[229]

2.111 Google said that it has four global objectives focused on: growing its cloud business, including attracting the 'largest enterprises and fastest growing digital natives'; developing and launching 'new products that meet evolving customer

[221] For example, see the CMA's Online Platforms and digital advertising market study final report at Final report (publishing.service.gov.uk) and the CMA's Mobile ecosystems market study final report at Final report (publishing.service.gov.uk).
[222] What is Cloud Computing? | Google Cloud, accessed 11 February 2024 and Google to migrate parts of YouTube to Google Cloud - DCD (datacenterdynamics.com), accessed 11 February 2024.
[223] According to Alphabet, revenues from Other Bets are generated primarily from the sale of health technology and internet services.
[224] GOOG Exhibit 99.1 Q1 2022 (abc.xyz), accessed 11 February 2024.
[225] Alphabet 2023 Form 10-K, page 35.
[226] CMA analysis of Alphabet 2023 Form 10-K, page 35; Alphabet 2019 Form 10-K, page 61.
[227] Google's response to the CMA's information request [✂].
[228] Google's response to the CMA's information request [✂].
[229] Google's response to the CMA's information request [✂].

needs'; 'offering differentiated customer services and developer experiences'; and building a robust enterprise cloud organisation.[230]

2.112 Google said that it focuses on 'designing new products/features' based on 'anticipating customers' evolving demands' and focusing on its 'differentiating strengths'.[231] Google provided two specific examples of this:

(a) Google said it has improved 'security features [✂] and it is focusing on developing its offering in this area.[232]

(b) Google said that it recognises 'that customers can only unlock the full potential of cloud computing if they are able to deploy their workloads and data flexibly across environments.' Google said this means offering multi-cloud is a 'key pillar' of its global strategy.[233]

2.113 Google said that it generally has a pay-as-you-go pricing structure whereby products are priced based on usage.[234] Google said that its pricing strategy [✂] Google said it also [✂][235]

2.114 Google also said that its 'pricing, discounting and free credits allow customers to test [its] products for free before they commit to using [Google Cloud Platform's] services more extensively or migrate workloads away from on-premises or other cloud environments.'[236]

2.115 Google said its [✂].[237] Despite this, Google said that it [✂].[238]

2.116 When asked to list its understanding of the factors that UK customers value the most when choosing Google over other cloud providers, Google said that the factors the customers consider generally depend on their specific business needs. However, it also said that its customers 'particularly value' Google's:[239]

(a) 'Innovative product and services offering, optimised for multi-cloud'. Google said that it 'believes that one of the primary reasons' is its 'range of products

[230] For example, attracting developing and retaining high-performing employees. Google's response to Ofcom's information request [✂].
[231] Google's response to the CMA's information request [✂].
[232] Google said an example of its efforts here is Google Cloud Firewall Plus which it recently added to its 'suite of security solutions to give users best-in-class threat protection.'
[233] Google said this is why it 'has pioneered technologies such as BigQuery Omni and Anthos, both of which are agnostic of the underlying technology' (ie technology that can work across different public clouds and on-premise). Google also said that '[i]n the past three months alone, Google has introduced the next evolution of Kubernetes to help customers scale new workloads with containers, announced BigQuery Studio to provide customers with a unified interface to perform data tasks across different cloud environments, and launched Cross-Cloud Network which is an open networking platform that enables connectivity between cloud and on-premises environments.' Google's response to the CMA's information request [✂].
[234] Google's response to the CMA's information request [✂].
[235] Google response to Ofcom's information request [✂].
[236] Google response to Ofcom's information request [✂].
[237] Google response to Ofcom's information request [✂].
[238] Google response to Ofcom's information request [✂].
[239] Google's response to the CMA's information request [✂].

and services […] that can be deployed in a multi-cloud environment.' Google also specifically said its 'multi-cloud enabled data analytics products' (eg Big Query Omni) as an area 'valued by our UK customers.'

(b) Security and resilience. Google said that 'another key reason' it is chosen is its 'ability to offer high levels of security at scale.'

(c) Commercial proposition and being a strategic business partner. For example, Google said that it dedicates 'time and resources to understanding our customers' needs so that we can act as a strategic and trusted partner.' Google also said that it 'adopts a flexible and transparent approach to contracts, offering customers, irrespective of size, the opportunity to negotiate contract terms, including volume/committed spend discounts, to meet their individual business and/or legal requirements.'

Evidence from analyst reports and customers

2.117 Google is described as being smaller than AWS and Microsoft in the analyst reports we have seen, although these reports also recognise that it is expanding its capabilities across IaaS and PaaS and that in some areas it has been influencing the rest of the industry (eg Kubernetes[240]). While one of the analyst reports said that Google was making revenue gains, two of the analyst reports said that Google is sustaining large operating losses to try and win share.[241]

2.118 More generally based on our review of analyst reports:[242]

(a) Google is identified as having strengths in areas such as: its offerings in sovereign cloud, multi-/hybrid-cloud through Anthos and data insights and analytics and AI; its sales execution; its ability to attract cloud native applications; its ability to offer its services globally; its industry-focused solutions; and its storage and network capabilities.

(b) Google is identified as having weaknesses such as: its recent increases in price (despite its history of aggressive pricing relative to competitors); a lack of strategic clarity in some areas (eg around its multi-cloud offering); a relatively weak position with enterprise customers; some customers having trust concerns (eg around privacy); and a lack of differentiation compared to rivals.

2.119 We asked large customers to rate the suitability of a list of public cloud providers as alternatives to their main public cloud provider (see paragraph 2.83 above).

---

[240] Kubernetes, also known as K8s, is an open-source system for automating deployment, scaling, and management of containerized applications. It groups containers that make up an application into logical units for easy management and discovery.
[241] [✂].
[242] [✂].

Google was the main provider for a small number of customers we spoke to and among other customers some identified Google as an effective alternative, but most identified it as neither an effective or ineffective alternative (ie rating of 3) such that its average rounded to 3.

2.120 The most common reasons for customers not seeing Google as having an effective offering (ie a rating of 3 or less) were Google not having as advanced or as broad a range of functionality / features / services as the customers' main providers (AWS or Microsoft)[243] or having no direct experience of Google's offering or a lack of internal skills to use Google's cloud services.[244]

IBM

2.121 IBM first started providing cloud services around 2008, with the launch of a SaaS offering. It has subsequently expanded to provide a range of IaaS, PaaS and SaaS products.[245] IBM's website identifies it as having over 170 products[246] and that it is active in 10 regions.[247]

2.122 IBM also has an offering, IBM Cloud Satellite, through which it provides certain cloud services on other public clouds and on-premises.[248] IBM is also active in hardware for on-premises solutions.[249]

Public cloud strategy and views of customers

Evidence from IBM

2.123 IBM said that its 'strategic areas' include:

(a) Hybrid multi cloud. IBM said that it offers 'a range of hybrid multi-cloud solutions'[250] which are available across IBM's public cloud, a customers' private cloud or data centres of other public cloud providers.[251]

(b) Regulated industries. IBM said that it is 'open to work with any industry but has a stronger presence in certain industries'.[252] IBM said this is because 'IBM can build foundations with inherent regulation, security, and controls that lower risk and accelerate time to value.'[253]

---

[243] Eg [✂].
[244] Eg [✂].
[245] What is IBM Cloud? Services Offered, Features & Pricing (datamation.com), accessed 11 February 2024.
[246] Cloud Products | IBM, accessed 11 February 2024.
[247] Locations for resource deployment | IBM Cloud Docs, accessed 11 February 2024.
[248] IBM Cloud Satellite, accessed 11 February 2024.
[249] Enterprise Business Server Solutions | IBM, accessed 11 February 2024.
[250] IBM response to the Issues Statement, paragraph 1.2.
[251] IBM response to Ofcom's information request [✂].
[252] IBM said that it '[✂].' response to Ofcom's information request [✂].
[253] IBM response to Ofcom's information request [✂].

(c) Digital transformation. IBM said that [✂] IBM has a proven track record of proactively solving multi-dimensional technical challenges including those affecting legacy or proprietary technologies that carry an extra degree of risk when they are being adapted in any way, shape, or form.' [254]

2.124 [✂].[255]

Evidence from analyst reports and customers

2.125 IBM is identified as being a smaller provider when compared to both the largest, AWS and Microsoft, but also Google. In particular, one of the analyst reports said that 'IBM has shifted its emphasis away from trying to keep pace with the hyperscalers toward a more systematic focus on key sectors […]. This approach recognises the reality that IBM can't compete with the leaders on every front'.[256]

2.126 More generally based on our review of analyst reports:[257]

(a) IBM is identified as having strengths in areas such as: its focused strategy on regulated industries; its capabilities relative to other providers around certain workloads such as mainframe workloads focused on testing and development; its container management software and its multi-cloud/hybrid cloud offering;[258] its security offerings and data integration components; and its price/performance offering.

(b) IBM is identified as having weaknesses such as: its historic reliability issues impacting on customer confidence; its sovereign cloud offering which is less comprehensive than some other providers; its relatively low market profile and the lack of a strong third party ecosystem; and its limited number of data centres internationally with some variation of available services;

2.127 We asked large customers to rate the suitability of a list of public cloud providers as alternatives to their main public cloud provider (see paragraph 2.83 above). IBM was not the main provider for any of the customers we spoke to.

2.128 Amongst other customers none identified IBM as an effective alternative (ie rating of 4 or 5) with most identifying it as ineffective or very ineffective alternative (ie 2 or 1) with its average rounding to 2.

2.129 The two main reasons given for these ratings for IBM (or for not giving a rating) were that IBM has more limited services or capabilities compared to larger

[254] IBM response to Ofcom's information request [✂].
[255] IBM response to Ofcom's information request [✂].
[256] [✂].
[257] [✂].
[258] One analyst described this as its ability to accommodate vendor-neutral cloud strategies. [✂].

providers[259] and a lack of any experience or knowledge of IBM's offering on the part of the customer.[260]

Oracle

2.130 Oracle entered cloud services with Oracle Cloud Infrastructure (OCI) in 2016 and has expanded to provide a wide range of services across IaaS, PaaS and SaaS.[261] Oracle's website identifies OCI as having over 100 services and that it is active in 48 commercial and government regions.[262]

2.131 In 2022, Oracle's top product category was databases which accounted for around [✂] of its revenue in the UK. The next largest category was compute which accounted for [✂].[263]

2.132 Oracle has historically been the 'world's largest' database management company[264] and is also active in on-premises solutions (both hardware and software).[265]

Public cloud strategy and views of customers

Evidence from Oracle

2.133 Oracle said that it has 'facilitated a multi-cloud strategy to help customers take advantage of each [providers'] architectural innovations, even when that innovation creates fundamental differences in engineering.' Oracle said that in part this is facilitated via a 'strategic partnership' with Microsoft.[266]

2.134 Oracle also said that it is differentiated from other providers because it optimises for speed and performance[267] and because it is designed to provide better security.[268]

2.135 Oracle said that it typically charges 'a prepaid fee that is decremented as the Oracle Cloud Infrastructure (OCI) services are consumed by the customer over a stated time period.'[269] Oracle, like others, offers a free trial and free credits to attract new cloud customers, this is done via 'Oracle Cloud Free Tier'.[270]

---

[259] Eg [✂].
[260] Eg [✂].
[261] Oracle Cloud Infrastructure Platform Overview, accessed 11 February 2024.
[262] Public Cloud Regions and Data Centers | Oracle United Kingdom, accessed 11 February 2024.
[263] Oracle's response to the CMA's information request [✂]. [✂].
[264] About Oracle | Company Information | Oracle United Kingdom, accessed 11 February 2024.
[265] Hardware | Oracle United Kingdom, accessed 11 February 2024.
[266] Oracle response to the Issues Statement, date 23 November 2023 and Oracle's distinct approach on hybrid and multicloud, accessed 11 February 2024.
[267] Oracle response to the Issues Statement, date 23 November 2023.
[268] Oracle response to Ofcom's information request [✂].
[269] Oracle response to the CMA's information request [✂].
[270] Cloud Free Tier | Oracle United Kingdom, accessed 11 February 2024.

2.136 In relation to its customer base, Oracle said that it has customers of 'many sizes across a range of industries, government agencies, educational institutions and resellers'. Oracle also said that it had recently purchased a leading provider of digital information systems used within hospitals and health systems and Oracle's new healthcare business unit plans to accelerate that work of that provider.[271]

2.137 When asked to list its understanding of the factors that UK customers value the most when choosing Oracle rather than other cloud providers, Oracle provided the following reasons:[272]

(a) Ease of migration to Oracle from on premises infrastructures when compared to other cloud providers.

(b) Oracle's lack of fees charged when moving data in or out of Oracle Cloud Infrastructure.

(c) Oracle's innovations such as 'off box virtualisation' for complete instance isolation and RDMA networking for compute intensive workloads.

(d) Oracle's security which is built in, on by default and available for no extra charge.

(e) Oracle's superior price / performance when compared to larger providers.

(f) Oracle's support for Hybrid Cloud deployments[273] and multi-cloud deployments.[274]

Evidence from analyst reports and customers

2.138 Analysts say that Oracle is a smaller provider when compared to both the largest, AWS and Microsoft, but also Google.[275] In particular, one of the analyst reports said that the Oracle's approach 'implicitly recognizes [sic] that hyperscaler status is out of reach for [Oracle].' [276]

2.139 More generally based on our review of analyst reports:[277]

(a) Oracle is identified as having strengths in areas such as: its position in databases; its innovation in areas of emerging enterprise needs including

[271] Oracle response to Ofcom's information request [✂].
[272] Oracle response to the CMA's information request [✂].
[273] Oracle said that with 'Oracle the customer has a wide range of deployment options. Customers can run an entire OCI region from within their own datacentre, move on premises VMWare environments to the public cloud regions, or deliver OCI services like Exadata and roving edge right to the site of the work needed.'
[274] Oracle said its 'low latency interconnects like the one with Microsoft Azure which facilitate a multi cloud architectural approach' was an example of this.
[275] [✂].
[276] [✂].
[277] [✂].

sovereign cloud; its improving capabilities that bring it closer to the larger providers; its market momentum (ie increases in revenue) and its enticements around pricing and fees such as first 10TB of data egress being free.

(b) Oracle is identified as having weaknesses such as: gaps in its offering; a perception that it is database focused and thus not seen as an IaaS provider by many users; a negative perception among some customers due to its past tough enforcement (eg of its IP rights) and inconsistent sales and support; not being positioned for adoption by all customers, but aimed at large enterprise customers with established IT expertise; and an immature ecosystem.

2.140 We asked large customers to rate the suitability of a list of public cloud providers as alternatives to their main public cloud provider (see paragraph 2.82 above). Oracle was not the main provider for any of the customers we spoke to.

2.141 Amongst other customers, none identified Oracle as an effective alternative (ie rating of 4 or 5) with most identifying it as ineffective or very ineffective alternative (ie 2 or 1) with its average rounding to 2.

2.142 The main reasons customers gave for these ratings (or for not giving a rating) were that Oracle is, or is perceived to be, behind the largest providers in relation to factors such as breadth of service, capabilities, etc.[278] Some customers also cited a lack of experience with Oracle.[279]

Other providers

2.143 There are also a range of smaller providers offering IaaS and PaaS products, such as OVHcloud and Scaleway.

(a) AWS provided a list of providers that have begun offering cloud services and said that both the number and size of these competitors are rapidly increasing.[280] For example, AWS said that there has been recent entry by providers (who may be considered SaaS providers)[281] entering to offer what some would label as IaaS/PaaS services and making significant inroads.[282]

(b) Similarly, Microsoft said that in recent years many providers of cloud services have emerged by focusing on certain customer groups or functionalities,

[278] Eg [✂].
[279] Eg [✂].
[280] AWS response to the Issues Statement, paragraphs 9 to 10.
[281] AWS said that it does not agree with the terms IaaS and PaaS as it does not believe that these distinctions reflect the actual competitive dynamics for cloud and other IT services.
[282] AWS response to the CMA's information request [✂].

include some focused on providing access to basic infrastructure at a low cost.[283]

2.144 While there are smaller providers active and there is evidence of entry, we note that:

(a) Of the analyst reports (which were global rather than UK focused) we reviewed:

(i) One did not identify any of these smaller players.[284]

(ii) One identified OVHcloud as a 'major player' and in the same grouping as Oracle and IBM (AWS, Microsoft and Google were identified as 'leaders'). More generally it identified DigitalOcean, Akami, Vultr and Zadara as 'contenders'. The report indicates that these players have a much smaller presence than any of the main providers outlined above.[285]

(iii) One identified OVHcloud, Salesforce and SAP as 'contenders and in the same grouping as Oracle and IBM (AWS and Microsoft were identified as 'leaders' and Google as a 'strong performer'). An additional provider Rackspace Technology was identified as a 'challenger'.[286]

(b) When asked about alternatives to their main provider (see paragraph 2.83 above) none of the large customers we spoke to identified OVHCloud as a suitable alternative[287] or identified any of the other providers set in the analyst reports. Only one 'other' public cloud provider ([✂]) was identified by a customer and was not rated as an effective alternative.[288]

2.145 There are other cloud providers who may have a large presence globally such as Alibaba, Huawei and Tencent and all of these have been identified as a competitor by at least one cloud provider.[289] However, these providers have not been identified as suitable alternatives to the main providers by UK customers we have spoken to.[290]

---

[283] Examples provided by Microsoft included CoreWeave and Paperspace who have offerings focused on AI workloads and Scaleway which is focused on providing cloud services to small businesses by offering very cheap cloud instances. Microsoft response to Ofcom's Interim Report, Annex 1, paragraphs 27 to 32.
[284] Although it did identify Alibaba, Huawei and Tencent, we discuss these providers at paragraph 2.147. [✂].
[285] [✂].
[286] [✂].
[287] Amongst customers OVHcloud's average rating rounded to 1. Reasons for OVHcloud being identified as an ineffective alternative included a lack of experience or knowledge of OVHCloud's offering [✂] and that OVHCloud has more limited services or capabilities compared to larger providers [✂].
[288] [✂] response to the CMA's information request [✂].
[289] Responses to the CMA's information requests [✂]; [✂] response to Ofcom's information request [✂].
[290] For example, Alibaba received an average rating that rounded to 1 and was not identified by any customer as an effective alternative. Reasons for these ratings or not providing a rating included a lack of experience or knowledge of Alibaba's offering [✂], that using Alibaba would be a supply chain risk or not appropriate (eg for data sovereignty or

Emerging views on different cloud providers

2.146 Based on the evidence we have seen to date, our emerging view is that AWS and Microsoft have the widest commercial offerings with many large customers considering them suitable alternatives to each other (although we note that just because they may be suitable alternatives does not mean customers can easily switch between them – we consider some factors that may influence customers' ability to switch above and in later working papers).

2.147 Google is seen as a suitable alternative by some large customers and it has expanded its offering over time, even if it is considered by some to be more limited than the offerings of AWS and Microsoft.

2.148 Large customers do not see Oracle, IBM or other providers as suitable alternatives to their main providers (generally AWS and Microsoft) with customers citing a weaker offering or a lack of experience with these customers.

2.149 While smaller providers are not seen as an effective alternative to large customers' main providers, we note that they may be seen as suitable alternatives for certain workloads with evidence showing that the smaller providers all have strong offerings in relation to certain market segments or types of customer.

Independent Software Vendors (ISVs)

2.150 ISVs are suppliers of cloud services, typically PaaS and/or SaaS, that do not usually own the underlying infrastructure.[291]

2.151 There are many ISVs present in the UK providing PaaS. Our share of supply analysis suggests that ISVs accounted for up to [30-40]% of UK PaaS revenue in 2022. They compete in specific product categories rather than across the entire range of PaaS products. Examples of ISVs include:

(a) VMWare, now part of Broadcom, which was founded in 1998[292] and specialises in providing virtualisation technology which allows users to run multiple operating systems, as virtual machines, on a single physical machine.[293] Its services are currently available on a range of cloud providers.[294]

security reasons [✂], that Alibaba has more limited services or capabilities compared to larger providers [✂] and that they are only considered for demand relating to China [✂].
[291] There are some exceptions to this – for example, Salesforce is an ISV, but operates its own infrastructure.
[292] VMware, Inc. - Company Profile - California Explore, accessed 20 May 2024.
[293] What is virtualization technology & virtual machine? | VMware, accessed 20 May 2024.
[294] Provider Search (vmware.com) and Hosted VMware - VMware Cloud on AWS - AWS (amazon.com), accessed 20 May 2024.

(b) MongoDB was founded in 2007. It specialises in database management and document databases.[295] MongoDB provides a service Atlas which gives developers the ability to run their databases across several cloud providers and provides them with access to a range of features and tools, enabling users to access, query and analyse data.[296] Given this, its services are available across a range of cloud providers.[297]

(c) Snowflake was founded in 2012.[298] It specialises in providing data warehouses (SaaS offering) which provides users with the ability to store and access structured and unstructured data. Snowflake also offers a 'cloud data platform' which is capable of supporting multiple data workloads from data warehousing to data engineering, across several cloud providers.[299] Given this, its services are available across a range of cloud providers.[300]

(d) Yugabyte was founded in 2016. It specialises in providing database technology. The YugabyteDB offering provides customers with open-source distributed databases (typically used to store data across multiple sites), and access to enterprise database features.[301] Its services are currently available on a range of cloud providers.[302]

2.152 As ISVs do not usually own the underlying infrastructure, they tend to rely on cloud providers as an input to their own cloud services. That is, they may use cloud providers' IaaS as an input to develop and offer their own PaaS and/or SaaS.

2.153 As such, within our analysis we have often considered ISVs as akin to customers of cloud providers' IaaS offerings. For example, in our market definition assessment, when considering IaaS and demand-side substitutability, we have considered ISVs alongside other customers. Similarly, when gathering evidence in relation to the four theories of harm outlined in our issues statement[303] we have, where relevant, sought evidence from ISVs as customers as well as non-ISV customers.

2.154 In some contexts, we have also considered the extent to which ISVs may act as competitors to the main cloud providers in relation to their PaaS offerings. For example, MongoDB offers database and management and document databases

---

[295] About Us - Our Story | MongoDB | MongoDB, accessed 20 May 2024.
[296] MongoDB Atlas | Multi-cloud Developer Data Platform | MongoDB and Advantages Of MongoDB | MongoDB, accessed 20 May 2024.
[297] MongoDB Atlas on AWS—Partner Solution (amazon.com), MongoDB Atlas on Azure | Microsoft Azure and MongoDB Atlas | Google Cloud, accessed 20 May 2024.
[298] The Snowflake Story - TechStory, accessed 20 May 2024.
[299] The Data Cloud Explained (snowflake.com), accessed 20 May 2024.
[300] Snowflake Cloud Partners | Snowflake, accessed 20 May 2024.
[301] About Yugabyte and YugabyteDB—the Distributed SQL Database, accessed 20 May 2024.
[302] YugabyteDB on Amazon Web Services (AWS), YugabyteDB on Google Cloud and YugabyteDB on Microsoft Azure, accessed 20 May 2024.
[303] Issues statement (publishing.service.gov.uk)

and, although their capabilities may differ,[304] so do AWS (DocumentDB)[305] and Microsoft (CosmosDB).[306] Therefore, in any assessment of market power in relation to PaaS services it is important to consider ISVs as well as the cloud providers outlined in the previous section.

2.155 In other situations, an ISV's product may be a complement to a cloud provider's wider offering. This is the case if an ISV offers a PaaS or SaaS service where a cloud provider does not have an equivalent offering, thereby allowing customers to use the ISV's product alongside the services offered by a cloud provider. More generally, the number of ISV PaaS offerings available within a cloud provider's ecosystem may influence customers' choice of cloud provider, and we consider this in section 7 below.

2.156 Finally, ISVs and cloud providers may also have a relationship as ISVs may rely on cloud providers as a distributor of their services. This could be through the cloud providers directly selling an ISV's services, offering ISVs a platform through which to sell their services (such as a marketplace) or access to customers.

[304] For example, MongoDB vs. DocumentDB: A Comprehensive NoSQL Database Comparison (sprinkledata.com), accessed 20 May 2024.
[305] Json Document Database - Amazon DocumentDB - AWS, accessed 20 May 2024.
[306] Azure Cosmos DB for NoSQL | Microsoft Learn.

3. Switching and multi-cloud

Switching clouds and use of multi-cloud by customers

3.1 In this section, we set out evidence we have gathered to date on the switching process and types of multi-cloud. We also set out evidence on the prevalence of the use of multiple clouds and switching.

3.2 At its most basic, switching between cloud providers involves a customer moving one or more workloads, or parts of workloads, from one provider's cloud to another provider's cloud.

3.3 Similarly, at its most basic, a multi-cloud approach involves the placement by a customer of at least one workload on one provider's cloud, and at least one workload on an alternative provider's cloud.

3.4 There is also some interaction between the two – for example, if it is possible to use multiple clouds then it is possible to switch individual workloads (or parts or workloads), but if it is not possible to multi-cloud then it is only possible to switch all workloads.

3.5 Both of these concepts are important to understand as they can significantly influence the competitive dynamics in the supply of Public Cloud Infrastructure Services.

3.6 Below we set out the implications of switching and using multiple clouds at the extremes (ie where either you cannot switch/multi-cloud at all or where you can switch any workload and use any form of multi-cloud). In doing this we have abstracted away from the detail of these concepts, such as how different types of multi-cloud may be possible.

Figure 3.1: Switching, multi-cloud and type of competition

	No multi-cloud	Full multi-cloud
No switching	**Competition just for customers new to the cloud**: Existing customers cannot move existing workloads or place new workload with a new provider	**Competition for all new workloads**: Existing customers cannot move existing workloads, but everyone can place new workloads with any provider.
Full switching	**Competition for customers**: Existing customers cannot split their workloads between providers, but can move all their workloads.	**Competition for all workloads**: Existing customers can switch any existing workloads and everyone can place new workloads with any provider.

Source: CMA

3.7 Understanding the implications of these extremes and actual behaviour in the markets informs our analysis of the scope and focus of competition.

(a) For example, in a world where customers can easily switch and multi-cloud then it is more likely that there will be competition between providers to attract and retain customers, as well as for specific workloads. In this scenario, competition would, all else being equal, be less likely to decline over time and smaller cloud providers may be able to more easily enter and expand by, for example, focusing their offer on certain niche products.

(b) In contrast, if both switching and the use of multiple clouds are prevented or restricted by features of the market(s) then this will result in competition among providers being centred on capturing and retaining new customers. This may also suggest that, at least to some extent, competition is for the market, which is likely to decline over time as the number of customers making the initial transition to public cloud decreases, rather than competition in the market.

(c) If customers can easily multi-cloud, but switching is prevented or restricted by features of the market(s) then this will result in competition between providers for new workloads, even after a customer has chosen their initial cloud environment. In this scenario, smaller providers may be able to adopt a strategy of targeting specific types of workload by specialising and as such make it easier for smaller providers to enter on the basis of a specific niche segment of the market. This may also suggest that competition is likely to

decline over time in the event that the number and size of new workloads were to decline over time.[307]

(d) If customers can easily switch, but use of multiple clouds is prevented or restricted by features of the market(s) then this will result in competition for customers rather than workloads. In this scenario it may be more difficult for smaller providers with a smaller product offering to compete especially for larger customers who will have diverse needs.

3.8 Evidence we have seen to date suggests that customers have some ability to switch and some ability to multi-cloud but that these abilities differ amongst customers. This may mean that for one set of customers there is only competition when they first choose a public cloud provider, for a second set of customers there is also competition for all new workloads, for a third set of customers there is ongoing competition at the customer level and for a fourth set of customers there is competition for all workloads.

3.9 This means that the overall nature of competition in the markets will be influenced by the mixture of customers – for example, if it is not feasible for the vast majority of customers to switch or multi-cloud then the focus of competition will be on customers new to cloud even if on the margins there is scope to compete for existing customers/workloads.

3.10 In the rest of this sub-section we set out evidence on the switching process and types of multi-cloud. We also set out evidence on the prevalence of the use of multiple clouds and switching.

Switching process

3.11 We asked cloud providers to explain the steps that customers need to go through in order to switch cloud provider. In this section we set out what providers have told us about the switching process, however we note that this may not reflect all of the nuances a customer may face. For example, one customer we spoke to that had switched[308] said that they had to operate two different public clouds and migrate systems between them, including periods of dual running.[309]

3.12 We welcome views from other market participants on the process of switching.

[307] We note that at present evidence set out above suggests that cloud services is still growing. While some of this may be due to the expansion of existing workloads, this will also include workloads that are new to the public cloud. While this may be the case, competition may still decline even if cloud services grows due to these new workloads if they represent a smaller proportion of total demand over time.
[308] [✂].
[309] Note of meeting with [✂].

3.13 Providers identified a similar set of steps that customers must take to switch between public clouds.[310]

(a) Assess existing environment. Most providers submitted that customers need to fully understand their current setup including interdependencies and contractual restrictions to identify what can be switched, the complexity of switching and how to order the switching process (eg which workloads to switch first).[311] We note that to the extent workloads are moved in stages this may lead to a customer needing to integrate across the cloud provider it is moving from and the cloud provider it is moving to.

(b) Develop a migration strategy. Two providers said that customers should develop a detailed migration plan taking into account their assessment of its existing environment and setting out risks, and timescales.[312] This includes identifying the specific cloud products and configurations on the target public cloud which are necessary for migrating customers' workloads from the origin public cloud.[313]

(c) Test integrity of current IT environment. Two providers said that, before workloads can be switched from one public cloud to another, customers may test the integrity of their existing IT architectures to ensure the switch can be successfully executed.[314]

(d) Deploy workloads in the target public cloud. Cloud providers said that, after completing the previous steps, customers can use their migration tools to move the chosen workloads to their target provider's cloud.[315] Most providers we contacted offer tools and technical assistance to accelerate and reduce the costs of switching to their clouds.[316] We note that some of the migration tools identified appear to be targeted more at migration from on-premises solutions than for moving between cloud providers.[317] We have also not sought to assess the effectiveness of any of these solutions as part of this working paper.

---

[310] Responses to the CMA's information requests [✂].
[311] Responses to the CMA's information requests [✂].
[312]Responses to the CMA's information requests [✂].
[313] [✂] response to the CMA's information request [✂].
[314] Responses to the CMA's information requests [✂].
[315] Responses to the CMA's information requests [✂].
[316] Responses to the CMA's information requests [✂].
[317] For example, Azure Migrate—Cloud Migration Services | Microsoft Azure and Oracle Cloud Migrations appear to targeted at migration from on-premises solutions rather than other cloud providers. Similarly some of AWS services appear to be more targeted at on migrating from on-premises solutions rather than between cloud providers (Free Cloud Migration Services - Free Data Transfer - AWS (amazon.com)), but others do appear to also be targeted at switching from other cloud providers (Cloud Application Migration Tool - AWS Application Migration Service - AWS (amazon.com)), all accessed 20 May 2024.

(e) Validate target cloud environment. One provider said that customers should validate their new cloud environment post-migration to ensure it is ready for use.[318]

3.14 Two providers said the process of switching remains the same regardless of whether a customer is switching to or from another cloud provider, to or from another deployment model (ie, private cloud), or to or from traditional IT.[319]

3.15 One provider said that when switching to or from a provider's public cloud, customers can migrate the necessary data themselves, with the assistance of a provider or using a professional IT services provider that partners with cloud providers.[320]

Types of multi-cloud

3.16 At its most basic, multi-clouding is where a customer places at least one workload on one provider's cloud, and at least one workload on an alternative provider's cloud (ie where a customer uses two or more public clouds simultaneously). A handful of customers we spoke to, as well as one provider, said that internet-facing APIs and third-party intermediaries can be used to interconnect different clouds, facilitating the use of multiple clouds.[321]

3.17 We asked cloud providers to explain the types of multi-cloud that customers can adopt. In this section we set out what providers have told us about different forms of multi-cloud, which may not reflect all the nuances a customer may face when using multi-cloud which are discussed in the Technical Barriers working paper.

3.18 Providers identified three models of multi-cloud that customers may choose to adopt, as well as a set of benefits and disadvantages specific to each of them.[322]

(a) Cloud duplication, where customers mirror the entirety or part of their IT architecture on several clouds, enabling their workloads and applications to run equivalently on all of them.[323] One provider said that this form of multi-clouding is relatively uncommon, and is preferred by customers who require resilience over specific components of their cloud architecture.[324]

(b) Integrated multi-cloud, where customers can mix and match Cloud Services from different public Cloud Providers and there is a degree of integration between these services (ie, data is stored on one public cloud but analysed

[318] [✂] response to the CMA's information request [✂].
[319] Responses to the CMA's information requests [✂].
[320] [✂] response to the CMA's information request [✂].
[321] Responses to the CMA's information requests [✂].
[322] Responses to the CMA's information requests [✂].
[323] Responses to the CMA's information requests [✂].
[324] [✂] response to the CMA's information request [✂].

on a different one). One provider said that there is 'real customer appetite for integrated multi-cloud strategies', but certain practices 'restrict or hinder many customers from adopting a truly integrated multi-cloud approach'.

(c) Siloed multi-cloud, where customers run different workstreams on different public clouds with no or minimal integration between these cloud environments. One provider said that this type of multi-clouding is the most prevalent amongst customers that have a multi-cloud architecture, as many of them are still at a relatively early stage of their cloud journey.[325]

3.19 Several providers said that certain customer groups are more likely to adopt a multi-cloud strategy than others; namely:[326]

(a) Two providers said that digital native customers[327] are better positioned to adopt multi-cloud strategies, with one of them clarifying that this is because they do not depend on legacy on-premises services.[328] One of these providers said that traditional enterprises may find the adoption of integrated multi-cloud more difficult, partly because of the various technical and cost challenges of migrating away from an on-premises solution, and partly because of certain practices of their legacy providers.[329]

(b) Three providers said that larger enterprises are more likely to multi-cloud vis-à-vis smaller customers, as they have the skills and scale to do so and are more likely to consider over-reliance on one cloud provider to be a risk.[330] One provider added that large companies tend to have distinct departments with their own independent workloads, facilitating the adoption of a siloed multi-cloud approach.[331]

(c) Two providers said that enterprises with complex regulatory compliance obligations and/or a requirement for enhanced resilience and stability (eg financial institutions) often opt for multi-cloud strategies.[332]

(d) One provider said that merged entities may – temporarily – adopt a multi-cloud approach to manage each company's distinct cloud strategy. However,

[325] [✂] response to the CMA's information request [✂].
[326] Responses to the CMA's information requests [✂].
[327] Digital natives are businesses that exist primarily or entirely online and use cloud technologies throughout their operation.
[328] Responses to the CMA's information requests [✂].
[329] This provider identified licensing practices as a specific concern and these are discussed in a separate Working Paper. [✂] response to the CMA's information request [✂].
[330] Responses to the CMA's information requests [✂].
[331] [✂] response to the CMA's information request [✂].
[332] Responses to the CMA's information requests [✂].

it clarified that these enterprises typically choose one primary provider after the merger.[333]

Types of multi-cloud architectures

3.20 As the evidence set out above shows, there are different types of multi-clouding and customers can use multiple public clouds in many ways, with varying services, architectures and levels of interdependence involved.

3.21 Consistent with the input from providers, we have identified some broad categories of integration lying on a spectrum and we will set these out in more detail in later working papers. For this working paper we simply refer to:

(a) Siloed multi-cloud where there are no interdependencies between clouds being used: the customer's clouds are completely independent and there is no communication between them.

(b) Integrated multi-cloud where there are some interdependencies between clouds being used. We recognise that there are many different forms of integration which we will consider in further detail in a later working paper.

3.22 Customers can also use 'duplicated multi-cloud'. This is where customers aim to mirror their cloud architecture on two or more public clouds, so that all or some of their applications and data can run equivalently on all of them.

Customer switching and use of multi-cloud

3.23 In the rest of this section, we consider the potential benefits and disadvantages of multi-cloud and factors relevant to moving workloads which are relevant to both switching and the use of multiple clouds.

3.24 This is because customers considering whether to switch or multi-cloud face a trade-off between the expected benefits and expected costs of doing so.

(a) The expected benefits of switching or using multiple clouds may include lower spend on cloud services, higher flexibility, and the ability to access better-quality and better-fitted services or new innovations.

(b) Expected costs may include the time and cost of moving workloads, the cost of reconfiguring the cloud architecture, the cost of retraining staff and the increased management complexity for customers using multiple clouds.

[333] [✂] response to the CMA's information request [✂].

3.25 Where the benefits of switching (or using multiple clouds) exceed the costs, we would expect customers to switch (or multi-cloud). Where the costs exceed the benefits, customers may choose not to use an alternative provider, even though the offer may be better for that customer. The greater the switching cost, the more likely it is that customers will refrain from switching or using multiple clouds despite the availability of products that would otherwise represent a better overall offer in terms of price, quality, range, service or innovation.

Multi-cloud benefits and disadvantages

3.26 We asked cloud providers to explain the benefits and disadvantages of multi-cloud. We have also asked customers about their use of multi-cloud and, in responding to such questions, some customers have provided evidence on disadvantages of operating a more integrated multi-cloud. We set out this evidence below.

Evidence from providers

3.27 Providers identified certain general benefits and disadvantages associated with multi-clouding:[334]

(a) Benefits: one provider said that multi-clouding allows customers to leverage the threat of moving individual workloads to competing providers to increase their bargaining power with their primary Cloud Provider.[335] One provider also said that multi-clouding minimises the risk of downtime, allows customers to comply with specific legal or regulatory requirements, and can help customers achieve global coverage.[336]

(b) Disadvantages: one provider said that, because of the greater complexity and increased staffing requirements, customers may incur higher costs when using multiple clouds.[337] Finally, two providers said that multi-clouding can increase management complexity.[338]

3.28 Providers also identified benefits and disadvantages in relation to each type of multi-cloud they identified.

(a) Cloud duplication:

[334] Responses to the CMA's information requests [✂].
[335] [✂] response to the CMA's information request [✂].
[336] [✂] response to the CMA's information request [✂].
[337] [✂] response to the CMA's information request [✂].
[338] Responses to the CMA's information requests [✂].

(i) Benefits: one provider said that cloud duplication gives the highest degree of resiliency and downtime protection.[339]

(ii) Disadvantages: one provider said this approach creates challenges around data latency and data governance.[340] Two providers said that having to duplicate services across several clouds increases the costs faced by customers.[341]

(b) Integrated multi-cloud:

(i) Benefits: some providers said that this approach gives customers the flexibility to use their preferred services from different public cloud providers.[342] One provider said that this multi-cloud strategy could also minimise customers' costs, as they are able to select the most cost-effective solution for each individual workload.[343]

(ii) Disadvantages: Two providers said that integrated multi-cloud also creates multiple points of failure across different clouds, increasing the risks customers are exposed to, including security ones.[344] One of those providers also said that integrated multi-cloud 'increases the complexity of building, maintaining and securing applications'.[345] Finally some providers said that integrating between different cloud environments is often a complex process, with one saying that integration between public clouds may in some cases be infeasible.[346]

(c) Siloed multi-cloud:

(i) Benefits: one provider said that the concerns around data latency and security are less prevalent for siloed multi-cloud as compared to integrated multi-cloud. That same provider added that siloed multi-clouding improves customers' bargaining position across all workloads because cloud services providers need to offer competitive prices to attract new workloads from existing customers.[347]

(ii) Disadvantages: one provider said that siloed multi-cloud may be costly and inefficient, as each workload requires interaction with a different cloud provider. The provider also said that this type of multi-cloud makes it challenging for workloads in different clouds to interact with

[339] [✂] response to the CMA's information request [✂].
[340] [✂] response to the CMA's information request [✂].
[341] Responses to the CMA's information requests [✂].
[342] Responses to the CMA's information requests [✂].
[343] [✂] response to the CMA's information request [✂].
[344] [✂] response to the CMA's information request [✂].
[345] [✂] response to the CMA's information request [✂].
[346] Responses to the CMA's information requests [✂].
[347] Responses to the CMA's information requests [✂].

each other if needed compared to if they were all run by the same cloud services provider.[348]

Evidence from customers

3.29 We have received more limited evidence from customers on the use of multiple clouds.

3.30 Customers (as well as providers) have told us that multi-clouding is used, especially by larger customers.[349] Therefore, there is clearly demand from customers for some form of multi-cloud. However, our evidence on the types of multi-cloud that customers use is more limited and our analysis suggests that often customers have one primary cloud provider accounting for the vast majority of expenditure.

3.31 In relation to highly integrated multi-cloud, in a handful of cases customers said that they currently have, at most, a limited use case.[350] Reasons for this included that there are challenges with adopting this approach.[351] However, other responses from these customers suggest that they do have some form of integration or are considering some form of additional integration.

3.32 The responses from a few customers suggest that they may be more open to adopting an integrated multi-cloud approach in the future. One customer said that while it currently has 'no business reason to have full integration', it is feasible that they may want to integrate in the future.[352] Another customer noted that, while its legacy applications migrated from on-premises are not suitable for integrated multi-cloud, the same is not necessarily true for its new cloud-native workloads.[353]

3.33 Customers had somewhat differing views on the disadvantages of multi-cloud. For example, one customer said that when utilising a multi-cloud strategy, consideration must be applied to risks associated with a business process traversing multiple cloud providers, to prevent the impact of a single provider failure from introducing increased operational resilience risk to the business process.[354] However, another customer said that it ran two related applications on two different clouds, and this did not negatively affect the resiliency of either application.[355]

---

[348] [✂] response to the CMA's information request [✂].
[349] [✂].
[350] Responses to the CMA's information requests [✂]; Note of call with [✂]
[351] [✂] response to the CMA's information request [✂]; Note of call with [✂]
[352] [✂] response to the CMA's information request [✂].
[353] Note of call with [✂].
[354] Note of meeting with [✂].
[355] [✂] response to the CMA's information request [✂].

Moving workloads

3.34 Both switching and the use of multiple clouds can involve moving workloads. In this sub-section we set out some evidence from providers and customers about factors that may deter customers moving workloads between cloud providers.

3.35 Providers said that there are certain financial and time costs associated with moving workloads between providers. These providers identified several factors influencing the timeframe and resources needed for a given migration process, such as:

(a) Customer's existing cloud set-up: one provider said that the contractual and technical restrictions associated with a customer's existing cloud set-up will affect the complexity, cost and timeline of the switching process.[356] Similarly, two providers said that dependencies on third parties can make switching more difficult. One of them said that this is the case as 'each dependency needs to be compatible with the new cloud environment, correctly configured, and installed for the switch to be successful'.[357] We consider technical barriers in detail in a separate working paper.

(b) Licensing restrictions: one provider submitted that certain artificial licensing barriers created by legacy software providers make switching workloads more challenging.[358] We consider licensing practices in a separate working paper.

(c) Complexity and size of the workloads being migrated: several providers said the greater the size and complexity of the workloads and applications being transferred, the more time and resources the migration or switching process will demand.[359] We consider technical barriers in detail in a separate working paper.

(d) Customers' priorities and chosen deployment method: one provider said that customers' choice of a deployment method (ie, manual versus automated) will involve a different degree of complexity and automation, affecting the ease of migration between clouds.[360] Similarly, another provider said that customers' priorities, as well as the skillset and level of their engaged resources, influence switching costs.[361]

(e) Data egress fees: two providers mentioned egress fees, with one provider saying that egress fees alongside other networking costs may hinder

[356] [✂] response to the CMA's information request [✂].
[357] Responses to the CMA's information requests [✂].
[358] [✂] response to the CMA's information request [✂].
[359] Responses to the CMA's information requests [✂].
[360] [✂] response to the CMA's information request [✂].
[361] [✂] response to the CMA's information request [✂].

integration and the other saying that these charges on outbound data transfers are applied when a customer is switching away from their cloud.[362] We consider such charges (egress fees) in a separate working paper.

3.36 Evidence we have seen from customers we spoke to indicates that two other factors may influence the extent to which customers move workloads between different cloud providers.

3.37 First, customers may not move workloads if they consider that there is a lack of differentiation between at least some providers[363] and in some cases customers have said that the effort involved outweighs the benefits.[364] For example:

(a) One customer said that the differences in capabilities between the three largest cloud providers is smaller now than it was five years ago.[365]

(b) Another customer said that cloud providers offer broadly the same capabilities at similar prices such that there is limited incentive to switch.[366]

(c) Another customer said that although switching is feasible, the necessary investment in terms of cost and time would resemble that of the initial migration to cloud.[367] The same customer said that they would need to have a good reason to consider switching workloads between providers, and that currently they do not have one.[368]

3.38 While customers may currently view providers' capabilities as broadly similar, there is evidence suggesting that AI's growing importance may change this. Namely, two customers said that AI capabilities are becoming an increasingly important source of differentiation between providers.[369]

3.39 Second, customers may not move workloads if they consider a switch needs to occur at the right moment within the application lifecycle. In particular, some customers' responses show that they will not consider switching until their workload has been running on the cloud for a certain period of time. For example, two customers said that decisions around switching are constrained by their applications' lifecycle.[370]

[362] Responses to the CMA's information requests [✂].
[363] Responses to the CMA's information requests [✂]; Notes of meetings with [✂].
[364] Responses to the CMA's information requests [✂]; Note of meeting with [✂].
[365] [✂] response to the CMA's information request [✂].
[366] [✂] response to the CMA's information request [✂].
[367] [✂] response to the CMA's information request [✂].
[368] Note of meeting with [✂].
[369] Responses to the CMA's information requests [✂].
[370] [✂] response to Ofcom's information request [✂]; [✂] response to the CMA's information request [✂].

Prevalence of multi-cloud and switching

3.40 As outlined above, the ability to multi-cloud and switch between cloud providers influences the nature of competition in the markets and at the extremes:

(a) If all customers are freely able to switch or use more integrated forms of multi-cloud then cloud providers have a greater incentive to make their offerings competitive with their rivals. This is because if customers are able to switch or multi-cloud then they would be able to switch all or part of their workloads away to, or place new workloads with, rivals in response to their incumbent cloud provider becoming less competitive (eg due to higher prices or lower quality) relative to its rivals.

(b) Conversely if all customers are unable to switch or use any kind of multi-cloud then cloud providers have a lower incentive to make their offerings competitive with their rivals. This is because customers would not be able to switch existing workloads or place new workloads with a rival in response to their incumbent cloud provider becoming relatively less competitive than its rivals.

3.41 As such, the ability to multi-cloud and switch is particularly important to maintain competitive pressure on providers in relation to existing cloud workloads.

3.42 However, evidence on the extent to which customers switch or use multiple clouds can have multiple interpretations:

(a) On the one hand, we would generally consider that low levels of switching or use of multiple cloud providers are consistent with a lack of ability or incentive to switch or multi-cloud, and therefore with weaker competition;

(b) On the other hand, a lack of switching or multi-cloud would be consistent with suppliers responding proactively to the credible threat of switching or multi-cloud with competitive prices and level of quality.

3.43 This alternative interpretation may be more compelling where the level of switching and use of multiple cloud providers is not particularly low (as particularly low switching or multi-cloud may be inconsistent with customers being able to make a credible threat to switch or multi-cloud), and where other evidence on the ease of switching and on the broader market context (such as profitability) are consistent with that interpretation.

3.44 To the extent that there are low levels of switching or multi-cloud driven by barriers to switching or multi-cloud then competition may be more ‘for the market’ rather than ‘in the market’ as providers compete to attract customers when they initially pick a cloud provider rather than competing to attract existing customers away from rivals.

3.45 To the extent that there is competition 'for the market' rather than 'in the market', it may be argued that cloud providers would, in theory, have an incentive to lower their prices or increase the quality of their offering to new customers but not for existing customers. This is because the cloud provider has an incentive to compete harder to acquire new customers in order to capture future profits when those customers are subsequently unable or less able to switch away or multi-cloud. The argument may imply that barriers to switching or multi-cloud lead the customer to benefit initially because there is more aggressive up-front competition for its custom; this may compensate for any later lack of competition.

3.46 However, while we recognise that there may be some benefits upfront, we are cautious about attaching weight to an argument that this fully compensates customers for any later lack of competition. This is because it relies on a number of assumptions.

(a) One such assumption is that initial negotiations for new customers or new workloads are already subject to fully effective competition and, therefore, customers are able to 'extract' the full long-term value they represent to suppliers in those initial negotiations. If competition is not fully effective upfront then customers may only be able to 'extract' some of the long term value. Whether initial negotiations for new customers or new workloads are subject to fully effective competition depends on many factors, including those which are the subject of our investigation.[371]

(b) In addition, the extent to which customers are able to achieve competitive prices that reflect their long-term value to the supplier may be limited by any inability of customers to fully anticipate their own future needs or demand for the relevant services, as well as by any limit on their ability to foresee any cost savings or other changes in the competitiveness of suppliers over time. Such uncertainty might be expected in markets such as this where customers' requirements and needs and suppliers' production technologies are likely to evolve over time.

Multi-cloud prevalence

3.47 This section sets out the evidence on the prevalence of multi-cloud use amongst customers. We consider the following:

(a) cloud providers' submissions including some quantitative analysis submitted by a cloud provider;

[371] In some contexts these effects are referred to as 'waterbed effects' and they occur even where there is a monopolist operating on a particular market. Genakos and Valletti (2011), Testing the "waterbed" effect in mobile telephony, Journal of the European Economic Association, 9(6), 1114-1142.

(a) publicly available survey data and the results of a survey that Ofcom commissioned during its market study;[372] and

(b) our analysis of customer data provided by cloud providers.

Cloud providers' submissions

3.48 The cloud providers generally submitted that there is a high prevalence of multi-cloud in the market. In particular:

(a) Two cloud providers submitted that independent surveys and industry reports show that using multiple clouds is common.[373]

(b) Microsoft said that its working assumption is that all customers multi-cloud at least to some degree,[374] and that multi-cloud is now the de-facto standard.[375] Of customers that multi-cloud, Microsoft said that it estimated that there is roughly a [✂] split between "single app, single cloud", "same app, any cloud" and "app spans multiple clouds" architectures, respectively.[376] Microsoft also said that the lack of use of 'integrated' multi-cloud is because it does not yield significant customer benefits and that there may be good reasons why customers concentrate their spend around a primary provider and/or to run different workloads separately in different clouds.

(c) Another cloud provider submitted that use of multi-cloud is widespread.[377]

(d) Google said that open cloud is [✂]. The provider said that it is a challenger to AWS and Azure, which hold 80% of the market share, so in order to win customers it has to convince customers with workloads already in the cloud to move any future workloads to its cloud.[378]

(e) Three cloud providers said that they have developed services that enable or promote multi-cloud;[379] one of these identified that usage of these services can be an indicator of the level of multi-cloud and another identified it could be an indicator of a multi-cloud set-up.[380]

(i) AWS said that data from its services indicate a high level of multi-cloud use, including integrated multi-cloud. It explained that customers can manage their users in another on-premises or cloud directory and then

[372] For more information on the research that Ofcom commissioned, see Cloud Services Market Research - Summary of Findings March 2023 (ofcom.org.uk)
[373] [✂]; [✂] response to the CMA's information request [✂].
[374] Microsoft response to the CMA's information request [✂].
[375] Microsoft (ofcom.org.uk)
[376] Note of meeting with Microsoft [✂].
[377] [✂] response to the CMA's information request [✂].
[378] [✂].
[379] [✂]; [✂] response to the CMA's information request [✂].
[380] [✂]; [✂].

connect them into that provider's cloud through the provider's Identity Access Management solution. It said that approximately [✂]% of customers with annual spending greater than $[✂]annual spend use third party identity sources. It added that, of these customers, approximately [✂] use AzureAD and a further [✂] use Google IAM.[381]

(ii) Similarly, Google said that the fact that a customer is using BigQuery Omni can give an indication that the customer is deploying a multi-cloud strategy.[382] This product allows customers to query data on other cloud providers' storage services (Amazon Simple Storage Server or Azure Blob Storage), and is therefore designed for multi-cloud data management.[383] It said that it is a pioneer of tools designed to remove operational friction typically associated with using multiple clouds.[384]

(f) One cloud provider submitted that [✂]% of the tenders it participated in between [✂] were issued by its existing customers, indicating that customers do not view themselves as locked into a single cloud provider.[385] This cloud provider said that its win rate in tenders for customers with existing workloads in its cloud is [✂] to its win rate for other customers. It said that this shows it does not enjoy a significant advantage as an incumbent cloud provider.[386]

(g) Microsoft said that there is a high prevalence of multi-cloud and it is likely to increase due to customer requirements such as regulation, resilience, ability to take advantage of new applications and privacy.[387]

3.49 Some cloud providers also said that certain types of customers were more likely to multi-cloud:

(a) Three cloud providers said that large customers are, or could be more likely to, multi-cloud.[388]

(b) One of these said that this is because larger companies are better placed to adopt multiple clouds, in terms of their size, internal resource and wider variety of use cases, and are also more likely to recognise the risk from relying on one cloud. It added that larger companies also often have distinct departments/divisions which may run their own workloads independently

[381] AWS, submission to CMA [✂].
[382] Google's response to the CMA's information request [✂].
[383] Introduction to BigQuery Omni | Google Cloud, accessed 20 May 2024.
[384] Google's response to the Issues Statement, 17 October 2023, paragraph 15.
[385] [✂].
[386] [✂].
[387] Microsoft's response to the Issues Statement, 17 October 2023, paragraph 29.
[388] Responses to the CMA's information requests [✂].

from other parts of the organisation, meaning a siloed multi-cloud approach is more suitable.[389]

(c) Another cloud provider said that larger enterprises have the skills and scale to deploy across multiple clouds in a manner where the same application can run on any clouds.[390]

(d) One cloud provider said that financial services and banking firms are more likely to leverage a multi-cloud approach due to regulatory guidelines.[391]

3.50 As set out in these submissions, cloud providers generally submitted that they consider that multi-cloud architectures are common and that enabling customers to multi-cloud is part of their business strategy. Some cloud providers said that using multiple clouds is particularly prevalent among large customers.

3.51 However, we note that in general these submissions provide limited evidence on the prevalence of multi-cloud.

3.52 Some cloud providers did highlight surveys that sought to estimate prevalence of multi-cloud, and we consider these surveys below. One cloud provider also submitted a quantitative analysis which we consider in the next sub-section.

Quantitative analysis from a cloud provider

3.53 A cloud provider submitted quantitative analysis of its UK customers' usage and opportunity data to us.[392] It said that the analysis showed that, weighted by revenues, [✂]% of customers that appear in its opportunity data awarded tenders to [✂] cloud provider between 2017 and 2022, indicating that these customers were using multiple clouds.[393]

3.54 The cloud provider said that its analysis likely understates the prevalence of multi-cloud because:

(a) many customers acquire IT services without a tender process. These customers would not be recorded in the opportunity dataset;

(b) the provider did not participate in all tenders issued by customers;

(c) it is often not clear who the tender the cloud provider lost was awarded to. Conservatively, the analysis only flags customers as having awarded a

[389] [✂] response to the CMA's information request [✂].
[390] [✂] response to the CMA's information request [✂].
[391] [✂] response to the CMA's information request [✂].
[392] [✂] response to the CMA's information request [✂].
[393] Analysis based on data used for [✂] and Ofcom's analysis in its Final Report, Annex 3, paragraph A3.37. There are [✂] customers in the dataset, of which [✂] use multiple clouds.

tender elsewhere if the cloud provider knows the identity of the other competitor that won the tender; and

(d) some customers may have awarded tenders before or after the sample period.[394]

3.55 This cloud provider also submitted analysis on the distribution of revenue share of customers in its opportunity data[395] by the number of cloud providers the customers awarded tenders to between 2018 and 2022.[396]

3.56 We consider that the analysis has several limitations, and therefore needs to be interpreted with care in light of these caveats and our initial view is that we should place less weight on it than our own assessment set out below:

(a) In the first instance, the analysis only includes customers that went through a tender process for their workload(s). While the provider said this may understate the prevalence of multi-cloud, it is not clear this is supported by any evidence of higher prevalence among customers that do not tender.

(b) Moreover, one potential explanation for demand being placed without a tender is that the customer may consider its current provider the only credible option. Therefore, the exclusion of this demand could bias the analysis in favour of suggesting that the use of multiple clouds is prevalent (by restricting the analysis to customers that have indicated a willingness to multi-home by running a tender).

(c) We understand that the majority of customers do not use a competitive tender process to acquire cloud services (see section 2).

3.57 Further, we agree with the limitations raised by Ofcom about such analysis:[397]

(a) Weighting customers by their spend on the provider does not take into account how much customers may have spent on other cloud providers. To the extent that weighting by spend is informative, the results should be weighted by customers' total spend on public cloud.[398]

(b) Customers are counted as using multiple clouds if they put one workload on another cloud, irrespective of the size of the workload.[399] In this regard, our initial analysis of providers' data (set out in paragraphs 3.73 to 3.86 below) suggests that customers typically only use secondary providers for a minority

[394] [✂].
[395] [✂].
[396] [✂].
[397] Ofcom's Final Report, Annex 3, paragraph A3.45-A3.48.
[398] We note that our analysis only includes spend on AWS, Microsoft and Google, and therefore will not capture spend on smaller cloud providers.
[399] This is also relevant to our analysis. We set this out, and have estimated the average revenue split across splits as one way to account for this factor.

of workloads. This indicates that the fraction of the provider's customer revenues lost to rival cloud providers could be limited.

(c) The provider's analysis includes the workloads of any customer whom it records as having allocated even a single workload to a rival cloud provider. Therefore, it implicitly assumes that all of these workloads face competition from rival providers. We are considering evidence from customers[400] that suggests that customers who multi-cloud tend to only do this for workloads that are sufficiently siloed from each other (ie do not need to interoperate) – meaning most workloads are not contestable.

(d) Datasets used by the provider may not be reliable. The main dataset used, consists of data where 'opportunities are manually made by members of the provider's sales team. As a result, the data is not always comprehensive and may contain errors and inconsistencies'.[401] The provider noted that the data is 'often comprised of anecdotal feedback from the customer, particularly in relation to competitor information' and that the dataset 'does not cover all the provider's customers. The provider is therefore not able to accurately assess how representative and comprehensive the dataset in percentage terms'.[402]

Quantitative surveys to measure multi-cloud and switching

Quantitative surveys in the public cloud market

3.58 In general, quantitative surveys have some advantages: if the sample is representative, the results are generalisable to the whole population. This means that in some cases they can be an important part of our evidence base.[403]

3.59 However, we have concerns about using surveys to assess the prevalence of multi-cloud and switching due to:

(a) Lack of validity: quantitative surveys require respondents to make judgements about category responses (ie respondents must choose one or more options in a given list), but there is little room to follow up or clarify. We consider that the public cloud markets are technical and, in this case, it is difficult to reduce the complexities of these markets to simple, mutually exclusive questionnaire categories.[404]

---

[400] We are considering this evidence as part of our analysis of potential technical barriers.
[401] [✂] response to Ofcom's information request [✂].
[402] [✂] response to Ofcom's information request [✂].
[403] The CMA's good practice on customer surveys for its casework can be found at Survey_good_practice.pdf (publishing.service.gov.uk).
[404] For example, as identified in Appendix A, the Public First survey asked 'How many different cloud infrastructure providers does your company currently use?'. However, we consider that respondents may have interpreted this in as including any of:

(b) Vulnerability to the quality and coverage of respondents: results from quantitative surveys are crucially dependent on the representativeness of the sample and likely response rate. We have concerns about both of these factors in this market. In particular, customer decision-making on public cloud infrastructure services within firms can be complex, especially within larger cloud customers. Multiple individuals can be involved from a technical, commercial or strategic perspective, which makes defining and identifying a clear target population from which to draw a representative sample problematic. Further, the quality and accuracy of customer record-keeping within public cloud providers is highly variable, meaning their use as a sample frame for a quantitative survey becomes challenging. In particular, this issue can suppress response rates and increase non-response bias, reducing the representativeness of the final achieved sample for any quantitative survey. Alternative robust sample frames for this target population do not appear to us to be available. For example, commercial online panels that some industry surveys have used as a sample frame in these markets are not consistent with the CMA's good practice on customer surveys for its casework.[405]

3.60 For these reasons, we consider that there are significant weaknesses with using a quantitative research method to estimate the prevalence of multi-cloud in this market. We therefore treat results from these surveys with caution and place limited evidential weight on them. We have chosen to use qualitative research methods for our customer research in this investigation.

Public surveys and our initial views

3.61 Several cloud providers submitted that independent surveys and industry reports show that using multiple clouds is common. We note that, as set out above, we have concerns about using quantitative surveys to assess the prevalence of multi-cloud and therefore our initial view is that limited evidential weight should be placed on the findings.

3.62 We have presented the results of these surveys in Appendix A. We have also set out supplemental points that apply to the individual surveys and, as a general point, we note that the survey providers for all these surveys did not provide their

---

- Using multiple cloud providers for the underlying infrastructure layer;
- Using multiple first- and third-party providers hosted on the same cloud;
- Using both public and private cloud providers for infrastructure; and/or
- Using multiple private cloud providers.

For the purposes of estimating the prevalence of multi-cloud in this market, we are only interested in the first category.

[405] Sample bias is also a concern when respondents are drawn from a panel, in particular from an online panel, where sample recruitment does not rely on randomisation methods. Whilst a panel can be made to look like a random, representative cross-section of consumers in terms of its demographic profile, the characteristics of people who join a panel may be very different from other consumers. The CMA tends to place less evidential weight on surveys involving customer recruitment from panels, though each case is treated on its individual merits. See paragraph 2.29 in Survey_good_practice.pdf (publishing.service.gov.uk).

sampling and methodology, and we do not have access to the underlying data. As a result, we cannot assess the representativeness of the sample.

Our analysis of cloud providers' customer data

3.63 Overall, the quantitative evidence on multi-cloud is subject to uncertainty and it is unclear whether it is capturing customers using multiple clouds as we define it.

3.64 For this reason, we have built our own estimate of multi-cloud prevalence using customer data from cloud providers. We requested customer datasets from AWS, Microsoft and Google that identified customer names and annual spend on their respective clouds for 2020, 2021 and 2022. By analysing these datasets to identify customers using multiple cloud providers, we avoid any potential issue of customers misunderstanding what it is to multi cloud, as we define it for the purposes of the investigation. For example, customers using both private cloud and public cloud would not be counted as using multiple clouds in our analysis, but such customers may have responded in surveys that they use multiple clouds.

Methodology[406]

3.65 We matched customers' names across the customer datasets from AWS, Microsoft and Google. We used two types of matching:

(a) Perfect matching: exact matches of customer names across datasets.

(b) Fuzzy matching: matches based on similar but non-identical strings in customer names. Fuzzy matching produces a similarity score based on how good the match is, with 0 meaning the two are not a match and 1 meaning a perfect match. We chose to use fuzzy matching to capture additional matches where customer names may have been recorded slightly differently across the providers' datasets (eg 'Company A' in one dataset but 'Company A LTD' in another).

3.66 We excluded customers that spent less than $1,000 a year on a provider. In the first instance, this was because one provider provided their dataset on the basis of customers spending at least $1,000 on their cloud. Further, we consider this approach to be appropriate because it eliminates customers that are spending relatively little on a cloud, and therefore more likely only to be trialling that provider, as opposed to using multiple clouds in a material way.

3.67 We note that this method counts customers as using multiple clouds in a binary manner: customers are counted as using multiple clouds if they spend over $1,000 on another cloud, irrespective of the size of that workload. To address this, we

[406] We set out a more detailed methodology in Appendix A.

have also calculated the proportion of customers that have at least 30% of their spend on their secondary cloud (or secondary and tertiary cloud combined). This provides an indication of the prevalence of customers who multi-cloud for more than just a minor workload.

3.68 We also calculated revenue splits because we think that it is possible that customers with a revenue split closer to 50/50 are more likely to have a more integrated multi cloud set up. In particular:

(a) With a 50/50 split, it is more likely that the customer has distributed its workloads (including their stored data) more evenly between the relevant clouds. As such, it is more likely that some integrations are needed between these workloads.[407]

(b) At the other end, with a less even split (eg 95/5), it more likely that the customer is just experimenting with a second cloud or backing up its most critical workloads.

3.69 As such, we consider that revenue splits are a potential indicator of the extent of integration. However, we acknowledge that there is considerable ambiguity in what revenue splits imply: for example, it is also possible that customers with a 50/50 have siloed workloads, potentially due to different business units using different, separate clouds. Therefore, we are considering revenue splits in the context of other evidence we have received on integration across clouds. We expect to present this evidence in a separate working paper.

3.70 For the reasons set out above, we have set out the proportion of customers that spend at least 30% of their spend on their secondary cloud (or secondary and tertiary cloud combined). However, we also plan to conduct sensitivity checks using other thresholds in due course.

Limitations

3.71 Our analysis is subject to the following limitations and should be interpreted in light of these caveats.

(a) The analysis is sensitive to the threshold chosen. As discussed above, fuzzy matches are assigned a similarity score based on how good the match is. We manually checked the quality of matches at different thresholds and therefore

[407] If integrated multi-cloud is possible, customers are more likely to be able to choose any cloud provider for any incremental workloads, meaning cloud providers have the incentive to compete (for example, by lowering prices or innovating) to win all workloads. In comparison, if only siloed multi-cloud is possible, it may be the case that customers are only to choose any cloud provider for incremental workloads *provided* those workloads are sufficiently separate and has little or no integration with existing workloads. Therefore, cloud providers would have the incentive to compete for customers' initial workloads and for the subset of incremental workloads that are sufficiently separate.

chose a cut-off of 0.99[408] similarity score for the purposes of the analysis. If this threshold is too high, it would mean we miss 'true' matches, leading to an underestimate of multi cloud prevalence. Conversely, if the threshold is too low it would mean we match 'false' matches, leading to an overestimate of multi cloud prevalence. We are continuing to consider appropriate assumptions for this part of the analysis and we plan to conduct further sensitivity checks of the chosen threshold in due course. This may lead to different results from our analysis.

(b) Customers may have been recorded under different names in different datasets. If so, the fuzzy matching would not identify these customers, even though they are using multiple clouds, resulting in an underestimate of multi-cloud prevalence.

(c) New smaller customers may not be paying much to a cloud provider if the initial cloud credits cover most of their needs. These customers will not be identified in the matching exercise if their recorded spend is less than $1,000 even if in subsequent years the same activity would lead to a spend over $1,000.

(d) Our analysis is based on the datasets from AWS, Microsoft and Google. This means that customers that use other cloud providers, such as Oracle or IBM, as an alternative cloud will not be identified in the matching exercise. We note that all other cloud providers in the UK are many times smaller than the large cloud providers (see Section 5 on shares of supply). Therefore, even if all the customers of these smaller cloud providers use multiple clouds, it would not substantially affect our estimates.

(e) Given the dataset, we cannot tell where customers that use multiple clouds lie along the spectrum of siloed multi-cloud to integrated multi-cloud. For example, if firms have different subsidiaries that use different clouds, but the clouds do not communicate, we will record them here as customers that use multiple clouds in the same way we would a customer that has fully integrated clouds.

Results

3.72 In this section, we set out:

[408] We tested a range of different thresholds to understand the extent to which the identified fuzzy matches identified true matches or false matches. In doing this we took a random sample of matches that were identified to be in certain ranges and manually inspected if they were true matches or not. In doing this we identified that the majority of matches around this 0.99 threshold were false matches. Further, when the threshold was lowered to 0.98 a random sample of 150 observations showed that 80% were false matches. On this basis, and given that the 0.99 threshold already captures some false positives, we consider 0.99 the most appropriate threshold at this point even if it excludes some true matches. We will continue to consider the appropriateness of this assumption.

(a) The prevalence of multi-cloud, unweighted and weighted, by total annual cloud spend[409] across all customers;

(b) The average spend split across clouds across all customers that multi-cloud;

(c) The proportion of customers that spend at least 30% of their total cloud spend on their secondary (or secondary and tertiary) cloud;

(d) The prevalence of multi-cloud by spend band; and

(e) The average revenue split by spend band.

All customers

Prevalence of multi-cloud

3.73 Table 3.1 below shows the results of our initial analysis on the prevalence of multi-cloud, both unweighted and weighted by spend.

Table 3.1: prevalence of multi-cloud, unweighted and weighted by spend, 2020-2022

Prevalence of multi-cloud	*2020*	*2021*	*2022*
Unweighted (%)	6.7	7.1	7.1
Weighted by revenue (%)	31.3	34.0	34.4

Source: our analysis of customer data provided by AWS, Microsoft and Google

3.74 Table 3.1 shows that based on our initial analysis:

(a) approximately 7% of customers in the dataset use at least two of AWS, Google and Microsoft (unweighted); and

(b) around a third of all spend is by customers that multi-cloud.

3.75 We consider that these initial results indicate that customers using multiple clouds is uncommon, but that large customers are more likely to multi-cloud. The initial analysis also indicates that the prevalence of multi-cloud has increased slightly between 2020 and 2022, particularly when considering weighted shares.

3.76 However, we acknowledge that our analysis likely underestimates the true prevalence of multi-cloud due to the caveats in the methodology set out above. We are continuing to consider the appropriate assumptions for this analysis.

[409] This weighting gives those with higher total annual spends on cloud services a greater 'weight' to reflect their increased importance relative to those with lower spends.

Average spend split of customers that multi-cloud

3.77 Based on our initial analysis the average spend split across clouds when operating a two-cloud architecture is around 80/20 – that is 80% on the primary cloud and 20% of spend on the secondary cloud.

3.78 In a three-cloud architecture, the average spend split is approximately 75/15/10. As such, the primary cloud generally accounted for about three quarters of spend, while the secondary and tertiary cloud saw roughly similar levels of spend.

Prevalence of at least a 70/30 spend

3.79 As discussed above, we consider that customer architectures with a more even split may be more likely to have integrations across their clouds. For the purposes of this working paper, we have set out the proportion of (i) all customers and (ii) customers that multi-cloud at least 30% of their spend on their secondary (or secondary and tertiary) cloud. We present the results of our initial analysis in table 3.2 below.

Table 3.2: prevalence of at least 70/30 multi-cloud for all customers and customers that multi-cloud, 2020-2022

Prevalence of at least 70/30 multi-cloud	*2020*	*2021*	*2022*
All customers (%)	2.1	2.1	2.0
Customers that multi cloud (%)	31.2	28.9	28.6

Source: our analysis of customer data provided by AWS, Microsoft and Google

3.80 Table 3.2 shows that, based on our initial analysis, around 2% of all customers and just under a third of customers that multi-cloud have at least a 70/30 spend split. If the assumptions we made in the analysis hold, we consider this evidence indicates that an integrated multi-cloud architecture is rare.

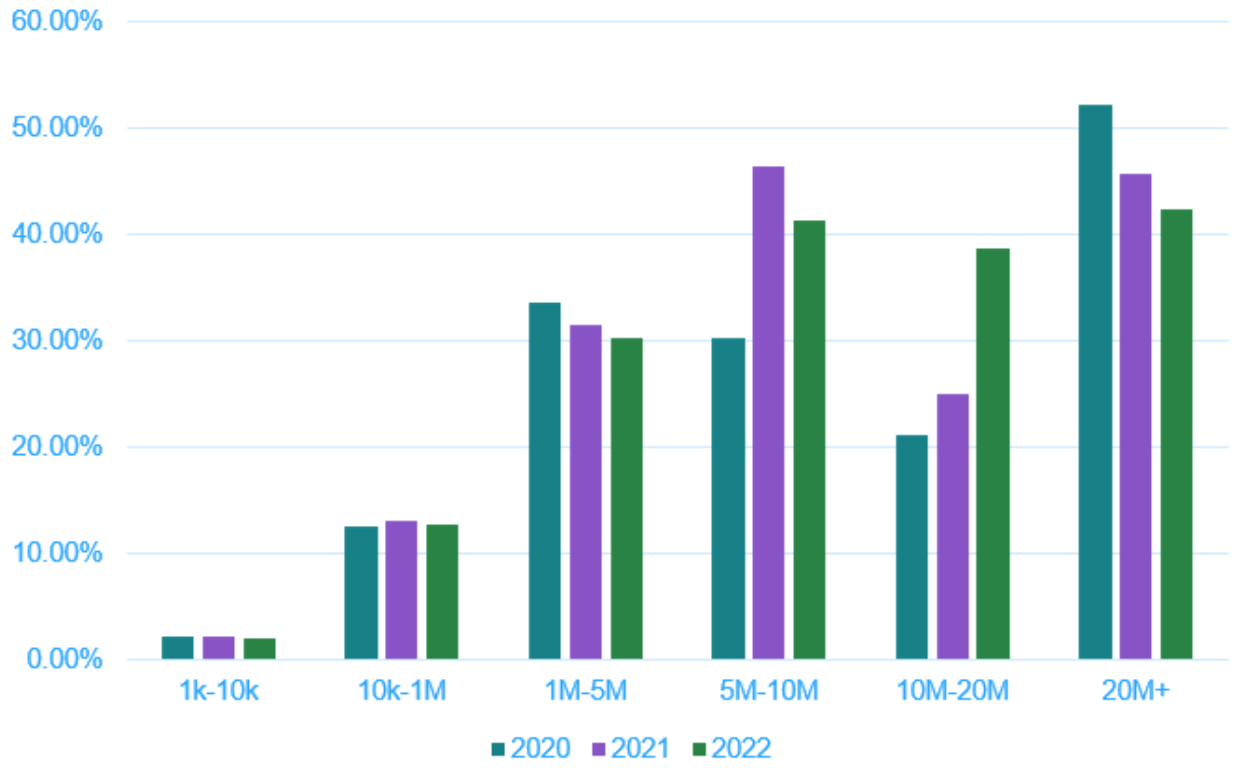
By spend band

3.81 We have also considered how some of these metrics differ when splitting customers into different spend bands (eg $1,000 to $10,000, $10,000 to $1 million). This is to better understand what is driving the differences between unweighted and weighted estimates of the prevalence of multi-cloud (see Table 3.2 above) and the extent to which the average spend split of customers that multi-cloud differs based on the size of the customer.

Prevalence of multi-cloud by spend band

3.82 Figure 3.2 presents the unweighted prevalence of multi-cloud by spend band.

Figure 3.2: prevalence of multi-cloud by spend band, 2020-22



Source: our analysis of customer data provided by AWS, Microsoft and Google

3.83 Figure 3.2 indicates that the prevalence of multi-cloud tends to increase with the amount of total spend on cloud.

3.84 In particular, our initial analysis indicates that only around 2% of customers that spend less than $10,000 on cloud use multiple clouds, compared to 45-55% of customers that spend over $20 million.

3.85 We note that the populations in the higher spend bands are often very small (in some cases less than five observations), and therefore highly sensitive to the actions of very few customers. Therefore, if we consider it appropriate to adjust the assumptions we made in our analysis, it may significantly change the results.

Average spend split of customers that multi-cloud by spend band

3.86 Table 3.3 presents the average spend split of customers that multi-cloud by spend band (the first number is the average proportion of spend on the primary cloud provider and the second number is the average proportion of spend on the secondary cloud provider).

Table 3.3: Average spend split of customers that multi-cloud by spend band, 2020-2022

Spend band	*2020*	*2021*	*2022*
Less than 10k	65/35	65/35	65/35
10K – 1M	80/20	80/20	85/15
1M – 5M	90/10	90/10	90/10

5M – 10M	85/15	95/5	90/10
10M – 20M	95/5	85/15	85/15
Over 20M	**80/20**	**85/15**	**85/15**

Source: our analysis of customer data provided by AWS, Microsoft and Google

3.87 Table 3.3 shows that, in general, customers in higher spend bands who use multi-cloud spend a greater proportion on their primary cloud. In comparison, lower-spend customers who use multi-cloud have a more even split across clouds.

3.88 We note that the fluctuations in average spend split in the higher spend bands is likely to be due to the low number of observations in those bands. As such, if those few customers change their behaviour year-on-year, it will be reflected in the overall spend band averages changing.

Emerging views on the prevalence of multi-cloud

3.89 There is a range of evidence and views on the prevalence of multi-cloud and each evidence source is subject to limitations. We have interpreted the evidence in light of these caveats. Overall our initial findings are that:

(a) Most cloud providers submitted that multi-cloud is common, but only one submitted an analysis of the actual prevalence of multi-cloud. This analysis has several limitations, including being based on the small sub-set of customers that use tenders, and therefore needs to be interpreted with care in light of these caveats. We currently place less weight on this provider's analysis than our own analysis of multi-cloud.

(b) There is some survey evidence on the prevalence of multi-cloud. However, we consider that there are significant limitations associated with using surveys in cloud services and would expect to place limited weight on this evidence; and

(c) Our analysis indicates that there is some degree of multi-cloud, but it may be quite limited in scope and mostly found in larger customers. However, we note that we are still considering the appropriate assumptions for this analysis.

Switching prevalence

3.90 This section sets out the evidence we have seen to date on the prevalence of switching by customers.

Cloud providers' submissions

3.91 Providers' responses on switching were mixed, with some saying it was low and others saying it was higher than Ofcom had suggested.[410]

3.92 However, limited data on switching was provided.

(a) One provider provided the number of UK customers that decreased their total annual spend across all of its services by at least 25%, 50%, 75% and 100% in the following year. The provider said that, while this does not necessarily indicate customer switching, it implies a reduction in spend or a halt in the use of some workloads on its cloud which may constitute switching. The provider added that the analysis would not capture customers simultaneously switching specific workloads from its cloud and other workloads to its cloud, and combined with the increased customer needs over time, this analysis likely understates switching.[411,412]

(b) Another provider gave an estimate of the number of customers that were not billed in that year, but billed the year before. The provider stated that this method does not conclusively evidence customer switching, as there are a number of reasons why customers might record revenues in one year but not the next, including testing its services, temporarily pausing use and the timing of invoice. [✂].[413]

(c) This provider also provided its lost global customers in 2022 (based on a customer whose stored data had essentially reduced to zero by December 2022). [✂], but there are limitations to what we can take from this give we do not have the provider's number of global customers and do not know how many of these were UK customers in line with the definitions we have used.[414] We note that almost all of the lost customers (97%) spent less than [✂] on the provider's cloud, and there were no lost customers that spent over [✂].

3.93 One provider said that low levels of switching are not necessarily evidence of weak competition as it may suggest that the gains from switching do not outweigh the costs of doing so. The provider also said that:[415]

[410] Ofcom's final report stated 'we remain of the view that switching levels are low in the cloud market', based on the evidence it had considered. However, it did not provide an exact estimated level of switching. See paragraphs 4.58 to 4.61 of Cloud services market study final report (ofcom.org.uk)
[411] [✂] response to the CMA's information request [✂].
[412] Responses to the CMA's information requests [✂].
[413] [✂] response to the CMA's information request [✂].
[414] [✂].
[415] [✂] response to the CMA's information request [✂].

(a) The CMA's music streaming market study showed that positive outcomes can arise notwithstanding limited switching and some barriers to switching.

(b) The market being considered is different from others the CMA has considered as customers are sophisticated, the cloud market is not mature and has a large number of new customers and it is not a digital multi-sided platform with network effects.

3.94 A provider said that low levels of switching are not caused by contractual restrictions, that customers have the ability to switch and to multi-cloud and that, given the prevalence of multi-cloud, it is much less common for customers to 'switch' fully than it is for them to scale up consumption of services on one cloud, which may or may not also result in scaling down of consumption on another. It said the key indication of whether it is 'losing business' to a competitor is [✂].[416]

3.95 A provider said that switching is far more prevalent than Ofcom's Final Report suggested and that, due to the way cloud services are designed and priced (pay-as-you-go) compared to previous IT environments, it has never been easier for customers to switch IT provider.[417]

3.96 Two providers said that they had an incentive to make switching easier to challenge larger players.[418]

3.97 Another provider also highlighted its own migration programmes as well as the transfer services of rival cloud providers and other IT companies that offer migration services (eg Accenture, BMC, Capgemini, and Deloitte).[419]

3.98 Some providers identified practices that may restrict customer choice and make switching more difficult such as certain licensing practices.[420] We consider these practices in separate working papers.

Quantitative analysis from a cloud provider

3.99 A cloud provider submitted quantitative analysis of its customer data and said that the results show that customers can and do switch.[421]

3.100 The cloud provider submitted that in 2020 and 2021, approximately [✂]% of its customers churned on an annual basis, considering all the cloud services it offered. This share would be approximately equal to [✂]%, if one only considers

[416] [✂] response to the CMA's information request [✂].
[417] [✂]
[418] [✂] response to the CMA's information request [✂]; [✂].
[419] [✂].
[420] [✂]; [✂].
[421] [✂] and Ofcom's analysis in its Final Report, Annex 3, paragraph A3.37.

cloud compute services in the analysis. According to the provider, these figures would exceed the churn rate from traditional on-premises IT providers.

3.101 Further, the provider said that [✂]% of its customers, accounting for [✂]% of the provider's revenues, [✂] their spend between the first half of 2020 and the first half of 2022. The provider said that these figures [✂] would be evidence of significant switching.

Emerging views on this quantitative analysis

3.102 Consistent with the limitations raised by Ofcom,[422] we consider that there are limitations to [✂] quantitative analysis on switching. We set out these limitations below:

(a) [✂] defined customers as churned if the customer reduced its spending by at least 80% for three consecutive months and did not return to 80% of their original spend within six months. This may only imperfectly capture the actual churn – a customer may still continue to buy from [✂] and/or significantly increase its spending on [✂] services after six months. [✂] measure of churn may reflect normal seasonal fluctuation – it is difficult to draw conclusions from the short time frame.

(b) The dataset includes many small customers who only spend a minimal amount on [✂] cloud services during the entire period considered. These customers may only have been trialling [✂] or only used the services occasionally. We note that:

(i) Removing customers that spent less than $100 reduces the churn rate from [✂]% to [✂]%.

(ii) Removing customers that spent less than $500 reduces the churn rate to [✂]%.

(c) Customer spend in [✂] dataset is heavily skewed towards the largest customers. The top 5% of [✂] customers account for [✂]% of [✂] total revenue and the top 10% account for [✂]%. The churn rate varies for different customer spend deciles:

(i) For customers above the 7th decile in spend, the churn rate is below [✂]% in each decile.

(ii) For customers in the top 10% of spend, the churn rates are approximately [✂]%.

[422] Ofcom's Final Report, Annex 3, paragraph A3.45-A3.48.

(d) With regards to [✂] calculations on customers that decreased spending between the first half of 2020 and the first half of 2022, we consider that a reduction in spending alone is not a meaningful measure of switching. There are many reasons a customer may have reduced spending; it does not necessarily follow that the customer is switching away all or part of its workloads.

3.103 For these reasons, we consider that the results from [✂] quantitative analysis should be interpreted in light of these caveats. We note that, in any case, [✂]% and [✂]% switching rates are not inconsistent with finding low switching levels in the cloud infrastructure services market.

Public surveys and Ofcom research

3.104 Cloud providers submitted one survey (Public First) that included an estimate of the prevalence of switching in the market. In addition, the prevalence of switching was considered as part of Ofcom's quantitative survey (Context Consulting).

3.105 There are also publicly available estimates of the prevalence of switching in the market. Our views on quantitative research in cloud services are set out in paragraph 3.60 above. As discussed, we consider that there are significant limitations to such estimates and therefore consider it appropriate to place very limited evidential weight on the quantitative evidence we have received thus far. Nevertheless, we present this evidence in Appendix A.

Ofcom research

3.106 The research[423] Ofcom commissioned with Context Consulting included both a qualitative research phase and a quantitative online survey. The quantitative online survey suffers from the same limitations outlined earlier about the available public survey evidence in this market, relating to issues of validity and sample quality. The qualitative research is not subject to the same methodological limitations that apply to its quantitative survey.[424]

3.107 We note that these qualitative results also indicate that switching is lower than the quantitative results reported. In particular, the former found:

(a) In some cases, firms were adding additional platforms, rather than switching;

[423] Cloud Services Market Research - Summary of Findings March 2023 (ofcom.org.uk)
[424] Context Consulting conducted 50 hour-long depth interviews with a further 14 follow-up interviews with current and potential cloud customers in their qualitative research. Qualitative research methods such as in-depth interviews allow for more discussion, clarification and explanation of customers behaviour compared to quantitative research methods, which use more structured research instruments such as questionnaires that seek to measure very defined category responses. Qualitative research methods can work well in complex and technical markets such as public cloud infrastructure services.

(b) Few examples of organisations switching away from AWS, Microsoft or Google to another cloud services provider;

(c) The switching that was described by customers tended to involve a relatively small portion of data and workloads, moving from one minority provider to another;

(d) It is still relatively early in the adoption journey for most companies, and they were evaluating progress rather than looking to make significant changes;

(e) In most cases, firms were still on their way into, not out, of their IaaS/PaaS environments.

Our analysis

3.108 As set out above, AWS, Microsoft and Google provided customer datasets that identified customer names and annual spend on their respective clouds for 2020, 2021 and 2022. We are currently considering the extent to which these datasets could be used to identify switching between at least these three providers.

3.109 In doing this we are considering whether we could use the matching exercise we have conducted (see the section on our analysis of multi-cloud prevalence above) to identify where customers appear to reduce spending on one public cloud while increasing it on another public cloud providers. Our current view is that this will be a better indicator of switching than just observing a reduction in spend on one public cloud. However, as our dataset only covers three public cloud providers it will exclude any switching to smaller providers.

3.110 In doing this, important considerations will include:

(a) The threshold used in matching customers across datasets.

(b) The revenue threshold(s) used to identify when a customer has switched. In particular, it would need to be sufficient to avoid fluctuations in usage that may occur absent any switching of workloads.

4. Market definition

4.1 Our guidelines state that defining the market helps to focus on the sources of any market power and provides a framework for the assessment of the effects on competition of features of a market.[425]

4.2 Further, our guidelines state that market definition is a useful tool, but not an end in itself, and that identifying the relevant market involves an element of judgement. The boundaries of the market do not determine the outcome of our competitive assessment in any mechanistic way. The competitive assessment takes into account any relevant constraints from outside the market, segmentation within it, or other ways in which some constraints are more important than others.[426]

4.3 There are normally two dimensions to the definition of a market:

(a) a product dimension where the relevant product market comprises a set of substitute products; and

(b) a geographic dimension, where the relevant geographic market may be national (or wider), regional or local.

4.4 We consider each of these aspects below.

4.5 The market definition(s) used by the CMA need not always correspond with the market for the goods or services described in the Terms of Reference ('relevant market(s)').[427] The CMA may conclude that the market definition goes wider or narrower than those goods and services.[428]

4.6 The willingness and ability of customers to switch to other products is a driving force of competition. This means in forming its views on market definition, the CMA will consider the degree of demand-side substitutability for a focal product. As noted in the guidance, in some markets supply-side constraints will also be important and we consider this below where relevant.[429] In determining whether there is supply-side substitutability the CMA may consider factors such as whether: [430]

(a) suppliers supply a range of different products in the same broad category, using the same set of assets and capabilities; and

[425] CC3, paragraphs 94 & 132.
[426] Ibid., paragraph 133.
[427] Ibid., paragraph 26.
[428] Ibid., paragraph 131.
[429] Ibid., paragraphs 134.
[430] Ibid., paragraph 134 and footnote 75.

(b) suppliers regularly introduce new products or reposition existing ones within the category.

4.7 Our guidelines also state that we may treat a group of product markets together for the purposes of assessing competitive effects. This can be the case where a feature manifests itself in a similar way across several different markets and the CMA is able to reach a view about the effects of the feature on competition across the group of markets as a whole.[431]

Product market

4.8 This section considers whether:

(a) IaaS is the relevant focal product;

(b) PaaS is substitutable for IaaS;

(c) Alternative IT models are substitutable for IaaS;

(d) PaaS is the relevant focal product;

(e) Alternative IT models are substitutable for PaaS; and

(f) SaaS is substitutable for PaaS.

IaaS

IaaS as the focal product

4.9 This section explains why we consider that IaaS is the relevant focal product. In particular, while IaaS consists of three services that provide access to raw computing resources for processing workloads and storing data, namely compute, storage, and network. Our current view is that it is reasonable not to use these narrower segments as our focal products for the purposes of our market definition assessment.

4.10 We set out below the cloud providers' submissions in relation to market definition for IaaS before explaining why we consider that IaaS is the relevant focal product.

Cloud providers' submissions

4.11 Most of the cloud providers' relevant submissions relate more generally to the similarities and differences between IaaS, PaaS and SaaS. However, we have set these out here as we also consider them relevant to the potential sub-

[431] CC3, paragraph 152.

segmentation between the IaaS elements. In particular, the cloud providers' submissions indicate that they consider that the market definition should be wider than IaaS, and therefore IaaS should at least not be subdivided:

(a) Providers said that there is not a single clear cut IaaS/PaaS/SaaS categorisation, and that different people might disagree over how a particular solution should be classified.[432] Providers said that they do not operate their businesses according to a strict IaaS/PaaS/SaaS segmentation.[433]

(b) One provider said that customers start by defining their objectives and needs and then look at a broad set of options that might help meet them, and do not typically evaluate products or services based on IaaS, PaaS or SaaS categories.[434] The same provider said that all models can be used to deploy similar solutions and are substitutable with each other, and that comparisons between categories are insufficient to reflect customer behaviour and the competition facing any given product.[435] Another provider said the level of control that an enterprise has over its workloads can be scaled up and down far more flexibly than the IaaS/PaaS/SaaS segmentation suggests.[436]

(c) Accordingly, one provider said that IaaS and PaaS should not be distinguished for the purposes of market definition.[437] Two providers said suppliers are innovative and a supplier present at a particular layer may be a competitive constraint on other infrastructure layers.[438]

4.12 On the relationship between compute, storage and networking, one cloud provider said that customers typically look to solve a specific IT problem, which may involve one or more different services, such as compute, storage and networking, working together in a specific way.[439]

Emerging views on IaaS as the relevant focal product

4.13 Based on the evidence we have seen to date we consider that each element of IaaS serves a different function, generally relies on each other and is not used in isolation by customers such that there may be no distinction made by customers. As such it is not clear that each element of IaaS are demand-side substitutes.

4.14 On the supply side, we understand that all IaaS suppliers supply each of compute, storage and networking. Further, we have not received any evidence to date that

[432] Responses to the CMA's information requests [✂].
[433] Responses to the CMA's information requests [✂].
[434] [✂] response to the CMA's information request [✂].
[435] [✂] response to the CMA's information request [✂].
[436] [✂] response to the CMA's information request [✂].
[437] [✂] response to the CMA's information request [✂].
[438] Responses to the CMA's information requests [✂].
[439] [✂] response to the CMA's information request [✂].

suggests that competitive conditions are different for each element. As such, we consider that there is likely to be supply-side substitution between IaaS elements.

4.15 However, we currently consider that we can leave open whether the market should be subdivided into IaaS elements. As identified above, our guidance states that we may treat a group of product markets together for the purposes of assessing competitive effects; for example, where a feature manifests itself in a similar way across several different markets.

4.16 We consider that the features that we are exploring in the course of this investigation are relevant across IaaS services, and as such consider that it is reasonable to consider IaaS services in aggregate for the purposes of this market investigation.

4.17 Having said this, in our analysis we will bear in mind where relevant that there are likely to be different segments within IaaS and, while we have received no submissions suggesting that IaaS should be segmented further, we remain open to any submissions from market participants that aggregating IaaS services would have implications for our assessment.

Including PaaS in the market

4.18 This section considers whether PaaS is a substitute for IaaS.

4.19 As set out in section 1, PaaS provide access to a virtual environment for customers to develop, test, deploy and run applications. These include application development computing platforms and pre-built application components and tools which customers can then use to build and manage full applications. The customer has less control over the cloud stack compared to IaaS – they still manage applications and data but not the PaaS computing platform (including its operating system) and the pre-built application components and tools.

Cloud providers' submissions

4.20 As set out in paragraph 4.12, cloud providers generally submitted that IaaS and PaaS were substitutes and thus should be included in the same market.

Evidence from customers

4.21 We asked large customers about the mix of IaaS and PaaS they use and the main factors that influence their usage of these public cloud services models. Most customers we spoke to said they use a mix of IaaS and PaaS for their cloud

workloads.[440] In general, customers expressed their preference for either IaaS or PaaS but they may use the other in specific cases.

4.22 We asked customers about the extent to which they regard IaaS and PaaS as substitutes for each other:

(a) Many customers said that IaaS and PaaS are not substitutes or, even if IaaS and PaaS are technical substitutes, they would not switch between them because they have different characteristics.[441]

(i) Some of these customers said that IaaS offers more control and flexibility compared to PaaS. For example, one customer said that it frequently creates its own solutions based on IaaS because PaaS does not offer sufficient scale or SLAs, is higher cost and has an even higher level of lock-in.[442]

(ii) Similarly, another customer explained that PaaS is not a perfect alternative because it wants to control its own services.[443]

(iii) Another customer said that, while it is possible to rearchitect and refactor an IaaS workload for PaaS, it uses PaaS sparingly because it is more constraining in terms of flexibility, propriety and specific ways of working as well as being more subject to lock-in.[444]

(b) A handful of customers said that IaaS and PaaS could be substitutes for specific workloads.[445]

(i) For example, one customer said that substitutability is largely dependent on the service in context, but overall there is minimal substitutability.[446]

(ii) Similarly, another customer stated that the capabilities of IaaS and PaaS are often not interchangeable and that it uses IaaS and PaaS to achieve different goals, but they can be substitutes in some cases.[447]

(iii) Another customer said that PaaS is a good substitute for IaaS when there is no customisation required at the operating level.[448]

---

[440] Responses to the CMA's information requests [✂].
[441] Responses to the CMA's information requests [✂].
[442] [✂] response to the CMA's information request [✂].
[443] [✂] response to the CMA's information request [✂].
[444] [✂] response to the CMA's information request [✂].
[445] Responses to the CMA's information requests [✂].
[446] [✂] response to the CMA's information request [✂].
[447] [✂] response to the CMA's information request [✂].
[448] [✂] response to the CMA's information request [✂].

(c) A few customers said that IaaS and PaaS are easily substitutable for most workloads.[449]

(d) Another customer said that it intends to switch from IaaS to PaaS, though due to a strategy shift to benefit from the lower management overhead and consumption based cost model associated with PaaS.[450]

(e) Similarly, one customer said that PaaS is first or second choice for most applications, including application refreshes for systems previously implemented as virtualised infrastructure. The customer explained it prefers the cost and operating advantages of managed services.[451]

Emerging views on including PaaS

4.23 Although providers said that they do not necessarily segment their products along the IaaS/PaaS/SaaS categorisation, and there is a spectrum of products that do not necessarily fit neatly into these categories, our emerging view, based on the evidence to date, is that this does not mean that the underlying products and services are in fact substitutes.

4.24 Our guidelines refer to using the hypothetical monopolist test (HMT) to identify effective substitutes.[452] We consider that, in the context of the HMT, PaaS is unlikely to be an effective alternative to IaaS. In particular, the HMT asks whether a hypothetical monopolist of IaaS could profitably increase price by 5-10%. We can consider two types of affected customers for IaaS: (a) ISVs who use IaaS as an input into the PaaS products they supply; and (b) other IaaS customers:

(a) ISVs that supply PaaS products cannot switch to using a PaaS product to avoid a price rise on infrastructure; they need the underlying infrastructure as an input to their product.[453] While some ISVs may be relatively large customers of cloud providers and thus able to negotiate generally to attract better terms from specific cloud providers, they also rely on the underlying infrastructure to supply their own PaaS products so would not be able to switch away from the price rise of a hypothetical monopolist IaaS provider. This is particularly the case as the number of customers using a specific cloud provider's infrastructure is important to ISVs (see the discussion of network effects in Section 7 below) such that ISVs have relatively less bargaining power in relation to larger IaaS providers.

[449] Responses to the CMA's information requests [✂].
[450] [✂] response to the CMA's information request [✂].
[451] [✂] response to the CMA's information request [✂].
[452] CC3, paragraph 138.
[453] See paragraph on the nature of ISVs' business models.

(b) Other IaaS customers face the choice of absorbing the price increase or switching to PaaS. However, as discussed, ISVs also face an increase to their input costs. As IaaS is a variable cost to ISVs, it is likely that this will be passed through to the price they charge for supplying PaaS. As such, both options for IaaS customers result in increased price and would essentially mean that they continue purchasing IaaS despite 'switching away'.

4.25 We note that the fact that some customers shift from IaaS to PaaS does not necessarily indicate that demand is contestable between IaaS and PaaS and may instead indicate changing requirements causing migration, rather than substitution between two products which serve the same need.

4.26 We recognise that for some customers, for some workloads, IaaS and PaaS are substitutes. However, we consider that evidence we have seen to date from customers indicates that PaaS is not a good substitute for IaaS for most customers and workloads and that most customers expressed that they are unwilling to substitute between the two, even if it may be technically possible to do so.

4.27 In light of the above, our emerging view is that it is unlikely that there would be a sufficient degree of demand-side substitutability to warrant widening the market to include PaaS.

4.28 We also currently consider that supply-side substitution is unlikely to be sufficient to warrant aggregating IaaS and PaaS together. As set out in paragraph 4.6 above, to aggregate markets on the basis of supply-side substitution, we consider factors such as whether (i) suppliers supply a range of different products in the same broad category, using the same set of assets and capabilities; and (ii) suppliers regularly introduce new products or reposition existing ones within the category. It is true that larger cloud providers offer both IaaS and PaaS. However, we consider it unlikely that ISVs would be able to quickly or easily start providing IaaS, given the substantial investment in infrastructure that would be required (see section 7 below). Further, we consider that the competitive conditions in IaaS and PaaS are different, for example see our estimates of shares of supply in Section 5.

4.29 As such, based on the evidence we have seen to date, our emerging view is that the market should not be widened to include PaaS. Where relevant, PaaS would be considered as an out-of-market constraint.

PaaS

PaaS as the relevant focal product

4.30 This section explains why we consider that PaaS is the relevant focal product, and that it should not be further sub-divided into segments.

4.31 Although PaaS consists of hundreds of individual products, many of which perform different functions, we believe that it is reasonable not to use these narrower segments as our focal products for the purposes of our market definition assessment.

4.32 We set out below the cloud providers' submissions in relation to market definition for PaaS before explaining why we consider PaaS is the relevant focal product.

Cloud providers' submissions

4.33 As set out in paragraph 4.12 cloud providers generally submitted that different elements of PaaS were substitutes and therefore should be in the same market.

Emerging views on PaaS as the relevant focal product

4.34 PaaS consists of hundreds of individual products, many of which perform different functions; they are not precisely defined in the industry and each cloud provider uses different categories.[454] These categories are not mutually exclusive or collectively exhaustive and therefore are not straightforwardly subjected to demand- and supply-side substitution questions. Although PaaS could theoretically be subdivided into more granular segments we consider that for the purposes of this market investigation, it is not reasonable to do so in identifying the relevant focal product(s).

4.35 For example, on the demand-side: (i) within PaaS categories there are products that are unlikely to be substitutable and (ii) there are products in different PaaS categories that may be substitutable with each other. Therefore, in this context, taking a rigid approach to applying demand-side substitution considerations to a large number of subcategory focal products which are both arbitrary and do not have clear boundaries may not be helpful to the overall assessment of the wider supply of Public Cloud Infrastructure Services.

4.36 In addition, we believe that supply-side substitution is unlikely to be helpful in aggregating together product categories into distinct focal products without blurred boundaries. As set out in paragraph 4.6 above, to aggregate markets on the basis of supply-side substitution, we consider factors such as whether (i) suppliers supply a range of different products in the same broad category, using the same set of assets and capabilities; and (ii) suppliers regularly introduce new products or reposition existing ones within the category. On these we consider that, while it may be the case that these factors apply to larger cloud providers, it is unlikely that the many ISVs that operate in PaaS have the capabilities or assets to quickly and easily redirect production from their current area of focus within PaaS to another area of PaaS. In considering supply-side substitutability, we are also likely to

[454] Categories include, for example, databases, analytics, containers and machine learning.

consider whether competitive conditions are similar across markets. In relation to this, ISVs are present in particular niches of PaaS to varying degrees, suggesting that competitive conditions are not similar across PaaS categories.

4.37 Given this, as well as the fact that none of the evidence received from cloud providers to date supports narrower focal products, our initial view is that it would not be appropriate to consider a narrower focal product than PaaS.

4.38 In any case, as set out in paragraph 4.8 above, our guidance states that we may treat a group of product markets together for the purposes of assessing competitive effects; for example, where a feature manifests itself in a similar way across several different markets.

4.39 We consider that the features that we are exploring in the course of this investigation are relevant across PaaS services, and as such consider that it is reasonable to consider PaaS services in aggregate for the purposes of this market investigation.

4.40 Having said this, in our analysis we will bear in mind where relevant that there are likely to be different segments within PaaS and, while we have received no submissions suggesting that PaaS should be segmented further, we remain open to any submissions from market participants that aggregating PaaS Services would have implications for our assessment.

Including SaaS in the market

4.41 This section considers whether SaaS are a substitute for PaaS.

4.42 SaaS are complete applications hosted in the cloud as described in section 1.

Cloud providers' submissions

4.43 As set out in paragraph 4.12 cloud providers generally considered IaaS, PaaS and SaaS to be in the same market.

4.44 In addition to the above, one provider said that there are grey areas between PaaS and SaaS services. In particular, it said that PaaS services can have some elements of SaaS solutions, but while SaaS-like functions are delivered by PaaS services on its platform, it does not believe those products belong in the SaaS service model.[455]

[455] [✂] response to the CMA's information request [✂].

4.45 One provider said that the SaaS layer is even more fragmented than the IaaS or PaaS layers, and another provider said that SaaS is a very crowded field with many large to small companies, some of which also offer PaaS.[456]

Evidence from customers

4.46 Only a handful of the customers we spoke to mentioned their use of SaaS. These customers identified that they use a mix of PaaS and SaaS services. For example:

(a) One customer said that SaaS solutions are given first preference where business requirements can be fulfilled. Where a SaaS solution is not available, the customer said it will consider other models iteratively.[457]

(b) Similarly, another customer said that its first guiding principle is to adopt SaaS for services and applications that are commodity in nature.[458]

(c) One customer said that the decision between SaaS and PaaS depends on the requirements of the individual product;[459] and another customer said that its choice of SaaS, PaaS and IaaS depends on use case and requirements.[460]

(d) One customer said that it makes use of each of SaaS, PaaS and IaaS, which allows it to benefit from the reliability and security of SaaS, high-availability of PaaS and flexibility of IaaS.[461]

4.47 In addition, one customer said that PaaS is fundamentally not a substitute for SaaS and explained that each layer of cloud represents a need to have skilled personnel to manage it and develop against it.[462]

Emerging views on including SaaS

4.48 The evidence we have seen to date shows that customers use a combination of PaaS and SaaS (and IaaS) products. While there is some evidence that showed that the choice of where to place the workload reflected the requirement of that particular workload (or a business strategy to place their workloads in that layer) where possible, in general the evidence on the extent to which these are substitutable is mixed and limited.

4.49 In relation to supply-side substitution, based on the evidence we have seen to date, our emerging view is that competitive conditions for PaaS and SaaS are

---

[456] Responses to the CMA's information requests [✂].
[457] [✂] response to the CMA's information request [✂].
[458] [✂] response to the CMA's information request [✂].
[459] [✂] response to the CMA's information request [✂].
[460] [✂] response to the CMA's information request [✂].
[461] [✂] response to the CMA's information request [✂].
[462] [✂] response to the CMA's information request [✂].

significantly different, in terms of the number of firms, such that our initial view is that we do not consider there to be supply-side substitution between the two layers.

Alternative IT models

4.50 This section considers the extent to which either (i) traditional IT (ie dedicated computing resources on-premises) or (ii) private cloud services (ie a cloud which is exclusive to one customer) are substitutes for any of IaaS, PaaS or SaaS.

Cloud providers' submissions

Cloud providers' submissions on traditional IT

4.51 Two cloud providers said that, from the customer's perspective, on-premises and public cloud are substitutable for most use-cases.[463] One provider also said that customers switch regularly between cloud and on-premises and provided examples of customers switching from cloud: one customer moved its data out of the provider to a hybrid on-premises and cloud solution, and another customer moved from the cloud to a hybrid solution by building its own network of servers to "switch seamlessly between cloud providers and its own servers".[464]

4.52 One provider added that customers typically look to solve a specific IT problem and then look at a broad set of options that might meet them; they rarely, if ever, look simply to use the cloud as an end in itself.[465]

4.53 Another provider said that the key difference between the two solutions are flexibility and scalability, rather than technical differentiation. It also said that traditional IT, unlike cloud services, is costly and can be inefficient.[466]

4.54 One provider said that traditional IT solutions suit customers with requirements for a high degree of control over their IT, or that have to comply with security or regulatory requirements.[467] It also said that traditional IT can be more resource-intensive to manage and maintain. [468] The same provider said that, in comparison, public cloud meets a broad and diverse array of customer needs, including customers that do not want to make large investments in their IT solutions or companies that are looking for access to the most cutting-edge computing capabilities.[469]

---

[463] [✂] response to the CMA's information request [✂].
[464] [✂] response to the CMA's information request [✂].
[465] [✂] response to the CMA's information request [✂].
[466] [✂] response to the CMA's information request [✂].
[467] [✂] response to the CMA's information request [✂].
[468] [✂] response to the CMA's information request [✂].
[469] [✂] response to the CMA's information request [✂].

4.55 From the supplier perspective, three cloud providers said that there are very few technical differences between on-premises and cloud services.[470] One provider added that the difference is in method of delivery, not the underlying technology.[471]

Cloud providers' submissions on private cloud

4.56 As set out above, one provider said that, from a customer's perspective, IT services are substitutable for most use-cases, regardless of their delivery method.[472]

4.57 Similarly, a provider said that customers multisource and switch between different deployment models and technologies, which highlights that these hosting methods are largely interchangeable from a technical perspective.[473]

4.58 Several cloud providers said that private cloud offers more control, security and reduced latency, while public cloud has greater scalability, flexibility and may be cheaper.[474]

4.59 One provider added that even for customers with latency or security requirements there remains strong competition for and substitutability between the different deployment models and technologies.[475] The same provider provided an example where a customer chose instead to invest in a private cloud platform.[476] Another provider provided an example of a customer switching from its public cloud to private cloud, though the customer ultimately brought workloads back to public cloud to benefit from the continuous investment and innovation in hardware and other services.[477]

Evidence from customers

4.60 We asked large customers about the extent to which they consider non-public cloud IT environments as a substitute for public cloud computing and about the main factors that influenced their choice between the two. Most customers we spoke to answered by referring to non-public cloud without specifying traditional IT or private cloud, but where possible we have differentiated their answers between options.

4.61 Most customers said that they would not switch from public cloud to non-public cloud, or that doing so would involve significant resources.[478] Most of these

---

[470] Responses to the CMA's information requests [✂].
[471] [✂] response to the CMA's information request [✂].
[472] [✂] response to the CMA's information request [✂].
[473] [✂] response to the CMA's information request [✂].
[474] Responses to the CMA's information requests [✂].
[475] [✂] response to the CMA's information request [✂].
[476] [✂] response to the CMA's information request [✂].
[477] [✂] response to the CMA's information request [✂].
[478] Responses to the CMA's information requests [✂].

customers identified that public cloud has different characteristics, such as advanced functionality, cost, elasticity, scalability, flexibility, and resiliency.[479]

(a) For example, one customer said that, while it uses on-premises for legacy applications, it would take significant effort to construct a stack with the full security and observability capabilities as virtual machines, and even then would still miss many elements of a public cloud value proposition.[480]

(b) Similarly, another customer said that the public cloud is materially more cost-effective, gives it flexibility and scale, allows it to rapidly respond to changes in demand, and provides a level of innovative services that non-public clouds do not.[481]

4.62 A handful of customers said that non-public clouds are a substitute for public cloud for certain workloads.[482]

(a) For example, one customer said this is the case in a limited setting, such as hosting virtual machines and layering other components on top, such as SQL Server to mimic SQL PaaS.[483]

(b) Similarly, another customer said that substitutability depends on the application stack development.[484]

4.63 A handful of customers said that non-public cloud environments are a substitute for public cloud.[485]

(a) One customer said that most of its workloads can be relatively easily migrated to non-public cloud environments.[486]

(b) However, some of these customers qualified their responses:

(i) one customer said that applications that use purely basic IaaS services can be run on non-public cloud environments, but even in this scenario there are complexities due to networking and security configurations.[487]

(ii) another customer said that, while currently its applications could be run in non-public cloud, it is feasible that it will evolve to an extent that those applications may not be suitable for on-premise hosting.[488]

---

[479] Responses to the CMA's information requests [✂].
[480] [✂] response to the CMA's information request [✂].
[481] [✂] response to the CMA's information request [✂].
[482] Responses to the CMA's information requests [✂].
[483] [✂] response to the CMA's information request [✂].
[484] [✂] response to the CMA's information request [✂].
[485] Responses to the CMA's information requests [✂].
[486] [✂] response to the CMA's information request [✂].
[487] [✂] response to the CMA's information request [✂].
[488] [✂] response to the CMA's information request [✂].

4.64 There were cases where customers also said that they have adopted a cloud first policy that specifically targets moving their workloads to cloud services. For example, one customer's technology strategy aims for 'SaaS first, cloud second, on prem only where necessary'.[489]

4.65 We also asked customers whether they had ever switched any workloads hosted on a public cloud to be hosted on non-public cloud IT environments instead, such as private cloud or traditional IT, what they switched and the reasons why.

4.66 Most customers said that they had not switched to on-premises,[490] although two of these were exploring switching.

4.67 All of the customers that did switch only did so for certain workloads, and some only did so because the characteristics of public cloud did not meet the requirements for those workloads, rather than as a response to a price rise.[491] For example:

(a) one customer switched because of application latency constraints and poor client application architecture;[492]

(b) another customer was unable to get the right level of performance and price in public cloud, and so switched back to traditional data centre services.[493]

4.68 One customer said that it had switched to private cloud and another had switched to a non-public cloud environment, but it was unclear whether this was to private cloud or traditional IT. The former customer said that it had moved some workloads to its private cloud, but it is extremely complicated to move workloads from one location to another, and it does not wish to do so frequently.[494] No other respondents identified moving to the private cloud.

4.69 A handful of customers specifically submitted evidence on the substitutability of private cloud for public cloud: these customers said that the public cloud can technically be replicated on private cloud.[495]

(a) Of these customers, most qualified that there would be cost and/or effort involved in switching.[496]

[489] [✂] response to the CMA's information request [✂].
[490] Responses to the CMA's information requests [✂].
[491] Responses to the CMA's information requests [✂].
[492] [✂] response to the CMA's information request [✂].
[493] [✂] response to the CMA's information request [✂].
[494] [✂] response to the CMA's information request [✂].
[495] Responses to the CMA's information requests [✂].
[496] Responses to the CMA's information requests [✂].

(b) A few of these identified that their operations are subject to significant variation in loads over time, some of which may be better suited to a public cloud environment.[497]

(c) Further, one customer said that, while it could do everything it needed on a private cloud, it would take a huge amount of time, effort, and re-engineering and would be incredibly disruptive.[498]

(d) Another customer said that private cloud potentially results in the loss of shared development, but private and public cloud use fundamentally the same technology.[499]

Emerging views on alternative IT models

4.70 Evidence we have seen to date suggests that there is a significant migration of workloads from traditional IT to cloud services, and indeed cloud services have been growing at a substantially higher pace than the traditional IT services.[500]

4.71 However, to the extent that this migration reflects a long term trend in the industry, it is not reflective of whether customers would switch from public cloud to traditional IT in response to a 5-10% increase in price, and therefore does not provide evidence on the extent to which public cloud is constrained by traditional IT.

4.72 In addition, the migration trend from the traditional IT services to public cloud indicates that any constraint is likely to be asymmetric: while public cloud may be a credible alternative for traditional IT customers, the evidence we have received suggests that the reverse is generally not true.

4.73 Based on the evidence set out above, our emerging view is that traditional IT is unlikely to be a close substitute for public cloud for a high proportion of customers.

4.74 Although customers do not choose public cloud as an end in itself, public cloud has distinct characteristics, such as flexibility and improved functionalities. Customers choose public cloud for these characteristics that traditional IT cannot replicate well. While there are some workloads that could be migrated back to traditional IT, in general customers have only done so when the workloads were not well suited for public cloud in the first place. Customers switching back to traditional IT in such cases reflects a learning process around which workloads

[497] Responses to the CMA's information requests [✂].
[498] [✂] response to the CMA's information request [✂].
[499] [✂] response to the CMA's information request [✂].
[500] Responses to the CMA's information requests [✂].

work well in the cloud, rather than necessarily the extent to which the solutions are substitutable.[501]

4.75 Regarding private cloud, examples of investing in private cloud rather than public cloud may represent its better suitability for a customer's use case, rather than substitutability between public cloud and private cloud.

4.76 The evidence we have seen to date indicates that public cloud workloads could technically be hosted on private cloud. However, evidence from customers shows that there would be significant costs and time associated with switching to private cloud and only two customers indicated that they had previously switched.

4.77 Based on the evidence we have seen to date, including that the majority of customers would not switch from public cloud to private cloud or traditional IT, we consider that whatever degree of substitutability is observed today is likely to decline over time. Although large customers may be better able to react to a price increase by switching to private cloud or traditional IT, we consider that the evidence we received indicates that, even for large customers, such a switch would be unlikely due to the specific reasons they place workloads on public cloud and costs and time associated with doing so.

4.78 Based on the evidence we have seen to date, our emerging view is that:

(a) traditional IT should be treated as separate from the markets for IaaS, PaaS and SaaS for the purposes of this investigation; and

(b) private cloud should be treated as separate from the markets for IaaS, PaaS and SaaS for the purposes of this investigation.

4.79 We will consider traditional IT and private cloud as an out of market constraint to the extent applicable.

Emerging views on the product market

4.80 The evidence we have seen to date suggests that there is a relevant product market for the supply of IaaS, but PaaS is not part of the same relevant market and, where relevant, PaaS would be considered as an out-of-market constraint. We have considered the extent to which there is a relevant product market for the

[501] AWS said that the Public First survey shows that cloud service providers face significant competitive pressure from on-premises providers. Public First conducted a survey of 1,001 current or potential users of IaaS/PaaS/SaaS in the UK. The survey found that 7.5% [54 out of 716] IaaS/PaaS users switched to an on-premises solution. However, we do not consider that it is appropriate to place any evidential weight on the survey from Public First. In particular, the sampling and methodology of the survey is not clear, and it is not clear what type of switching it captured. It is possible that some of the switching may have been between private cloud and on-premises solutions. Further, the reasons for switching are not clear: as identified above, customers may be switching workloads they discovered were not well suited to the cloud. Indeed, the survey results indicate that respondents use cloud services and on-premises for different reasons. Finally, the survey was funded by CCIA, which has Amazon and Google as two of its members. AWS' response to the Issues Statement, 17 October 2023, paragraph 17.

supply of PaaS and the evidence on the extent of substitutability between PaaS and SaaS is mixed and limited.

4.81 Further, the evidence to date indicates that, even for large customers, switches from public cloud to traditional IT and private cloud would be unlikely due to the specific reasons they place workloads on public cloud and the costs and time associated with doing so. Therefore, our emerging view is that traditional IT and private cloud should be considered as out of market constraints where applicable.

Geographic market

4.82 Our guidelines state that geographic markets can be defined based on the location of either suppliers or customers[502] by considering the degree of substitutability, ie the extent to which suppliers can switch their areas of supply and the extent to which customers in one area may be served in another area.[503]

4.83 The Terms of Reference in this case concern the supply of public cloud infrastructure services in the UK. As noted above, the market definition(s) used by the CMA need not always correspond with the relevant market(s) described in the Terms of Reference; specifically the CMA may conclude that the market definition goes wider or narrower than those goods and services.[504]

4.84 In that context, in this section we consider whether the market for public cloud infrastructure services is national or whether it should be expanded to Europe-wide (ie UK and EEA) or global basis.

Cloud providers' submissions

4.85 Most cloud providers said that it is not necessarily a requirement to have physical infrastructure based in the UK to compete effectively.[505] One added that a cloud provider can rely on physical infrastructure outside but nearby the UK and still compete effectively[506] and a different provider said that many of its UK customers use infrastructure located outside of the UK.[507] This is consistent with the capacity shares relative to the revenue shares, which are both set out in Section 5.

4.86 However, one provider said that, while it cannot say for definite, having physical infrastructure in the UK is important because the UK no longer being part of the EU means data sovereignty is even more important.[508] Two other cloud providers said that UK-based infrastructure may be important for certain customers such as

---

[502] CC3, paragraph 145.
[503] Ibid., paragraph 147.
[504] Ibid., paragraph 26 and 131.
[505] Responses to the CMA's information requests [✂].
[506] [✂] response to the CMA's information request [✂].
[507] [✂] response to the CMA's information request [✂].
[508] [✂] response to the CMA's information request [✂].

those with specialised security needs.[509] Another provider said that, while infrastructure does not need to be located near customers, providers may choose to do this to decrease latency or to address customer preferences for data location.[510]

4.87 Two cloud providers said that significant UK-specific investments are not necessary to competitively provide cloud infrastructure services to UK customers.[511] Three other cloud providers said that a cloud provider would need to make some UK-specific investments:

(a) One said that costs would include legal, sales, marketing and administrative roles and that a cloud provider may choose to buy or lease data centres, servers and networking equipment;[512]

(b) Another said that costs would include sales, marketing (though noting marketing could rely on corporate resources), and access to data centre infrastructure;[513]

(c) A different provider said that there is a multitude of UK-specific investments required to compete effectively, and that it has invested in sales force, ecosystem network, marketing, UK-based data centres and business managers.[514]

4.88 Although three providers said that they tailor certain aspects on narrower geographical basis, where appropriate, most cloud providers set their pricing, advertising and marketing strategies globally.[515]

Evidence from customers

4.89 We asked large customers to rate the importance of a list of factors their organisation considers when choosing their main public cloud. In answering this question, customers rated the importance of (a) the number and location of datacentres and (b) data sovereignty, alongside other factors. They were asked to rate these factors from one to five, with one being not important at all, and five being very important.

4.90 Number and location of data centres [✂] and data sovereignty [✂] were identified as some of most important factors by customers we spoke to when choosing their

[509] Responses to the CMA's information requests [✂].
[510] [✂] response to the CMA's information request [✂].
[511] Responses to the CMA's information requests [✂].
[512] [✂] response to the CMA's information request [✂].
[513] [✂] response to the CMA's information request [✂].
[514] [✂] response to the CMA's information request [✂].
[515] Responses to the CMA's information requests [✂].

main public cloud provider attaining an average rating that round four (see paragraph 2.83)

4.91 In relation to the number and location of data centres, the majority of customers considered this to be important or very important.

(a) A handful of customers said that it was important that cloud providers were able to cover their operating regions, both in the UK and more broadly;[516]

(b) A few said that a European presence was important[517] and one customer said that its preference is to host personal data in the UK or a country deemed to be "adequate" by the UK Government, eg EU countries;[518] and

(c) A few said that they require data centres in the UK.[519] One said that this was due to GDPR implications.[520]

4.92 In relation to data sovereignty, the majority of customers considered this to be important or very important.

(a) Most of these customers identified the need for regulatory compliance, data protection and/or security.[521]

(b) One customer specified that having a cloud provider that operates in multiple regions of the world means that it can meet data sovereignty requirements within its regions of business operation.[522]

4.93 As set out above, we asked large customers to rate the suitability of a list of public cloud providers as alternatives to their main public cloud provider based on their perception or any direct experience. They were asked to rate these providers from one to five, where one was not effective as an alternative and five was fully effective as an alternative. Customers could also identify and rate public clouds not listed by the CMA in their response.

4.94 Alibaba is heavily concentrated in China[523] and it was not identified as an effective alternative to UK customers' main providers by respondents. In particular, Alibaba received an average rating that rounded to 1 and was not identified by any respondent as a suitable alternative.

---

[516] Responses to the CMA's information requests [✂].
[517] Responses to the CMA's information requests [✂].
[518] [✂] response to the CMA's information request [✂].
[519] Responses to the CMA's information requests [✂].
[520] [✂] response to the CMA's information request [✂].
[521] Eg responses to the CMA's information requests [✂].
[522] [✂] response to the CMA's information request [✂].
[523] Alibaba Cloud Global Locations | Deploy Around the World Including Mainland China, accessed 20 May 2024.

4.95 Reasons that customers gave for these ratings or not providing a rating included a lack of experience or knowledge of Alibaba's offering,[524] that using Alibaba would be a supply chain risk or not appropriate (eg for data sovereignty or security reasons,[525] that Alibaba has more limited services or capabilities compared to larger providers[526] and that they are only considered for demand relating to China.[527]

Emerging views on the geographic market

4.96 Overall, the evidence we have seen to date suggests that the relevant markets are wider than the UK, but not as wide as global for the following reasons.

4.97 First, it suggests that the markets are wider than the UK because:

(a) customers can theoretically choose datacentres globally and do choose datacentres outside of the UK;

(b) some customers identified that having datacentres across their operating regions was an important factor when selecting a public cloud provider;

(c) the main cloud providers to UK customers (AWS, Microsoft and Google) are active globally and set their strategies globally; and

(d) most cloud providers said that UK infrastructure was not necessary to compete effectively for UK customers.

4.98 Second, some of the evidence suggest that it is not as wide as global because:

(a) certain customers require UK datacentres for regulatory or security purposes;

(b) customers may prefer data centres that are located relatively close to reduce latency;

(c) customers rated the number and location of datacentres and data sovereignty requirements were rated as important factors in selecting a cloud provider; and

(d) Alibaba, a Chinese provider, was universally rated as an ineffective alternative by customers.

4.99 Based on the evidence we have seen to date, our emerging view is that the geographic scope of the markets is more likely to be Europe-wide (ie UK and EEA). To the extent relevant, we will take into account non-European providers as

[524] Responses to the CMA's information requests [✂].
[525] Responses to the CMA's information requests [✂].
[526] Responses to the CMA's information requests [✂].
[527] Responses to the CMA's information requests [✂].

out-of-market constraints. We will also take into account the way in which providers' strength in UK datacentres may be important in competing for some customers.

5. Market shares

5.1 In this section we consider the structure of the markets based on our shares of supply analysis. We have calculated shares of supply using various metrics to give an overall picture of the market structures and an indication of how those structures are likely to evolve over time.

Framework for our assessment

5.2 In a market investigation, the calculation of market shares of the suppliers of the reference products provides useful background data for the assessment of the levels of firms' market power and may be relevant to our assessment of the theories of harm.[528] For example:

(a) High market shares can be a sign that a firm faces weak constraints from rivals. If its market shares have been stable or even increasing over time, especially in the face of demand or supply shocks, this could indicate that the firm has market power and that market outcomes are worse than they could be.[529]

(b) Certain practices are more likely to be of concern when a firm has market power. The CMA's guidance sets out how market power may be considered in relation to certain vertical relationships. For example, the guidance says that if a vertically integrated firm has significant market power in an upstream market, it may have an incentive to refuse access to the input or to raise its prices, and consequently increase the costs of competing downstream firms.[530] More generally if market concentration is increasing, this may be evidence that the certain practices are restricting, preventing or distorting competition.

5.3 As we noted in our Issues Statement,[531] Ofcom was concerned that the cloud market could become more concentrated over time, allowing the market leaders to entrench their position and avoid competing vigorously in the future.[532] For this reason, among others outlined below, we have taken a particular focus on forward-looking share of supply metrics.

5.4 In our analysis we have calculated shares of supply based on three different metrics: (i) shares by revenue; (ii) shares by capacity; and (iii) shares based on the

[528] CC3, paragraph 100.
[529] Ibid., paragraph 187.
[530] Ibid., paragraph 268-270.
[531] Issues statement (publishing.service.gov.uk), paragraph 16.
[532] Ofcom Cloud Services Market Study Final Report, paragraph 1.1

flow of new business. In the rest of this section we set out our methodology and initial results for each of these measures of shares of supply in turn.

5.5 We have largely gathered data on a UK basis, which is informative of providers' success in competing for UK customers and therefore useful in assessing their strength in the market. However, given our initial view that the geographic markets are wider than the UK, we plan to gather data on a wider basis in due course.

Shares of supply by revenue

5.6 Shares by revenue are typically the most direct measure of the distribution of customer demand in a market as they take account of differences in the prices and quality of firms' offerings.[533] High and stable or increasing shares of supply can be a strong indicator of market power, although they should be considered alongside other indicators, such as:

(a) High barriers to entry and expansion. Where barriers to entry and expansion are high, larger firms are less constrained from exercising market power by the threat of competition from smaller firms and entrants. We consider the extent to which there are barriers to entry and expansion in cloud services in section 4 of this working paper;

(b) high profitability. Absent market power, rivals should typically be able to enter and compete for high profits until they have been reduced to competitive levels. We consider profitability in section 6 below; and

(c) high barriers to switching. Where barriers to switching are high customers' ability to exercise choice and move to suppliers with better offerings is limited, giving rise to market power. We expect to consider barriers to switching further in subsequent working papers.

Our approach

5.7 We have calculated shares of supply by revenue using the annual UK revenue[534] data from AWS, Microsoft, Google, Oracle, and IBM.

5.8 We asked providers to supply revenues by individual cloud services or service categories. We then mapped these into IaaS, PaaS, and SaaS and calculated totals for IaaS, PaaS, and for IaaS and PaaS combined.[535] We used Ofcom's approach to mapping individual cloud services or service categories into IaaS,

[533] CC3, Annex A, paragraph 2.
[534] For the purposes of this analysis, we defined UK Revenues as revenues generated from UK Customers in the UK, and we defined UK Customers as Public Cloud Infrastructure Service customers that are operating or trading in the UK. We defined Annual Revenues as revenues generated within a calendar year.
[535] We converted AWS, Microsoft, Google, and IBM revenue data from USD to GBP using the Bank of England's average annual exchange rates for each year respectively.

PaaS and SaaS and have so far received no evidence to suggest this was incorrect.[536]

5.9 In relation to the data provided directly to us by providers, we note that:

(a) One provider said it records relevant data in several different systems, some of which may not readily reconcile with each other. The cloud provider has therefore made certain adjustments to reconcile data across these systems. [✂] and we have adjusted the IaaS and PaaS totals that we calculated for based on the individual revenues by cloud service from one system such that their combined total matches the aggregate revenues from another system. This had the effect of increasing that provider's shares by less than [✂] pp in each category.

(b) Oracle said that it was unable to provide annual revenue data based on calendar years and therefore it provided data according to its fiscal year, which ends on 31 May.[537] This means that Oracle's shares in any given year could be slightly overestimated or underestimated.

(c) IBM and Oracle did not provide revenue data segmented by individual cloud services. Oracle provided revenues grouped into service categories rather than segmented by individual cloud services – these service categories appear to follow broadly the mapping into IaaS and PaaS that we applied elsewhere.[538] [✂].[539] Given the shares of Oracle and IBM set out below, any differences in the categorisation of their cloud services are unlikely to have a material effect on the final shares.

(d) The revenue data did not align with our definition of UK revenues, ie revenues generated from customers that are operating or trading in the UK.

(i) One provider identified UK revenues as revenues generated by customers with a UK tax address, billing address, or customer address associated with the Account ID.[540]

(ii) One provider identified UK revenue as revenues that were billed to the customers' billing address(es) in GB.[541]

---

[536] One provider said it was concerned that categorising cloud services in this way distorts the assessment of the competitive dynamic in cloud computing because these categories will not be applied consistently due to the lack of a common definition and it is likely to exclude alternatives such a traditional on-premises solutions, dedicated data centres, and collocated data centres that compete for the same customers. In relation to the first of these points we have not seen evidence to suggest there are large discrepancies such that this would distort our overall share of supply estimates, although we have estimated shares based on IaaS and PaaS combined which will control for this to some extent. In relation to the second point, we consider the extent to which such alternatives are part of the relevant markets above.
[537] Oracle [✂].
[538] Oracle response to the CMA's information request [✂].
[539] IBM response to the CMA's information request [✂].
[540] [✂] response to the CMA's information request [✂].
[541] [✂] response to the CMA's information request [✂].

(iii) Three providers identified UK revenues as generated by customers with billing addresses in the UK.[542]

5.10 For the total revenue across all providers (ie the denominator for our shares) we used the market size estimates for IaaS, PaaS, and IaaS and PaaS combined that Ofcom published in its cloud market study final report.[543] These were based on an average of estimates from two third-party data providers, IDC and Synergy, adjusted according to the first-hand revenue data that Ofcom collected.[544]

5.11 We are exploring collecting additional data to calculate our shares of supply by revenue that will both allow us to extend the estimates to 2023 and verify the market size.

5.12 AWS was the first provider to supply cloud services in 2006[545] with Microsoft entering in 2010,[546] and Google in 2011.[547] As first entrant, AWS historically had a very high share which has declined over time as others entered and providers competed for customers migrating to public cloud for the first time.

5.13 In recent years the markets have been growing significantly, for example, in the UK IaaS and PaaS revenues have more than doubled from £[2.5-3.0] billion in 2019 to £[7.0-7.5] billion in 2022. This has been accompanied by significant increases in capacity by cloud providers (see Figure 2.1 above).

5.14 The analysis set out here has focused on how the relative position of cloud providers has evolved during this recent period of growth and how their positions may be expected to change in the near future.

UK shares of supply for IaaS

5.15 The evidence we have seen to date shows that the IaaS segment is highly concentrated, and concentration is increasing over time.

(a) AWS and Microsoft's combined share of supply is increasing steadily year on year while the shares of smaller cloud providers are generally declining overall.

(b) Of the smaller cloud providers Google holds the largest share but it is [at most a quarter] of the size of AWS and if the current rate at which its share of

[542] [✂] response to the CMA's information request [✂]; [✂], [✂].
[543] Cloud services market study final report (ofcom.org.uk), paragraph 4.93.
[544] [✂].
[545] About AWS (archive.org), accessed 10 February 2024.
[546] Windows Azure Platform Now Generally Available in 21 Countries | Blog Azure | Microsoft Azure, accessed 11 February 2024.
[547] The History of Google Cloud Platform (pluralsight.com), accessed 11 February 2024.

supply is increasing continues it will be a long time before it catches up with the AWS and Microsoft.

5.16 Our analysis also shows that:

(a) AWS is the largest provider of IaaS, although its share has declined recently from [40-50]% in 2019 to [40-50]% in 2022;

(b) Microsoft is the second largest provider of IaaS and its share has increased from [30-40]% in 2019 to [30-40]% in 2022;

(c) Google is the third largest IaaS provider and has a significantly lower share of IaaS revenues than AWS and Microsoft: it held just a [5-10]% share in 2022 which was [✂] higher than its share of [5-10]% in 2020;

(d) For IBM and Oracle, shares have remained in the [0-5]% range from 2019 to 2022.[548]

5.17 While the market is growing overall (and we consider evidence on where new revenue is generated from below) and this may provide opportunities for smaller providers as there is still new business to be won, this figure also shows that concentration in IaaS is increasing with the combined share of AWS and Microsoft increasing over time from [✂] in 2019 to [✂] in 2022. In addition, Microsoft has narrowed the gap to AWS with the difference in shares falling from [✂]pp in 2019 to [✂]pp in 2022.

UK shares of supply for PaaS

5.18 The PaaS segment is less concentrated than IaaS, but concentration is also increasing over time:

(a) AWS and Microsoft are both relatively smaller in PaaS than IaaS while Google is relatively larger.

(b) Google is much smaller than AWS or Microsoft and growing at a slower rate.

(c) The combined share of all other smaller cloud providers and ISVs has declined significantly in recent years.

5.19 Our analysis also shows that:

(a) AWS is the largest provider of PaaS, although its share has increased slightly from [20-30]% in 2019 to [20-30]% in 2022;

[548] Source: CMA analysis of first-party revenue data and market values published by Ofcom. Responses to the CMA's information requests [✂].

(b) Microsoft is the second largest provider of PaaS and its share has increased from [20-30]% in 2019 to [20-30]% in 2022;

(c) Google is larger in PaaS than IaaS and is exhibiting stronger growth: its share has increased from [5-10]% in 2020 to [5-10]% in 2022;

(d) For IBM and Oracle, shares, have remained in the [0-5]% range from 2019 to 2022.[549]

5.20 Our analysis also shows that concentration in PaaS is increasing with the combined share of AWS and Microsoft increasing over time from [40-50]% in 2019 to [50-60]% in 2022.

UK shares of supply for IaaS and PaaS combined

5.21 The overall cloud services sector is concentrated, and concentration is increasing over time. Google is much smaller than AWS or Microsoft and growing at a slower rate. The combined shares of all other smaller cloud providers and ISVs are also, in aggregate, falling year on year.

5.22 Our analysis shows the shares of supply in cloud services by revenue – that is IaaS and PaaS in combination. Our analysis shows that:

(a) AWS is the largest provider of IaaS and PaaS and it share has remained broadly stable: its share was [30-40]% in 2019 and [30-40]% in 2022;

(b) Microsoft is the second largest provider of IaaS and PaaS and its share has increased from [20-30]% in 2019 to [30-40]% in 2022 as it gains ground on AWS;

(c) Google is the third largest provider of IaaS and PaaS and has a significantly lower share than AWS and Microsoft: Google is [at least a quarter] of the size of AWS and has [✂] increased its share from [5-10]% in 2020 to [(5-10]% in 2022;

(d) For IBM and Oracle, shares have remained in the [0-5]% range from 2019 to 2022.[550]

5.23 The analysis also shows that concentration in IaaS and PaaS is increasing, with the combined share of AWS and Microsoft increasing over time from [60-70]% in 2019 to [70-80]% in 2022.

[549] Source: CMA analysis of first-party revenue data and market values published by Ofcom. Responses to the CMA's information requests [✂].
[550] Source: CMA analysis of first-party revenue data and market values published by Ofcom. Responses to the CMA's information requests [✂].

5.24 Conversely, the combined share of the other smaller cloud providers and ISVs has fallen from [20-30]% in 2020 to [10-20]% in 2022, which is further evidence that the markets are concentrating.

Shares of supply by capacity

5.25 Shares by capacity look at the market structure from the supply side. At a basic level they show us the relative strength of each provider in terms of their production capability. Absent barriers to competition or switching, the greater a firm's capacity the greater the constraint it can impose on rivals by competing for business.

5.26 In markets with high barriers to entry and expansion, firms with high shares of overall capacity may be less constrained by firms with smaller shares of overall capacity, because the latter will tend to reach their capacity – and the point at which they experience a diminished ability and incentive to win new demand – more quickly.

5.27 In this sense, capacity shares provide useful context to shares by revenue in that a firm with a high revenue share may be competitively constrained by a firm with a substantial share of capacity. To assess this, it may be appropriate to consider shares of spare capacity (ie the level of capacity not currently being used) as they give an indication of the strength of firms' incentives to compete. Specifically, a firm with a higher share of spare capacity is likely to have a greater ability and incentive to compete for more demand to utilise that capacity. In contrast, a firm with a low share of spare capacity with have a lower ability and incentive to compete.

5.28 Forward-looking shares of future capacity based on firms' investment and expansion plans may tend to reflect firms' own beliefs about their prospects for growth and future competitive significance, particularly where expansions in capacity involve significant sunk costs.

Our approach

5.29 We calculated the shares of supply by capacity using data from AWS, Microsoft, Google, IBM, and Oracle on their datacentre capacity in megawatts (MW) within UK+EEA,[551] globally, and in the UK. We note, therefore, that these shares do not include the capacity of other smaller IaaS providers and as such each provider's share is an overestimate across all providers and should be interpreted as an indicator of relative share between the largest cloud providers. Subject to

[551] For the purpose of this analysis, we defined the Europe region as the EEA plus the UK.

availability, we plan to gather datacentre capacity data from other IaaS providers that serve UK customers.

5.30 We gathered these data directly from the cloud providers and requested this capacity data in MW based on our understanding that this is the industry standard metric.

5.31 In relation to this data provided directly to us by providers it is important to note that:

(a) AWS provided datacentre capacity within the EMEA region rather than the EEA.[552] According to AWS' website, it has some datacentre capacity in South Africa as well as in Israel, Bahrain, and the UAE.[553] Therefore, the shares of European capacity that we present for AWS are slight overestimates while the shares of the other cloud providers are slight underestimates.

(b) Google provided [✂].

(c) Google provided [✂].[554] [✂].[555] [✂].

(d) Oracle provided data on its datacentre capacity in the UK and EEA going back to 2021 and global capacity going back to 2020.[556] To ensure that the missing data point for Oracle would not artificially inflate the shares of all other cloud providers in 2020, we included an approximation of what Oracle's capacity in the UK and Europe (ie UK + EEA) would have been in 2020 if Oracle's growth rate from 2020 to 2021 were in line with its average growth rate implied by its capacity data and projections 2021 to 2026. We did this by calculating the average year-on-year capacity growth rate over the years 2021-2026 in the UK and EEA respectively and divided the 2021 capacity figures by these average growth rates.[557]

5.32 We believe that the European shares give the best reflection of the market structure with respect to UK customers.

UK and European shares of supply by datacentre capacity

5.33 UK and European Economic Area (EEA) shares of supply by capacity appear to be consistent with other evidence about the market structure: AWS and Microsoft are the largest providers and Microsoft's share has been increasing relative to

[552] AWS' response to the CMA's information request [✂].
[553] Global Infrastructure - AWS (amazon.com), accessed 20 May 2024.
[554] Google's response to the CMA's information request [✂]; Google said that [✂].
[555] Google's response to the CMA's information request [✂].
[556] Oracle's response to the CMA's information request [✂].
[557] [✂].

AWS' share. However, in contrast to revenue shares Microsoft was the largest provider in the UK and EEA based on capacity in 2022. Smaller cloud providers do not have a similar scale. Google's share of capacity is significantly higher than IBM and Oracle's [✂].[558] [559]

Global shares of supply by datacentre capacity

5.34 Global shares of supply by capacity indicate that AWS and Microsoft have the largest capacity, followed by Google and then by Oracle and IBM. These positions appear unlikely to change in the next few years.[560]

UK shares of supply by datacentre capacity

5.35 UK shares of supply by capacity suggest that Microsoft is by far the largest cloud provider in the UK, followed by AWS.[561,562]

5.36 Some inconsistency between these shares and the shares of supply by revenue suggests that cloud providers serve UK customers using overseas datacentres.

Shares of supply by flows of new business

5.37 In markets characterised (or potentially characterised) by switching costs, shares of supply based on the 'stock' of customers (eg shares by 'installed base' or share of revenues, where revenues tend to be 'recurring' and have relatively low churn) may not reflect recent changes in the relative competitive position of suppliers. In such cases, it is useful to consider evidence on shares of supply on a 'flow' basis (eg shares of new customers or new revenues).

5.38 In this section we present the following shares based on the flow of new business:

(a) Shares by new customers acquired – this involves calculating the number of customers that each provider acquired as a proportion of the total customers acquired in each year.

---

[558] [✂].
[559] CMA analysis of datacentre capacity in megawatts submitted by the cloud providers. Responses to the CMA's information requests [✂].
[560] CMA analysis of datacentre capacity in megawatts submitted by the cloud providers. Responses to the CMA's information requests [✂].
[561] Based on the inconsistency between these shares and the shares of supply by revenue, as well as other evidence set out in Section 4 on market definition, we do not consider that these shares give an accurate representation of the structure of the market
[562] CMA analysis of datacentre capacity in megawatts submitted by the cloud providers. Responses to the CMA's information requests [✂].

(b) Shares by revenue from newly acquired customers – this involves calculating the revenues that each provider earned from customers acquired in each year as a proportion of the total revenues from new customers.

(c) Shares by overall revenue growth – this involves calculating the year-on-year revenue growth of each provider as a proportion of the total revenue growth.

(d) Shares by new revenues from existing customers – this involves calculating the growth in spend from each provider's existing customers as a proportion of the total growth spend from existing customers. We calculated the growth in spend from existing customers by subtracting each provider's revenue from new customers from its overall year-on-year revenue growth in each year.

5.39 The data available at present is at an aggregate level – this has important implications for interpreting the analysis as it means we cannot distinguish between the following:

(a) If the new customers a provider acquires are: (i) customers completely new to the cloud (representing competition for customers); (ii) customers that are only new to that provider and placing a new workload (representing competition for new workloads); or (iii) customers that are only new to the that provider and switching an existing workload (representing competition for existing workloads).

(b) If changes in a provider's revenue from existing customers is caused by: (i) some existing customers decreasing/increasing their spend on existing workloads without switching (eg cost optimisation, business expansion); (ii) some existing customers switching existing workloads to or from another cloud provider (representing competition for existing workloads); or (iii) some existing customers placing new workloads with that provider.

5.40 We still consider the analysis is informative in understanding in general terms the relative importance of revenue growth from existing and new customers and understanding which providers are gaining new customers. However, it means that we cannot identify the drivers of changes. We welcome views on the value of this evidence.

Our approach

5.41 We calculated the various shares by business flows using data from AWS, Microsoft, Google, Oracle, and IBM on the number of new UK customers they acquired in each of the calendar years 2020-2023, the revenues generated from those new customers, and their overall annual UK revenues. We note that these

shares are between these cloud providers only and do not include other cloud providers active in the UK.

5.42 In relation to this data each cloud provider submitted data on a slightly different basis and in some cases it may have captured customers who were only using a provider on a trial basis such that we plan to test the sensitivity of the results using different thresholds:

(a) AWS provided data based on defining a newly acquired customer as one that spent at least $100 with it for the first time in a year.[563]

(b) Microsoft provided data based on defining a newly acquired customer as one that spent at least $100 on its Public Cloud Infrastructure Services for the first time in a year.[564]

(c) Google provided data based on defining a newly acquired customer as [✂].[565]

(d) Oracle defined a new customer as one that appeared in its revenue data for the first time in a year.[566]

(e) IBM defined a new customer [✂].[567]

5.43 One cloud provider provided data on UK customer flows according to [✂].[568] However, the cloud provider provided data on UK revenues by [✂].[569] While these definitions are not entirely consistent, they are unlikely to have a material effect on the analysis.

5.44 We currently only have the necessary data to present shares by new business flows for the years 2021 and 2022. We expect to extend this analysis to include 2023 data after the working paper.

5.45 In the remainder of this sub-section we first set out shares by year-on-year revenue growth. We then consider the extent to which this split between new customers and existing customers, as this provides important context for the following shares we set out based on newly acquired customers, revenue from newly acquired customers and revenue from existing customers. For example, if new customers make up a significant proportion of year-on-year growth then competition is more likely to be determined by those new customers (although as

[563] AWS response to the CMA's information request [✂].
[564] Microsoft response to the CMA's information request [✂].
[565] Google response to the CMA's information request [✂].
[566] Oracle response to the CMA's information request [✂].
[567] IBM response to the CMA's information request [✂].
[568] [✂] response to the CMA's information request [✂].
[569] [✂] response to the CMA's information request [✂].

outlined above the data does not allow us to fully disentangle the nature of that competition).

Shares of supply by year-on-year UK IaaS and PaaS revenue growth

5.46 Shares of supply by year-on-year revenue growth show that AWS and Microsoft are the fastest growing cloud providers. Google's growth is consistent with its position in the market.

5.47 We have analysed UK shares of supply by year-on-year revenue growth of the largest UK cloud providers. Our analysis shows that:

(a) Microsoft has become the fastest growing cloud provider as of 2022: its share of revenue growth increased from [30-40]% in 2021 to [40-50]% in 2022;

(b) AWS' share of overall revenue growth [✂] over 2021 to 2022: it won [40-50]% and [40-50]% of overall new revenues in 2021 and 2022, respectively;

(c) Google's share of overall growth fell slightly over 2021 to 2022 from [10-20]% to [10-20]%;

(d) For IBM and Oracle, shares of growth have remained in the 0-5% range.[570]

5.48 When looking at the split of year-on-year revenue growth between new and existing customers it can be seen that the vast majority of year-on-year revenue growth across these cloud providers is driven by existing customers. In particular, across all providers existing customers accounted for [✂]% of year-on-year growth in the UK in 2022 and [✂]% of year-on-year growth in 2021.

5.49 This means that new customers make up a relatively small proportion of year-on-year growth and thus competition is less likely to be driven by new customers. Rather provider level revenue growth is occurring either because existing customers are expanding their existing workloads, adding new workloads or switching workloads between cloud providers they already use. We will consider this alongside any further evidence we gather, for example, on the prevalence of switching to understand what is driving this revenue growth.

Shares of supply by newly acquired UK customers

5.50 Shares of supply by newly acquired customers show that Microsoft is winning customers at a significantly higher rate than other cloud providers. It won [more

[570] CMA analysis of revenue and customer acquisition data from Microsoft, AWS, Google, Oracle, and IBM. Responses to the CMA's information requests [✂].

than twice] as many customers in each of 2021 and 2022 as AWS, and [at least three times] as many as Google.

5.51 We have analysed the shares of supply by newly acquired UK customers. Our analysis shows that:

(a) Microsoft has won more than [60-70]% of new customers in each of 2021 and 2022;

(b) AWS' share of new customers has been around [20-30]%.

(c) Google's share of new customers is consistent with its other shares of supply presented above: it won [10-20]% of new customers in 2021 and [5-10]% in 2022;

(d) Oracle and IBM's share of new customers remained in the 0-5% range in 2021 and 2022.[571]

Shares of supply by revenue from newly acquired UK cloud services customers

5.52 Shares of supply by revenue from new customers show that Microsoft is winning significantly more completely new business than other providers, while Google and Oracle also appear to be winning relatively large shares. Relative to its overall position in the market, AWS' share of revenue from new customers is low.

5.53 We have analysed the shares of supply by newly acquired UK customers. Our analysis shows that:

(a) Microsoft earned [50-60]% and [60-70]% of revenues from new customers in 2021 and 2022, respectively;

(b) AWS' position here is relatively weaker than in the other shares presented above: it earned [10-20]% and [10-20]% of revenues from new customers in 2021 and 2022, respectively;

(c) Google appears to be relatively strong in winning new business: it held a [10-20]% share in 2021 and [5-10]% in 2022;

(d) Oracle and IBM's shares ranged from [0-5]% to [10-20]% during this time period.[572]

[571] CMA analysis of revenue and customer acquisition data from Microsoft, AWS, Google, Oracle, and IBM. Responses to the CMA's information requests [✂].
[572] CMA analysis of revenue and customer acquisition data from Microsoft, AWS, Google, Oracle, and IBM. Responses to the CMA's information requests [✂].

Shares of supply by increased spend from existing UK cloud services customers

5.54 Shares of supply by increased spend from existing customers show that in 2021 AWS' customers were increasing their cloud spend year-on-year at a higher rate than other cloud providers' customers but the gap between Microsoft and AWS closed in 2022.

5.55 We have analysed the shares of supply by new revenues from existing customers by provider. Our analysis has shown that:

(a) AWS' share of increased spend from existing customers [✂] from [40-50]% in 2021 to [40-50]% in 2022;

(b) Microsoft's share increased from [30-40]% in 2021 to [40-50]% in 2022;

(c) Google's share of increased spend from existing customers fell slightly from [10-20]% to [10-20]% in 2022;

(d) Oracle's share of new revenue from existing customers remained in the [0-5]% range in 2021 and 2022.[573]

[573] CMA analysis of revenue and customer acquisition data from Microsoft, AWS, Google, Oracle, and IBM. Responses to the CMA's information requests [✂].

6. Market outcomes

6.1 In a market investigation the CMA will normally consider outcomes of the competitive process such as prices, innovation, quality and profitability. This is because outcomes of the competitive process can provide evidence about the functioning of a market.[574]

6.2 Evaluating these outcomes helps the CMA determine whether there is an adverse effect on competition and, if so, the extent to which customers may be harmed by it, ie the degree and nature of 'customer detriment'.[575] In particular, considering whether market outcomes are in line with what we may expect in a well-functioning market[576] is informative of whether competition is working well.

6.3 In this section we set out our initial analysis in relation to the following market outcomes in the supply of public cloud infrastructure services: prices, quality, innovation, and profitability.

6.4 In assessing these market outcomes, we consider the following:

(a) Current outcomes may not reflect future outcomes;

(i) We have seen that the supply of cloud services has evolved over time. When AWS launched its first service in 2006 it was the only supplier and was seeking to attract customers to public cloud for the first time (eg from traditional IT solutions). Over time, others such as Google and Microsoft entered and cloud services grew as cloud providers competed to attract customers to the public cloud for the first time.

(ii) Evidence suggests that the number of customers moving to public cloud for the first time has decreased and in 2021 and 2022 new customers made up a small proportion of year-on-year growth. In particular, based on data we have gathered so far, although cloud services is still showing strong growth, growth in these years has mainly been from existing customers rather than those migrating to the public cloud for the first time.

(iii) This suggests that over time the focus of competition may be expected to shift towards existing customers and either their existing workloads on the public cloud or workloads they are considering migrating to the

[574] CC3, paragraph 103.
[575] Ibid., paragraph 103.
[576] The CMA uses a well-functioning market as its benchmark in market investigations. A well-functioning market is one that displays the beneficial aspects of competition, rather than an idealised, perfectly competitive market. Generally the well-functioning market is the market envisioned without the features, but there may sometimes be reasons to depart from that general concept, for example, if features are intrinsic to the market but nevertheless have anticompetitive effects (as in the case of a natural monopoly) or if the nature of competition in the market is defined by arrangements put in place by Government. See CC3, paragraphs 30 and 320.

public cloud from alternative IT solutions – with the latter category potentially also shrinking in relative importance as more workloads are migrated.

(iv) We have also seen evidence that there are barriers to entry and expansion in relation to public cloud services and that there is a high level of market concentration which has increased in recent years and some evidence that this trend is likely to continue.

(v) In this context, current providers with large shares of supply are likely to have a stronger incentive, than they will have in the future, to compete for customers new to the public cloud and workloads that are more generally new to public cloud in order to maintain a strong position in the future.

(vi) This means that future competition may look quite different to current competition. This is likely to be particularly the case if existing customers face barriers to switching existing workloads or using multiple clouds (which affects their placement of new workloads) as cloud providers' incentives to compete will reduce as the proportion of demand from existing customers increases.

(b) For at least some indicators, there is no clear counterfactual to compare outcomes with what they might be in a well-functioning market. For example, it is particularly difficult to identify a counterfactual when considering indicators showing levels of innovation or switching. This may make it difficult to evaluate such market outcomes in relation to whether or not they indicate a competition issue.

(c) Some outcomes are difficult to observe and measure. To the extent this is the case it may be difficult to assess these outcomes.

(d) Finally, to put our profitability assessment into the wider context of our overall assessment of the market investigation, profitability will be one of the outcomes of the competitive process we will be taking into account. We note that while informative, findings that price-cost margins are wide or profitability is high in a market do not on their own provide conclusive evidence that the market could be more competitive; such findings are not in themselves causes of competitive harm – ie they are not features of the market for the purpose of the AEC test.[577]

[577] CC3, paragraph 126.

Pricing, quality and innovation

Pricing trends

6.5 Prices tend to be observable and measurable outcomes and an analysis of prices may be useful in assessing the extent and nature of competition. Prices can be analysed in various different ways depending on the theories of harm being considered and where relevant we are considering pricing in the context of individual theories of harm.

6.6 As part of our work on market outcomes we have sought to understand the trends in prices over time among the largest cloud providers and whether that can tell us anything about competition in cloud services. In doing this we note that prices can be driven by both competitive constraints and other factors such as costs and changes in quality. This means any analysis of price alone is only indicative and needs to be considered alongside other evidence such as costs and this is why we also consider profitability below.

6.7 Pricing trends can be informative in that a trend of increasing prices over time in a sector may be consistent with a competition concern subject to other supporting evidence. For example, it may not be profitable for a firm to increase or even hold prices stable over time when competition is working effectively, as we might expect rivals to undercut incumbent firms in a bid to win customers. If a firm is maintaining share and increasing prices (or holding prices constant while costs are falling, for example, due to economies of scale) over time then this could suggest that that firm has market power.

6.8 We can also compare pricing trends across providers. If large providers either have materially higher prices than smaller providers or if large providers' prices are trending upwards when smaller providers' prices are not, then this may indicate that large providers are able to exert some degree of market power such that there is a competition concern.[578] Diverging pricing trends between larger and smaller providers suggest that any increase in prices by larger providers is not an effect of supply or demand shocks that affect the market as a whole.

6.9 During its market study Ofcom assessed analysis submitted by both AWS and Microsoft on their prices. This included separate analysis by both AWS and Microsoft of their list and average net of discount prices, an analysis by Microsoft of the change in the list prices of new cloud services and an analysis by Microsoft of the proportion of revenue associated with products which recorded,

[578] This would be subject to other evidence such as if those price differences were driven by quality differentials.

respectively, no price changes, price increases, price decreases, or both price increases and decreases over 2016-2023.[579]

6.10 We have also received an analysis of pricing by Microsoft the results of which it said 'show that pricing trends are inconsistent with customer lock-in and are instead consistent with strong customer choice.' It added that, in its view, 'pricing trends support the reality that multicloud is predominant and that the ability for customers to switch results in competitive and decreasing real prices over time.'[580] We are currently considering this submission in detail including the nature of the analysis and the assumptions used.

6.11 We have requested pricing data, both list and net prices, from cloud providers for their top 20 cloud services based on revenue for each year since 2018. This is in order to undertake a more detailed and UK-specific assessment of net and list prices and how they have changed over time when compared to the analysis submitted to Ofcom. We also plan to extend this analysis beyond AWS and Microsoft to include Google.

Quality and innovation

6.12 As part of its assessment of market outcomes in past market investigations, the CMA has considered the extent to which there might be factors which lead to less innovation and evidence of quality being lower than what might be expected in a well-functioning market.[581]

6.13 As set out in Section 4, we have received evidence that cloud services are chosen over traditional IT services as they have different characteristics such as advanced functionality, elasticity,[582] scalability, flexibility and resiliency. These characteristics suggest that cloud services are therefore of a higher quality and more innovative, at least in these areas, than traditional IT services. In this regard, cloud providers have had an incentive to innovate in order to expand the market.

6.14 We have also received evidence, for example, in terms of the number of patents and new products which show that cloud providers are continuing to innovate over time. There is also likely to be further innovation in the future, for example, as cloud providers increase the use of AI based products in their offering.

6.15 While this evidence suggests that cloud services are in some ways of higher quality/more innovative than traditional IT services, this does not directly tell us about the quality of or level of innovation in cloud services themselves and

[579] Cloud services market study final report (ofcom.org.uk), paragraphs 4.125 to 4.140.
[580] [✂].
[581] For example, see Energy market investigation: Final report (publishing.service.gov.uk), paragraph 10.131.
[582] In this context elasticity refers to the ability to quickly increase or decrease the cloud computing resources being used.

whether the quality of these services or level of innovation reflects what may be expected in a well-functioning market.

6.16 To more directly consider quality and innovation in the markets we are considering, we have collected a range of evidence relating to quality and innovation[583] across cloud providers and are considering this evidence closely. This includes evidence on quality and customer satisfaction metrics, software updates, number and uptake of new products, patents and R&D investment.

6.17 More generally, and as set out in the following paragraphs, it is difficult to evaluate evidence on quality and innovation.

6.18 First, quality and innovation tend to be less readily quantifiable than prices and measures of quality and innovation are sometimes difficult to observe. Evidence about quality, in particular, tends to be subjective, coming particularly from surveys, questionnaires or discussions with customers, investors, or other market observers.[584] While more quantitative evidence on innovation may be available these are proxies for the level of innovation rather than actual measures of the level of innovation.[585]

6.19 Second, for quality and innovation there is no clear counterfactual to compare outcomes with what they might be in a well-functioning market. That is, it is not clear what the level of quality or innovation would be in a well-functioning market to assess if current quality / innovation is higher than, the same as or lower than that level of quality / innovation.

6.20 Subject to supporting evidence it may still be possible to infer something about how well competition is working. For example: changes in the quality of services over time could be an indicator of changes in the strength of competition, subject to other supporting evidence;[586] if we find that quality tends to be higher among smaller providers than the largest providers, then this could suggest that competition may be weak as market shares are not reflective of differences in quality;[587] and if we were to observe slowing rates of innovation over time this could indicate that providers lack the incentive to continue innovating substantially

[583] Microsoft submitted some analysis of innovation in cloud services that it had conducted and we are still considering this submission as part of our assessment of innovation in cloud services.
[584] See CC3, paragraph 127.
[585] For example, while the evidence we have gathered could tell us about the absolute numbers of software updates and patents, whether they appear high and if they differ across cloud providers, not all updates or patents are of the same relative importance such that absolute numbers do not tell us about the actual level of innovation and how they differ across providers. For instance, one provider may engage in lots of relatively small updates whereas another may engage in a few large updates. Overall the underlying nature of all these updates may be the same across the two providers, but this would not be apparent from data on the number of updates alone.
[586] For example, if a firm maintains market share and/or profitability while degrading the quality of its services, this suggests that either (i) there are no viable alternatives that customers can switch to, or (ii) customers face barriers to switching.
[587] This may be because customers are not able to access and/or assess information about their available alternatives, or that there are barriers to switching provider. See CC3, paragraphs 295 to 318.

to attract and retain customers and in turn that competition is not functioning effectively.

6.21 Moreover, while comparing quality metrics across cloud providers may allow us to assess relative quality levels (eg if they provide comparable metrics), it may not provide a full picture. For example, if competition is relatively weak, then cloud providers may still have an incentive to provide a certain level of quality to attract customers who are new to cloud services or new workloads from existing customers. However, that quality may be lower than what the quality would be if there was strong competition between providers. Similarly, even if existing customers were not able to switch or multi-cloud, cloud providers may still have an incentive to innovate to increase the opportunities for generating additional revenue from their cloud services.[588] As such the presence of innovation is unlikely on its own to imply strong competition.

6.22 As set out above, cloud providers have had an incentive to innovate to expand the market and there is evidence of ongoing innovation in cloud services which we expect to continue at least to some extent (eg as AI offerings expand). However, it is difficult to assess the exact level of quality and innovation in cloud services (both over time and currently), and we do not currently consider that the evidence we have reviewed to date allows us to accurately compare quality and innovation across cloud providers. Moreover, it is difficult to assess if the current levels of quality and innovation will persist and whether what we observe reflects what might be expected in a well-functioning market.

Profitability

6.23 Profitability tends to be a more observable and measurable outcome and an analysis of profitability may be useful in quantifying the extent and nature of competition by examining the outcomes of that market in terms of the financial performance of the participating firms.[589] Profitability can provide a more holistic approach to assessing outcomes as it takes into account not only prices, but also other factors such as quality and innovation and firms' relevant costs.

6.24 Firms in a competitive market would generally earn no more than a 'normal' rate of profit. We have sought to consider whether profits among certain providers reflect a 'normal' rate of return based on the nature of competition in the supply of public cloud infrastructure services.[590]

[588] For example, such innovations may involve developing additional features that can then be charged at a premium or products and services that do not overlap with existing products and services such that they generate incremental revenue. In this respect there are different types of innovation and each type may have different implications for competition.
[589] See CC3 paragraphs 104, 114 and 116.
[590] Ibid., 104, 114 and 116.

6.25 We do not regard 'excess' profitability in itself to be a problematic feature of a market, but instead a market outcome that provides an indication of limitations in the competitive process. Profitability findings may also be used in the context of determining the scale of the harm or detriment that might arise, for example in the form of higher prices.

6.26 In this section we set out the role of profitability analysis and summarise our preliminary assessment of profitability for cloud providers. We will provide additional details on our analysis in due course.

Role of profitability analysis

Purpose of profitability analysis

6.27 The aim of profitability analysis is to understand competitive conditions within a market, by examining the outcomes of that market in terms of the financial performance of the participating firms.

6.28 We consider that firms in a competitive market would generally earn no more than a 'normal' rate of profit – the minimum level of profits required to keep the factors of production in their current use in the long run, ie the rate of return on capital employed for a particular business activity would be equal to the opportunity cost of capital for that activity.[591]

6.29 The purpose of conducting profitability analysis, therefore, is to understand whether the levels of profitability (and therefore prices) achieved by cloud providers are consistent with the levels we might expect in a competitive market.

Interpretation of profitability analysis

6.30 In interpreting the results of our analysis, we take into account a number of factors. First, we recognise that at particular points in time the profitability of some firms may exceed what might be termed the 'normal' level. There could be several reasons for this, including cyclical factors, transitory price or other marketing initiatives, and some firms earning higher profits as a result of past innovation, or superior efficiency.[592]

6.31 Where firms have been generating profits above their cost of capital for a sustained period, this could indicate limitations in the competitive process.[593]

[591] Ibid., paragraph 116.
[592] Ibid., paragraph 117.
[593] Ibid., paragraph 118

6.32 On the other hand, a finding of low profitability does not necessarily signify that competition is working well, for example, low profitability may be concealing ineffective competition.[594]

6.33 The trend in profits over the period of review will be an important consideration as an indicator of improvements or deteriorations in the competitive environment. For example, where profitability has increased over a number of years, this may indicate a worsening of the competitive situation or weakening of competitive pressures in the relevant markets.[595]

Our assessment of profitability

6.34 We have looked at the relevant revenues, costs, and capital base of the main public cloud infrastructure services providers operating in the UK. We examine the profitability of cloud services for the largest cloud providers in the UK (AWS, Microsoft and Google) as well as financial information on smaller cloud providers who have been identified as global and UK competitors and where we have been able to obtain cloud services profit margin figures.[596]

6.35 We examined the global profitability of providers in our analysis due to (i) the global nature of the cloud services they provide, and (ii) the global nature of their financial reporting, asset base and capital investment.

6.36 We have analysed gross margins and earnings before interest and tax (EBIT) margins for cloud providers as indicators of financial performance.

6.37 We have also analysed and compared the return on capital employed (ROCE) for AWS and Microsoft's Azure and Cloud & Enterprise business segments, as these are the two largest providers in the UK market, to our estimate for the weighted average cost of capital (WACC) to assess their profitability. We do not include Google in our ROCE analysis as it has only recently (in 2023) reported profits[597] and we are primarily interested in assessing the profitability of the largest incumbent providers in the market.[598]

6.38 We compare the ROCE for AWS and Microsoft's cloud businesses to the WACC, to assess the extent to which these providers earn a 'normal' rate of profit. Where

[594] Ibid., paragraph 125
[595] Ibid., paragraph 124
[596] We have analysed AWS, Microsoft, Google, Oracle and IBM as we have identified these firms as relevant (and largest, based on cloud services revenues) providers of public cloud infrastructure services in the UK. We also analysed OVHcloud as they publicly report their public cloud business performance and are Europe-focussed, and we consider them to be a reasonable proxy for a mid-sized competitor operating in the UK.
[597] Alphabet 2023 Form 10-K, page 87.
[598] CC3, paragraph 114. As set out in our analysis of shares of supply in Section 5, Google has a significantly lower share of cloud services revenues compared to AWS and Microsoft.

firms persistently earn in excess of a normal return, this signals that there may be limitations in the competitive process.[599]

6.39 We benchmark margins as it provides useful context and insight into the comparative profitability of cloud providers, as well as trends in profitability over time, but note that this often has limitations when seeking to determine excess profits which makes other measures, such as ROCE, preferable where possible.

6.40 Below we summarise the emerging findings of our profitability analysis. The evidence that we have seen to date indicates that:

(a) Gross margins

(i) In the most recent financial year for which we have data, the three largest providers have all generated gross margins which are higher than the smaller providers for which we have been able to estimate relevant cloud segment margins.

(b) EBIT margins

(i) The EBIT margins for AWS have consistently been between 25% and 30% for the last eight financial years.[600]

(ii) The EBIT margins for Microsoft's Intelligent Cloud business segment have consistently been between 33% and 44% for the last seven financial years.[601]

(iii) Google Cloud became profitable in FY23 and is reporting growing EBIT margins (averaging a 9 percentage point improvement in margin in the last two financial years),[602] albeit the EBIT margins for Google Cloud are currently significantly lower than AWS and Microsoft's Intelligent Cloud business segment.

(c) ROCE analysis

(i) AWS ROCE has consistently been above our estimated WACC for the last nine years;

(ii) Microsoft Cloud & Enterprise ROCE has been above our estimated WACC for the last five years; and

[599] CC3, paragraphs 118 and 119.
[600] CMA analysis of Amazon Form 10-Ks.
[601] CMA analysis of Microsoft Form 10-Ks.
[602] CMA analysis of Alphabet Form 10-Ks.

(iii) Microsoft Azure ROCE has been above our estimated WACC since financial year 2021.

6.41 Our ROCE analysis shows that the ROCE for AWS and Microsoft Cloud & Enterprise ROCE has been falling slightly since financial year 2021, and Microsoft Azure ROCE was flat between financial years 2022 and 2023.[603]

6.42 The evidence that we have received to date suggests that these recent trends are, in large part, a result of increased levels of investment in cloud infrastructure, which are likely to be aimed at supporting the development of AI services.

6.43 This would not indicate a reduction in profitability as a result of competitive forces in the cloud services market, but instead would represent an upfront investment ahead of developments that would be expected to generate their own returns in the future.

6.44 Therefore, we do not consider these recent trends to be robust evidence that ROCE for the provision of cloud services for the largest cloud providers will decline to a 'normal' competitive level (ie to the cost of capital).

6.45 Based on the evidence available, our analysis indicates that AWS and Microsoft have been generating returns from their cloud services above their cost of capital, and this could be expected to continue in the future.

[603] We will disclose further information on the basis of our analysis to the relevant providers soon.

7. Barriers to entry and expansion

Introduction

7.1 Our guidance sets out that entry or expansion by firms can increase competition in a market and that the prospect of entry or expansion can countervail against factors that may be harming competition.[604]

7.2 In this section we consider potential barriers to entry and expansion. To do this we set out evidence we have gathered to date and our initial analysis of:

(a) whether larger cloud providers benefit from economies of scale when compared to smaller competitors;

(b) whether having a large portfolio of cloud services gives cloud providers strategic advantages over their competitors; and

(c) regulatory barriers in cloud services.

7.3 This analysis is focused on the underlying structure of the cloud services market. We will consider other potentially relevant features in other working papers, in particular with regard to cloud service providers' conduct and/or incentives. For example, some third parties raised concerns relating to the functioning of the markets and framed them as issues that raise barriers to entry, including:

(a) One cloud provider said that software licensing practices by legacy software providers, in particular Microsoft, present a significant barrier, as they harm customers' ability to switch and smaller cloud providers' ability to access new customers.[605]

(b) A second cloud provider said that software licensing practices can present a significant barrier, as they negatively impact customer choice and the ability of cloud providers to provider services or expand the provision of cloud services.[606]

(c) A third cloud provider said that barriers limiting its ability to grow arise primarily from (i) technical barriers and (ii) commercial/contractual barriers stemming from business and licensing practices.[607]

(d) One small cloud provider said that there are reputational barriers to entry and expansion as customers perceive the large cloud providers to be superior and are seen as a safe choice for a Chief Information Officer when choosing

---

[604] CC3, paragraph 205.
[605] [✂] response to the CMA's information request [✂].
[606] [✂] response to the CMA's information request [✂].
[607] [✂] response to the CMA's information request [✂].

a cloud provider.[608] Some customers also commented that they consider the large cloud providers to be more credible and capable than smaller cloud providers. We invite further views on reputational barriers. [609,610,611]

Economies of scale

7.4 Economies of scale arise where average costs fall as the level of output rises over a range of output volume. Economies of scale, in combination with sunk investment costs, can constitute a barrier to entry or expansion.[612]

7.5 In markets where economies of scale are significant, entry on a smaller scale may not be profitable and entry on a large scale will often entail a high risk as it will generally be successful only if a firm can expand the total market significantly or substantially replace one of more of the existing firms.[613]

7.6 In this section, we set out our analysis of whether large cloud providers benefit from economies of scale. The evidence that we have gathered suggests that economies of scale are achieved through:

(a) sunk cost investments in cloud infrastructure;

(b) bulk purchasing servers, components and network equipment;

(c) operating efficiencies; and

(d) spending on research and development.

7.7 Economies of scale are most applicable to IaaS. However, they may also apply to PaaS in some circumstances, for example around research and development.

7.8 To understand how large cloud providers realise economies of scale, we first consider the costs incurred operating data centres. A cloud provider and a global third party data centre provider said that the main resources required to operate a data centre are: [614,615]

(a) land and site costs;

(b) servers, components and network equipment;

---

[608] Note of call with [✂].
[609] Note of call with [✂].
[610] Note of call with [✂].
[611] [✂] response to the CMA's information request [✂].
[612] CC3 paragraph, 212.
[613] Ibid., 213.
[614] [✂] response to the CMA's information request [✂].
[615] Note of call with [✂].

(c) energy supply; and

(d) labour.

7.9 These costs may be incurred either directly or indirectly by cloud providers.[616] Any economies of scale that large cloud providers benefit from will be a function of comparatively lower costs in one or more of these areas.

Investment by cloud providers in cloud infrastructure

7.10 Below, we comment on the size of cloud providers' investments in fixed asset infrastructure and the extent to which some or all of these fixed costs are unavoidable 'sunk' costs that cannot be recovered upon exit.

7.11 We focus on large, unavoidable, 'sunk' costs because they could deter new entry into the market, as new entrants would need to make significant upfront investments in infrastructure without any certainty of return. If their plans to enter the market failed, they would be unable to recoup the initial investment, further increasing the risk of entry for them.

7.12 AWS said that one of the benefits for its customers of using cloud services is that customers can replace fixed, capital IT expenditure (CapEx) with variable, operating expenditure (OpEx). This is because cloud providers invest in the fixed asset infrastructure required to deliver computing resources and customers only pay when they consume these resources.[617]

7.13 The corollary of customers moving from CapEx to OpEx is that the cloud providers must make the CapEx investments in the fixed asset infrastructure for their customers. The cloud providers' CapEx therefore represent the unavoidable 'sunk' costs which we discuss further below.

Investment in fixed assets

7.14 Evidence we have seen to date suggests that cloud providers use their fixed asset infrastructure to deliver IaaS, rather than PaaS.

7.15 PaaS requires IaaS, but PaaS providers can use third party IaaS so there is no requirement for PaaS providers to own infrastructure. Hence, investment in infrastructure is a requirement for entry and expansion into IaaS, but not PaaS.

[616] For example, servers, components and network equipment will be incurred directly by cloud providers regardless of whether they own or co-locate. Energy and labour costs can be incurred directly or indirectly depending on whether the cloud provider owns the data centre or the terms of the co-location agreement.
[617] Six advantages of cloud computing - Overview of Amazon Web Services, accessed 20 May 2024.

7.16 Our findings on fixed asset infrastructure and economies of scale should be viewed in this context.

7.17 Cloud providers group their fixed assets under three broad headings: [618,619,620]

(a) Data centre assets: these include the shell of the data centre, building the data centre, the infrastructure for energy and water and the land on which the data centre is built. Data centres can be owned outright by cloud providers or they can be owned by third party data centre providers. If the data centre is owned by a third party, cloud providers can co-locate in a third party's data centre, which involves the third party data centre provider offering a managed service or cloud providers can lease space in a third party data centre.

(b) Network assets: cloud providers connect their data centres using network assets, which connect all server racks to each other, to other datacentres, to customers and to the internet. Network assets include IP Transit, Cross Connects, Backbone and Metro Fibre.

(c) Servers / components: servers run the software to enable cloud related services. Servers and components include racks and the specific components on the racks, such as CPUs,[621] GPUs[622], DRAM,[623] Motherboards, SSDs[624] and HHDs[625].

7.18 Cloud providers depreciate their investments over an asset's useful economic life. This means that the carrying value of an asset, as reflected in a cloud provider's annual financial statements, is less than the upfront cost of the asset.[626] This also indicates that if a cloud provider had to sell its assets, for example if it exited the market, the recoverable amount would likely be less than the initial investment, as the value of the asset decreases over time.

7.19 We have considered evidence on the carrying value of the global fixed assets of AWS, Microsoft Azure, Google, Oracle Cloud and IBM Cloud as at the end of financial year 2022 (the most recent year that we have data for these cloud providers).[627] This showed that the larger cloud providers have made significant investments in infrastructure.

---

[618] [✂] response to the CMA's information request [✂].
[619] [✂] response to the CMA's information request [✂].
[620] [✂] response to the CMA's information request [✂].
[621] Central Processing Units
[622] Graphics Processing Units
[623] Dynamic Random Access Memory
[624] Solid State Drives
[625] Hard Disk Drives
[626] This is standard accounting practice.
[627] Source: Responses to the CMA's information requests [✂].

7.20 Public statements from cloud providers indicate that they expect to accelerate their investments in fixed assets in the next couple of years. [628] [629] An internal document from a cloud provider indicates that this will include investment in data centres and cloud infrastructure, including significantly increased GPU spend to build AI capacity.[630]

Co-locating vs owning data centres

7.21 Two cloud providers said that it is possible to enter the markets without making the same levels of investment and that many cloud providers can start small and scale up as their business grows. They mentioned OVHcloud, eCloud VPC, Outscale, Scaleway and Brainboard as examples.[631]

(a) One cloud provider said that smaller cloud providers are able to compete despite the existence of scale advantages, noting that there is no need for a cloud provider to build its own global network of data centres and that infrastructure can be owned, leased or outsourced.[632]

(b) Another cloud provider said that small and mid-scale cloud providers can enter the market by using capital-efficient strategies such as leasing and co-location and scale out their capacity as their business grows.[633]

7.22 In the UK, we have seen that cloud providers, including larger ones, co-locate or lease most, if not all, of their data centre capacity:

(a) one cloud provider said that [✂] its currently operating data centre sites in the UK are co-located;[634]

(b) another cloud provider said that all of its data centres in the UK are owned and operated by third parties;[635] and

(c) a third said that it does not typically build or operate data centres but it leases them (globally).[636]

[628] AMZN-2024.03.31-EX99.1 (q4cdn.com), MSFT_FY24Q3_10Q.docx (live.com): 44.
[629] [✂].
[630] [✂].
[631] Responses to the CMA's information requests [✂].
[632] [✂] response to the CMA's information request [✂].
[633] [✂] response to the CMA's information request [✂].
[634] [✂] response to the CMA's information request [✂].
[635] [✂] response to the CMA's information request [✂].
[636] [✂] response to the CMA's information request [✂].

Purchasing efficiencies

7.23 If larger cloud providers are able to achieve bigger discounts than smaller cloud providers it will allow them to realise economies of scale by operating at a lower cost.

7.24 We set out the evidence we have gathered to date on the extent to which cloud providers can achieve discounts when buying servers, components and network equipment.

(a) One cloud provider said that due to its size, it benefits from certain economies of scale derived from its global cloud infrastructure in relation to certain inputs, notably servers and components;[637]

(b) another cloud provider said that purchasing hardware such as processors, servers, cooling units and network infrastructure in bulk allows it to achieve lower cost per MW of computing resources;[638] and

(c) a third cloud provider said that bulk purchasing and supply chain management resulted in direct savings in its unit costs for both IaaS and PaaS services.[639]

7.25 This shows that cloud providers can negotiate discounts on capital spending that reduce costs. A new entrant is unlikely to benefit from the same levels of discounts.

Operating efficiencies

7.26 Below, we consider the evidence that we have gathered to date and our analysis of the extent to which large cloud providers are able to achieve operating efficiencies, which lead to comparatively lower costs, by operating larger networks of data centres.

7.27 Microsoft has noted publicly that its cloud businesses benefit from three economies of scale:

(a) Data centres that deploy computational resources at significantly lower cost per unit than smaller ones.

(b) Data centres that coordinate and aggregate diverse customer, geographic and application demand patterns, improving the utilisation of computing, storage, and network resources.

---

[637] [✂] response to the CMA's information request [✂].
[638] [✂] response to the CMA's information request [✂].
[639] [✂] response to the CMA's information request [✂].

(c) Multi-tenancy locations that lower application maintenance labour costs.[640]

7.28 Microsoft states that larger cloud providers can achieve economies of scale by having larger data centres and aggregating demand across data centres to achieve high utilisation rates.[641]

7.29 Below, we describe types of operating efficiencies that cloud providers can realise in terms of: energy requirements; data centre capacities; and utilisation rates.

Energy requirements

7.30 Evidence we have seen to date suggests that energy is the largest variable cost incurred by a cloud provider when operating data centres, to the extent that data centre capacity is measured in megawatts, ie the amount of energy that it consumes. Any efficiencies that large cloud providers achieve in their energy consumption will give them a competitive advantage compared to smaller cloud providers.

(a) A cloud provider said that the supply of electricity comprises the largest portion of the total costs of operating a data centre, noting that energy supply is required for running the servers and network devices, security systems, heating, ventilation and air conditioning mechanical and cooling systems, and lighting.[642]

(b) The same cloud provider said that larger data centres will have greater energy overheads than smaller sites, but larger data centres also tend to be more energy efficient compared to smaller ones.[643]

(c) A second cloud provider said that large data centres achieve a lower cost per MW of computing resources through lower energy costs, through even distribution of workloads between servers and optimised cooling solutions.[644]

(d) A global data centre provider also said that larger data centres achieve economies of scale through more efficient use of energy. [645]

(e) The same global data centre provider said access to energy can be a barrier to the provision of new data centre capacity, as uninterrupted and reliable sources of energy are not always available, particularly in locations with high demand due to a high density of networks. [646]

---

[640] Microsoft 2023 Annual Report
[641] Microsoft 2023 Annual Report
[642] [✂] response to the CMA's information request [✂].
[643] [✂] response to the CMA's information request [✂].
[644] [✂] response to the CMA's information request [✂].
[645] Note of call with [✂].
[646] Note of call with [✂].

Data centre capacities

7.31 We have reviewed evidence on global cloud data centre capacity in megawatts for the largest cloud providers for the past few years and forecasts for future years.[647] [648] This showed that the largest cloud providers forecast an increase in capacity. This is consistent with the increase in planned investment in fixed assets, most notably investments to meet demand for accelerated compute services.[649]

7.32 We considered the sizes of the data centres for six cloud providers: AWS, Microsoft, Google Cloud, Oracle, IBM and OVHCloud to understand which cloud providers have larger data centres which may allow them to use energy more efficiently.[650] We note that there are differences in composition of the data centre networks between the different cloud providers, for example, some providers having many smaller data centres, other providers having very large or no large centres.

7.33 While larger data centres may be more energy efficient, this is not the only consideration for a cloud provider. For example, we understand that physical proximity can often affect latency for customers, which might incentivise having more locations.[651]

Utilisation

7.34 Cloud providers aim for high levels of utilisation of their data centres in order to maximise cost efficiencies and also to allow for a capacity buffer in case of spikes in demand.[652]

7.35 We have analysed the utilisation rates that cloud providers consider to be 'high'. We note that providers' view on this varies quite widely and the actual average utilisation rates that cloud providers achieve are between 10 and 20% below these 'high' levels.[653]

Investment in research and development

7.36 Below, we set out the evidence we have gathered to date on the cloud providers' spend on research and development.

7.37 We recognise that spend on research and development can improve quality and increase efficiency. We also recognise that customers stand to benefit from better

---

[647] [✂]].
[648] Data from cloud providers. [✂].
[649] [✂].
[650] Sources: Data from cloud providers. [✂].
[651] https://aws.amazon.com/blogs/architecture/what-to-consider-when-selecting-a-region-for-your-workloads/, accessed 20 May 2024.
[652] [✂] response to the CMA's information request [✂].
[653] Sources: Data from cloud providers. [✂]

quality, more innovative cloud services and from lower prices (assuming any benefits and/or efficiencies that are realised are passed on, eg through lower pricing).

7.38 However, when spend on research and development is significant, it could be difficult for smaller firms to achieve comparable levels of spend, and so could contribute to barriers to entry and expansion.

7.39 We have received the following evidence on cloud provider spend on research and development:

(a) A cloud provider said that it spent USD 27bn on research and development on cloud computing in financial year 2023 and that the other large cloud providers made similar investments in research and development. It also said that high levels of research and development spend are required to attract new customers and limit the number of customers switching to rivals.[654]

(b) Another cloud provider also said that it incurred significant annual expenditure on research and development exceeding USD 6bn across its global technology businesses and that investment results in innovation benefitting products both in and out of cloud.[655]

(c) Another cloud provider said that its internal research and development entails a degree of fixed and sunk costs. The same cloud provider also said that it is not necessary to make significant research and development investments or incur costs to begin to offer basic services (compute, storage and networking).[656]

(d) Another cloud provider said that it works with its partners to custom-develop server hardware that it deploys into its data centres. This approach helps it to reduce the time to market for new innovative products and enables cost reductions for these inputs.[657] The provider told us that deploying customised hardware, combined with scale related purchasing discounts (the effects of which are discussed more broadly under 'purchasing efficiencies' above) resulted in [✂] reduction in server capital expenditure to meet an additional [✂] of demand in the period from 2018 to 2022. [658, 659]

7.40 In summary, this evidence shows that cloud providers make significant investments in research and development.

---

[654] [✂] response to the CMA's information request [✂].
[655] [✂] response to the CMA's information request [✂].
[656] [✂] response to the CMA's information request [✂].
[657] [✂] response to the CMA's information request [✂].
[658] [✂].
[659] [✂] response to the CMA's information request [✂].

7.41 We currently consider that, while new entrants may be able to enter the markets without making significant investments in research and development initially, the scale of investment in research and development required to remain competitive may represent an ongoing barrier. For example, competitors and potential competitors are likely to consider this when weighing up whether to seek to enter, or invest to expand, within these markets.

Emerging views on economies of scale

7.42 The evidence we have seen to date shows that:

(a) IaaS requires significant capital investment in fixed assets. These investments are mainly sunk costs that would not be recovered in full on exit.

(b) The levels of up-front investment can be reduced by co-locating or leasing data centres, allowing new entrants to scale their business as they grow their customer base. However, data centres only make up part of the value of large cloud providers' fixed assets and a new entrant would still need to invest in the servers, components and network equipment for a co-located or leased data centre.

(c) The largest cloud providers are making significant further investment in their cloud infrastructure, in particular in accelerator chips (eg GPUs), to meet growing demand for AI services (see section 8 below). This increases the capital investment required by a new entrant, should they choose to offer customers accelerated compute capacity.

(d) Larger cloud providers benefit from bulk purchasing discounts on necessary equipment.

(e) Most cloud providers have submitted that larger data centres are more energy efficient, although not all cloud providers consider these efficiencies to be significant. There is also evidence that cloud providers can distribute workloads between data centres to achieve efficiencies in how they use energy.

(f) Some cloud providers have larger data centres in their networks and target slightly higher rates of utilisation, which should allow them to achieve economies of scale.

7.43 In view of the above, our emerging view is that large cloud providers benefit from economies of scale. When considered individually, some of the factors represent marginal incremental benefit. However, when considered together they represent material barriers to entry and expansion.

Product portfolios

7.44 In this section, we assess whether having a large portfolio of cloud services may give cloud providers advantages over their competitors.

7.45 Cloud providers typically offer a range of first party cloud services through their platforms and some also offer third party services provided by ISVs through a marketplace accessible by customers on the provider's platform. Marketplaces are used by eligible ISVs to offer their own services to the customers of those providers.

7.46 In assessing cloud providers' product portfolios, we look at both first and third party services, as well as both IaaS and PaaS.

7.47 We break down product portfolios by discussing the following aspects in turn:

(a) Importance or range of services

(b) Economies of scope

(c) Network effects

Importance of range of services

7.48 If customers value having access to a large portfolio of cloud services, then providers with a large portfolio may be more attractive to customers relative to smaller providers, such that smaller providers may find it harder to compete for customers.

7.49 We have gathered customer views on the importance of range of first and third party products.

Evidence from customers

7.50 We asked large customers to rate the importance of a list of factors their organisation considers when choosing their main public cloud. In answering this question, customers rated the importance of the range of cloud infrastructure services alongside other factors. They were asked to rate these factors from one to five, with one being not important at all, and five being very important.

7.51 Range of cloud infrastructure services was identified as one of the most important factors by customers we spoke to when choosing their main public cloud provider, attaining an average rating that rounds up to either five or four. This was alongside other factors such as price (including discounts or cloud credits), service quality and data sovereignty requirements and number and location of data centres.

7.52 Customers rating it as either important or very important (ie 4 or 5 out of 5) gave various reasons for this.

(a) A few customers said having a broad range of services was important in allowing them to build their own end products effectively and with flexibility, and meet their business needs.[660]

(b) A few other customers said that a provider offering a large range of services could be good signal that the provider is introducing new and improved services.[661]

7.53 A handful of customers said that the range of cloud infrastructure services is less important because most providers offer the same services.[662] However, we note that this is somewhat inconsistent with what customers we spoke to said when asked about whether smaller providers (eg IBM, Oracle, OVHcloud) were suitable alternatives to their main providers.

7.54 As set out in the discussion of providers in Section 2, customers generally considered these providers were not suitable alternatives, with some of the main reasons given being the more limited services or capabilities compared to larger cloud providers and a lack of any experience or knowledge of that provider's offering. This suggests that these customers may have been considering the offerings of AWS, Microsoft and, possibly,[663] Google rather than smaller cloud providers.

Emerging view on the importance of range of services

7.55 The evidence we have seen to date suggests that the range of first party products is an important factor for customers when choosing which cloud providers to use.

Economies of scope

7.56 Economies of scope arise when producing two (or more) services is less costly for a single firm than for two (or more) firms each to produce the services separately. Where economies of scope are significant, an entrant, if it is to be successful, might have to produce a range of services from the outset, adding to the costs of entry. Economies of scope might be relevant if, for example, R&D and operations spend can be spread over a wider range of services.

---

[660] Responses to the CMA's information requests [✂].
[661] [✂] response to the CMA's information request [✂]; [✂] response to the CMA's information request [✂];
[662] Responses to the CMA's information requests [✂].
[663] More customers identified Google as being a credible alternative, but others still said that Google did not have as advanced or as broad a range of functionality / features / services as their main providers (AWS or Microsoft). Eg [✂].

7.57 We asked cloud providers to explain whether they could benefit and have in the past benefitted from any efficiencies (eg reduction in costs) as a result of increasing their range of services.

7.58 Cloud providers said that these efficiencies can arise and some also said they have benefitted from them. However, the views of cloud providers varied as to the strength of these efficiencies. Some also said that expanding the range of services can sometimes lead to inefficiencies.

7.59 In relation to the efficiencies that can arise:

(a) One provider said that it has benefitted from efficiencies through expanding its number of cloud offerings. It said that these efficiencies have benefitted its business both in the development and operation of services. On the development side, this provider said that it has been able to build new services using existing expertise and technology across different layers of the cloud stack. On the operation side, it said that it has been able to optimise the utilisation and allocation of its resources.[664]

(b) Another provider said that cloud providers benefit from increasing their range of services. It said that cloud providers also use their own services, just like customers do. Therefore, cloud providers also directly benefit from new services and features. The provider said that any new service likely builds upon the fundamental services of compute, storage, networking, identity, and security. It also said that it has benefited from increasing the range of services.[665]

(c) Another provider said that an increase in the number of services offered does not necessarily increase overall efficiency for a cloud provider.[666]

(d) The same provider said it benefitted from simplifying operations and sharing learnings technologies across teams where there are opportunities to do so that would benefit customers. It said that some of its products share commonalities across codebases, so if one of the service teams discovers a vulnerability or bug in the code, they will quickly share the learnings with the other team to remedy the issue for customers. The provider also said that its product teams can use the same services for commonly used functionality like logging, billing, and authentication, meaning they do not need to build those tools from the ground up to launch a new service and can pass on benefits to customers in the form of a consistent user experience.[667]

---

[664] [✂] response to the CMA's information request [✂].
[665] [✂] response to the CMA's information request [✂].
[666] [✂] response to the CMA's information request [✂].
[667] [✂] response to the CMA's information request [✂].

(e) Another provider said that cloud providers can benefit from efficiencies as they increase the number of customer-facing cloud services by developing processes and systems to support those services. It said that having centralised teams that build and manage base runtimes, pipelines, automation and architecture allows the service teams to focus on the higher-order functions / customer-facing aspects of their service. In general, centralisation of resources allows cloud providers to better scale and drive down costs.[668]

(f) Another provider said that any efficiencies that may be derived from developing a broader range of services are, in its experience, limited (albeit there may be exceptions).[669]

7.60 In relation to the inefficiencies that may arise:

(a) One provider said that many of the services it offers are discrete and distinct, and the learnings and functionalities that apply to one may not necessarily translate or be useful to others. It also said that increasing the number of service offerings introduces inefficiencies due to increased complexity. The provider said that customers generally expect that its services will work together and offer a consistent experience. Therefore, this provider said, as the number of its service offerings grows, its service teams must design, maintain, and adapt their services to satisfy an increasing number of parameters tied to its other services.[670]

(b) Another provider said that having a large service portfolio requires a broad and deep set of operational resources and expertise to manage, debug, and secure a cloud comprised of so many different offerings. On the other hand, it said, entities that focus on just one or a limited set of cloud solutions also have their own efficiencies that can be generated by such specialisation.[671]

(c) Another provider provided an example of a new service it built which required that teams learn a new domain specific language, a new ecosystem and build the integration with existing systems. It said that, although these factors will lead to long term efficiencies, that is not the case in the short term.[672]

Emerging views on economies of scope

7.61 The evidence we have seen to date suggests that there may be some economies of scope in supplying a range of services, but this may not be the case for all cloud

[668] [✂] response to the CMA's information request [✂].
[669] [✂] response to the CMA's information request [✂].
[670] [✂] response to the CMA's information request [✂].
[671] [✂] response to the CMA's information request [✂].
[672] [✂] response to the CMA's information request [✂].

providers and for all services. In some cases, increasing the portfolio of services might instead lead to inefficiencies.

Network effects

7.62 If customers value having access to a large portfolio of third party services (ISVs) through the cloud infrastructure providers they use, providers with a large pool of ISVs being hosted on their platform and/or listed on its marketplace would have an advantage over other providers such that other providers may find it harder to compete for customers.

7.63 This might be compounded if there are indirect network effects between cloud infrastructure providers, ISVs and customers. Indirect network effects may arise as follows:

(a) The more customers a cloud infrastructure provider has, the more attractive it becomes to ISVs (eg because it provides them with access to a larger customer base) and the more likely ISVs are to use that cloud provider's platform. This could just be ISVs running their services on the cloud infrastructure of this provider but also ISVs listing their services on this provider's marketplace.

(b) The more ISVs available on a cloud infrastructure provider's platform, the more attractive the provider becomes to customers (eg because they can access more ISV services on that platform) and the more likely customers are to use that provider's platform.

Evidence from ISVs

7.64 To assess the strength of these network effects, we collected evidence from ISVs and customers.

7.65 To assess the first limb, that is whether ISVs are more attracted to cloud providers with more customers, we asked some ISVs how important, if at all, is it that a public cloud provider has a large user base in their choice of which cloud provider to be available on.

7.66 Two ISVs said that large user base is a relevant consideration:

(a) One ISV said that a cloud provider having a large user base is quite important. It explained that given limited resources, putting resources into preparing to work with or on a public cloud has to have sales opportunities downstream attached to it. It further explained that one way of helping guarantee future revenue is to make sure there is a large user base on that

cloud and that many of the European clouds have sufficient user bases to be of interest to it.[673]

(b) The other ISV said that its customers typically want to use its services on the cloud provider and in region(s) where their data resides. It explained that, because of this, the size of the cloud provider's customer base is relevant when it decides which public cloud providers to support.[674]

7.67 We note this evidence is consistent with the evidence collected by Ofcom in its Market Study. In particular, some ISVs highlighted the size of user base of the different cloud providers, eg as proxied by its market share, as an important consideration when deciding which cloud providers to run on and/or which to run on first.[675] Some ISVs also noted that there are significant costs associated with running their services to additional public cloud providers besides the largest ones.[676]

Evidence from customers

7.68 To assess the second limb, that is, whether customers value cloud providers with more ISVs, we included range of services offered by ISVs as a criteria when asking large customers about factors in their choice of main public cloud. As set out above we also asked about the range of cloud infrastructure services. To some extent there may be some overlap between these two criteria if customers do not always distinguish between services they receive from their main public cloud provider and those they receive from ISVs on that public cloud.

7.69 As set out above, large customers were asked to rate the importance of the selection criteria they considered in choosing their main public cloud and provide an explanation for their rating. Customers rated each factor on a scale of 1 to 5, with 1 being not important at all, and 5 being very important.

7.70 Range of cloud infrastructure services was identified as one of the most important factors by customers when choosing their main public cloud provider attaining an average rating that rounds to either five or four. This was alongside other factors such as price (including discounts or cloud credits), service quality and data sovereignty requirements and number and location of data centres.

7.71 Range of services offered by ISVs was identified as one of the least important factors relative to the other selection criteria, attaining an average rating that

[673] Note of meeting with [✂].
[674] Note of meeting with [✂].
[675] [✂] response to Ofcom's information request [✂]. Note of meeting between Ofcom and [✂]
[676] Responses to Ofcom's information requests.

rounds to three. Examples of reasons given for the lack of importance of this factor were:

(a) Some customers said that most ISV services were accessible via all cloud providers and therefore were not very important when deciding which main cloud provider to use.[677]

(b) A few other customers said that the range of ISVs is not a significant differentiator when choosing a main provider and first party services are a more important factor.[678]

7.72 However, a few customers said that the range of cloud infrastructure services offered by ISVs was becoming more important and may be important to their medium and long-term strategy.[679]

Emerging views on network effects

7.73 The evidence we have seen to date suggests that there may be some indirect network effects, in that ISVs find public cloud providers with a larger customer base more attractive.

7.74 However, the evidence on the impact in the other direction is less clear, that is the attractiveness to customers of public cloud providers with a large range of ISVs. Therefore, while there may be indirect network effects, the evidence to date does not suggest that there are strong indirect network effects between ISVs, customers, and public cloud providers.

Regulatory barriers

7.75 Cloud providers identified a number of regulations they have to comply with in order to run their operations including data security (for example the Network and Information Systems Regulations) and data privacy (for example UK General Data Protection Regulation). [680]

7.76 However, no cloud providers flagged these as being particular barriers to entry or expansion. For example, one provider said it does not believe that there are significant legal or regulatory barriers which restrict new entrants from developing and offering services.[681]

[677] Responses to the CMA's information requests [✂].
[678] Responses to the CMA's information requests [✂].
[679] Responses to the CMA's information requests [✂].
[680] Responses to the CMA's information requests [✂].
[681] [✂] response to CMA's information request [✂].

7.77 Generally, cloud providers were also unable to accurately provide the costs of compliance and how this might differ globally and in the UK.

7.78 Many cloud providers commented on their overall regulatory burden, including European Union legislation (most notably the Data Act[682]), and the impact this had on their business. But providers focussed on the broader consequences of this for their operations, for example diverting resource from delivery and innovation as opposed to representing barriers to entry or expansion.

7.79 To serve financial services customers, cloud providers in the UK face more regulatory obligations:

(a) Financial services firms themselves face certain regulations and these regulations place additional burdens on any cloud providers that provide products or services to financial services firms.[683]

(b) All providers recognised the tailored offerings they were required to offer financial services firms.

7.80 The UK regulatory landscape is changing as sectors and regulators respond to the increased adoption and importance of cloud services across the economy. For example, the Bank of England, Prudential Regulation Authority and Financial Conduct Authority have proposed requirements for 'Critical Third Parties in Financial Services'.[684] Their cost benefit analysis recognises that, when applied to 'a market that is already highly concentrated, the proposals could have the impact of further entrenching market power'. However, they find that due in part to a number of mitigations, the risk of such impact is considered to be very small.

7.81 None of the providers we have engaged with indicated otherwise. We therefore still consider that there are no particular regulatory barriers to entry or expansion in the UK. [685]

Public sector procurement

7.82 Some stakeholders, in responses to our Issues Statement, raised concerns about the impact that public sector procurement practises are having on competition in the market. [686] We continue to gather evidence to assess these points further.

[682] Data Act | Shaping Europe's digital future (europa.eu), accessed on 21 March 2024.
[683] For example, SS2/21 Outsourcing and third party risk management | Bank of England in the UK and the Digital Operational Resilience Act (DORA) - European Union (europa.eu) in the EU, both accessed on 02 April 2024.
[684] CP26/23 - Operational resilience: Critical third parties to the UK financial sector | Bank of England, accessed on 14 February 2024.
[685] Cost benefit analysis (bankofengland.co.uk), accessed on 14 February 2024.
[686] Prolinx, Oracle and former UK Cloud employees' responses to the Issues Statement, 17 October 2023.

8. Artificial intelligence and public cloud infrastructure services

8.1 In our issues statement we said that we will consider the potential impact of artificial intelligence (AI) on how competition works in the cloud services market.[687] This is because access to cloud services underpins the development and deployment of AI[688] as it provides the computing resources and infrastructure needed to train and deploy AI models at scale.[689]

8.2 We continue to gather evidence to understand the importance to cloud providers of demand from AI developers and the extent to which there are any differences in the nature of competition for AI developers compared to other sets of customers.

8.3 In this section we set out our assessment of evidence we have gathered to date on:

(a) the role of AI developers as customers of cloud providers;

(b) the importance for cloud providers of providing accelerated compute for AI development and its deployment amongst customers; and

(c) competition between cloud providers to supply accelerated compute.

8.4 In addition to the evidence we have gathered for this inquiry, the CMA's ongoing AI foundation models (FMs) review has informed our understanding of developments in the FM sector and the use of compute as a key input to the development of FMs.[690]

The role of AI developers as customers of cloud providers

8.5 Organisations that develop and/or deploy AI-powered products and services, particularly FMs, are a fast-growing source of demand for cloud services.

(a) Boston Consulting Group predicts that the total addressable market for uses of generative AI (which is powered by FMs) will increase from $18 billion in 2023 to $121 billion in 2027;[691] and

[687] Issues statement (publishing.service.gov.uk), paragraph 13.
[688] AI has a wide variety of forms and applications. Here, our use of the term generally refers to foundation models, which typically require substantial specialist compute resources. Foundation models (including, among others, language models or LLMs) are a type of AI technology that are trained on vast amounts of data that can be adapted to a wide range of tasks and operations.
[689] See AI Foundation Models: initial review - GOV.UK (www.gov.uk).
[690] AI Foundation Models: initial review - GOV.UK (www.gov.uk). See the Update Paper published on 11 April 2024 and the Technical Update Report published on 16 April 2024.
[691] BCG Executive Perspectives (2023) The CEO's Roadmap on Generative AI, page 8, accessed 19 April 2024.

(b) IDC forecasts that worldwide core IT spending on generative AI will grow from $40.1 billion in 2024 to $151.1 billion in 2027.[692]

8.6 These figures relate to a broader category of spending than just FM-related demand for IaaS and PaaS. However, they are indicative of the anticipated speed of growth of demand.

8.7 Most FM developers rent the accelerated compute that they need from cloud providers rather than build it themselves, due to the large up-front cost and complexity involved in doing so, which results from the size of FMs and vast amounts of data needed to train them.[693,694] Other options for organisations that train FMs are to invest up-front in building their own infrastructure or to access a public super-computer.[695]

8.8 Evidence we have seen to date suggests that the following FM developers and platforms are customers of public cloud services:[696]

(a) AWS: AI21 Labs,[697] Anthropic,[698] Hugging Face,[699] Runway ML,[700] and Stability AI.[701]

(b) Microsoft: [✂],[702] Inflection AI,[703] Meta,[704] Mistral,[705] OpenAI.[706]

---

[692] IDC (December 2023) Generate Growth in Your Markets with the GenAI Opportunity, accessed 19 April 2024.
[693] Note of meeting with [✂].
[694] Not all uses of FMs require cloud computing/large-scale data centres. Running inference on FMs generally requires far less compute than training them; furthermore, the development of smaller models and more efficient chip performance has enabled the possibility in some cases of running inference on smaller clusters ('workstation') or even on consumer devices such as laptops and phones (known as 'edge AI'). Best NVIDIA GPUs for Serving Inference on CoreWeave — CoreWeave, accessed 19 April 2024. Computation at the edge (on device) reportedly provides users with faster inference and greater privacy and security, as data does not need to be sent to and from the cloud. On-Device Processing and AI Go Hand-in-Hand | MIT Technology Review, accessed 19 April 2024.
[695] One provider gave two examples of Dell collaborations to offer on-premises solutions for FM deployment (with Hugging Face and Meta, respectively). [✂] response to CMA's information request [✂]. [✂]; see [✂].
[696] Note that this list is illustrative rather than exhaustive: there are likely thousands of FM developers in existence, and there may be examples where cited FM developers have, or have had, other cloud suppliers beyond those mentioned here. Where a customer is associated with more than one cloud provider, in some cases this is due to switching over time, in other cases due to multi-clouding.
[697] [✂] response to CMA's information request [✂]; AI21 Labs Accelerates Generative AI Model Adoption Using Amazon SageMaker | Case Study | AWS, accessed 19 April 2024.
[698] Amazon will invest up to $4B in Anthropic to advance generative AI (aboutamazon.com), accessed 19 April 2024.
[699] [✂] response to CMA's information request [✂]; Note of meeting with [✂]. Hugging Face and AWS partner to make AI more accessible, accessed 19 April 2024.
[700] [✂] response to CMA's information request [✂]; Scaling our in-house research infrastructure with AWS | Runway Blog (runwayml.com), accessed 19 April 2024.
[701] Stability AI builds foundation models on Amazon SageMaker | AWS Machine Learning Blog, accessed 19 April 2024.
[702] [✂] response to CMA's information request [✂].; AI21 Labs Accelerates Generative AI Model Adoption Using Amazon SageMaker | Case Study | AWS, accessed 19 April 2024.
[703] [✂] response to CMA's information request [✂].After raising $1.3B, Inflection is eaten alive by its biggest investor, Microsoft | TechCrunch, accessed 19 April 2024.
[704] [✂] response to CMA's information request [✂]; Meta selects Azure as strategic cloud provider to advance AI innovation and deepen PyTorch collaboration | Microsoft Azure Blog, accessed 19 April 2024.
[705] Introducing Mistral-Large on Azure in partnership with Mistral AI | Microsoft Azure Blog, accessed 19 April 2024.
[706] [✂] response to CMA's information request [✂]; OpenAI and Microsoft extend partnership, accessed 19 April 2024. With a systems approach to chips, Microsoft aims to tailor everything 'from silicon to service' to meet AI demand - Source. What is Azure OpenAI Service? - Azure AI services | Microsoft Learn, accessed 19 April 2024.

(c) Google: AI21 Labs,[707] Anthropic,[708] Character AI,[709] Cohere,[710] Midjourney,[711] and Runway ML.[712]

(d) Oracle: Adept AI,[713] Character AI,[714] Cohere,[715] and MosaicML.[716]

(e) CoreWeave: EleutherAI,[717] Inflection AI,[718] and OpenAI.[719]

8.9 FM developers will typically access accelerated compute from cloud providers in one of three ways:

(a) purchasing at commercial on-demand rates. However, this may be cost-prohibitive and has no guarantee of availability at any given time;

(b) entering an agreement (usually one to three years) to purchase compute at a reduced rate. This might guarantee availability; however, potentially with a lead time of six to nine months; or

(c) entering into a commercial partnership with a cloud provider which may involve the cloud provider using the model for its own services or making the FM available for others to use via its cloud platform.[720]

Importance of accelerated compute

8.10 The evidence set out above shows that cloud providers are an important source of accelerated compute for organisations developing and/or deploying FMs (including FM developers) and that demand from such customers is likely to increase.

8.11 In this context, we have sought to understand the importance to cloud providers of demand for accelerated compute now and their view of the importance of demand for this going forward. We note that our main focus so far has been on demand from FM developers. However, we understand that there is increasing demand for

---

[707] Generative AI startups choose Google Cloud | Google Cloud Blog, accessed 19 April 2024.
[708] Anthropic Partners with Google Cloud \ Anthropic, accessed 19 April 2024.
[709] [✂] response to CMA's information request [✂]. Character.AI and Google Cloud Partner to Build the Next Generation of Conversational AI - May 10, 2023 (googlecloudpresscorner.com), accessed 19 April 2024.
[710] Generative AI startups choose Google Cloud | Google Cloud Blog, accessed 19 April 2024.
[711] Midjourney Selects Google Cloud to Power AI-Generated Creative Platform (prnewswire.com), accessed 19 April 2024.
[712] Runway to Make Content Creation More Accessible with Google Cloud's Generative AI (yahoo.com), accessed 19 April 2024.
[713] AI Startups Find an Unlikely Friend: Oracle — The Information, accessed 19 April 2024.
[714] AI Startups Find an Unlikely Friend: Oracle — The Information, accessed 19 April 2024.
[715] Oracle to Deliver Powerful and Secure Generative AI Services for Business, accessed 19 April 2024.
[716] MosaicML Trains Generative AI Models Faster with Oracle, accessed 19 April 2024.
[717] CoreWeave Partners with EleutherAI & NovelAI to Make Open-Source AI More Accessible — CoreWeave
[718] Announcing our collaboration with NVIDIA and CoreWeave on MLPerf (inflection.ai) , accessed 19 April 2024.
[719] [✂] response to CMA's information request [✂].Microsoft inks deal with CoreWeave to meet OpenAI cloud demand (cnbc.com), accessed 19 April 2024.
[720] CMA AI Foundation Models: Initial Report, page 35. In line with this, a [cloud provider] [✂] internal document compares pricing of H100 instances across [✂], based [✂]. See also EC2 Reserved Instance Pricing – Amazon Web Services, accessed 19 April 2024.

AI services from other customers that will rely on accelerated compute. For example, as set out in Section 2, multiple customers we spoke to said that providers' AI capabilities are becoming an increasingly important consideration.[721]

8.12 Cloud providers' views differ somewhat on the potential impacts of demand from companies developing and training FMs on competitive dynamics in the cloud market. For example:

(a) AWS said: 'The rapid and evolving demand for innovative, generative AI solutions, and the corresponding demand by FM developers for compute capacity, drives competition among IT providers. As such cloud providers face competitive pressure from all types of IT service providers, and this continue to grow.' It said that 'demand from FM developers has and will continue to create opportunities for a variety of IT providers, including on-premises and smaller cloud providers, to enter and expand in supplying IT resources to power the development of FMs and AI more broadly.'[722]

(b) Microsoft said: 'Demand for cloud computing services from companies developing, training, using and tuning Foundation Models is expected to grow rapidly.' It said that, since the accelerator chip capacity that provides the compute required for FM development is also used to provide the compute required for fine-tuning and inference,[723] demand for cloud services from FM developers has an influence in the cloud market more generally.[724]

(c) Microsoft said that it expects competition to intensify from the following sources:[725]

(i) Existing cloud providers (eg AWS, Google, Oracle, IBM, OVHcloud) serving FM developers.

(ii) Competitive pressure from new entry and expansion catering to demand for the cloud infrastructure required to train and run FMs. Microsoft provided a list of smaller providers, including new speciality providers, and said that there were aggregators that would offer capacity across different cloud providers and publicly owned supercomputers. Microsoft also said that NVIDIA has a 'virtual monopoly' on the accelerator chips to supply compute for FM training and inference, NVIDIA is actively promoting new cloud providers by, for example, ensuring allocations of hardware (ie accelerator chips) to new

[721] Responses to the CMA's information requests [✂].
[722] AWS response to CMA's information request [✂].
[723] Fine tuning is an optional process that can be applied to pre-trained models to add specific capabilities or improvements using particular datasets. Inference is the process of an AI model making predictions from new inputs. This is done by feeding the model new data and then using the model's parameters to generate a prediction.
[724] Microsoft response to CMA's information request [✂].
[725] Microsoft response to CMA's information request [✂].

speciality providers. Microsoft also said NVIDIA was expanding its own offering.

(iii) AI workloads being expected to increase the prevalence of multi-cloud strategies. It said that in general multi-cloud strategies are prevalent and only likely to increase. Specifically in relation to FM developers, it said that there are many cases where FM developers have used different cloud providers for different workloads (either switching between providers or using multiple providers for different workloads at the same time) and provided a non-exhaustive list of examples.

(iv) Customers switching from cloud to on-premises provision of AI accelerator capacity. It said that there is growing evidence of customers switching from cloud computing services to on-premises provision of accelerator capacity. Microsoft gave the examples of Meta, xAI, Zoom and Zoho.

(d) A cloud provider said [✂].[726]

8.13 Cloud providers' internal documents were more consistent in highlighting the importance of growing demand from generative AI.[727] [728] [729] [730] [731]

(a) For example, one cloud provider's internal documents characterise AI as a 'significant growth opportunity for Cloud', noting that 'AI is driving an unexpected acceleration in revenue growth in 2024', and that there is a 'land grab in AI' (among [✂] cloud competitors).[732, 733]

8.14 Other evidence also points to the importance of supplying accelerated compute. For example:

(a) Nvidia said: '[O]ver the next 5 years … data centers across the world will be reconfigured as accelerated computing data centers, moving away from traditional hardware and software solutions towards an infrastructure that can also effectively deploy generative AI.'[734]

(b) In recent quarterly earnings calls, Microsoft, AWS, and Google have all stated publicly that significant portions of their capital spending are going

[726] [✂] response to CMA's information request [✂].
[727] Some statements in internal documents refer generally to opportunities to cloud providers arising from generative AI – this likely includes opportunities across various levels of the stack, ie increased demand for accelerated compute at the infrastructure level, but also revenue opportunities in infusing generate AI into cloud providers' other platforms, services, and software products.
[728] [✂] response to CMA's information request [✂].
[729] [✂] response to CMA's information request [✂].
[730] [✂] response to CMA's information request [✂].
[731] [✂] response to CMA's information request [✂].
[732] [✂] response to CMA's information request [✂].
[733] [✂] response to CMA's information request [✂].
[734] Nvidia submission to the CMA [✂].

towards generative AI systems.[735] We note that not all of this investment will relate to increased demand for accelerated compute from FM developers and deployers.

Competition between cloud providers to supply accelerated compute

8.15 We have considered evidence of how cloud providers are competing to supply accelerated compute services. In particular, we have looked at how they are able to access the resources which are necessary to supply accelerated compute and, in particular, accelerator chips.

The supply of accelerated compute: access to accelerator chips

8.16 The provision of accelerated compute requires access to large numbers of AI accelerator chips[736] to work together in parallel, embedded into servers, stacked in racks or cabinets within a data centre or supercomputer and networked together.

8.17 In particular, cloud providers need access to state of the art (SOTA) AI accelerator chips as these are most efficient for FM developers to use.[737] While the upfront cost of SOTA AI accelerator chips may be high, potentially raising the cost passed onto compute customers, the ability to train FMs more efficiently (and hence bring better products to market faster) is important to FM developers' ability to compete. A cloud provider must therefore be able to acquire SOTA AI accelerator chips in sufficient quantities.

8.18 We have seen some evidence that it is important to cloud providers to be able to access AI-specialist hardware (both chips and other hardware such as interconnects) in sufficient volume. For example, one cloud provider tracks in

[735] Financial Times (5 November 2023) 'Tech giants pour billions into cloud capacity in AI push' Tech giants pour billions into cloud capacity in AI push (ft.com), accessed 20 May 2024. Microsoft Q42023 investor earnings call, accessed 20 May 2024. Amazon quarterly earnings calls. Alphabet quarterly earnings calls, accessed 20 May 2024. One provider also submitted an internal document which states what each of the major cloud providers ([✂]) announced in their 2023 Q4 earnings calls, including expected increase in CapEx in 2024 driven by generative AI, [✂] response to CMA's information request [✂].

[736] While all the hardware and software components comprising AI compute – and the way they are configured – are important in influencing computational performance, AI accelerator chips are a particularly key component (this was noted, for example, by Amazon's CEO in Amazon's Q3 2023 earnings call, see page 4 03 - 20 - 2023 (q4cdn.com)). Accelerator chips are well-suited or specifically designed to accelerate the computation processes underpinning FMs. Accelerator chips are generally one of three types: general processing units (GPUs), field programmable gate arrays (FPGAs), or application-specific integrated circuits (ASICs). Developing (and in some cases, deploying) FMs typically requires thousands of accelerator chips working together – for example, it has been reported that OpenAI used a system of 25,000 GPUs to train its GPT models Analysts positive on Nvidia's ChatGPT, AI opportunity By Investing.com, accessed 19 April 2024.

[737] Note of meeting with [✂]. It has been estimated that building an AI model on non-SOTA AI chips would be at least 33 times more expensive than using SOTA AI chips. See Computational Power and AI - AI Now Institute, accessed 19 April 2024.

detail estimates of AI-related compute capacity being installed by its larger competitors.[738]

8.19 In the past couple of years, the supply of AI accelerator chips has not kept pace with the rapid growth in demand and we have seen some evidence that cloud providers and other customers of chip providers have found it challenging to access them in sufficient quantity and/or without long lead times.[739] [740] [741]

8.20 We understand that recent shortages of GPUs are a result of the recent surge in demand, coupled with the complexity of accelerator chip fabrication, meaning that supply cannot be quickly ramped up to meet demand.[742]

8.21 We have also seen some evidence that chip supply capacity is being expanded and is expected to continue expanding over the coming years. This should ease some of the constraint on the availability of AI accelerator chips.[743]

8.22 Nevertheless, there are currently limited options for cloud providers to acquire AI accelerator chips. The main sources of supply are: (i) established chip providers; and (ii) self-supply (of custom silicon).[744] We discuss each of these below.

---

[738] For example, in an internal document a cloud provider [✂] uses public statements and publicly disclosed financial information from other firms in the AI chip supply chain (designers and manufacturers) to estimate in some detail another cloud provider's [✂] capacity in AI compute. [✂] response to CMA's information request [✂].
[739] New York Times (16 August 2023) 'The desperate hunt for the A.I. boom's most indispensable prize' (link); WIRED (24 August 2023) 'Nvidia Chip Shortages Leave AI Startups Scrambling for Computing Power' (link). Dell APJ chief says the industry won't wait for Nvidia H100 • The Register, accessed 20 May 2024. Nvidia sold half a million H100 AI GPUs in Q3 thanks to Meta, Facebook — lead times stretch up to 52 weeks: Report | Tom's Hardware (tomshardware.com), accessed 20 May 2024. Note of meeting with [✂].
[740] [✂] response to CMA's information request [✂].
[741] [✂] response to CMA's information request [✂].
[742] The manufacturing level of the supply chain is highly concentrated. There are only three companies capable of manufacturing leading node chips, the Taiwan Semiconductor Manufacturing Company (TSMC), Samsung, and Intel. Of these, TSMC has been reported to have a 70-80% share of supply for leading-node chip production; see Computational Power and AI - AI Now Institute.
[743] Note of meeting with [✂]. It has been reported that TSMC plans to double its chip-on-wafer-on-substrate (CoWoS) capacity by the end of 2024 (CoWoS is the type of technology used for leading AI chips such as Nvidia's H100 and A100 and AMD's MI300). TSMC Boosts CoWoS Production 20% to Meet Surging Demand | Tom's Hardware (tomshardware.com), accessed 20 May 2024. It was reported in August 2023 that Nvidia intended to triple shipments of H100s to 1.5 million in 2024 (from 500,000 in 2023). Financial Times (August 2023) 'Supply chain shortages delay tech sector's AI bonanza, accessed 20 May 2024. Over the next few years, manufacturing capacity may also increase due to international industrial policy initiatives to build new semiconductor foundries ('fabs'): UK Government National Semiconductor Strategy; US' CHIPS and Science Act of 2022; European Union's European Chips Act of 2023.
[744] A third option is procuring chips from specialist chip start-ups (such as Graphcore, Tenstorrent, and Cerebras); however, their entry has had a limited impact on the market – we are not aware of any larger cloud providers currently using, or planning to use, their AI accelerator chips (see cloud providers' submissions in next paragraph). However, we are aware of some small cloud providers, such as Paperspace and Gcore, having a relationship with at least one chip start-up – see Note of meeting with [✂].

Supply from established chip design providers

8.23 Evidence from cloud providers shows that the majority of their accelerated compute capacity uses Nvidia's GPUs.[745, 746, 747, 748, 749, 750]

8.24 Cloud providers' internal documents also indicate the importance of being able to offer Nvidia chips at scale. [751]

8.25 Market participants have told us that a greater variety of high-performing chips are becoming available (or are expected to become available over the next few years).[752] For example:

(a) In December 2023, AMD announced several major customers for its new MI300X GPU, including Microsoft and Oracle.[753]

(b) Microsoft said that it is currently testing the new MI300X and has announced that it will use the product in its new Azure ND MI300x v5 virtual machine series optimized for AI workloads.[754] Microsoft also said that the 'launch of the MI300X accelerator series has been welcomed by the industry as a key step towards developing 'alternatives to the expensive NVIDIA graphics processors that have been essential for creating and deploying artificial intelligence programs''.[755]

(c) Intel's Gaudi series of accelerator chips supplies at least one major AI lab (Stability AI).[756]

---

[745] [✂] response to CMA's information request [✂].
[746] [✂] response to CMA's information request [✂].
[747] [✂] response to CMA's information request [✂].
[748] [✂] response to CMA's information request [✂].
[749] [✂] response to CMA's information request [✂].
[750] [✂] response to CMA's information request [✂].
[751] [✂] response to CMA's information request [✂].
[752] Notes of meetings with [✂]; [✂]; [✂]; [✂]; [✂].
[753] (19 September 2023) AMD Delivers Leadership Portfolio of Data Center AI Solutions with AMD Instinct MI300 Series, accessed 20 May 2024.
[754] Microsoft response to CMA's information request [✂].
[755] Microsoft response to CMA's information request [✂]. Meta and Microsoft to buy AMD's new AI chip as alternative to Nvidia (cnbc.com), accessed 20 May 2024.
[756] In September 2023, Intel announced Stability AI as an anchor customer for its Gaudi2-powered AI supercomputer: Intel Innovation 2023: Empowering Developers to Bring AI Everywhere :: Intel Corporation (INTC), accessed 20 May 2024. Stability AI has publicly stated that the Gaudi2's immediate availability, 'top-tier performance' at a more economical price point, and scalability motivated the move: Behind the Compute: Building the New AI Supercomputer — Stability AI, accessed 20 May 2024. Intel has announced plans to release the next-generation Gaudi3 chip in 2024: Intel shows Gaudi3 AI accelerator, promising quadruple BF16 performance in 2024 | Tom's Hardware (tomshardware.com), accessed 20 May 2024. While Intel built its AI supercomputer itself (ie it is not available on public cloud infrastructure), this demonstrates the viability of the Gaudi2 chip for at least one leading AI lab.

Self-supply of accelerator chips

8.26 Three cloud providers have developed their own customised AI accelerator chips, known as ASICs or 'custom silicon'.[757]

(a) AWS offers access to additional chip options 'Trainium' for AI training workloads and 'Inferentia' for deploying AI.[758] Amazon's Trainium and Inferentia chips are available only on AWS. External customers include Anthropic[759] and Snap.[760, 761, 762] AWS said that it has invested a total of [✂] over five years (2019-23) in order to develop Trainium, Inferentia, and the accompanying Neuron software.[763]

(b) Microsoft announced its first custom AI accelerator chip, an ASIC called the 'Maia 100', in November 2023.[764] In 2024 the Maia 100 will power inference on Microsoft's services such as Copilot and Azure OpenAI Service.[765] Microsoft said that, going forward, it expects to [✂].[766]

(c) Microsoft said that between July 2021 and June 2024 (including forecast spend), it expects to have spent in total $[✂] in developing Maia.[767]

(d) Google has developed a series of ASICs called TPUs, designed for machine learning workloads, including FM training and inference.[768,769] TPUs are available only on the Google Cloud Platform. While the chips are used predominantly by Google for its internal workloads (such as training its PaLM and Gemini models), they are also used by external FM developers including AI21,[770] Anthropic[771] and Midjourney.[772]

(e) A Google internal document which benchmarks itself against other hyperscalers' commercialisation of AI, states that 'GCP has also differentiated itself through the release of its Tensor Processing Units (TPUs) These chips complement GCP's compute offerings, offsetting demand for

---

[757] The manufacture of the chips is contracted out to fabs such as TSMC.
[758] AWS response to CMA's information request [✂].
[759] AWS AI chips powering Amazon's partnership with Anthropic (aboutamazon.com), accessed 20 May 2024.
[760] AI Chip - AWS Inferentia - AWS (amazon.com), accessed 20 May 2024.
[761] [✂].
[762] AWS response to CMA's information request [✂]. [✂].
[763] AWS response to CMA's information request [✂].
[764] Microsoft response to CMA's information request [✂]. Microsoft Azure delivers purpose-built cloud infrastructure in the era of AI | Microsoft Azure Blog, accessed 20 May 2024.
[765] Microsoft response to CMA's information request [✂]. Microsoft Azure delivers purpose-built cloud infrastructure in the era of AI | Microsoft Azure Blog, accessed 20 May 2024.
[766] Microsoft response to CMA's information request [✂].
[767] Microsoft response to CMA's information request [✂]. Note these figures include the research, development and operating expenditure costs of the teams [✂]. Data for earlier years is not readily available.
[768] Tensor Processing Units (TPUs) | Google Cloud, accessed 20 May 2024. [✂], [✂] response to CMA's information request [✂].
[769] Google has also developed the 'Edge TPU' for inference at the edge (ie on device, rather than in the cloud). Edge TPU - Run Inference at the Edge | Google Cloud, accessed 20 May 2024.
[770] Building the most open and innovative AI ecosystem | Google Cloud Blog, accessed 20 May 2024.
[771] Cloud TPU v5e is generally available | Google Cloud Blog, accessed 20 May 2024.
[772] Building the most open and innovative AI ecosystem | Google Cloud Blog, accessed 20 May 2024.

standard GPU accelerators used in most AI workloads, and offering even more premium performance in most advanced neural net / LLM scenarios.'[773]

8.27 We have heard from other cloud providers that high levels of investment and lengthy timeframes are required to self-supply accelerator chips.[774]

8.28 Self-supply also requires development of the accompanying software (that is used to programme the AI accelerator chips), which we understand is highly challenging and chip suppliers have told us requires significant investment.[775]

8.29 Finally, we note that, while FM providers currently require substantial amounts of specialist compute resources, we have seen some evidence that there is a drive to develop smaller models that can still exhibit the capabilities required but require fewer resources (including compute) to develop and deploy.[776] Smaller and/or more efficient models would require less compute from cloud providers.[777]

Partnerships between cloud providers and FM developers

8.30 We have seen evidence that some cloud providers – including AWS, Google, and Microsoft – are entering into partnerships with, and making investments in, some FM developers.[778]

8.31 These partnerships often enable the FM developer to access scarce inputs, including accelerated compute, from the cloud provider.[779] In some cases, these partnerships may allow the cloud provider to add the partner's model(s) to their library or provide access to the partner's model(s) via their developer tools.[780]

8.32 The recent growth of such partnerships may indicate one emerging feature of competition in AI between cloud providers. For example, as noted above, many customers we heard from have told us that providers' AI capabilities are becoming

[773] Google response to CMA's information request [✂].
[774] [✂]. Responses to the CMA's information requests [✂]; [✂]; Notes of meetings with [✂].

[776] Notes of meetings with [✂].
[777] See CMA AI Foundation Models: Technical Update Report (16 April 2024) under 'FM development'. AI Foundation Models: initial review - GOV.UK (www.gov.uk)
[778] See Figure 7 in the CMA's AI Foundation Models: Technical Update Report (16 April 2024) AI Foundation Models: initial review - GOV.UK (www.gov.uk). Examples include Amazon's partnership with Anthropic, Google's partnership with Anthropic, and Microsoft's partnership with OpenAI. These partnerships and investments take on a wide range of structures, including those with and without an equity stake.
[779] Microsoft (2023) Microsoft and OpenAI extend partnership, accessed 20 May 2024. Anthropic (2023) Expanding access to safer AI with Amazon, accessed 20 May 2024. Anthropic (2023) Anthropic Partners with Google Cloud, accessed 20 May 2024.
[780] Examples include Amazon and Cohere, Amazon and Hugging Face, Google and Mistral, and Microsoft and Meta. Amazon (2023) Cohere brings language AI to Amazon SageMaker, accessed 20 May 2024. Reuters (13/12/2023) Google Cloud partners with Mistral AI on generative language models, accessed 20 May 2024. Microsoft (2023) Microsoft and Meta expand their AI partnership with Llama 2 on Azure and Windows, accessed 20 May 2024. Amazon Hugging Face on Amazon SageMaker, accessed 20 May 2024.

an increasingly important consideration.[781] These partnerships may be one way in which cloud providers compete to expand the AI capabilities that they offer.

Potential opportunities for smaller cloud providers to provide accelerated compute

8.33 The growth in demand for accelerated compute from customers may create new opportunities for competition amongst cloud providers if they can gain early access at scale to the specialist hardware needed.

(a) Some technology industry reports have characterised Oracle as a 'surprise success story in this segment',[782] an 'unlikely friend' to AI start-ups and an 'improbably early leader in the race to rent servers to these start-ups'[783] due to factors including its relatively late entry to cloud which enabled it to specifically configure its infrastructure to suit AI workloads.

(i) In line with this, Oracle's internal documents mention [✂].[784]

(ii) Microsoft said that in May 2023 it extended its pre-existing partnership agreement with Oracle to cover managed access to Oracle's Nvidia GPU compute capacity, which is being used to run inference on models that power Bing conversational searches (now Copilot).[785]

(b) There has been entry by smaller, specialist cloud providers (ie specifically offering compute to AI developers), including CoreWeave, Lambda Labs, and a number of others, which provide access to Nvidia's market-leading GPUs.[786]

(c) CoreWeave and Lambda Labs were two of the first cloud providers to get general access to Nvidia's H100 chip when it launched in 2023.[787] Both have supplied this compute capacity to Microsoft:

[781] Responses to the CMA's information requests [✂].
[782] Computational Power and AI - AI Now Institute, accessed 20 May 2024.
[783] AI Startups Find an Unlikely Friend: Oracle — The Information, accessed 20 May 2024. This article also notes that AI start-ups may be attracted to Oracle (over Azure, AWS or GCP) as it is not a rival in FM development.
[784] Oracle response to CMA's information request [✂].
[785] Microsoft response to CMA's information request [✂].
[786] AWS also identified Denvr Dataworks, G42, and Omniva AWS response to CMA's information request [✂]. Microsoft also identified Scaleway, Hewlett Packard Enterprise, Together AI, Nvidia (its DGX Cloud platform), LaminiAI, Paperspace, Crusoe Cloud, Cirrascale, Clever Cloud, Denvr Dataworks, FluidStack, Runpod, Oblivus, Vast.ai, Tensordock, TensorWave, Modal, and OctoAI. Microsoft response to CMA's information request [✂].
[787] Computational Power and AI - AI Now Institute, page 32, accessed 20 May 2024. Coreweave has also received a $100 million investment from Nvidia (ibid.). CoreWeave came 'out of nowhere.' Now it's poised to make billions off AI with its GPU cloud | VentureBeat, accessed 20 May 2024. Lambda Cloud Deploys NVIDIA H100 Tensor Core GPUs (lambdalabs.com), accessed 20 May 2024. See also [✂] response to CMA's information request [✂].

(i) Microsoft said that in February 2023 it signed an agreement with CoreWeave for the supply of Nvidia A100 and H100 accelerator capacity with InfiniBand networking in US data centres.[788]

(ii) Microsoft said that in April 2023 it signed an agreement with Lambda Labs for the supply of Nvidia H100 accelerator capacity deployed across various data centres in Lambda's network.[789]

Emerging views on AI and public cloud infrastructure services

8.34 Based on the evidence to date, our emerging view is that:

(a) Cloud providers are an important source of accelerated compute capacity for organisations developing and/or deploying FMs. These organisations are in turn becoming an increasingly important source of revenue for cloud providers and this is expected to continue. Partnerships between the large cloud providers and FM model developers are extensive and are likely to play an important role going forwards.

(b) Accelerator chips are vital for cloud providers seeking to supply accelerated compute for FM development and/or deployment. A shortage of these chips has meant that cloud providers have been competing to secure them. The three largest cloud providers have also invested in self-supply.

(c) We are still gathering and assessing evidence on this, which will enable us to judge whether the growth in AI workloads will allow smaller cloud providers to expand or whether it will result in a further barrier to entry and expansion.

8.35 We welcome the submission of any further evidence on any of the points above, and on any other aspects of the impact of AI (either in terms of the supply of accelerated compute or AI services provided by cloud providers) on the nature of competition in cloud services, and whether competitive conditions in relation to the supply of accelerated compute may differ to competitive conditions more generally in cloud services and the implications of any differences.

[788] Microsoft response to CMA's information request [✂]. Reportedly, this agreement was worth billions of dollars, and was made in order to secure adequate compute access to meet the needs of Microsoft's partner, OpenAI Microsoft inks deal with CoreWeave to meet OpenAI cloud demand (cnbc.com), accessed 20 May 2024.
[789] Microsoft response to CMA's information request [✂].

9. Overview of our emerging views

9.1 This section sets out an overview of the emerging views set out in this paper and our initial conclusions based on these.

9.2 In this paper we have considered:

(a) the competitive landscape of cloud services including: (i) the nature of the customer base and trends in the usage of public cloud; (ii) the customer journey; (iii) customer preferences and parameters of competition between public cloud providers; and (iv) the main providers focusing on the vertically integrated suppliers of cloud services (which we refer to as 'cloud providers');

(b) how customers switch and multi-cloud, including the prevalence of this;

(c) the relevant markets that we are considering as part of this investigation;

(d) shares of supply relating to the markets that we are considering as part of this investigation;

(e) market outcomes including profitability;

(f) barriers to entry and expansion in cloud services; and

(g) the potential impact of AI on cloud services.

Nature of competition

9.3 In relation to the competitive landscape, the evidence we have seen to date suggests that cloud services are increasingly important inputs to many businesses and organisations across the UK economy and across a range of different industries.

9.4 A relatively small number of high-spend customers account for a large proportion of cloud providers' UK revenues and a relatively large number of low-spend customers are responsible for a small proportion of their revenue. In particular, the top 10% of customers account for a very large majority of revenues and the top 1% account for over half of revenues.

9.5 Customers buy the large majority of cloud services directly from cloud providers. Most customers have standard contracts that have been agreed without negotiation, but larger customers either engage in bilateral negotiations or tenders and are able to negotiate terms that depart from standard contracts.

9.6 There are different models of multi-cloud use, and we cannot accurately measure the full extent of switching by customers, nor the extent to which customers use multiple clouds. However, the evidence suggests that, while there is some degree

of multi-cloud use, it may be quite limited in scope and mostly found amongst larger customers. We have also not seen strong evidence that switching between cloud providers is common. We are continuing to consider the evidence on the prevalence of switching and use of multiple public clouds by customers.

9.7 Cloud providers compete on a range of factors and the factors that seem to be the most important to customers when choosing their main public cloud provider are service quality, price (including discounts or cloud credits), data sovereignty requirements, range of services and the number and location of data centres.

9.8 The supply of cloud services has evolved over time: AWS launched the first public cloud service in 2006; Microsoft, Google and others then entered, and cloud services grew as cloud providers competed to attract customers to the public cloud for the first time.

9.9 In recent years the markets have grown significantly both in terms of revenues (UK IaaS and PaaS revenues more than doubled during the period from 2019 to 2022) and datacentre capacity in the UK and Europe (more than doubling in both between 2020 and 2023).

9.10 The evidence suggests that the number of customers moving to public cloud for the first time has decreased and in 2021 and 2022 new customers made up a small proportion of year-on-year growth. Demand for cloud services is growing because existing customers are expanding their existing workloads and/or adding new workloads. If existing customers face barriers to switching and multi-cloud, then the strength of competition is likely to be weaker than when competition was focused on customers moving to public cloud for the first time. We will consider the extent to which there may be particular barriers to switching and multi-cloud in our later working papers.

Market definition

9.11 In relation to the relevant markets, the evidence we have seen to date suggests that there is a relevant product market for the supply of IaaS, but PaaS is not part of the same relevant market and, where relevant, PaaS would be considered as an out-of-market constraint. We have considered the extent to which there is a relevant product market for the supply of PaaS and the evidence on the extent to which PaaS and SaaS are substitutable is mixed and limited.

9.12 While some large customers of public cloud services may be able to react to a price increase by switching to private cloud or traditional IT, the evidence to date indicates that, even for large customers, any such switching would be unlikely due to the specific reasons they place workloads on public cloud and the costs and time associated with doing so. Therefore, our emerging view is that traditional IT

and private cloud should be considered as out-of-market constraints where applicable.

9.13 The geographic scope of these relevant markets is likely to be Europe-wide (ie UK and EEA).

Shares of supply

9.14 In relation to the relative position of suppliers in the market, the evidence to date suggests that across the metrics we have considered (shares by revenue, capacity and flows of new business) AWS and Microsoft are the largest two cloud providers and significantly larger than Google, the next largest public cloud provider.

9.15 Both the IaaS and PaaS markets are concentrated and both have become more concentrated from 2019 to 2022 as the collective shares of AWS and Microsoft have increased.

(a) In IaaS, AWS is the largest cloud provider with a share of [40-50]% in 2022; Microsoft is the second large cloud provider with a share of [30-40]%. Both were much larger than Google's share of [5-10]%. In addition, the combined share of AWS and Microsoft increased over time from [✂] in 2019 to [✂] in 2022.

(b) In PaaS, AWS is the largest cloud provider with a share of [20-30]% [✂] in 2022; Microsoft is the second large cloud provider with a share of [20-30]% in 2022. Both were much larger than Google's 2022 share of [5-10]%. In addition, the combined share of AWS and Microsoft increased over time from [✂] in 2019 to [✂] in 2022.

(c) Based on the available capacity data, [✂].

9.16 Many large customers do not see any suitable alternatives to AWS and Microsoft as their main cloud provider. They do not perceive other smaller providers to have comparable offerings to AWS and Microsoft, albeit Google is perceived as being closer than Oracle and IBM and other smaller providers. Smaller providers may still be seen as suitable alternatives for certain workloads as they all have strong offerings in relation to certain segments or types of customer.

Market outcomes

9.17 In relation to market outcomes, our assessment of the profitability of cloud providers, based on evidence received to date, indicates that AWS and Microsoft have been generating returns from their cloud services above their cost of capital, and that this could be expected to continue in the future.

9.18 Our assessment of prices, quality and innovation is less developed and requires careful interpretation for the reasons given in Section 6. For example, it is particularly difficult to understand what the counterfactual may be in relation to quality and innovation and whether the current level of quality or innovation reflects the level of quality or innovation that would be expected in a competitive market.

Barriers to entry and expansion

9.19 In relation to barriers to entry and expansion, the evidence to date suggests that economies of scale (which include high levels of capital investment), the importance of the range of services, economies of scope and to a lesser extent network effects represent a significant barrier to entry and expansion in the public cloud infrastructure services market. In particular:

(a) The supply of IaaS requires significant capital investment in fixed assets. These investments are mainly sunk costs that would not be recovered in full on exit. While the upfront investments can be reduced by co-locating or leasing data centres, a new entrant or a competitor looking to expand would still need to invest in the servers, components and network equipment for a co-located or leased data centre.

(b) The largest cloud providers are planning significant further investment in their cloud infrastructure, in particular in accelerator chips (eg GPUs), to meet growing demand for AI services. This increases the capital investment required by a new entrant, should they choose to offer customers accelerated compute capacity.

(c) Large cloud providers benefit from economies of scale such as benefiting from bulk purchasing discounts on necessary equipment and from more efficient use of energy associated with large data centres and more efficient use of assets by targeting higher rates of utilisation.

(d) There is evidence that having a large portfolio of cloud services gives cloud providers advantages over their competitors. In particular, the range of first party products is an important factor for customers when choosing which cloud providers to use, there may be some economies of scope in supplying a range of services and there may be some indirect network effects between ISVs and customers.

Impact of AI on competition in cloud services

9.20 In relation to AI and public cloud infrastructure services, the evidence to date suggests that cloud providers are an important source of accelerated compute capacity for organisations developing and/or deploying Foundation Models (FMs),

who in turn are becoming an increasingly important source of revenue for cloud providers and this growth is expected to continue. Partnerships between the large cloud providers and FM model developers are extensive and are likely to play an important role going forwards.

9.21 Accelerator chips are vital for cloud providers seeking to supply accelerated compute for FM development and/or deployment. A shortage of these chips has meant that cloud providers have been competing to secure them. The three largest cloud providers have also invested in self-supply of accelerator chips.

9.22 We are still gathering and assessing evidence on this which will enable us to judge whether the growth in AI workloads will allow smaller cloud providers to expand or whether it will result in a further barrier to entry and expansion.

Conclusions

9.23 The evidence and analysis set out in this working paper, alongside our other working papers, will inform our assessment of whether one or more cloud providers hold significant market power.[790]

9.24 Based on the evidence we have seen to date, our emerging view is that there are indicators of significant market power being held by the largest two providers, AWS and Microsoft. This is because:

(a) They both have high market shares and the collective share of all other providers in these markets is falling;

(b) Potential rivals face significant barriers to entry and expansion, including high levels of capital investment and economies of scale and scope; and

(c) Whilst assessing current market outcomes is complex given the current stage of market development, our profitability assessment indicates that AWS and Microsoft have both been generating returns above their cost of capital.

9.25 We will continue to investigate the extent to which any such market power is likely to endure into the future.

9.26 Our evidence gathering and assessment is ongoing and we will consider any further evidence from market participants on the contents and emerging views set out in this working paper.

[790] The term ‘market power’ is used to denote the ability of a firm to influence aspects of competition; there are gradations of market power with many firms having limited or transitory market power and others having ‘significant market power’ which endures over time. CC3, paragraph 9.

Appendices

Appendix A: Multi-cloud and switching prevalence

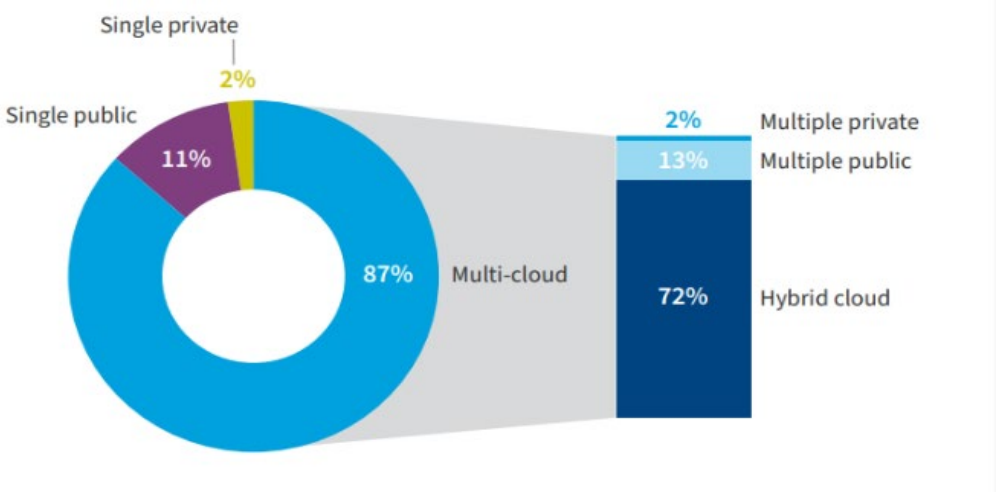
Public surveys and Ofcom research – multi-cloud

A.1 The section sets out the results of public surveys submitted to us by cloud providers and the research for Ofcom. In addition to the limitations of these set out in Section 3 above, we also identify specific limitations of each survey.

Flexera 2023 State of the Cloud

A.2 Two providers submitted that the results of from the Flexera 2023 State of the Cloud report[791] indicate a high prevalence of multi-cloud.[792] We present the results in Figure 9.1 and 9.2. We consider that these results should be interpreted with caution. In particular, there are uncertainties in how respondents have interpreted the questions, making assessing the relevance or usefulness of it very challenging. We do not have access to the underlying data and do not know the extent to which the questionnaire was cognitively tested with respondents, to ensure the questions are understood and being answered as intended. We think this type of testing and research process is particularly important in this technical market, where language and terms can vary significantly depending on customers experience of using public cloud infrastructure services. For example, it is possible respondents may have recorded themselves as using multiple public clouds if they used both first and third party PaaS, but the same underlying IaaS provider.

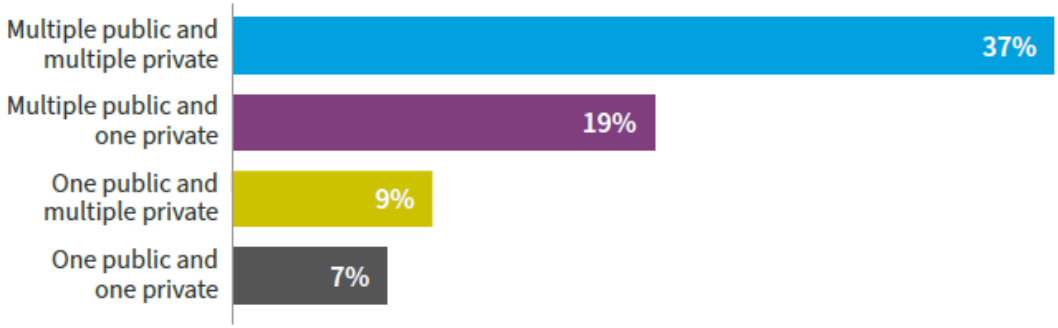
Figure 9.1: Multi-cloud prevalence estimated by Flexera 2023 State of the Cloud report



Source: Flexera 2023 State of the Cloud

[791] Flexera 2023 State of the Cloud Report Flexera-State-of-the-Cloud-Report-2023.pdf
[792] [✂]; [✂].

Figure 9.2: Hybrid cloud strategies estimated by Flexera 2023 State of the Cloud report



Source: Flexera 2023 State of the Cloud

A.3 Figure 9.1 shows that customers that use only private clouds are included in the sample. For the purposes of this market investigation, we are concerned with the prevalence of multi-cloud among those that use public cloud. As such, we can remove customers that are only using private clouds from the sample (2% of respondents use a single private cloud and 2% of respondents use multiple private clouds).

A.4 Further, in the Flexera report, multi-cloud is defined as using at least two clouds, regardless of whether the clouds are public or private. In contrast, we define multi-cloud as using at least two public clouds. Figure 9.2 shows that we are therefore interested in the 37% that use multiple public clouds and multiple private clouds, and the 19% that use multiple public and one private cloud. The remaining respondents we consider as customers that are only using a single public cloud.

A.5 Overall, using our definition of customers that multi-cloud, the Flexera report suggests that 72% of customers multi-cloud.

Oracle's survey

A.6 Two providers submitted that the results from Oracle's "Multi-cloud in the Mainstream" survey[793] show that there is a high prevalence of multi-cloud.[794] We presented the results in Figure 9.3 below.

A.7 As above, we consider that these results should be interpreted with caution because respondents may not have interpreted the question in a manner that is consistent with our definition of multi-cloud.

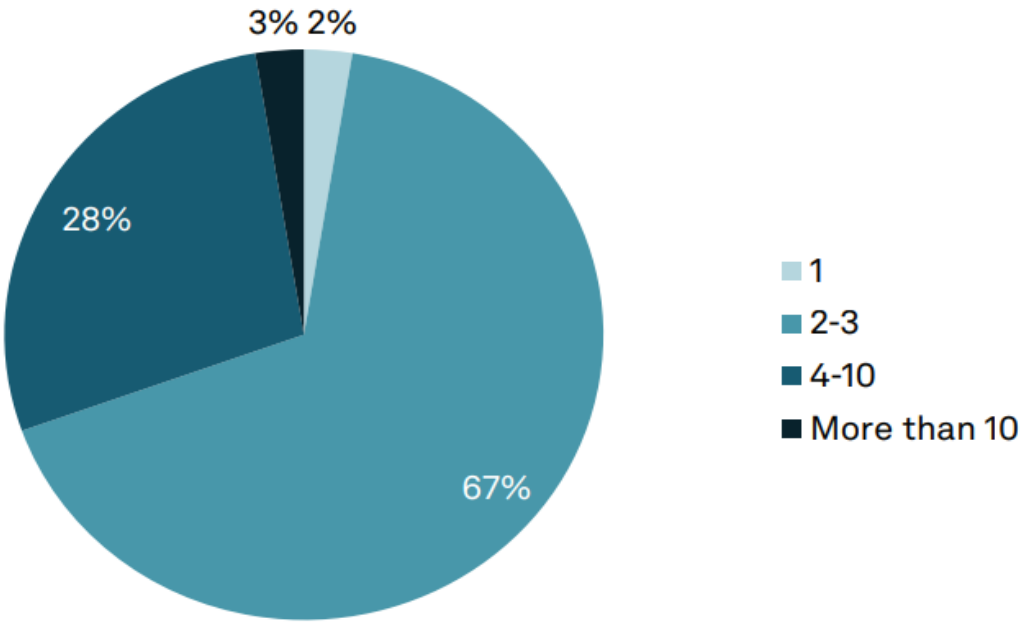
A.8 In particular, we note that the survey does not distinguish between use and planned use, and we cannot split the two out. It is unclear based on the question asked what the time frame over which customers plan to start using multiple clouds is and we do not know how the likelihood of that planned usage translates into actual usage. It is

[793] S&P Global Discovery Report: Multicloud in the Mainstream (oracle.com)
[794] [✂]; [✂].

feasible that customers encounter barriers to using multiple clouds that prevents them from enacting their plans.

A.9 Further, we note that respondents were asked 'How many IaaS/PaaS public cloud providers (such as AWS, Azure, Google Cloud Platform [GCP], etc.). do you use/plan to use in your IT environment?'. We consider this wording to be ambiguous, especially in the context of a quantitative survey where one cannot follow up with respondents for details, and respondents may have interpreted using multiple clouds as using one cloud provider for IaaS in addition to an ISVs. We also consider that this is consistent with the relatively high proportion of respondents that said that they use four or more providers.

Figure 9.3: Oracle's Multi-cloud in the Mainstream, Number of current or planned cloud infrastructure providers (IaaS/PaaS)



Source: Oracle's Multi-cloud in the Mainstream (figure 1)

Public First survey

A.10 One provider submitted that the results from Public First[795] show a high prevalence of multi-cloud and submitted further analysis of the result from the survey.[796] We set out the results from the Public First survey below. We consider that the results should be interpreted with caution and therefore we do not consider it appropriate to place evidentiary weight on the answers.

A.11 The Public First survey asked respondents 'How many different cloud infrastructure providers does your company currently use?'. We consider that this

[795] [✂].
[796] [✂].

phrasing is ambiguous and may capture multi-vendor architectures, hybrid cloud or multiple private clouds, which are not relevant for our investigation. In fact, we believe that this survey is particularly likely to be capturing hybrid cloud or multiple private clouds because of the prevalence of private cloud users in the study: only 17% of IaaS and PaaS respondents used public cloud only; compared to 32% that used private cloud only and 48% that used both private and public cloud.

A.12 The Public First survey reported that 71% of IaaS/PaaS respondents used more than one cloud infrastructure provider. Further, 56% of IaaS/PaaS respondents said that they would be likely to add an additional cloud infrastructure provider in the next few years.

A.13 In terms of where respondents lie along the siloed to integrated multi-cloud spectrum, the Public First survey reported that 52% of respondents that used more than one cloud had “a mix of integrated and independent” cloud architectures, and 10% of these respondents had clouds that were “largely integrated together”.

A.14 We have received a further submission from a cloud provider relating to the Public First survey, which we are considering further.

Ofcom survey

A.15 Ofcom commissioned a quantitative survey during its market study on the public cloud market. We consider that the results should be interpreted with caution and that no evidential weight should be put on them.

A.16 The agency, Context Consulting, conducted 20-minute online self-completion survey with 1,004 respondents that were existing users of IaaS and/or PaaS, or that were considering adoption within 12 months. The sample source was an online access panel, which is not in line with the CMA’s good practice on survey research because they do not rely on randomisation methods for sampling, and therefore may contain bias. Further, we consider that the limitations of quantitative surveys identified in Section 3 above apply to the Context Consulting quantitative survey.

A.17 In addition, we believe that the results may include respondents that are not solely using multi-vendor architectures or hybrid cloud. In particular, the survey asked whether respondents use more than one IaaS/PaaS provider and did not distinguish between whether those providers are the same public cloud or not.

A.18 The Context Consulting survey reported that 61% of respondents use more than one IaaS/PaaS provider, when respondents that only use private cloud are excluded. The survey also asked whether respondents would potentially use multi-cloud architecture in the future: the majority of respondents said that they would definitely (23%) or possibly (63%) multi-cloud in the future.

Prevalence of multi-cloud – methodology

A.19 This section sets out the methodology for our analysis of cloud provider data to assess the prevalence of multi-cloud. Specifically, we consider the prevalence of multi-cloud in 2020, 2021 and 2022 between customers of the three largest cloud providers, AWS, Microsoft, and Google.

A.20 This analysis builds off the customer matching exercise, the methodology and limitations of which are described in paragraphs 3.64 to 3.71. As such, it contains the same assumptions and procedural points (eg the $1,000 spend threshold).

Methodology, results, and caveats

Number of unique UK customers of cloud

A.21 We first append the 'cleaned' list of customers for each cloud provider in the years 2020-2022 generated as part of the customer matching exercise. This creates a large dataset containing every customer of the three large providers over this period.

A.22 We then scan the dataset for duplicates by name and year, then drop these duplicates – that is, if a combination of name and year appears multiple times in the dataset, we only keep one of these observations as we want a list of unique customers of cloud services in each year.

A.23 Collapsing the dataset by customer name and year then gives us a unique observation for each customer in each year they were active between the three large cloud providers, with an annual revenue figure for each. Where they were not a customer of a particular cloud provider, their revenue shows as 0.

A.24 We can count the number of observations in each year to find the number of unique customers of cloud services in the UK. These results are shown in Table 9.1.

Table 9.1: Number of unique UK cloud customers across the three large providers, 2020-2022

Year	*Number of unique UK cloud customers across the three large providers*
2020	38,969
2021	45,855
2022	52,454

Source: CMA analysis of AWS, Microsoft, and Google data

A.25 These results are likely to be overestimates of the true number of unique UK cloud customers. This is because at this stage we have not engaged in the fuzzy matching process described at paragraph 3.65 such that there will be some customers counted twice in the table above because different providers recorded their name in slightly different ways. We do not believe that this is a large overestimate as the number of fuzzy matches as a proportion of the entire customer database is fairly low.

Number of customers using multiple clouds

A.26 To estimate the number of customers using multiple clouds amongst AWS, Microsoft and Google clouds, we first load in and append the three 'pairwise' datasets generated as part of the customer matching exercise outlined in paragraphs 3.66. Each of these datasets contains customer names, years, revenues, and a similarity score rating the similarity of fuzzy matches across the two providers.

A.27 As this exercise relies on the customer matching exercise, the caveats and limitations outlined in paragraph 3.71 apply.

A.28 For each 'pairwise' dataset, observations with a similarity score of less than 0.99 are dropped so that we are left only with customers that are either exact matches or strong fuzzy matches. This is to try and capture customers where their names do not correctly match across cloud providers, but to reduce the risk of 'false' matches (see paragraph 3.72(a)). Any observations outside of the years 2020-22 are also dropped on the basis that we only had data for all of the providers for these years.

A.29 We then scan the dataset for duplicates by name and year, then drop these duplicates – that is, if a combination of name and year appears multiple times in the dataset, we only keep one of these observations as we want a list of unique customers of cloud services using multiple clouds in each year.

A.30 Collapsing the dataset by customer name and year then gives us a unique observation for each customer in each year they were active in at least two of the three large cloud providers, with an annual revenue figure for each. Where they were not a customer of a particular cloud provider, their revenue shows as 0.

A.31 We can count the number of observations in each year to find the number of unique UK customers using at least two of the three large providers' clouds between 2020 and 2022. Table 9.2 shows the results.

Table 9.2: Number of unique UK cloud customers using multiple clouds, 2020-2022

Year	*Number of unique UK cloud customers using multiple clouds*
2020	2,591
2021	3,246
2022	3,710

Source: CMA analysis of AWS, Microsoft, and Google data

A.32 Table 9.2 shows the number of UK customers using two or more of the three largest cloud providers' clouds. This is likely to be an underestimate of the true figure as not all accurate fuzzy matches will have been captured due to our similarity score threshold of 0.99, although as noted in Section 3 at this threshold some false positives (ie 'false' matches) are included. It may also be the case that some customers multi-cloud outside of the three largest providers (eg with AWS and Oracle) which would not be captured in this analysis.

A.33 Table 9.3 shows the unweighted proportion of UK customers who multi-cloud from 2020 to 2022 using the results from Table 9.1 and Table 9.2.

Table 9.3: Unweighted proportion of UK customers using multiple clouds, 2020-2022

Year	*Unweighted proportion of UK customers using multiple clouds*
2020	6.7%
2021	7.1%
2022	7.1%

Source: CMA analysis of AWS, Microsoft and Google data

A.34 Table 9.3 shows that 6.7% of UK customers used multiple clouds in 2020, rising to 7.1% in 2021 and 2022, unweighted by cloud spend. This is likely to be an underestimate of the true prevalence of multi-cloud as our denominator is likely an overestimate, as outlined in A.25, and because we miss examples of multi-cloud where a smaller cloud provider (ie not AWS, Microsoft, or Google) is involved.

A.35 To understand how using multiple clouds is influenced by total cloud spend, we also calculate the proportion of UK customers using multiple clouds weighted by cloud spend. Customers with higher total annual spend on cloud will be given more importance, and therefore if customers that spend more use multi-cloud more the weighted prevalence of multi-cloud will be higher than the unweighted prevalence.

A.36 Table 9.4 shows weighted proportion of UK customers who multi-cloud from 2020 to 2022.

Table 9.4: Weighted proportion of UK customers using multiple clouds, 2020-2022

Year	*Weighted proportion of UK customers using multiple clouds*
2020	31.3%
2021	34.0%
2022	34.4%

Source: CMA analysis of AWS, Microsoft, and Google data

A.37 Table 9.4 shows that, weighted by cloud spend, 34.4% of UK customers use multiple clouds. Multi-cloud appears much more common when weighted by cloud spend, suggesting that those with higher spends on cloud are more likely to use cloud services from multiple providers.

A.38 This is likely an underestimate of the true weighted prevalence of multi-cloud for the reasons outlined in paragraph A.34.

Public surveys and Ofcom research – switching

Public Surveys

Public First

A.39 We believe that very limited weight should be put on the results from the Public First survey. This is because of the limitations to the survey that we set out in Section 3,

and because, in addition, we consider that the switching question from the survey may be capturing types of switching that are not relevant to our investigation. In particular, the question asked 'Have you ever switched one of your cloud infrastructure providers in the past?'. Respondents may have included:

(a) Switching between private and on-premises solutions;

(b) Switching within public clouds – ie. between first- and third parties hosted on the same cloud; and/or

(c) Partial switching – ie. only switching a few workloads away.

(d) The question asked is phrased as 'Have you ever switched one of your cloud infrastructure providers in the past' (emphasis added) which creates uncertainty over interpreting the precise time period for this switching data.

A.40 We note that partial switching is relevant in maintaining competitive pressure for existing workloads. However, it is not clear if respondents would have the ability to switch all workloads, or are only able to switch sufficiently siloed workloads. If the former is the case, cloud providers have an incentive to compete for all existing workloads. However, if the latter is the case, cloud providers only have an incentive to compete strongly for siloed existing workloads. We are unable to assess the true situation given the data from this survey.

A.41 Despite the limitations set out above, we present the results from the Public First survey:

(a) 26% of IaaS/PaaS users said they had switched; 35% had considered switching, but had not switched in the end; 36% of IaaS/PaaS users had not considered switching.

(b) 51% of IaaS/PaaS users reported that they were likely to switch cloud providers in the future; 20% of IaaS/PaaS users were unlikely to switch in the future. The remaining 29% said that they were neither likely or unlikely to switch in future.

A.42 We have received a further submission from a cloud provider relating to the Public First survey, which we are considering further.

Context Consulting

A.43 We set out the methodology to Context Consulting's quantitative research in paragraph A.16 above.

A.44 In addition to the limitations set out in Section 3, we also note that respondents may not have understood the switching question. In particular, respondents may have

included switching within the same public cloud (ie. switching between first- and third-party services within a public cloud) or switching between on-prem/private IT and public cloud. Further, no time frame was given for switching, making it difficult to interpret the results. Overall, we do not consider it appropriate to place evidential weight on the results of this quantitative research. Nevertheless, we set out the results below.

A.45 The Context Consulting quantitative research reported that 21% of respondents said they had switched IaaS/PaaS providers in the past, 29% said they took on an additional IaaS/PaaS provider, 32% considered switching but did not switch, and 21% never considered switching.