THE CURIOUS TALE OF BIG DATA AND THE MERGER CONTROL REGIME

By Shahista Khan
LLM from National Law School of India University, Bangalore

Abstract
The advent of user-friendly technology and applications have led to huge amount of data collection at an immeasurable rate by the tech companies. This ‘Big Data’ not only helps in growth of the tech companies but also plays a major role in mergers and acquisitions. Data aggregation and data sharing is one of the reasons behind the big tech mergers such as Facebook/WhatsApp, Google/DoubleClick etc. The merger control regime in India is weakened by the high-threshold test and CCI’s reluctance to go into the data sharing issues in the review process. There is merit in looking at Big Data as an asset/input being transferred in big tech mergers. Further, the acquisition of big data by few big tech companies (GAFAM) can have exclusionary effect, thus, adversely affecting the competition in the market.

INTRODUCTION
“If you are not paying for the product, then you are the product” – The Social Dilemma

The following quote from the Netflix documentary ‘The Social Dilemma’ captures the nature of the digital platform in the present era. The technological market has transformed from a money value economy to a zero-price economy. The value of the services is no longer price-based but data-based. Data has become the new currency of the Internet. The vast amount of data generated through these multi-sided platforms is commonly referred to as ‘Big data’. The Big Data companies process these large datasets to give a personalized experience to users, thus enhancing the overall experience. This leads to a pro-competitive regime where the big techs are constantly competing for user attention by introducing new features and innovations.

However, as Sophocles once said, ‘nothing vast enters the life of mortals without a curse’, the big data analysis also comes with its evils. The flip side of the coin is that the data not only personalizes the user experience but it also subtly manipulates the choices and behavior of the user. One recent example is the famous Facebook-Cambridge Analytica data scandal wherein thousands of user-profiles were harvested for targeted political campaigning in the US elections. Similarly, social media platforms such as YouTube, Twitter, or Instagram works on algorithms that are designed to capture and

1Jeff Orlowski, The Social Dilemma (Netflix 2020).
retain the attention of a user by showing only similar opinions or views, and thus influencing the political, social, or cultural environment of a nation. J. Chandrachud, in the famous privacy case, highlighted the dangers of data mining on an individual choice and his preferences.6 These concerns are directly related to the antitrust/competition regime as some major big tech companies act in collusion with political parties to retain their dominant position in the market by these unsavory tactics.

Apart from the moral concerns, the accumulation of big data by few tech companies might also have the effect of establishing a monopoly in the market. Although one can argue that data is cheap and readily available, there is considerable evidence to suggest that big data owned by dominant companies cannot be replicated without considerate cost and resources. For example, the data available at google does not simply comprise of search history but is coupled with data from other sources, such as fitness data collected through its subsidiary Fitbit, thus making it unique in the market.

Further, major big tech mergers are highly influenced by the factor of data aggregation and data sharing. Though the negative implications of such mergers which are focused solely on big data are recognized by the competition authorities worldwide, there is little to no legal action taken to remedy the situation. In India, the merger control regime is further weakened by high threshold requirements and the exclusion of big data influence from the review process (exception being Jio-Facebook and Jio-Google combination).

Thus, in the present paper, the author proposes that: first, big data plays a significant role in big data mergers, and therefore, there is a need to define ‘big data’ in terms of input or asset that is being acquired. Second, the merger control regime in India is limited to turnover-based threshold requirements which have led to major big tech mergers falling outside the review process. Further, the Competition Commission of India has relied on post-review action to address the problem of data aggregation and collusion (E.g., Google-Jio, and Facebook-Jio). It is submitted that this position is problematic as it ignores the wide implications that data sharing/aggregation can have on the market and the sensitive nature of user data.

The paper is divided into four parts; the first section will try to define big data in an antitrust regime. The II section would deal with the role of big data in incentivizing mergers between big tech companies. The III section will look at the Indian merger control regime in India and the role of CCI in reviewing big data mergers. The IV section will deal with the various approaches adopted in the UK, EU, and the US merger control regime. In the V part of the paper, the author will try to conclude by giving out policy suggestions.

**DEFINING BIG DATA**

1.1. General Definition

The origins of the term ‘Big Data’ are uncertain. However, most literature on data analytics seems to agree that Mr. John

---

Mashey, a retired scientist at Silicon Graphics, first coined the term. He described ‘big data’ as massive unstructured datasets, which were required to be handled and analyzed by complex algorithms, later what came to be known as the field of data analytics. The more specific features of big data were provided by Doug Laney, a former Gartner employee. He gave three characteristics of big data, popularly known as the three Vs, Volume, Velocity, and Variation. Although many subsequent papers have tried to add other features to the definition of big data, such as value, veracity, continuity, and so on, the three Vs are the hallmark of every question on big data.

Volume: One of the characteristic features of big data is that it consists of large volumes of datasets. However, it is not just the volume that distinguishes big data from the data. Usually, data could be handled by analysts, but for big data, there are data scientists who are well versed in deep learning. Around 44 Zettabytes of data has been produced since 2020, and it is expected to grow by 463 exabytes by 2025.

Velocity: With the exponential growth of digital platforms and e-commerce, the data being generated daily is incomprehensible. Thus, it is not the volume but the velocity with which it is being produced. In 2020, the average data produced by an individual was around 2.5 quintillion bytes per day. In fact, the statistics suggest that there are around 4.66 billion active users of the Internet, i.e., 60 percent of the population. These numbers are increasing daily with the rise in new technological advances such as the Internet of things (IoT). The velocity of the data is further increased by the use of real-time data collection, which is commonly referred to as data-casting or forecasting.

Variation: Big data is also characterized by its heterogeneous nature. All kinds of unstructured information are produced due to the vast number of services offered by multi-

---

8 Data Analytics refer to the process of analysing raw data to derive information in relation to user preferences, behaviour and choices.
9 Ibid (n 7).
16 The term ‘Internet of Things’ refer to the network of physical objects which are interconnected to different objects, people or other devices through Internet. Eg. Fitness Bands, Google Home etc.
17 Supra Note at 14. See also, OECD, ‘Big Data: Bringing Competition Policy to the Digital Era – Background Note’ (2016).
Multi-sided platforms\textsuperscript{18}. For example, a social media platform such as Facebook, along with personal information, produces audio, video data as well. Further, the increase in the use of IoT devices has led to the production of varied sensory data.

1.2. Big Data in antitrust regime
There is a dearth of literature when it comes to defining big data in the antitrust regime. The Organisation for Economic Cooperation and Development (OECD) background note on big data defines it as the traditional three Vs described above. Further, it relies on Stucke and Grunes to add a fourth V, i.e., value.\textsuperscript{19} The fourth V is crucial when it comes to antitrust regimes as the value of the data would determine a tech company’s position in the market. The relationship between the three and Vs and volume can be best understood by De Mauro’s definition, “Big Data is the information asset characterized by such a high volume, velocity, and variety to require specific technology and analytical methods for its transformation into value.”\textsuperscript{20}

The value of any data is dependent on the method with its derived such as data-casting or the Internet of things. Thus, the big data in the competition regime is best understood as something so valuable that it is not easily substitutable and can have the effect of exclusion from the market.

The competition authorities abroad have been more proactive in defining big data than India. So, there is a dearth of academic literature and reports on the influence of big data companies on the market. However, there have been few instances where CCI has also tried to tackle the definition of Big Data. For example, in Samir Agarwal v. ANI technologies\textsuperscript{21}, the CCI briefly referred to big data as large data sets. In Matrimony LLC v. Google Inc\textsuperscript{22}, the CCI, while dealing with an allegation against google, defined big data as large volumes of data generated which can affect the consumer preferences and advertising industry. However, there is much-needed research in this area in India.

One aspect that is worth pondering is that merger review includes ‘acquisition of asset’\textsuperscript{23} and whether big data can be looked through the lens of an asset? As the next part of the paper would deal analyze that in big tech mergers, it is data that is being traded under the garb of acquisitions and mergers. So, the question the author leaves with and would come back to the end of this paper is whether the competition commissions should be examining big data under the head of the asset.

**DO BIG DATA INCENTIVIZE MERGERS?**

“The World’s Most Valuable Resource is No Longer Oil, But Data” - The Economist\textsuperscript{24}

\textsuperscript{18}Multi-sided platforms or two-sided platforms provides a platform for interactions of two distinct participants such as a buyer and a seller.
\textsuperscript{20}Ibid.
\textsuperscript{22}(2018) SCC OnLine CCI 1.
\textsuperscript{23}The Competition Act, 2002, s 5(1).
The economist writes that data is the new oil of economy, and like the Oil Industry, it is owned mainly by five big data companies that are dominant in the digital platform: Google, Amazon, Facebook, Apple, Microsoft, or more popularly known as the GAFAM. The merger trends of the big five suggest that this holds true in the global digital market. As of 2020, the GAFAM has acquired over 175 companies, from small start-ups to highly valued companies. Considering the dynamic nature of the digital market, it is possible that the big five might lose their dominant position in the market. However, for now, there is a need to scrutinize the GAFAM mergers closely, so they are not able to block rival entrants.

One way by which the GAFAM is already blocking out rivals is through acquisitions of nascent competitors from the market. For example, the acquisition of Instagram and WhatsApp by Facebook or the acquisition of Myntra by Flipkart and subsequent acquisition of Jabong by Myntra. However, such a practice is primarily attributable to horizontal mergers. There are acquisitions that are hard to be categorized under the three heads of horizontal, vertical, and conglomerate mergers, such as Acquisition of Waze and Fitbit by google, which primarily are for data sharing purposes.

We are attuned to thinking of acquisitions and mergers in price value terms. However, in digital platforms, it is not only the shares of the company that are being transferred but the vast collection of data as well. An important factor in most digital mergers is the sharing of unique sets of datasets possessed by each company. Data might be cheap and readily available, but each platform gathers a different set of data; for example, both Facebook and WhatsApp offer some similar services that giving a platform to connect with people. However, the sets of data produced by both companies would be different as the primary data in WhatsApp is the chat data and contact information of a person, but in Facebook, apart from chat data, visual and audio datasets are being produced as well. When we mix these two groups of data, there will be a unique set of data that the others competitors might find hard to replace.

Thus, based on the presumption that big data is being acquired as an asset, let us take a relook at some of the significant merger cases across the globe.

One of the first cases to address the question of data sharing was Google/DoubleClick Merger. Google, the dominant playing in search engine queries and advertising business, sought to enhance its wealth of data by acquiring DoubleClick’s ad serving technology, which targeted and monitored bought-instagram-as-it-was-a-threat-to-facebook-120073000324_1.html> accessed 18 August 2021.

Office of Fair-Trading Decision of November 11, 2013 in Case ME/6167/13—Motorola Mobility Holding (Google, Inc.)/Waze Mobile Limited.


the performance of ads. The combined data gave an edge to Google in the advertising business as with DoubleClick’s Ad data; it can better predict the behavior of users and provide user-tailored advertisements. Though both Federal Trade Commission (FTC) and the EU commission considered the privacy concerns arising out of sharing personal data of consumers, the potential dangers of Google’s dominant position in the advertising market were ignored.

The Facebook/WhatsApp Merger is also one of the cases where the acquisition of data was the driving force behind the merger. Facebook is a dominant player in the Social Media Network by acquiring WhatsApp not only eliminated the risk of future competition and also got access to tons of user data which is sensitive and private. Though in the initial stages, Facebook denied that it would combine the data of both platforms respecting the privacy policy of the WhatsApp users, it later changed the privacy policy to include mandatory data sharing in case of business accounts. Thus, cementing the position, it was the user data of WhatsApp that Facebook sought to acquire. The recent announcement of Facebook to reinvent itself as the metaverse company would mean tons of user data in the hands of a company with past defaults and loose privacy norms.

One of the interesting mergers that are missed in the mainstream literature is Google/Waze. Waze is an Israeli developer app that collects real-time information from users about accidents, traffic, bottleneck, and so on. Based on the data provided by the users, it then re-routes the map to a faster route. This unique idea which already has 50 million users, captured the attention of big tech companies such as Facebook, Google, and Apple, since most navigation still struggled with accurate and real-time data on route mapping. Google acquired it for 1.1 billion dollars to combine its online data with the offline data of users.

There are several mergers, such as Google/DoubleClick, Apple/Shazam, Microsoft/LinkedIn, where data played an important role in incentivizing the merger.

**FINDING BIG DATA IN THE INDIAN MERGER CONTROL REGIME**

India has a fairly new merger control regime. The current framework came into force in 2011, nearly a decade after the passing of the Competition Act, 2002. Primarily, the merger regime is governed by Sec 5 and 6 of the Act and the Competition Commission of India (Procedure in regard to the transaction of business relating to combinations) Regulations, 2011. Like many jurisdictions, India adopts a turnover-based threshold assessment when it comes to merger control. Under the competition act 2002, every combination needs to be notified and pre-

---

36Google/DoubleClick (n 31).
37Waze/Google (n 29).
approved by the CCI. The term combination is defined under sec 5 of the act. It includes any merger, amalgamation, or acquisition of shares, control, voting rights, etc., by an enterprise that exceeds the threshold requirements of prescribed turnover in a financial year. The threshold requirements are different for an individual acquirer and groups. The section also covers the acquisition of any indirect or direct control in the target company. Every combination that adversely affects the competition in the market is deemed void by the CCI. However, the CCI is yet to completely prohibit a merger, acquisition, or amalgamation under the act. In most cases, the CCI would give some remedial measures and ask for refiling of notifications.

Schedule 1 of the act provides that certain transactions are exempted from the notification requirement. One such exemption is the de minimus or target exemption. If the target company has less than Rs 3.5 billion assets in India or less than Rs 10 billion annual turnover in India, the combination does not need to be mandatorily notified. This exemption poses a problem in the digital economy, especially in cases of multi-sided platforms such as WhatsApp and Google, which provide free services. The Zero price digital platform, thus, is rarely able to meet the threshold requirement provided under the act. Therefore, when Facebook acquired WhatsApp, WhatsApp had less than 10 billion annual turnover, and thus the transaction was exempted and was never examined by the CCI. This raises some very critical issues in the merger control regime. First, the use of high threshold requirements is not sufficient to address the zero value digital mergers. Second, Non-price parameters such as data sharing and network effects are left out of the discussion on mergers.

However, it would be a fallacy to say that the big data mergers have been completely left out of the CCI review. The CCI under sec 3 and 4 of the act has looked into the possible data-sharing concerns arising out of the Facebook/WhatsApp Merger in the Harshita Chawla case. The recent privacy policy of WhatsApp was also reviewed by the CCI. Thus, the CCI is taking a roundabout way to examine merger issues which it is unable to do directly due to the high threshold test.

Big Data is mostly in Indian legislation, but the CCI did take a look at data integration in two of its recent merger orders. The orders reflect the relevance of data, even in a minority shareholding, and the misplaced trust in the big tech companies.

1.3. Analysis of Google-Jio and Facebook-Jio Merger Order

1.3.1. Facebook-Jio Platform Merger Order

In a 26-page long order, the CCI went into all the issues from data integration to net-neutrality and horizontal overlaps. However, the order was a hit and a miss, as would be evidenced by a critical analysis of the order. In 2019, Facebook, through its wholly-owned subsidiary, Jaadu Holdings, acquired 9.99% of shares in the Jio platforms, a subsidiary of Reliance Industries Ltd. Though the acquisition was exempted from notification

---

38 The Indian Competition Act, 2002, s 6.
39 The annual threshold value of turnover or assets are notified by the Central government from time to time.
40 The Indian Competition Act, 2002, s.6.
41 Harshita Chawla, In re, 2020 SCC OnLine CCI 32
43 Jaadu Holdings LLC Combination Registration No. C-2020/06/747.
under the de minimis rule, Facebook submitted the acquisition for review by the CCI. It is to be noted that after the acquisition, Facebook became a majority shareholder in the Jio platforms. Facebook also has affirmative rights under SHA and the investment agreement. The author submits that the rights offered to Jaadu under the agreement are in the nature of strategic commercial transactions. Thus, there is an element of indirect control that should have been examined by the CCI. There have been multiple cases where the CCI has ignored the threshold requirement and looked at the transaction from the control test.

Further, the strategic arrangement between WhatsApp, Jio Platforms, and RRJ for the purposes of Jiomart suggests the potential for the enormous amount of data transfer. Both Facebook and Jio platforms are in the business of processing consumer data for targeted advertisements; thus, combining the data of the social media platform and the telecom authority raises some serious concerns about the data monetization and exclusion of rivals from the market. The position adopted by the CCI was to trust the undertaking provided by the Jaadu that no data would be shared between the parties. However, if we are to learn anything from the previous Facebook/WhatsApp merger case, Facebook tends to deviate from its word.44 The subsequent change of privacy policy which allowed for data sharing between Facebook and WhatsApp suggests that it would be imprudent to rely on only an undertaking by the parties.

1.3.2. Google-Jio Merger Order45
Not to be left behind in the race, soon after Facebook’s investment, google also announced a 7.7 % share in Jio platforms. Like the Facebook-Jio combination, the Google-Jio combination was also examined by the CCI on the grounds of net neutrality and data sharing. The CCI recognized the potential dangers from pooling the data of Google, which has access to rich data about user behavior, and Jio platforms, which is a dominant telecom authority, has access to sensitive user information. However, the CCI passed the transaction based on Google’s undertaking and believes that if there is any anti-competitive behavior through data sharing in the future, an action under sec 3 or 4 can be initiated against google. However, the future action implies the damage would have already occurred, and data, as have been seen from the recent Pegasus incident, if leaked, spreads like wildfire.46

1.4. The CLRC Report
The Competition Law Review Commission (CLRC) report47 does address some of the gaps in the merger control regime. It recommended that the deal-value threshold test should be adopted to capture the big tech mergers which are exempted due to the low turnover threshold. Further, the report adds

45Google International LLC Combination Registration No. C-2020/09/775.
46The Pegasus leak example suggests the ways in which tech companies can secretly collude with governments to exploit the user data. Concerns with Aarogya Setu app and Aadhar card data leak emphasis the need for good data protection laws.
that in addition to the deal-value threshold, local nexus to India must also be established for a merger to become notifiable.\textsuperscript{48} However, there are certain operational challenges to implementing the deal-value threshold which needs to be addressed first. The question then arises is what would be the parameters on which such mergers will be reviewed? Would big data factor in the review of such mergers?

However, as one can see from both the orders discussed above, the CCI has adopted a post-review action when it comes to data sharing. The merger control regime allows CCI to look into potential anti-competitive behavior caused by any transaction. However, when it comes to the potential dangers of data sharing, which can lead to data monopoly and exclusion, it has skirted away from the problem.

LESSONS FROM FOREIGN JURISDICTIONS

The limitations of the turnover-based threshold test are not unique to India. However, most other jurisdictions have found a way to skirt this by introducing alternate tests or exercising the residuary powers.

1.5. EU: To refer or not to refer:
The merger or ‘concentrations’ in the European Union is regulated by the EUMR regulations\textsuperscript{49}. Similar to India, the merger control regime is suspensory and based on prescribed turnover thresholds requirements.\textsuperscript{50} The threshold requirement has led to many mergers falling outside of the purview of the EU commission. However, this is remedied by one of the peculiar features of the EUMR regulations, i.e., the right to refer mergers. Reference can either be pre-notification referral\textsuperscript{51} or post-notification referral\textsuperscript{52}. A pre-notification referral is when the parties to the concentration, which is capable of being reviewed under more than three national competition laws, ask for examination by the EU commission on the basis of a reasoned submission. For example, the Facebook/WhatsApp merger. A member state makes a post-notification referral to examine the potential threats to competition caused by such a merger in the member state(s) market. A merger having no community dimension can only be transferred. For example, Austria referred the Apple/shazam merger to the EU commission, which was otherwise not caught by the turnover-based threshold requirements.

Recently, the EUMR regulations were amended, keeping in view the threat of ‘Killer acquisitions.’\textsuperscript{53} Now, there is no threshold (national or EU) requirement for a reference by a member state.\textsuperscript{54} Apart from this, countries like Austria and Germany have adopted the value-based threshold to ensure that big data companies with zero revenues

\textsuperscript{48}The deal value threshold recommendation is adopted in the Competition (Amendment) Bill, 2020.
\textsuperscript{50}EC Merger Regulation, art 1(2) and 1(3).
\textsuperscript{51}EC Merger Regulation, art 4(5).
\textsuperscript{52}EC Merger Regulation, art 22.
\textsuperscript{53}Killer Acquisitions is when a dominant company in the market acquires the nascent rival companies.
are exempted from the merger review process.  

1.6. UK: The Share of Supply Test  
In the UK, the merger control regime is governed by the Enterprise Act, 2002. Unlike India, in the UK, the requirement of notification is not compulsory; the parties are free not to notify. Sec 23 of the act lays down two conditions under which a transaction can be reviewed: first, when two or more entities cease to be distinct, and second, the turnover threshold is exceeded by the merger or if the ‘share of supply test’ is met. The first limb of the section is determined through the control test, whereas the second limb requires either high turnover or the determination of supply in the market. Ordinarily put, a merger can be reviewed if it increases the overall supply of a product or service by 25% in the territory of the UK.

1.7. USA: The Three Tests  
In US, a merger transaction becomes notifiable under the HSR Act if it fulfills any of the following three tests: first, the commercial test, i.e., whether either of the party to the transaction is involved in any business activity in the US; second, the size of the transaction test otherwise known as deal-value threshold test; and third, the size of the parties which requires the FTC to look at the annual turnover of acquired or acquirer company. One aspect of the US merger control regime is worth looking into is the definition of an asset. It is broadly defined to include both tangible and intangible goods. Although there is no precedent to suggest

The various tests suggests that India also need to look into adopting a merger control test which is efficient in capturing the big data mergers. The Competition Bill, 2020 seems to address this by giving the central government power to decide factors on which the combination could be reviewed. However, it remains to be seen whether big data integration would be one of the factors or not.

CONCLUSION  
The traditional price-based notions of the antitrust regime are no longer capable of addressing the complex nuances of digital markets. The digital platform is growing by leaps and bounds, and the current competition law is falling behind. We live in an era where a digital entity knows about you more than you know about yourself. It is almost impossible to contain the vast amount of personal data that will be produced as we move towards more technological innovations. However, the adverse effects of this can be controlled by appropriate law reforms. Some of the suggestions worth pondering over are:

1.8. Data as an input or asset  
Big Data is one of the main factors driving acquisitions and mergers in the digital market. Then why are most proposed amendments and laws of various jurisdictions more focused on the valuation of the

55CLRC Report, p. 130.  
56Miranda Cole and Rolf Ali, ‘The CMA’S Approach to Jurisdiction in Recent Merger Cases’ (Covington Competition, 2020) <https://www.covcompetition.com/2020/08/the-  
57Hart-Scott-Rodino Antitrust Improvements Act, 1976 (the HSR Act).
transaction or the parties? Why are we not looking at Big Data as an asset? One problem seen in all the big merger cases is the inability to define big data precisely. The author agrees that there are certain limitations to defining big data because of its dynamic nature. However, this can be remedied by looking at firstly, the substitutability of the data produced from proposed data integration/aggregation and second, by examining the data from an effects-based doctrine, i.e., whether the data combined through the proposed combination would have the effect of excluding competition from the market or establishing a monopoly in the market.

1.9. Data-Driven Theories of Harm
In recent years, the competition authorities have evolved various theories of Harm while evaluating the mergers and their possible consequences. Some of the theories of harm evolved by courts in data-driven mergers are: first, consumer harm caused by quality degradation and targeted advertisement; second, conglomerate effects caused by tying or bundling; thirdly, foreclosure or exclusionary effects caused due to data combination.58 Though the tests do pave a path for controlling the mergers, the interpretation of the court is wide enough to allow for such mergers. The courts have relied on the notion that data can be easily substituted. This assumption is somewhat based on limited knowledge of working digital platforms. Thus, this can be remedied by allowing experts who could testify as to how data algorithms could affect the competition market.

Further, the author would like to conclude by stating that there can be no hard and fast rule to evaluate the data-driven mergers considering the dynamic and evolving nature of digital platforms. For a constantly changing field, we need a law which is not static and can be evolved with the current and emerging demands of the market.

*****