Appendix D: Market structure and concentration methodology

Introduction

D.1 This appendix provides additional detail on the methodology for calculating shares of supply as set out in Chapter 3.

D.2 We present the methodology used to calculate shares of supply by revenue, including;

(a) the categorisation of provider revenues and limitations of provider data; and

(b) the calculation of the market size.

D.3 We also present notes about the data used to calculate:

(a) shares of supply by capacity; and

(b) shares of supply by flows of new business.

Shares of supply by revenue

Categorisation of provider revenues and limitations of provider data

D.4 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.[1] We reviewed the mapping of cloud services into IaaS, PaaS and SaaS used by Ofcom and adopted this in order to map individual cloud services or service categories into IaaS, PaaS and SaaS.

D.5 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. It therefore made certain adjustments to reconcile data across these systems.[2] We have adjusted the IaaS and PaaS totals based on the individual revenues by cloud service from one system such that the cloud provider's combined total matches the aggregate revenues from another system. This methodology had the effect of increasing the cloud provider's shares by less than [✂] in each category relative to just using revenues based on individual cloud services.

(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,

[1] We converted AWS, Microsoft, Google, IBM, [✂], Upcloud and Wasabi revenue data from USD to GBP using the Bank of England's average annual exchange rates for each year respectively. One Upcloud data point was submitted in Euros. For this we used the corresponding EUR to GBP exchange rate.
[2] [✂] response to the CMA's information request [✂].

which ends on 31 May.[3] 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.[4] [✂].[5] 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) We did not have access to the same level of information about each individual cloud service for some of the small UK IaaS providers (UpCloud, Civo, [✂], Hyve, Wasabi and Centerprise) as we did for AWS, Microsoft and Google. Therefore, for some providers we conducted desktop research to determine each cloud service's purpose and accordingly which subcategory it should fall into. We then applied the same mapping into IaaS, PaaS and SaaS as outlined above. As the small UK IaaS providers cumulatively totalled [<1%] of IaaS and PaaS combined revenues in 2023, any differences in categorisation into subcategories, and therefore IaaS and PaaS based on misinterpretation of cloud services' purpose should not have a material effect on the final shares.

(e) Some providers included some revenues from IaaS based on accelerated compute which could not always be distinguished in the Parties' data, but we understand the contribution of accelerated compute to the total revenues to be small.

D.6 Across all providers, 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.

(a) 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.[6]

(b) One provider identified UK revenue as revenues that were billed to the customers' billing address(es) in GB.[7]

(c) Three providers identified UK revenues as generated by customers with billing addresses in the UK.[8]

Calculation of the market size

D.7 In order to calculate the market size for IaaS and PaaS, we procured data from IDC and Synergy. We chose to procure data as it would not be practical to send

[3] Oracle's response to the CMA's information request [✂].
[4] Oracle's response to the CMA's information request [✂].
[5] [✂] response to the CMA's information request [✂].
[6] [✂] response to the CMA's information request [✂].
[7] [✂] response to the CMA's information request [✂].
[8] Responses to the CMA's information requests [✂].

statutory information requests to all companies providing cloud services in the UK. We procured data from two sources which allowed us to compare across data sets, and (if necessary) combine the reported figures based on our understanding of the underlying methodologies. We considered that using a combination of the two data sets would be the closest estimate of market size which matched our definitions of IaaS and PaaS. We chose IDC and Synergy as these were the only providers which were able to provide the data to our specification. Additionally, the selection of these two providers is consistent with the methodology Ofcom used to calculate shares by revenue for cloud services.[9]

D.8 We combined data from the two sources using a different methodology to calculate the market size for IaaS and PaaS respectively:

(a) IaaS: The two data sources combined reported revenues for ~70 firms. We considered that a number of these (across both data sets) did not supply services which aligned with our definition of IaaS, therefore we excluded revenue for some firms at this stage. We then took an average of the revenue for each firm that appeared in both data sets. If a firm was present in only one data set, we included the revenue value as reported. After these steps, in order to calculate the total market size, we aggregated across the total revenues.

(b) PaaS: The two data sources combined reported revenue for ~281 providers. We did not consider it practical to assess whether each individual firm provided services that aligned with our definition of PaaS. IDC had the most comprehensive list of firms, however we considered some firms included to provide SaaS. Synergy had a shorter list of firms, however, did not include some IDC firms which we considered provided PaaS. Therefore, after having replaced first party revenues where available, we took the average of the revenue for each firm that appeared in both data sets. Then, within each data set, we aggregated the non-first party revenues which did not appear in both data sets and took an average across both data sets 'other' category to calculate an estimate of the rest of the market.

Herfindahl-Hirschman Index

D.9 The Herfindahl-Hirschman Index (HHI) can reflect both the number of firms in the industry and their relative size. It is defined as the sum of the squares of all the market shares in the market, and thus gives proportionately greater weight to the larger market shares.[10]

D.10 AWS and CMA estimates of HHI are presented in the Market structure and concentration section of Chapter 3. Differences between these analysis and relevant caveats are explained here.

[9] Ofcom cloud services market study, Annex A1. Share of Supply for cloud infrastructure services in the UK.
[10] CC3 (Revised), Annex A, paragraph 7.

(a) We calculated the HHI for IaaS, PaaS and IaaS and PaaS combined using the results of the shares of supply by revenue analysis. We note that we have defined separate markets for IaaS and PaaS, therefore a HHI calculated across IaaS and PaaS combined, as in AWS' analysis, does not accurately reflect the boundaries of the relevant product markets.

(b) Our HHI estimates and those submitted by AWS will be an overestimate. This is due to a number of firms being grouped into an 'other' category and the 'other' category being treated as a single firm in the HHI calculations. The overestimate will be greater for PaaS, compared to IaaS as the 'other 'category is larger for PaaS. In addition, this effect in PaaS is larger in our estimates of HHI than AWS' estimates, as our 'other' category was larger due to the method by which we combined the third party data sources.

(c) AWS' analysis used third party data from a research firm. Our analysis is also based on third party research but we have supplemented this with first party revenue data from 11 UK cloud providers which is likely to be more accurate. Further, AWS' analysis includes more providers for both IaaS and PaaS than our own analysis, which we refined to more closely match the relevant product markets.

Shares of supply by capacity

D.11 We used formal information gathering powers to gather data from AWS, Microsoft, Google, IBM, Oracle and some smaller IaaS providers that serve UK customers[11] on their data centre capacity in megawatts (MW) within UK+EEA,[12] globally, and in the UK, for the period 2020 to 2026. In relation to this data we note the following.

(a) Google provided realised (historic) data centre capacity for all regions up until the end of 2023. Google provided [✂].[13]

(b) Oracle provided data on its data centre capacity in the UK and EEA going back to 2021 and global capacity going back to 2020.[14] To estimate Oracle's UK and EEA capacity in 2020, we included an approximation of capacity based on Oracle's growth rate from 2020 to 2021 being in line with its average growth rate implied by its capacity data and projections for 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.

[11] Centreprise, Coreweave, [✂], Hyve and Wasabi.
[12] For the purpose of this analysis, we have allocated providers' capacities to UK+EEA if they were classified by the provider as relating to Europe.
[13] Google's response to the CMA's information request [✂].
[14] Oracle's response to the CMA's information request [✂].

(c) Hyve provided [✂]. Centerprise provided [✂]. CoreWeave provided [✂]. [✂] provided [✂]. Wasabi provided [✂].[15]

(d) We consider that submissions on datacentre capacity include capacity from IaaS based on accelerated compute which could not be distinguished in the Parties' data, but we understand the contribution of accelerated compute to overall datacentre capacity to be small.

Shares of supply by flows of new business

D.12 We used formal information gathering powers to gather data on new customers from cloud providers. We asked them to define a new customer as one who spent more than $100 for the first time in a year (in that provider's revenue data). In responding to this each cloud provider submitted data on a slightly different basis.

(a) AWS provided data based on defining a newly acquired customer as one that spent at least $100 with it in a year, or alternatively for the purpose of specific questions, an acquired customer is defined as one that spent at least $100 with AWS in a year but nothing prior to that.[16]

(b) Microsoft provided data based on defining a newly acquired customer as one that spent at least $100 on its cloud services ('Azure Consumption Revenue') for the first time in a year.[17]

(c) Google provided data based on defining a newly acquired customer as one that spent at least $100 annual recurring revenue on its cloud services for the first time in a financial quarter starting from 2021.[18]

(d) Oracle defined a new customer as one that appeared in its revenue data for the first time in a year.[19]

(e) IBM defined a new customer as [✂].[20]

(f) UpCloud provided data based on credit usage, rather than revenue, which gives a slightly different result and noted that services were charged in USD between 2018–2022 and in EUR from 2023 onwards.[21]

(g) [✂] provided the total number of customers it believed to be in the UK that were acquired by the end of each calendar year.[22]

D.13 AWS provided data on UK customer flows according to how customers are classified within its customer relationship management system. This data therefore includes (i) [✂].[23] However, it provided data on UK revenues by reference to

[15] Responses to the CMA's information requests [✂].
[16] AWS' response to the CMA's information request [✂].
[17] Microsoft's response to the CMA's information request [✂].
[18] Google's response to the CMA's information request [✂].
[19] Oracle's response to the CMA's information request [✂].
[20] IBM's response to the CMA's information request [✂].
[21] UpCloud's response to the CMA's information request [✂].
[22] [✂] response to the CMA's information request [✂].
[23] AWS' response to the CMA's information request [✂].

[✂].[24] AWS said the 2023 figures were based on a different methodology for identifying a 'UK customer' to the 2018 to 2022 figures [✂].[25] While these definitions are not entirely consistent, they are unlikely to have a material effect on the analysis.

D.14 We tested the sensitivity of the assumption used to define a new customer by requesting from Microsoft, AWS and Google, data on the number of new customers and revenue from new customers, based on different spend thresholds ($500, $1,000, $10,000).[26] We compared this to the same data based on a threshold of $100 for 2023.

(a) A limitation with this sensitivity analysis was that we could not calculate the actual shares because we did not have the sensitivity results from all other cloud providers. Therefore, we compared the shares just between Microsoft, AWS and Google.

(b) The sensitivity analysis for the number of new customers showed that Microsoft continued to have the largest share of number of new customers in the first two scenarios ($500 and $1,000), though its share dropped as the threshold increases. In the final scenario ($10,000), AWS's share of new customer was the largest. This could suggest AWS has a larger proportion of existing customers who are increasing their spend.

(c) The sensitivity analysis for the revenues from new customers showed that increasing the threshold at which a new customer is defined increases the revenues from new customers. This is because while some low spend customers are now excluded, some higher spend customers are now included which would not have been counted at the lower threshold as they have already spent above the lower threshold amount. The results showed that as the threshold increases, Microsoft's share of revenues from new customers declined, and Microsoft continued to have the largest share in each scenario.

[24] AWS' response to the CMA's information request [✂].
[25] AWS' response to the CMA's information request [✂].
[26] CMA analysis of revenue and customer acquisition sensitivity data from Microsoft, AWS and Google. Responses to the CMA's information requests [✂].