Competition Commission of India
Journal on Competition Law and Policy

Doi: 10.54425/ccijoclp.v4.113
Vol. 4, No. 1, June 2023, pp. 1-25

Data Sharing for Contestability in Data-Driven Digital Markets: An Analysis

Priyanka Vinayak Bhat[1]

Abstract

This research paper explores the unique features of data-driven digital markets and investigates the feasibility of implementing a mandatory data-sharing regime for dominant entities. The context of the study arises from the recognition that competition in digital markets operates differently from traditional markets, prompting global efforts to regulate these markets and promote competition and consumer welfare. This paper employs a policy analysis methodology, examining expert reports on data-sharing strategies from the UK and the EU. Additionally, it analyses recent amendments to the Competition Act, 2002, and relevant reports and discussions within India's antitrust regulation framework. The paper identifies that data enables competition within these markets and recommends a mandatory data-sharing regime that optimises competing considerations, such as fairness to incumbents and maximising data efficiencies. Such a data-sharing regime would ensure that entrants in the digital market are not disadvantaged and can leverage their innovation to pose a genuine competitive threat. The paper also explores alternative approaches to promoting competition, concluding that data sharing is the most efficient solution.

Keywords: big tech, data-driven digital markets, data efficiency, ex-ante regulation, fair competition, mandatory data sharing

1. Introduction

In recent years, the rapid expansion of data-driven digital markets has raised concerns about big technology companies' dominance and

[1]Paralegal, Nishimura & Asahi; priyankavbhat1996@gmail.com

rent-extraction practices. The need for effective competition in these markets to ensure fair market conditions, prevent anti-competitive behaviour, and safeguard consumer welfare has become increasingly evident (Scherer, 2007). This research paper delves into the multifaceted data-related challenges that entrants face in these markets, which prevents them from emerging as a competitive threat. To overcome the barriers to competition, entrants require access to the vast data pools controlled by dominant players (Parker *et al.,* 2020).

This paper builds on existing literature that advocates for a new competition regime embracing mandatory data sharing. The success of such a data-sharing regime hinges upon striking a delicate balance between various competing interests and considerations. This paper attempts to achieve this balance within the existing Indian economic and legal landscape. Additionally, this paper explores the suitability of alternative techniques and existing legal regimes for levelling the playing field for entrants (Krämer & Schnurr, 2021). By addressing these aspects, the research aims to contribute to the discourse on promoting competition in data-driven digital markets, providing insights and recommendations for fostering fair and competitive markets in India in the digital age.

This paper uses a policy analysis approach by examining the data-sharing strategy proposed in expert reports from the UK (Furman *et al.,* 2019) and the EU (Montjoye *et al.,* 2019). Through analysing these and similar reports, valuable insights are gained into the recommendations and frameworks regarding data sharing extended by experts from other nations. Furthermore, Indian parliamentary and other reports have been examined to assess the need for changes to the country's antitrust regulation and to develop a feasible data-sharing regime for India. The suggested data-sharing regime incorporates the evolving regulatory landscape in India, including the Competition (Amendment) Act, 2023, and the challenges and considerations related to data sharing and competition in the Indian context. Finally, an examination of the suitability of alternatives to the data-sharing regime is made with the help of existing literature (e.g., Prufer & Graef, 2021; Feasey & Streel, 2020; Krämer *et al.,* 2020). Overall, this comprehensive methodology provides a robust framework for assessing the feasibility and effectiveness of the proposed data-sharing strategy outlined in the research paper.

Section 2 of this paper examines the importance of competition in data-driven digital markets to keep big techs in check and prevent them from engaging in rent extraction. It identifies the need for an ex-ante solution to enable entrants in these markets to overcome data-related barriers to competition. Section 3 advocates for mandatory data sharing as the required ex-ante measure. Section 4 explores ways of balancing various competing interests and considerations for the success of the data-sharing regime. Section 5 points out the limitations of existing data-sharing laws in some of the most prominent jurisdictions, followed by an analysis of alternatives to data sharing. Section 6 concludes the discussion.

2. Competition in the Data-Driven Digital Market—Conspicuous in its Absence

Consumer data is an essential resource for a firm competing in any market, particularly in digital markets. For example, the e-commerce or social media market is data-driven (Krämer & Schnurr, 2021), i.e., these markets supply services using big data. Large technology companies gather vast user data from their platforms and by using crawlers on third-party websites. They constantly align their services with individual consumers' interests using real-time and continuous consumer data.

The companies dominating the data-driven digital markets have steadily grown in both size and influence. The top five technology companies, commonly known as GAMAM (Google, Amazon, Meta, Apple, and Microsoft), are among the highest-valued companies in the world (Forbes, 2022). In India alone, over 50 crore individuals access the internet and the services of these companies using smartphones (IANS, 2020), and the role played by these "big techs" in people's day-to-day lives has only enlarged in scope and importance.

Given the significance of these companies, it is crucial to prevent any harm being caused to consumers by their actions. Such harm is difficult to identify in these markets due to the unique features of these services, zero-priced services being one of them. In digital markets, consumers barter their data for free services (Zingales & Lancieri, 2019). Consequently, other aspects of the exchange and not the price indicate the level of consumer welfare. More particularly, the volume of data collected from the user, the level of privacy and security in handling the data, the amount

of advertising, and the quality of the free service indicate welfare levels. Harm to consumers could come from lessening the quality of service provided to the consumer or lowering the levels of innovation below what would otherwise have existed in the face of competition. These non-price–based inefficiencies are difficult to identify, especially when there is no competition in the market with which to compare the incumbent.

Harm to consumers could also be indirectly caused by charging a higher price or engaging in unfair business practices against businesses that constitute the other side of this multi-sided market. The businesses either recoup the higher cost from their final consumers or accept their businesses becoming less profitable (Armstrong, 2006). Recently, such practices came to light in India where online travel platforms like MakeMyTrip and Goibibo were held liable for imposing price restrictions on hotel partners wishing to rent out their rooms through these platforms ("MakeMyTrip, Goibibo, OYO Fined Rs 392 Crore by CCI for Unfair Business Practices", 2022).

On an average, a firm in a digital-intensive sector had more markup than other firms during 2001–03 and 2013–14, which supports the inference of rent extraction (Calligaris *et al.,* 2018). Moreover, a report by the French Competition Authority estimated that advertisers and platforms share revenue from advertiser sales at a 40:60 ratio (L'autorité De La Concurrence, 2018). There has been a trend of higher markups in digitalised sectors, which has become stronger over time. This high markup can be attributed to the fact that few platforms can target customers with as much accuracy. The Competition Commission of India (CCI), in the 2021 case against Google, noted this disparity in bargaining power between big techs and businesses that are dependent upon the services of the big techs (Case No. 41/2021).

Such a high profit margin is a standard feature of markets lacking effective competition. An entity which is dominant in an imperfectly competitive market would naturally be tempted to maximise its returns by selling sub-standard goods or services at Supernormal profits if there were no repercussions. It would lose the incentive to invest in improvements or innovations. Corporations, after all, aim to make profits. Competition, whether strong enough to replace the incumbent or not, becomes vital in

such a market; the threat of competition will keep the incumbent in check, providing consumers with more choices, new features, and better quality.

However, competition does not occur naturally in the data-driven digital market due to high data-related barriers. Because of the role of data in improving quality and user experience, firms with access to continuous and real-time data enjoy a disproportionate advantage. Conglomerate digital platforms such as the GAMAM operate in multiple business verticals, including mail, maps, and search engines. They are in a unique position to exploit data-mining opportunities and economies of scope by collecting different dimensions of data on a user, such as identity, location, and purchase intent. The more users a platform receives, the more it can improve on its output quality and the relevance of its recommendations. The improved user experience, in turn, attracts more users in a positive loop (Furman *et al.,* 2019).

In a market where more data equals higher quality and accuracy, an entrant is disadvantaged and cannot compete on an equal footing with the incumbent. The entrant's business is plagued by a negative feedback loop of fewer users availing of its service because of low quality. Concurrently, it cannot improve the quality of its service for want of user data. Consumers further aggravate this negative loop through their behavioural bias towards the status quo as well as impatience. The incumbent's economies of scale and scope intensify high entry barriers (Ducci, 2020). Innovators may also face barriers to entry into markets adjacent to those dominated by data-rich incumbents due to the domino effect and envelopment strategies (Prüfer & Schottmüller, 2021). All these factors create high entry barriers that are near-impossible for an entrant to overcome on its own.

Even where competition does develop in the data-driven digital market, network effects cause the market to tip in favour of the most dominant entity (Zingales & Lancieri, 2019). In these markets, the value of service to both end users and business users increases with an increase in end users. Even if two firms start with equal levels of data and insights, the network effects in such a market will ultimately lead to users preferring one competitor over the other, thereby tipping the market in the favour of a single winner. In such markets, therefore, competition is "for the market rather than within the market" (Calvano & Polo, 2021). Hence, competition cannot naturally sustain itself in the market for long.

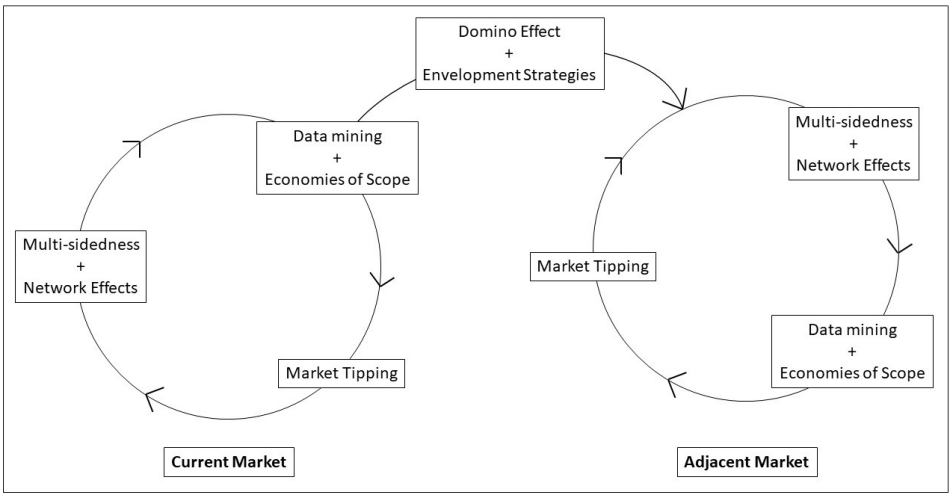

Figure 1. The positive loop leading to market tipping in current and adjacent markets.

Figure 1 represents how network effects enhance the incumbent's ability to gather and use data. Improved service through data efficiencies leads the market to tip in favour of the incumbent, attracting more users and data to the platform in a positive loop. The economies of scope of the data enable the incumbent to provide targeted services in the adjacent market, which also experiences the positive loop.

The UK Digital Competition Expert Panel's 2019 report (Furman *et al.,* 2019) shows that market concentration exists in various digital markets: Google was found to dominate the online search market, with some competition from Microsoft Bing; Meta dominated the social media market, with some competition from Twitter and Snapchat; Google and Meta dominated the digital advertising market; Apple and Google dominated the mobile app market; and Amazon dominated the e-commerce platforms market, with some competition from eBay.

Irrespective of the degree of innovation by an entrant, the incumbent's existing infrastructure and data-based insights enable it to replicate a non-patentable innovation at a fraction of the cost borne by the entrant. The incumbent can thus make the entrant's business model unprofitable (Furman *et al.,* 2019). The high probability of finding itself again in the "runners-up spot" with a considerable sunk cost for an attempted leap in

innovation is enough to take the incentive out of any prospective entrant (Prüfer & Schottmüller, 2021). Baumol (1982) refers to this phenomenon as the exit barrier, which undermined the probability of potential competition. With the high sunk cost of competing in the data-driven digital market, the threat of potential competition is naturally low.

The inability of entrants to sustain themselves in a data-driven market has also negatively impacted the willingness of venture capital investors to fund start-ups that plan to enter a market dominated by or adjacent to one dominated by data-rich incumbents (Krämer *et al.*, 2020). In this way, barriers created by sunk costs limit the potential for disruptive competition and lead to a lack of choice and innovation in the market.

It is, therefore, clear that, in most cases, one or two entities dominate data-driven digital markets, and these firms can and do unduly profit from their position of dominance. When high entry barriers and sunk costs exist, dominant entities can be secure in their market position, and hence, they tend to operate on higher profit margins. In the interest of consumer welfare and fairness to entrants, creating a legal regime that fosters competition in these markets is essential.

In a market where entrants are prone to experiencing a negative loop, it is desirable to intervene even in the absence of an abuse of dominance by the incumbent. Such an ex-ante regulatory framework will not be unique to data-driven digital markets. Such regulations were implemented in the telecom market in the UK, with mobile phone operators being required to connect calls from users of other services to enable interoperability and competition (Furman *et al.*, 2019). CCI has also recognised the need for ex-ante measures in the digital market, as recorded in the 2022 report of the Standing Committee on Finance.

3. The Case for Mandatory Data Sharing in Digital Markets

The ex-ante measures suggested in this paper are mandatory data-sharing regime for dominant entities in data-driven digital markets. Monopolistic data usage by dominant entities has been identified as a significant issue in the digital market in India (Standing Committee on Finance, 2022). Access to relevant data would enable entrants to predict the changing needs of the market, provide targeted products and services, and overcome data-related barriers to entry. A chief scientist at Google

even admitted that "we don't have better algorithms than anyone else. We just have more data" (Asay, 2010). The lack of data is akin to grappling in the dark, and with sufficient innovative capacity and infrastructure, entrants will be able to pose a genuine competitive threat to the incumbent. In this way, data sharing by those with exclusive access to valuable data in favour of entrants would diminish consumer welfare concerns. As Baumol (1982) explains, even in a monopolistic market like a tipped data-driven digital market, the mere threat of an entrant emerging as a competitor will create perfect competition-like conditions.

Data sharing would also increase overall efficiency in the market. Data can be used to gain different insights depending on the data user's needs. Such data in the hands of a single incumbent implies an inherent market failure, leading to efficiency losses of other utilities that could otherwise have been derived from the data (Krämer & Schnurr, 2021). As Acemoglu *et al.* (2019) point out, when a data incumbent has an extensive sample of user data, the marginal value of collecting more data begins to approach zero.

Moreover, datasets are non-rivalrous, i.e., sharing them among more data users does not diminish the benefits derived by the original incumbent. It is not a physical asset that can be used up or used only by one user at a time (Parker *et al.,* 2020). Businesses such as Citymapper are classic examples of how data sharing helps promote innovation and consumer welfare; Citymapper has generated up to GBP 130 million a year in economic benefits in the UK after Transport for London [TfL] began providing its real-time data, free of charge and in an open format in 2009. Hence, with the right approach, a data-sharing strategy can be implemented without unfairly prejudicing the original incumbent in the market.

Placing such additional burdens on dominant entities to preserve competition is not new in competition law (*United States v. Dentsply International, Inc, 399 F.3d 181 (3d Cir.2005)*). The competition law theory that the risk of false positives is more harmful than the risk of false negatives does not hold in digital markets. The theory presumes that any risks from under-enforcement will soon be balanced by a spontaneous increase in rivals, thereby limiting the supranormal profits of the incumbent (Caffarra *et al.,* 2020). However, the digital market is prone to tipping and does not automatically correct itself into a reasonable profit margin for the incumbent. Big techs have come under the antitrust radar

worldwide because of the recent unravelling of the undue advantages they have enjoyed for many years. Further, it is now widely accepted that competition policy must promote fairness and be armed with interventionist powers where appropriate (Furman *et al.,* 2019). Hence, it is better to err on the side of over-enforcement in the case of digital markets by implementing a mandatory data-sharing regime.

Multiple studies (Furman *et al.,* 2019; Montjoye *et al.,* 2019; Prufer & Graef, 2021; Feasey & Streel, 2020; Krämer *et al.,* 2020) have already suggested this ex-ante solution, although most recommendations have focused on modifying and changing governance frameworks. India could take the initiative and become the first nation to implement the data-sharing approach that experts in several nations have been contemplating. When implementing such a radical antitrust strategy, there would be a legitimate fear of its repercussions on trade and the economy. It is pertinent to note that the consumer market in India has grown to the extent that no business can ignore its value. Competition among big techs would also encourage big techs to consider the mandatory data-sharing regime. Moreover, as discussed earlier, data is non-rivalrous, and sharing it would retain the data efficiencies enjoyed by big techs.

However, India's current competition law regime is ill-equipped to provide such an ex-ante solution. In particular, the Competition Act, 2002, only provides ex-post measures. Only recently has India recognised the need for ex-ante measures to deal with the rapidly growing digital market. As the world progresses towards more efficient ways of producing and selling products and services, the law must similarly evolve to meet the new challenges. Legal evolution can be observed throughout history, including legal recognition and protection for new kinds of intellectual property and the recent debate around crypto-assets. Data-driven digital markets also possess features that make it imperative for competition law to evolve to ensure fairness in this market. Hence, there is a need for a new regulatory regime in the form of mandatory data sharing in India.

4. Striking the Right Balance Within the Data-Sharing Regime

The success of the proposed data-sharing regime would depend on its ability to maximise data-related efficiencies while minimising the costs incurred by each involved party. The structure of the regime would involve

the incumbents (data sharers) who make their data assets available, the entrants (data recipients) who benefit from the shared data, the individual consumers whose data is used for the sake of maximising data efficiency (data subjects), and the administrative unit of the government (Data Access and Competition Authority [DACA]) that oversees implementing and regulating the data-sharing regime. The interests and costs of each of these parties must be balanced to ensure the most optimal outcome from the data-sharing strategy.

4.1. Ensuring Privacy for Data Subjects while Maximising Data Efficiency

It is crucial to preserve the efficiencies of data. Protecting privacy by anonymising personal data so that the data subjects are unidentifiable is challenging (De Montjoye *et al.*, 2015) and counterproductive. If the data was successfully stripped of all identifiers such that it cannot be traced back to any individual, it would also be considerably stripped of its value (Prufer & Graef, 2021). Hence, the data-sharing strategy must use raw data. On the other hand, sharing personal data without consent violates the data subject's right to privacy.

To balance these competing interests, data pooling and data sandboxing by DACA must form the central feature of the data-sharing regime. DACA, as a data trust, shall receive data from data sharers in its "original and detailed form" (Prufer & Graef, 2021), thereby retaining the data value. The raw data will enable DACA to maintain a data pool of a scale and scope that could theoretically match those of the big techs over time. Once the data is consolidated (data pool) by DACA, transferring this raw data directly to the data recipient would not be prudent. Instead, the data recipients shall send their algorithm based on the analysis they wish to conduct in the data pool (Krämer & Schnurr, 2021). DACA would then run the algorithm on the data pool on behalf of the data recipient. The data recipient would receive only the algorithm's results (data outcome) (Figure 2).

The above modus operandi has multiple benefits. If a centralised body such as DACA were to receive the data from data sharers and consolidate them, the personal data of individuals would be transferred only once. This transfer through a central body is desirable to its alternative, i.e., data

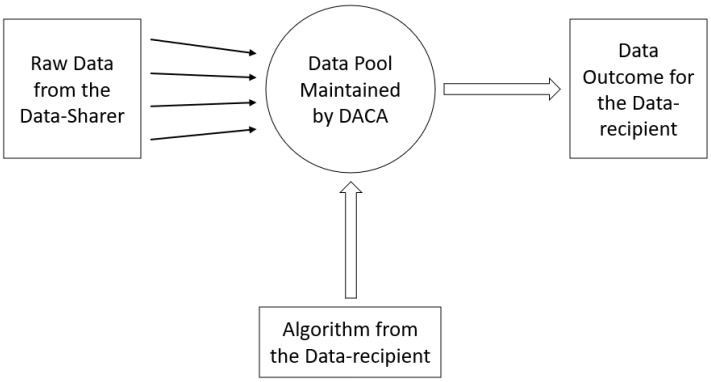

Figure 2. Overview of the data-sharing regime.

being transferred separately between every data sharer and data recipient. This intermediation would reduce the costs and risks of data leaks to a large extent (Krämer & Schnurr, 2021). Data sandboxing, i.e., creating a secure and safe environment for using data without unauthorised access, would further minimise the risks of data leaks or data misuse.

Here, it is prudent to have checks to identify and deny access to an algorithm that tries to gather raw data from the data pool. Further, the non-rivalrous nature of data that makes it possible to share its benefits easily among deserving data recipients also makes it easy to misuse such data. The existence of openly available data must not give rise to a black market for data outcomes. Checks in the form of mandatory firewalls are required to prevent data outcomes from being misused. Deterrence in the form of high fines may prevent data misuse. Ensuring protection for whistleblowers and leniency programs may bring incidents of misuse to light.

If implemented correctly, such a data-sharing regime will also be compatible with other laws. Though there will be information sharing between competitors, it will not result in collusion if the incumbent shares only raw user data and not the insights it gained from the data (Prufer & Graef, 2021). Data sharing under such a regime would not violate privacy laws as it would be in the public interest and part of a legal obligation (Parker *et al.,* 2020). Passing data through a centralised authority without direct interaction between the incumbent and the entrant further diminishes competition and privacy concerns (Krämer & Schnurr, 2021).

4.2. Retaining the Trust and Co-operation of Data Sharers

While it is true that big techs have benefited from the unique features of digital markets, it is essential to ensure that the data-sharing regime does not form disciplinary action against these incumbents. The regime is only a policy measure to encourage competition, innovation, and consumer welfare in the market. To make this explicit, big techs must be reasonably involved throughout the regime's planning, setting up, and enforcement stages.

Additionally, requiring the incumbent to share its "inferred or derived" data and its "core" data would be unfair and may disincentivise big techs from innovating or investing in data collection and analysis technology (Furman *et al.*, 2019). Therefore, mandatory data sharing shall be limited to "ancillary" data that is either "volunteered" by the user or passively "observed" by the incumbent (target data). In other words, DACA must require the sharing of only the user data acquired by the data sharer as a byproduct of its core business activity and collected by the incumbent passively using its existing infrastructure (Figure 3).

To illustrate, a video-streaming platform may collect data volunteered by a user, e.g., the ratings provided to shows, and data observed about the user, e.g., the genre of shows watched or the duration spent on a specific show. This data is collected passively by the platform. However,

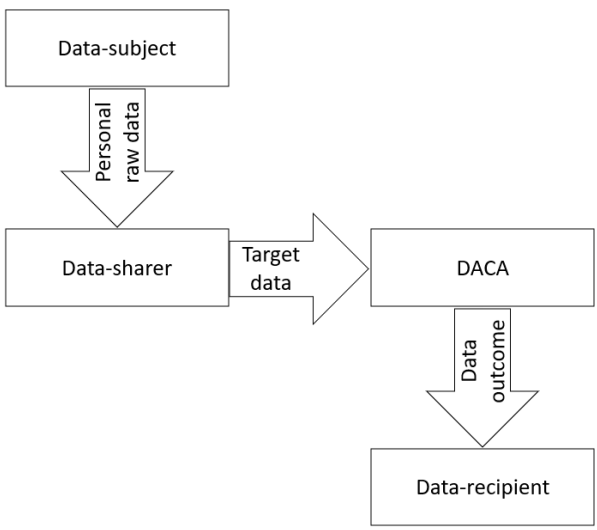

Figure 3. Flow of data among the major components of the data-sharing regime.

such information is not directly beneficial to the platform. The inferred or derived data provides insights that enable the platform to make personalised recommendations to the user, encouraging the user to prefer the platform over others and spend additional time on it. For example, the platform may use volunteered and observed data about a user to infer the kind of entertainment that the user prefers.

Data can also be core or ancillary to a platform. A weather-information application, for example, may collect the location data of its users to automatically update the current temperature at their location. However, this is not the primary business model of the application. In contrast, the "feels like" temperature information collected, calculated, and shared by the application would form part of its core activity and directly contribute to its success as an application.

The location data would be core data for a local event-discovery platform that provides personalised event information based on the user's latest historical location. The "feels like" temperature information would be ancillary data for such a platform. In this way, data ancillary for one market may be core data for another, and vice-versa. Data recipients will thus be able to take advantage of the economies of scope of data that have provided data sharers with an extraordinary advantage for decades.

In specific markets, it may become challenging for an entrant to compete without access to the core data of the incumbent. In the interest of contestability, it may be necessary to mandate the incumbent to share such core data on a case-by-case basis. Fair compensation would be required to justify collecting core data from the incumbent. Fair compensation involves the complex question of determining a price that is fair, reasonable, and non-discriminatory (FRAND) (Heim & Nikolic, 2019).

An incumbent may be concerned that mandatory data sharing would violate their intellectual property rights over the data. However, copyright protection is given only to the structure of a database and not to the data itself. Furthermore, intellectual property protection is not an absolute right, and legislation may compromise any protection given to data representing a trade secret in the interest of public welfare. Moreover, *sui generis* protection accorded to raw data involving substantial investments would not apply in the case of data collected passively, as a byproduct

of the core business of the incumbent (Prufer & Graef, 2021). A similar principle may be applied in the case of mandatory data sharing.

4.3. Building Confidence and Competence among Data Recipients

The success of the data-sharing strategy will depend significantly on the response of the data recipients. Data recipients must receive clarity on the procedural aspects of the data-sharing regime and the kind of information and assistance they can expect from DACA. The commercial terms and obligations on which data outcomes are shared with the data recipient must also be drawn up. Awareness and competence among stakeholders must be built through targeted workshops, industry discussions, helpdesks, guidance notes, and circulars.

The data-sharing strategy must also provide sufficient enforcement mechanisms against defaulting data sharers. There must also be a protocol in place in the case of any disruption in the supply of data outcomes to data recipients who rely on said data for their businesses. These measures will develop confidence among the data recipients to use the data outcome and compete with the incumbents. Without the assurance of quality, timeliness, and constancy of data outcomes, data recipients may not invest in innovation that relies on such data.

4.4. Instituting a Fit-for-Purpose Administration while Minimising Costs

It is suggested that the mandatory data-sharing regime be implemented in India within the larger structure of the proposed Digital Competition Act. Mandatory sharing of target data could be written into the recently deliberated ex-ante obligations on big techs. Consequently, the definition of "data sharer" must coincide with the definition of the proposed Systemically Important Digital Intermediaries (SIDIs). Bringing DACA under the ambit of the Digital Competition Act would be ideal as the Committee on Digital Competition Law is best suited to analyse and recommend the data-sharing strategy.

Adopting a rule-based definition for "data recipient" and "target data" may help control administrative costs. A rule-based definition will prevent ambiguity and speed up the assessment of an applicant's eligibility. The scope of "target data" changes depending on the business activity of the

data sharer. After formulating clear rules for defining what is and is not target data, it would be appropriate to leave the identification of target data to the entity that is most qualified to decide, i.e., the relevant data sharer. The annual report to CCI on measures taken by SIDIs to comply with various mandatory obligations could also include matters related to the data-sharing regime.

There is also a need to develop a set of common technical standards that DACA must implement upon the data sharers and data recipients to promote ease and security of data sharing and receiving.

4.5. Preventing Over-Reliance on the Data-Sharing Regime

The data-sharing strategy must be cautious of the risk of free-riding and over-reliance on its data pool. The data-sharing regime's convenience must not cause data recipients to lose all incentives to collect user data independently. The development of new ways of collecting and using data is vital for the continuous growth of the digital market. Therefore, there is a need to limit the reliance of firms on the data-sharing option.

There are several ways to limit the number and type of data recipients through a registration process. First, data recipients may be restricted to firms within a threshold of market power in relevant and adjacent markets. This restriction ensures that incumbents and other emerging strong firms cannot access the data pool. Second, registration may be limited to firms that intend to compete directly with dominant players through innovation or otherwise. This will ensure that the data-sharing regime is used exclusively to promote competition in the digital markets. Third, registration may be limited to firms with substantial venture capital funding as a preliminary sifting measure. Venture capitalists could thoroughly examine a company's credentials before endorsing it, serving a similar purpose that underwriters do in public offers. Finally, data-recipient registration may be limited to a certain number of years, after which there would be a need for re-registration. These restrictions would achieve multiple goals at once and would prevent over-reliance and misuse of data by recipients while also limiting the administrative burden on DACA.

By restricting the data pool to target data, data recipients retain the incentive to independently collect user data, ultimately adding to the

overall efficiency of the market (Krämer & Schnurr, 2021). It is suggested that data recipients must pay the price for the data outcome they receive from DACA as an additional deterrence against complacency. The price must depend on the volume and value of the data pool used and not just the cost of maintaining the data pool and running the algorithm. The intention is to incentivise data recipients to continuously assess their data needs and to terminate their reliance on the data-sharing arrangement when they can collect data on their own at a lower cost (Feasey & Streel, 2020).

4.6. The Road Ahead

The previous sections attempted to chalk out a probable data-sharing regime and suggest how to successfully implement it in India. However, there are several issues yet to be identified and smoothed out. Figure 4 illustrates the concerns of the critical components within a data-sharing structure and the policy considerations that the data-sharing regime must determine in addressing these concerns.

Moreover, datasets could be 'wide', i.e., comprise data about several consumers and used to infer consumer trends within and outside the dataset. Data could also be 'deep', i.e., containing in-depth information about each subject in the dataset and used to predict the needs of those consumers with greater accuracy, although at a diminishing rate. It is acknowledged that restricting access to raw personal data would limit the utility that the data recipient could derive from deep data. Additional research is needed to strike the right balance between data utility and user privacy.

It is suggested that policymakers and regulators refine and enhance the data-sharing regime by conducting empirical studies to assess the effectiveness of the proposed data-sharing regime and its impact on competition and consumer welfare. There is a need for a more in-depth examination of the legal and technical challenges associated with implementing and regulating the data-sharing regime. Once implemented, the impact of the data-sharing regime impact on market dynamics, innovation, and consumer outcomes must be continuously monitored and evaluated to determine the need for policy adjustments and improvements.

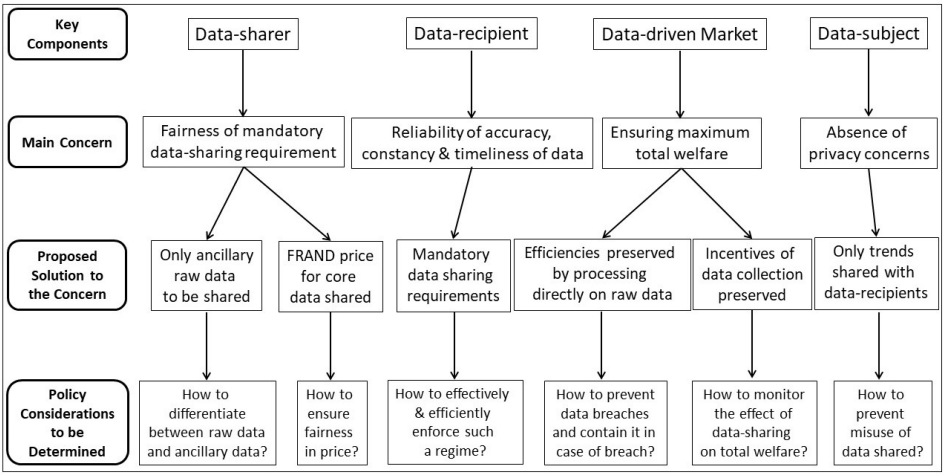

Figure 4. Questions that the data-sharing regime must address.

5. Evaluating Existing Legal Regimes and Techniques as Data-Sharing Alternatives

Before accepting data sharing as the ideal solution, it would be prudent to consider the viability of alternatives in existing legal regimes. First, the essential facilities doctrine is a competition law remedy when a monopolist or a dominant entity refuses access to a critical facility vital for competition. The argument of 'essential data' in our context of digital markets is not compelling because the data is not essential in the same sense as in the essential services doctrine. Though it would be nearly impossible to compete effectively in a data-driven market without this data, it is still possible to enter and attempt to compete using publicly or commercially available data. Hence, proving the indispensability of such data would be onerous (Feasey & Streel, 2020). Even if the required data is proven to be 'essential' to compete, the case studies analysed by Krämer *et al.* (2020) show that what is necessary to be competitive in data-driven digital markets goes far beyond the minimum requirements of 'essential data'. The quality of the product or service gradually improves through the direct and indirect benefits of the data. Hence, digital markets could not be genuinely competitive unless dominant firms share as much data as they are privy to. The essential facilities doctrine would therefore, not be sufficient. Moreover, this remedy is an ex-post one and it would take considerable time to resolve on a case-by-case basis. Furthermore, one

cannot request data based on the essential facilities doctrine if the data is proposed to be used to innovate in a new market where the incumbent is not yet active (Martens *et al.*, 2020). Hence, a small firm would never have the first-mover advantage if left to the essential facilities doctrine alone.

The UK ("Online Platforms and Digital Advertising Market Study", 2019) and the EU ("Proposal for a Regulation on Contestable and Fair Markets in the Digital Sector (Digital Markets Act)", 2020) also have express and proposed data-sharing requirements, where dominant search engine service providers must share click and query data with rival search engines. These regulations are promising. However, we are concerned with all data-driven digital markets, not just search engines. The EU Digital Markets Act (DMA) requires gatekeepers to provide business users with "effective, high-quality, continuous, and real-time access and use of aggregated or non-aggregated data" generated in the context of those business users using the platform.

The above DMA provision is very similar to the solution proposed by the Indian Parliamentary Standing Committee on Finance to the problem of advertisers' lack of bargaining power in the search advertising market. It is, however, also insufficient for our discussion as it does not provide any information to rivals who may wish to compete with the platform. The business user will only receive data about their own activities and will not have the breadth and depth of data available with the platform.

The proposed Platform Accountability and Transparency Act in the US would mandate data sharing, although for the limited purpose of research and not to promote competition (Nonnecke & Carlton, 2022).

The EU Data Governance Act facilitates "data altruism", where registered organisations may voluntarily make data available for the common good. However, the incentives gained from exclusive access to data by incumbents in data-driven markets are such that they would not typically volunteer the data and lose their competitive advantage.

Alternatively, individuals who are subjects of the data held by the incumbents may require such data to be shared with other entities under the EU's GDPR. This provision could allow entrants or rivals to gain the insights that incumbents enjoy. However, individuals rarely port data with this intention in mind. Moreover, no stable business model can

solely rely on the low probability of individuals availing their right to data portability (Prufer & Graef, 2021).

The Indian Open Network for Digital Commerce (ONDC) is seemingly the first-of-its-kind initiative to combat the dominance of e-commerce giants. It adopts an open-network framework where buyers and sellers of various e-commerce products and services 'from retail goods, food to mobility' can transact on a single platform. Though this initiative has considerable potential for digitalising brick-and-mortar businesses, it is not targeted at breaking the data-based disparities between incumbents and entrants. A press release by the Government of India has clarified that the ONDC will not require its participants to share any transaction-level data ("Shri Piyush Goyal Chaired Open Network for Digital Commerce (ONDC)", 2021). Though it intends to publish "anonymised aggregate metrics on network performance", the level of insights that may be gathered from such metrics is doubtful.

Additionally, several alternatives could reduce the data-based disparity in digital market competition. However, none of them can match the effectiveness of data sharing. Alternatives such as horizontal breakup, where a dominant firm is divided into segments to limit data access; vertical breakup, which prevents a vertically integrated firm from using information about demand or strategies of upstream competitors; or market-based breakup, which suggests splitting a dominant firm into segments based on the markets it operates in, would each have shortcomings (Parker *et al.,* 2020). Each approach can decrease the value derived from data and potentially make the market less efficient.

Data siloing is another alternative that is very similar to splitting up the firm. It limits an incumbent's insights by preventing data aggregation from different services (Krämer *et al.,* 2020). The Indian Parliamentary Standing Committee on Finance has suggested solutions incorporating data siloing in its recent report. The efficiency losses from adopting such an approach would be too significant to be desirable.

Setting aside the privacy concerns of comprehensive end-user tracking and profiling, collecting and using massive amounts of data has several advantages. Data mining has led to better-quality services and improved efficiency in digital markets. Hence, it is essential to retain the value derived from data. Competition must be promoted, not by weakening the

incumbent but by enabling the entrant to compete on an equal footing with the incumbent. A more level playing field for the entrants must be created while preserving data-related efficiencies. This is why any of the above alternatives to the mandatory data-sharing regime would not be desirable.

Shortening the data-retention period for analysis is ineffective because the incumbent continues to receive a steady stream of new data and hence, continues to benefit from data monopoly. Prohibiting incumbents from becoming the default option for consumers may somewhat restrict their dominance, but consumer preferences for familiar and perceived better services make this approach questionable. Privacy-enhancing technologies can be beneficial but also raise barriers for entrants. Further, they diminish the data-driven network effects of incumbents, reducing the efficiencies achieved by data (Krämer & Schnurr, 2021).

Treating data as labour and having platforms pay individuals for their data is a popular alternative, but it relies on independent data labour unions to negotiate fair prices. Moreover, the bargaining power of these unions and the value of data compared to the value received by consumers is uncertain. Taxation of digital platforms aims to redistribute value more evenly. However, taxation alone cannot prevent market tipping or directly promote contestability in the digital market (Posner *et al.,* 2018).

Each of the above alternative legal regimes and techniques has limitations and cannot replicate the benefits of data sharing in promoting competition and providing a level playing field for entrants. Hence, a mandatory data-sharing regime is the most appropriate solution to ensure contestability in digital markets.

6. Conclusion

The findings of this research underscore the distinctive nature of digital markets, highlighting the need for differentiated regulations that foster competition within this unique landscape. A key revelation of this study is the indispensability of data, which is predominantly held by incumbents in the digital market. Mandatory data sharing by incumbents emerges as a viable solution to address this concentration of data and promote contestability.

The research uncovered various pivotal findings that contribute to developing an optimal framework for the data-sharing regime. First, it was found that to strike a balance between privacy and data efficiency, the utilisation of raw data, rather than anonymised data, is crucial in retaining the value of the data. Data pooling and sandboxing were identified as effective mechanisms to achieve this balance, with the proposed involvement of DACA playing a central role.

Furthermore, the study highlighted the significance of retaining the trust of data sharers by framing the data-sharing regime not as a disciplinary action against incumbents but as a way to encourage competition and innovation. The sharing of ancillary data collected as a byproduct was identified as a viable approach to address concerns about fairness and innovation in the marketplace.

Building confidence among data recipients emerged as another crucial aspect. Clear procedures, commercial terms, and obligations must be established to ensure the competence and confidence of those accessing and utilising the shared data. Additionally, enforcement mechanisms and protocols should be in place to address any potential disruptions in data supply, ensuring the quality and reliability of data outcomes.

The findings further emphasised the need for a fit-for-purpose administration to implement the data-sharing regime effectively. Integrating the proposed regime within the Digital Competition Act in India, aligning DACA with the Act, and adopting rule-based definitions were identified as steps to minimise administrative costs. Developing common technical standards was also suggested to facilitate secure and efficient data sharing.

Finally, the study highlighted the importance of preventing over-reliance on data sharing to avoid free-riding. Implementing restrictions on data recipients based on market power, innovation intent, and registration duration can effectively achieve this goal while limiting administrative burdens. By adopting the recommendations presented, policymakers and regulators can foster competition, drive innovation, and ultimately create a more dynamic and consumer-centric digital market ecosystem.

Looking ahead, policymakers and regulators must focus on refining and enhancing the data-sharing regime. Further research in this direction

include empirical studies to assess the regime's effectiveness, legal and technical challenges, and continuous monitoring of its impact on market dynamics, innovation, and consumer outcomes. By pursuing these directions, policymakers can ensure that the data-sharing regime evolves to foster fair competition, innovation, and consumer welfare in data-driven digital markets.

References

Acemoglu, D., Makhdoumi, A., Malekian, A., & Ozdaglar, A. (2019). Too much data: Prices and inefficiencies in data markets. *Working Paper No. 26296*. National Bureau of Economic Research. https://doi.org/10.3386/w26296

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, *37*(3), 668–691. https://doi.org/10.1111/j.1756-2171.2006.tb00037.x

Asay, M. (2010). Tim O'Reilly: 'Whole web' is the OS of the future. *CNET*. https://www.cnet.com/culture/tim-oreilly-whole-web-is-the-os-of-the-future/

Baumol, W. J. (1982). Contestable markets: An uprising in the theory of industry structure. *The American Economic Review*, *72*(1), 1–15.

Caffarra, C., Crawford, G., & Valletti, T. (2020). *'How tech rolls': Potential competition and 'reverse' killer acquisitions*. https://cepr.org/voxeu/blogs-and-reviews/how-tech-rolls-potential-competition-and-reverse-killer-acquisitions

Calligaris, S., Criscuolo, C., & Marcolin, L. (2018). Mark-ups in the digital era. *OECD*. https://doi.org/10.1787/4efe2d25-en

Calvano, E., & Polo, M. (2021). Market power, competition and innovation in digital markets: A survey. *Information Economics and Policy*, *54*. https://doi.org/10.1016/j.infoecopol.2020.100853

De Montjoye, Y.-A., Radaelli, L., Singh, V. K., & Pentland, A. (2015). Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science*, *347*(6221), 536–539. https://doi.org/10.1126/science.1256297

Digital News Publishers Association V. Alphabet Inc. And Others, Case No. 41/2021 (Competition Commission of India). https://www.cci.gov.in/antitrust/orders/details/11/0

Ducci, F. (2020). *Natural monopolies in digital platform markets*. Cambridge University Press. https://doi.org/10.1017/9781108867528

Feasey, R., & Streel, A. de. (2020). *Data sharing for digital markets contestability: Towards a governance framework*. CERRE. https://cerre.eu/publications/data-sharing-digital-markets-competition-governance/

https://doi.org/10.2139/ssrn.3855489

Forbes. (2022). *The Global 2000*. https://www.forbes.com/lists/global2000/?sh=4aed00155ac0

Furman, J., Coyle, D., Fletcher, A., McAuley, D., & Marsden, P. (2019). *Unlocking digital competition, report of the Digital Competition Expert Panel*. Gov.uk. https://www.gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel

Heim, M., & Nikolic, I. (2019). A FRAND regime for dominant digital platforms. *SSRN Scholarly Paper No. 3326186*. Social Science Research Network. https://doi.org/10.2139/ssrn.3326186

IANS. (2020). Over 50 crore Indians now use smartphones, 77% on Internet. *The Hindu*. https://www.thehindu.com/sci-tech/technology/gadgets/over-50-crore-indians-now-use-smartphones-77-on-internet/article30691614.ece

Krämer, J., & Schnurr, D. (2021). Big data and digital markets contestability: Theory of harm and data access remedies. *Journal of Competition Law & Economics*. https://doi.org/10.1093/joclec/nhab015

Krämer, J., Schnurr, D., & Micova, S. B. (2020). *The role of data for digital markets contestability: Case studies and data access remedies*. CERRE. https://cerre.eu/publications/data-digital-markets-contestability-case-studies-and-data-access-remedies/

https://doi.org/10.2139/ssrn.3789510

L'autorité De La Concurrence [French Competition Authority]. (2018). *Portant sur l'exploitation des données dans le secteur de la publicité sur internet.* https://www.autoritedelaconcurrence.fr/fr/avis/

portant-sur-lexploitation-des-donnees-dans-le-secteur-de-la-publicite-sur-internet

MakeMyTrip, Goibibo, OYO fined Rs 392 crore by CCI for unfair business practices. (2022). *The Economic Times*. https://economictimes.indiatimes.com/industry/services/travel/makemytrip-goibibo-oyo-fined-rs-392-crore-by-cci-for-unfair-business-practices/articleshow/94970021.cms

Martens, B., de Streel, A., Graef, I., Tombal, T., & Duch-Brown, N. (2020). *Business-to-business data sharing: An economic and legal analysis*. European Commission. https://joint-research-centre.ec.europa.eu/publications/business-business-data-sharing-economic-and-legal-analysis_en

Montjoye, Y.-A. de, Schweitzer, H., & Crémer, J. (2019). *Competition policy for the digital era*. European Commission. https://data.europa.eu/doi/10.2763/407537

Nonnecke, B., & Carlton, C. (2022). EU and US legislation seek to open up digital platform data. *Science*, *375*(6581), 610–612. https://doi.org/10.1126/science.abl8537

Online platforms and digital advertising market study. (2019). Gov.uk. https://www.gov.uk/cma-cases/online-platforms-and-digital-advertising-market-study

Parker, G., Petropoulos, G., & Van Alstyne, M. W. (2020). Digital platforms and antitrust. *SSRN Scholarly Paper No. 3608397*. Social Science Research Network. https://doi.org/10.2139/ssrn.3608397

Posner, E. A., Weyl, G., & Naidu, S. (2018). Antitrust remedies for labor market power. *Harvard Law Review*. https://harvardlawreview.org/2018/12/antitrust-remedies-for-labor-market-power/

Proposal for a regulation on contestable and fair markets in the digital sector (Digital Markets Act). (2020). EUR-Lex. https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A842%3AFIN

Prüfer, J., & Graef, I. (2021). Governance of data sharing: A law & economics proposal. *SSRN Scholarly Paper No. 3774912*. Social Science Research Network. https://doi.org/10.2139/ssrn.3774912

Prüfer, J., & Schottmüller, C. (2021). Competing with big data. *The Journal of Industrial Economics*, *69*(4). https://doi.org/10.1111/joie.12259

Scherer, F. M. (2007). Technological innovation and monopolization. *SSRN Scholarly Paper No. 1019023*. Social Science Research Network. https://doi.org/10.2139/ssrn.1019023

Shri Piyush Goyal chaired Open Network for Digital Commerce (ONDC). (2021). PIB. https://pib.gov.in/pib.gov.in/Pressreleaseshare.aspx?PRID=1745611

Standing Committee on Finance. (2022). *Anti-competitive practices by big tech companies*. https://loksabhadocs.nic.in/lsscommittee/Finance/17_Finance_53.pdf

United States v. Dentsply International, Inc, 399 F.3d 181 (3d Cir.2005).

Zingales, L., & Lancieri, F. M. (2019). *Stigler Committee on Digital Platforms: Final report*. https://www.chicagobooth.edu/-/media/research/stigler/pdfs/digital-platforms---committee-report---stigler-center.pdf