WORKING PAPER NO: 677

Industrial Concentration in India

Gaurav Ghosh
KREA University
gaurav.ghosh@krea.edu.in

Subhashish Gupta
Economics Area
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 560 076
sgupta@iimb.ac.in

Industrial Concentration in India

Abstract

Concentration of economic power is usually frowned upon because its social and political effects are substantial and undesirable. On the flip side one can argue that concentration provides the necessary scale to compete in international markets. Researchers use concentration measures like the HHI or Lerner's Index, but coverage is typically restricted to the specific industries under study. Wouldn't it be useful to have measures for the universe of industries, and at different levels of aggregation? And then aggregate these to calculate concentration for a country? This would allow us to say, for example, that India is more concentrated than the USA.

We address these questions in our paper. Using Prowess data (in this early draft), we find that one can estimate concentration ratios at different levels of industrial aggregation, ranging from the NIC Section to the NIC Sub-class for the universe of firms. However, these estimates can be unreliable because of heterogeneous data coverage across industries. This is more of a problem at lower levels of aggregation, but higher levels are not unaffected. We also find that country-level ratios lead to overestimates of concentration unless appropriate weights are used.

Keywords: Concentration, indices, aggregation, industry classification

1. Introduction

Societies are increasingly worried about of inequality of late. Most of their disquiet stems from the growing inequality of incomes. Newspapers regale us with tales of the fortunes of Zuckerberg, Bezos and Adani, and million-dollar salaries are common for CEOs, especially of technology companies, while the rest of us scrape together a living. Sometimes the glare of attention focuses on economic or market concentration. Firms like Google, Facebook and Amazon stride their respective industries like giants obliterating all challenge while taking consumers for a ride. This perception, true or false, has led to competition authorities to conduct investigations of wrongdoing and to impose penalties if evidence can be found. Most of the concern with market concentration seems to be directed towards high technology businesses usually within the information technology sector. These are prone to monopolization and dominance according to commentators. However, there has been consolidation within other industries as well, in Europe and North America. The economic consequences of concentration could include increased prices for consumers and reduced prices for inputs. Besides concentration of economic power is usually frowned upon. The oldest competition law, the Sherman act, was specifically introduced to curb the power of trusts (a conglomeration of businesses) in the USA. Some countries in Europe and Asia resorted to nationalization of some industries to reduce concentration. The social and political effects of concentration are substantial and undesirable. On the flip side one can argue that concentration provides the necessary scale to compete in international markets.

India has had a convoluted history when it comes to inequality and economic concentration. The guiding principles of state policy implores the state to reduce inequality and concentration of economic power in the constitution that came into effect in 1950. Since then, successive governments tried to reduce inequality through taxation policies and also to reduce market concentration through industrial licensing and the Monopoly and Restrictive Trade Practices Act. To what extent these policies reduced inequality and concentration is unknown. There are certainly anecdotes about how individuals got around them, often with the help of conniving officials. At any rate India had a change of heart in the 1990s and ushered in regime of liberalization. Industrial licensing was abolished, tax rates lowered and the MRTPC Act was replaced with the Competition Act that sought to rely more on "rule of reason" rather than declare a whole slew of activities "per se" illegal as under the old MRPTC Act.

To what extent this change reflected a change in the mindset of Indians in general is unclear. Did they transform themselves from firm believers in the power of the government over markets to free marketeers and believers in competition? It seems very unlikely. In fact, there is no corresponding word for competition in any Indian language. Also, it took a long time for the competition act to become operational which probably reflects the ambivalence of the Indian public. At any rate the competition act has been operational for a while and the country has registered fairly high rates of growth in the recent past. Industrial licensing and the prevalence of business groups probably resulted in excessive diversification during the pre-liberalization regime. Liberalization and the unleashing of market forces should have led to consolidation as businesses divested themselves of peripheral activities leading to an increase in concentration. At the same time growth should have spawned a whole host of new firms willing to challenge the incumbents. What has been the result of these two opposite forces? An increase or a decrease in competition? What about market concentration? Are the two synonymous? It is worth finding out.

Our research will lead to detailed information on market concentration in India, as well as supplying a mechanism for modeling the impact of changes in market structure (e.g., through entry, exit or merger).

There is a large literature dealing with concentration. We have highlighted some of these in the second section. There are both theoretical and empirical issues as well as microeconomic versus macroeconomic concerns (Todd 2008, Covarrubias, Gutierrez, and Philippon 2019). There are also many industry studies (e.g., Borenstein and Bushnell 1999, Kim and Singhal 1993, Nevo 2001). Two papers which are of particular interest to us are Saraswathy (2019) and Jacob, Paul and Sen (2021). Saraswathy calculates HHIs and concentration indices for industries in India. Jacob, Paul and Sen (2021) estimates markups for industries in India through the methods suggested by De Locker and Warzynski (2012) and compares them to HHIs. Saraswathy (2019) finds that some industries are marked by strong concentration and that overall, there has been some decline in concentration. Jacob, Paul and Sen (2021) confirm that HHIs have been declining but find that markups have been rising, providing inconclusive evidence about the nature of competitive pressures in the economy. Both of these studies do not look deeper into the data and base their observations about aggregate concentration by a casual look at the data. We are aware of the limitations of the data and also provide a measure of aggregate concentration.

The issue of concentration and rise of market power in the Indian economy is worthy of robust investigation, which we do in the following pages. The rest of this paper is organized as follows: We begin by contextualizing our research question by considering the pros and cons of concentration in Section 2. We then move on to formally describing measures of concentration often discussed in the literature in Section 3. The data that is used to estimate our measures of concentration for India are described in Section 4. Section 4.2 presents results and Section 6 concludes.

2. Is concentration bad?

This may seem a strange question to start this section. The obvious answer, to most commentators, would be undeniably so. The answer may not be that straightforward, as we will find out later. Let us first start with listing the many ways concentration may be inimical to an economy. We will divide them into micro and macro effects, though the divide may not always be that clear.

The most prominent charge against concentration is that concentration leads to a rise in in market power and consequently a rise in prices and is detrimental to consumer welfare (White 2013, Syverson 2019, Wright 2018). To reach this conclusion one can rely on textbook economic theory which shows that a monopolist will typically charge higher prices than a more competitive market. Usually, as the number of firms in a market declines the price tends to go up as does the price-cost margin, or the mark-up. However, as is usual for theoretical models, the result depends on the assumptions made about the features of the market and the competitors. The other approach is empirical where researchers investigated the effect of concentration on prices and profitability. This approach relied on the structure-conduct-performance (SCP) paradigm popularized by Bain and others. Here it is assumed that market structure (monopoly, oligopoly etc.) determines conduct (pricing and output decisions), which in turn determines performance (profitability etc.). Looking across industries in the USA it was found that the hypothesis was supported by data (Bain 1959).

However, there are problems with the SCP paradigm. How do we know the direction of causation? Does structure cause conduct or is it the other way around? It is possible that predatory behavior

(conduct) can lead to monopolization (structure). Also, predatory behavior may take the form of keeping prices deliberately low which results in lower profitability, but only temporarily. In the same vein with extreme price competition (Bertrand competition) one can produce competitive results with just two firms. Finally, there is the issue of contestability. For some industries prospective entry may act as a disciplining device. Firms within the industry may know that supernormal profits will attract entry and result in lower prices and profits and possibly even exit. It would be much better to play safe and stick to average-cost pricing even with zero economic profits. In an extreme situation there may be average cost pricing with only one firm in the industry. The theory of contestable markets requires the absence of sunk costs which may preclude its usefulness. However, it should be clear that economic concentration does not necessarily imply the presence of market power.

Another source of profitability could be production efficiencies and managerial effectiveness. Empirical studies of profitability within industries show a wide variety in profits. Obviously, some firms are better than others at generating profits and consistently so. Thus, firms may possess resources or capabilities which cannot be easily imitated or purchased in the market.

Given the problems with the SCP approach one may look for alternative approaches to measuring market power. One such approach is to estimate the Lerner's index. This is the price minus the marginal cost divided by the price. It measures by what percentage does the market price exceed the marginal cost. Standard economic theory suggests that with perfectly competitive markets price would equal marginal cost. Then the Lerner index would be zero for such markets. The higher the Lerner's index the more market power the firm possesses.

However, the problem here is that it is difficult to estimate marginal costs from accounting data. One can, however, use methods from the New Empirical Industrial Organization (NEIO) literature to estimate versions of the Lerner's index. Typically, researchers rely on the techniques developed by Breshnahan (1989) or the H statistic developed by Pazner and Rosse (1987). One problem with this approach seems to be the wide range of estimates that have been reported for the same industries.

There are other microeconomic effects of concentration besides rise in market power and increased prices. Fewer firms would imply lower levels of output and consequently lower levels of input demand, including labor. This in turn would lead to lower wages and increasing inequality. In addition, there is the vexing question about the effect of market power on innovation. At its heart the question is about whether a more concentrated industry structure is more likely to foster innovation than a more competitive one. Schumpeter took the view that big firms would have the wherewithal to spend on research and development and would be more likely to spur innovation. Arrow and others argued that a competitive firm would have more to gain from innovation than a monopolistic firm that would merely indulge in a form of economic cannibalism through innovation. The evidence on this issue is mixed and Aghion et al (2005) find that the relation seems to be shaped like an inverse U.

Another strand in the literature suggests that increased competition may be accompanied by increased concentration. The empirical literature on the evolution of markups in the USA finds increased markups over time and that bigger firms have higher markups (De Loecker, Eeckhout and Unger 2020). A possible explanation for this phenomenon is that bigger firms are more efficient and that over time smaller firms who are less efficient tend to get weeded out. In the presence of increased substitutability, the result turns out to be different. Lower trade costs or other factors may increase competition between products and as a result the price-cost margin reduces but since inefficient firms

leave the industry concentration increases. So, there is increased concentration accompanied by lower market power.

The macroeconomic effects of increased market power could explain some of the trends seen globally (IMF 2019) These include sluggish investment, a falling rate of return on government bonds even though rates of return on capital are stable, increased difference between wealth as measured by asset values and productive wealth, declining share of labor income, rising inequality and lower productivity growth. As suggested by the theory increased market power could reduce the need for further investment in innovative activity thereby reducing productivity growth. The return on capital would still be relatively high because of higher profits compared with a more competitive market. Also, shareholders assets would grow in value compared to their productivity as measured by their marginal product. This would lead to a lower share of workers incomes. India has reported a decline in the rate of capital formation over the last few years and there is evidence of rising inequality. Rising market power could be the reason behind it.

3. Measuring Concentration

A simple way of measuring concentration is to use the K-firm concentration index ("C_K"). This measures the relative contribution or market share of the top K firms within an industry. Suppose there are N firms in the relevant market, indexed by $i = 1, \dots, N$. Let the market share of the ith firm be s_i. We find the set of $K \subseteq N$ largest firms and sum across the relative contributions of all firms within K to obtain C_K (see equation **Error! Reference source not found.**). Higher values of C_K indicate more concentration.

$$C_K = \sum_{i \in K} s_i \tag{1}$$

Another common method is to compute the Herfindahl-Hirschman Index ("HHI"), which is the sum of squares of market shares of all firms in a relevant market. Market shares are expressed as percentages, which means that an HHI varies between 0 and 10,000, with higher HHIs implying greater market concentration. If firms are indexed by $i = 1, \dots, N$ and the market share of firm i is s_i, then the HHI is calculated as shown in equation (2).

$$HHI = \sum_{i=1}^{N} s_i^2 \tag{2}$$

The third approach could be through the Lerner's index, which, looks at the profit margin, which is the difference between price and marginal cost (see equation (3)). The idea is that the higher the profit margin, the larger the market power of the firm. This follows from microeconomic theory and has ample empirical support. Competitive markets are those where firms price close to market cost and do not wield market power. As a corollary, one concludes that firms with high margins must have market power.

Commented [GG1]: SG: indirect measure

$$\mathcal{L} = \frac{P - MC}{P} \tag{3}$$

Finally, one can use several other measures to look at the level of rivalry in the industry. This could include entry and exit rates. Market power is often maintained because of barriers to entry by competitors. These barriers may be regulatory (e.g., public utilities) or technological (e.g., some two-sided marketplaces). Incumbents tend to be profitable and are therefore unlikely to exit. Therefore, we may expect lower rates of entry and exit in industries with high market power.

The most direct measure of market power is the Lerner's index. The other measures of market power using concentration indices can be misleading. The degree of concentration would be the result of nature of competition that is played out in the market. However, as we have remarked earlier Lerner's index is difficult to calculate since it is difficult to calculate marginal costs using accounting data. New methods have been developed to calculate marginal costs indirectly, but they all rely on certain assumptions. Thus, the usage of concentration ratios and HHIs continue.

There are some other issues to keep in mind. Use of concentration ratios firstly comes with an implicit assumption of Cournot-like competition (Syversion 2019). If the nature of interaction among the players in the market is different using HHI can lead to misleading information. A closer look at individual industries may be warranted. Another issue is that of relevant markets (Benkard, Yorokuglu and Zhang (2021). Measures of concentration are often used to decide on antitrust issues. In all such situations it is important to define the relevant market by the degree of substitutability of different products. Consider air travel. The relevant market here is for different sectors. The Bangalore to Delhi route cannot be substituted by the Bangalore-Mumbai route. So, considering all passengers flying on all airlines everywhere in the country would be a wrong measure of concentration. One should look only at individual markets and its close substitutes.

A more difficult feat is to measure the degree of competition and market power in individual countries or across regions (Bai, Mao, Zie and Zhang 2014). This has become necessary because of the economic consequences of greater market power. Governments and the public are worried about the level of market power in their country comparatively (Carbo, Humphrey, Maudos, and Molyneux 2009). Thus, the construction of an index of overall market power assumes some importance. One should point out that a relatively simple method of comparing competitive benchmark prices may provide some indications with lower computational and theoretical problems (Hausman and Sidak 2007).

One can use measures of concentration that are commonly used by researchers and practitioners alike to calculate concentration within a specific industry, but how does one calculate concentration for a country? Is it meaningful to say that, for example, India is more concentrated than the USA? We can theoretically calculate a market concentration measure for all the industries and add them up in some way to arrive at a measure of overall concentration. This raises a plethora of issues. Some of them include: At what level of aggregation do we measure? Should we calculate concentration geographically and then add up? How do we add up? What are the variables that we should use to calculate market shares? Finally, there are methodological problems with the construction of the overall concentration index.

Consider the question of adding up or aggregation. Simple averages imply that each firm within an industry group gets the same weight. But is this always appropriate? One may argue that firms located close [not close] to the firm under investigation should be given weight [not given weight] because they are part of the same [not part of the same] geographic market. Or the question of the market share variable. Industry-wide sales (or income) or revenues are commonly used, but in some cases, it may make sense to look at assets, or headcount, or other variables.

4. Data discussion

Apart from these conceptual issues, we must also select our data sources. Two sources that are widely used by practitioners and researchers for quantitative analysis of Indian firms are Prowess and the Annual Survey of Industries ("ASI"). The former is maintained by the Center for Monitoring the Indian

Economy, a private organization, and the latter by the Ministry of Statistics & Program Implementation.

4.1 Preparing the data

We began our analysis by looking at the market share of firms included in Prowess. We considered the universe of firms in Prowess over the period stretching from Financial Year ("FY") 2011-12 to FY2021-22. We used the database of vintage "Sep 2022" to download data on firms in the Prowess database in any of these years. This was the newest vintage on our date of data download – which was 31.10.2022. This yielded a dataset of 313,544 observations covering 40,908 companies.

Preliminary analysis of the data revealed that some firms had filed two financial statements in a given FY.[1] When we calculate a market concentration index for a given industry each year, we aggregate across all observations associated with that industry-year combination. This leads to double counting of all firms that have filed two annual statements in that year. As a result, these firms receive twice the weightage of other firms in the calculation of the market concentration index. This double weightage arises only because of an administrative decision by the double-filing firms and should therefore be removed. We remove this double counting by deleting the earlier financial statement in the case of all firms filing two statements in a year. This removes 521 observations, leaving a dataset of 313,023 observations.

These observations are distributed across financial years as shown in Table 1. There are fewer observations for the most recent financial year, FY2021-22, than for earlier ones. There are two plausible reasons for this. Many firms may have not released their annual reports by the release date of the Sep 2022 vintage of Prowess. Or, firms may have released their annual reports, but they may not have been included in Prowess as of the release date of the Sep 2022 vintage. Irrespective of the specific reason, there may be patterns in the omission – such as, perhaps, certain types of firms releasing their annual reports earlier and other types releasing their annual reports later – that lead to biases in the estimation of market concentration indices for FY2021-22. To avoid such possibilities, we remove all 6,320 observations for FY2021-22, leaving 306,703 observations.

Table 1: Prowess data extract, observations per year

Financial year	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22
Number of obs.	27,147	26,038	28,661	31,761	33,833	33,809	33,554	32,747	31,029	28,124	6,320

Source: Prowess, own calculations

Finally, we find that some data points for the variables that we shall use to construct our market concentration indices are missing. After removing the corresponding observations, we are left with 268,627 observations over a decade of financial data extracted from Prowess. These are the observations that we shall use to estimate concentration ratios. The missing data pose a potential problem. To the extent that these data are missing in some systematic fashion, working with the remaining data will lead to biased and inconsistent estimates of the "true" concentration ratios. This

---

[1] *Inter alia, Prowess stores standalone annual financial statements of Indian companies. We will extract data from these financial statements to calculate market concentration indices.*

may be the case if, for example, smaller firms are more likely to have missing data. If, on the other hand, the data are randomly missing, then our estimates will uncover the "true" concentration ratios. We cannot be sure if data is missing systematically or not. Accordingly, we will proceed under the assumption that data is missing at random, and that our estimates will be unbiased and consistent. Table 2 shows how we arrived at our final data set.

Table 2: Steps taken to arrive at final data set

Step	Observations
Prowess data download	313,544
- **Double counts**	313,023
- **FY2021-22**	306,703
- **Missing data**	268,627

Source: Prowess, own calculations

4.2 Industry classifications

The intent of this paper is to measure market power by sector (or equivalently, industry). In Prowess, each firm's industry is indicated by its National Industry Classification ("NIC") code. The NIC is a classification nomenclature promulgated by the Ministry of Statistics of Programme Implementation. All economic activity in India can be classified into an NIC code (Central Statistical Organization, 2008). The structure of NIC classification is as follows:

1. At the top, we have 21 **NIC Sections**, labelled from "A" to "U".
    a. Example: Section A is Agriculture, Forestry and Fishing
2. NIC Sections are divided into **NIC Divisions**, which are identified by 2-digit codes.
    a. Example: Division 02 is part of Section A and is defined as "Forestry and Logging"
3. NIC Divisions are divided into **NIC Groups**, which are identified as 3-digit codes.
    a. Example: Group 022 is part of Division 02 and is defined as "Logging"
4. NIC Groups are divided into **NIC Classes**, which are identified as 4-digit codes.
    a. Example: Class 0220 is part of Group 022 and is defined as "Logging"
5. NIC Classes are divided into **NIC Sub-classes**, which are identified as 5-digit codes.
    a. Example: Class 02201 is part of Class 0220 and is defined as "Gathering and preparation of firewood"

All firms in our data set are in Sections "A" to "P". There are no firms in Sections "Q" to "U" in Prowess. This is likely because the latter sections pertain to nonmarket, non-corporate, and / or unorganized activities. Section Q, for example, refers to "Human health and social work activities". The distribution of observations across Sections A-P is shown in Table 3.

Table 3: Firms in Prowess by NIC sector, FY2012-21

NIC Sector	Number of Observations
Admin & support	14166
Agriculture	4288
Construction	23648
Education	6097
Finance	65142

NIC Sector	Number of Observations
Hospitality	4633
InfoComm	16192
Manufacturing	99936
Mining	2020
Power	7397
Public admin	391
Real estate	188
Skilled services	9644
Trade	50277
Transport	8799
Water & waste	205

Source: Prowess, own calculations

It is immediately apparent that economic activity in India, at least as captured by Prowess, is clustered in a few sectors: manufacturing, trade, finance, and construction. Some sectors, like real estate services and water & waste management, have only a handful of firms. Real estate services, for example, has only 188 observations in Prowess over a decade, meaning an average of less than 20 observations per year. We would expect such sparsely represented sectors to have high HHIs at all levels of aggregation, from the Section level to the Sub-class level.

Prowess provides no guidance about why certain NIC Sections have sparse representation. A couple of arguments present themselves. One reason may be that firms in certain NIC sections exist but are not appearing in Prowess. This may be the case for NIC Sections like real estate services and agriculture. Many such service providers tend to be small and unorganized: think of real estate and small-scale farmers. Prowess compiles firm-level data from public sources, including stock market filings, and submissions to the Ministry of Corporate Affairs. Only larger firms will be found in such databases. As a result, a large proportion of real estate services or agricultural activity will never appear in Prowess.

Another reason might be that some sections do indeed have very few firms within them, and that Prowess does accurately capture the state of the industry. This may be the case with the water & waste section, which has city-level public sector entities. Concentration in these industries may very well be the truth.

5. Results

We present some results on market concentration as measured by the HHI in Section 5.1. Further results on the HHI and other indices will be added to this paper later.

5.1 Herfindahl-Hirschman Index

As mentioned in Section 3, $HHI = \sum_{i=1}^{N} s_i^2$, where s_i is the market share of firm $i = 1, ..., N$. However, this definition does not specify how market share, s_i, is calculated. The literature usually makes use of two financial variables to construct market share. The first variable is the total income of a firm. The second is total assets. Both variables are understood to capture the size of a firm, and firm size is used to construct HHIs. There are advantages and disadvantages to using either variable.

Income is closely related to sales. It is a flow variable, measures current period market vigor, and therefore captures the current size of a firm. By this metric, firms with more income are considered larger than firms with lower income. It is relatively easy to define, and this definition does not vary much across different types of firms, products, and industry. Further, since definitions are quite standardized, measuring income is straightforward.

However, because income is a flow variable, it is affected by seasonality and business cycles. Some firms may see relatively higher income during downturns (e.g., pawn brokers), while others may see higher income during upturns (e.g., white goods retailers). The relative sizes of these firms will therefore modulate because of business cycle or seasonal effects and this modulation may have little to do with their actual business heft. Finally, income may not be appropriate when comparing certain types of firms. Trading firms, for example, have large income volumes when compared to other measures of financial size. Manufacturing firms, on the other hand, have lower income volumes when compared to their actual size. Therefore, when using income to compare manufacturers and traders, we may come away with a distorted picture of the relative sizes of these firms.

Assets are a stock variable. They, therefore, are affected less by seasonality and business cycle effects: we may expect them to be less variable over time. However, some assets, such as intangible assets, may be difficult to value. These may include patents, copyrights, trademarks, brands, trade dress, client lists, and goodwill. There are also challenges associated with the valuation of assets over time. Typically, asset values are distributed across multiple years on a balance sheet according to depreciation rules and the lifetime of the asset. However, there is some arbitrariness associated with depreciation rules and lifetimes. This means that asset values may – to some extent – be driven by valuation rules instead of the inherent worth of the asset or demand-supply fundamentals. Intangible assets may, in particular, be overvalued, and these form increasingly important components of firm balance sheets. Assets may be no better than income in inter-industry comparisons: just as income may overstate the size of services firms, assets may understate their size.

However, we observe that these variables complement each other (Centre for Monitoring Indian Economy, 2022).

- Income measures current performance, assets measure historic performance
- Income is affected by seasonality & business cycles, assets are not
- Income is easy to measure, assets are not
- Income is likely to overvalue service firms, assets are likely to undervalue them

Hence, it makes sense to construct a composite variable that is a function of assets and income. This variable would likely be less affected by seasonality than income, but less influenced by historic performance than assets. Further, it may be ideal for inter-industry comparisons since it balances the tendency of income to overvalue service firms against the tendency of assets to undervalue services firms.

With the previous discussion in mind, construct market shares and HHIs using three variables

1. Total income
2. Total assets
3. Composite = Total income + total assets

5.1.1   Total income

We begin by presenting HHIs across NIC sections over the FY2012-21 period in Figure 1, where HHIs are calculated based on total income as reported in Prowess. Only the NIC sections with the highest HHIs over the last decade have been colored and labeled. Other NIC sections remain in grey and are unlabeled. This is to increase the readability of Figure 1. Other figures have a similar color scheme.

Figure 1: HHI as measured by total income for NIC sections, FY2012-21
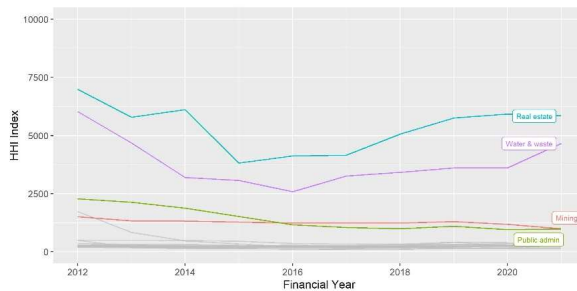
We observe that the Real Estate Services and Water & Waste sections have the highest levels of market concentration with HHIs consistently exceeding 2500 over the entire decade ending in FY2020-21. This indicates high levels of concentration, as defined by the US Horizontal Merger Guidelines (US DoJ; US FTC, 2010). Mining and Public Administration have HHIs that would lead them to being classified as having moderate levels of concentration[2] in some years in the last decade. Other NIC sections have low HHIs and have been shaded in grey.

HHIs across NIC sections over the FY2012-21 period in Figure 2, where HHIs are calculated based on total income as reported in Prowess. We have colored and labeled the top 10 percent of divisions by HHI. Division 75 (Veterinary Activities) has the highest HHI. Indeed, it has the highest possible HHI of 10000 from FY2013-14 onwards. This is because there is a single firm in NIC Division 75, which entered the database in FY2013-14. Is it likely that Veterinary Activities constitute a monopoly in India? Or is it more likely that we get this result because Veterinary Activities are offered by service providers too small to enter Prowess? Our guess is that the latter is more likely.

---

[2] *HHIs lying between 1500 and 2500 imply moderate concentration under the US Horizontal Merger Guidelines.*

Figure 2: HHI as measured by total income for NIC divisions, FY2012-21
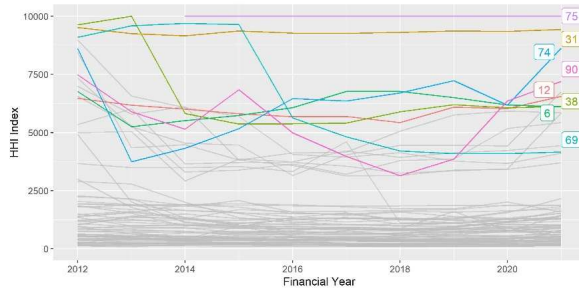
Other divisions with high HHIs are Division 31 (Manufacture of furniture), Division 74 (Other professional, scientific, and technical activities), Division 90 (Creative, arts, and entertainment), Division 12 (manufacture of tobacco products), Division 38 (Waste collection, treatment and disposal activities; materials recovery), Division 6 (Extraction of crude petroleum and natural gas), and Division 69 (Legal and accounting activities). Divisions 12, 38, and 6 may indeed have high levels of market concentration. Division 69 may see a Stackelberg fringe with an oligopolistic center and many small firms. Other divisions in this list may seem to be concentrated because most service providers in these divisions do not appear in Prowess.

Finally, we plot HHIs across NIC classes in the manufacturing sector in FY2012-21 in Figure 3. We have colored and labeled the top 10 percent of divisions by HHI.

Figure 3: HHI as measured by total income for NIC classes associated with manufacturing, FY2012-21
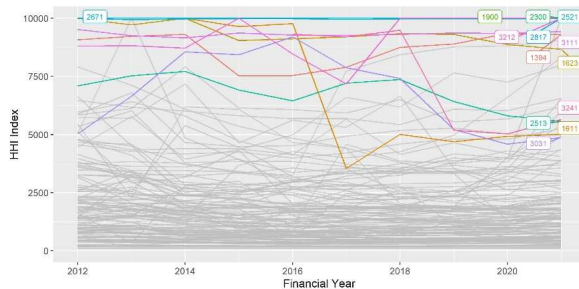
There are a few classes with HHIs of 10000, indicating monopolies at the NIC class (or 4-digit) level. These include 1900 (Coke & refined petroleum), 2300 (Other non-metallic mineral products), 2521 (weapons & ammunition), and 2671 (Optical instruments). There are many other classes with high HHIs. We omit further details. Interested readers can look up 4-digit codes in the NIC classification documentation (Central Statistical Organization, 2008).

Looking at Figure 1 to Figure 3, some observations can be drawn. First, market concentrations increase as we move from broadly defined markets to narrowly defined markets. We see this in the dispersion of the grey lines (representing the industry sectors with lower HHIs) across the three figures. In the

NIC section plot (Figure 1), which represents the highest level of aggregation, almost all HHIs are below 1000. In the NIC division plot (Figure 2), which represents the intermediate level of aggregation, many grey lines lie above the 2500 mark (indicating high concentration), and many more lie in the 1500-2500 range (indicating moderate concentration). In the NIC class plot (Figure 3) there is a much larger presence of industry sectors in the high concentration and moderate concentration ranges. None of this is surprising: broadly defined markets have more firms within them.

A second observation is that market concentrations seem to be falling over the years. This seems especially evident in the narrowly defined markets in Figure 3. However, this observation is muddled by the COVID-19 year of FY2020-21 when we observe an increase in HHIs in many sectors. It remains to be seen whether this uptick in market concentration is transitory or indicates a fundamental change in industry structure.

5.1.2   Total assets

HHIs across NIC sections in FY2012-21, where HHIs are calculated based on total assets, are shown in Figure 4. The results are similar to those in Figure 1, where HHIs are calculated at the level of NIC sections but using total income. As in that analysis, Water & Waste and Real Estate Services have the highest HHIs, followed by Mining and Public Admin. However, the overall level of concentration is lower, with lower HHI values. Unlike in the previous analysis (where HHI was based on total income) Real Estate Services are only moderately concentrated (HHI between 1500 and 2500) in some years. Mining is moderately concentrated throughout, but Public Administration has low concentration in many years.
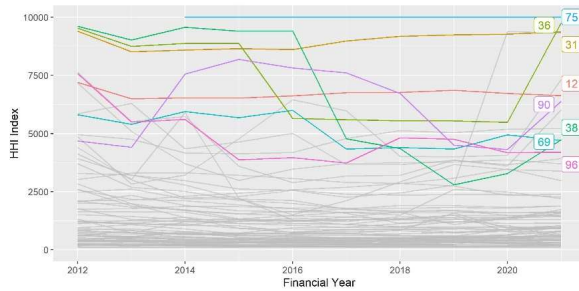
Figure 4: HHI as measured by total assets for NIC sections, FY2012-21
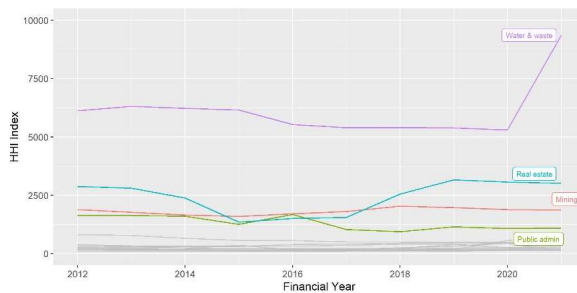
HHIs across NIC divisions in FY2012-21, where HHIs are calculated according to total assets, are shown in Figure 5: HHI as measured by total assets for NIC divisions, FY2012-21

. We have colored and labeled the top 10 percent of divisions by HHI.

Figure 5: HHI as measured by total assets for NIC divisions, FY2012-21

Divisions 75, 31, 12, 90, 38, and 69 have high levels of concentration, just as they did when concentration was measured according to total income (see Figure 2). Other divisions with high HHIs are Division 36 (Water collection, treatment, and supply) and Division 96 (Other personal service activities). These divisions were not among the top 10 percent of divisions when measured by total income. Hence, there are some differences between the market concentration measures.

HHIs across NIC classes in the manufacturing sector in FY2012-21, where HHIs are calculated according to total assets, are shown in Figure 6. We have colored and labeled the top 10 percent of divisions by HHI. Most of the highly concentrated classes were also highly concentrated when measured by total income. Classes that appear in Figure 6 but not in Figure 3 are Class 1812 (Services related to printing), Class 2681 (Magnetic and optical media), and Class 1211 (Related to tobacco products).

Figure 6: HHI as measured by total assets for NIC classes associated with manufacturing, FY2012-21
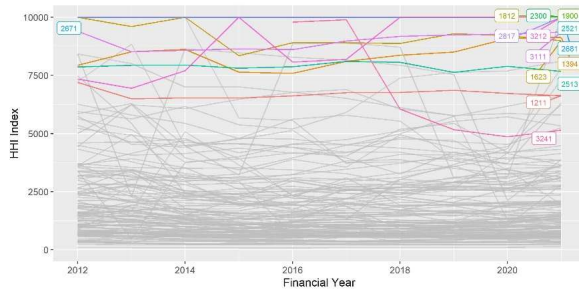
Comparing our HHI measurements using 2 variables (total income and total assets) and at 3 different scales (Section, Division (2-digit) and Class (4-digit), we find that results look substantially similar. Most NIC sections that with high HHIs under total income also had high HHIs under total assets. Analogously, most NIC divisions and classes in the top ten percent of HHIs under total income are also in the top ten percent of HHIs under total assets.

5.1.3   The composite measure of size

As discussed in Section 5.1, total assets and total income have advantages and disadvantages when used to calculate market shares. Particularly, service firms will tend to have higher HHIs under total income calculations and manufacturing firms will tend to have higher HHIs under total asset calculations. Our composite measure of firm size is simply the sum of total assets and total income and should lead to a measure that neither favors service or manufacturing firms. Figure 7 presents HHIs across NIC sections, division, and classes, in FY2012-21, where HHIs are calculated based on our composite measure.

We have already observed that HHI estimates using total income (see Figure 1 – Figure 3) are similar to HHI estimates using total assets (see Figure 4 – Figure 6). While exact HHI numbers may vary, the *ranking* across sectors at the top end of the distribution is similar. The same sectors that have high HHIs under total assets also have high HHIs under total income. Hence, we would expect the high-HHI sectors under total income or assets to also have high HHIs under our composite measure, which is just an unweighted sum of the other variables.

If we look at estimated HHIs under the composite measure in Figure 7, we see that this is indeed the case. When we look at the sectors at a high level of aggregation (i.e., NIC sections) it is the same four NIC sections that have the highest HHIs: Water & Waste, Real Estate Services, Mining, and Public Administration. When we look at lower levels of aggregation, once again, the set of high-HHI sectors (whether measured by Division or Class) remain unchanged.

Figure 7: HHI as measured by total assets + total income for NIC sections, NIC divisions & NIC classes, FY2012-21
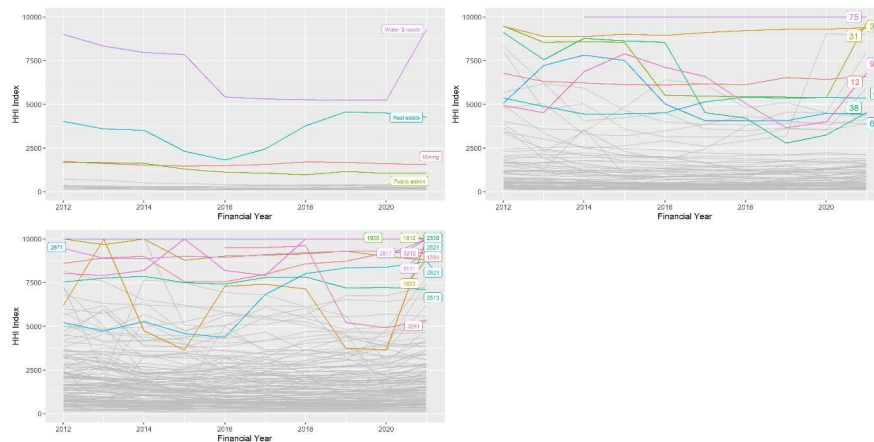
5.2 Aggregate Indices of Concentration

As we can see different industries have different levels of concentration and have exhibited different kinds of behaviour over the years. So it would be difficult to describe the behaviour of the economy as a whole in terms of concentration. It would be foolhardy to suggest an increase or decrease in overall concentration. Yet, we shall endeavour to do just that. The reason being that while it be

difficult conceptually to create an overall index of concentration it would nevertheless provide us with some interesting information. Just as knowing the average height and weight of all Indians provide us with some indications of wellbeing, knowing the "average" concentration could reveal something about industrial concentration.

We calculate aggregate indices of concentration in four ways. One is the simple average across industry-level concentration indices. The other three are weighted averages, which differ in the types of weights used. Details are below.

1. Simple average: If there are J sectors, $j = 1, ..., J$ indexes a specific sector, and C_j is the concentration index for sector j, then the simple average measure of aggregate concentration is $\sum_{j=1}^{J} C_j/J$. Effectively, each sector receives equal weight in the estimation of aggregate concentration. This measure will lead to biased estimates if each sector does not have the same weight in the overall economy: if, for example, construction is a more important component of the Indian economy than water & waste.
2. Weighted average method 1 (WA_J): where the weight is the number of firms in a given sector. For example, if there are 100 firms in agriculture and 200 in mining, then the mining sector will receive twice as much weight in the construction of the aggregate index. If w_j is the weight of sector j, then the weighted average is calculated as $\sum_{j=1}^{J} w_j C_j / \sum_{j=1}^{J} w_j$.
3. Weighted average method 2 (WA_{TI}): where the weight is the total income of a given sector. Firms which generate more income will get higher weight in index construction.
4. Weighted average method 3 (WA_{TA}): where the weight is the total assets of a given sector. Firms which generate more assets will get higher weight in index construction.

The weighted averages each give more weight to larger sectors, whether measured by number of firms within a sector (WA_J) , total sales / revenues / income in a sector (WA_{TI}), or total assets in a sector (WA_{TA}).

We construct aggregate concentration indices using the three variants of the HHI described in Section 5.1. Results are provided below.

Table 4 shows the numbers for aggregate concentration (weighed by assets) for the period 2012-2021. It is difficult to interpret the absolute numbers but we can look at the trends and differences. The first observation is that the unweighted HHIs are much larger than the weighted HHIs. This is simply an artificial effect. Weighting involves arithmetic operations on some measure of size (see explanations for WA_J, WA_{TA}, WA_{TI} in Section 5.2) and in the case of fractions will necessarily be smaller. It is interesting to note that the aggregate concentration index using simple averages decreases till 2016 across indices and increases thereafter. The increase from 2020 to 2021 is also large. This observation is not reproduced when weighied by assets or income. The weighed indices exhibit similar trends to that of the unweighted index. Figure 8 shows the same information graphically.

Table 4: Aggregate HHI using total assets (under 4 weighting schemes)

YEAR	SIMPLE AVERAGE	WA_J	WA_{TI}	WA_{TA}
2012	1228.15	237.65	236.10	304.29
2013	1161.09	219.51	222.84	291.58
2014	1085.78	207.78	207.93	283.21
2015	1013.62	197.53	200.26	279.33
2016	790.64	202.98	214.64	295.80
2017	826.93	219.66	230.41	316.56
2018	895.40	250.15	256.93	354.70
2019	969.15	274.47	278.52	360.10
2020	977.85	321.31	325.21	398.42
2021	1244.06	274.47	281.83	411.33

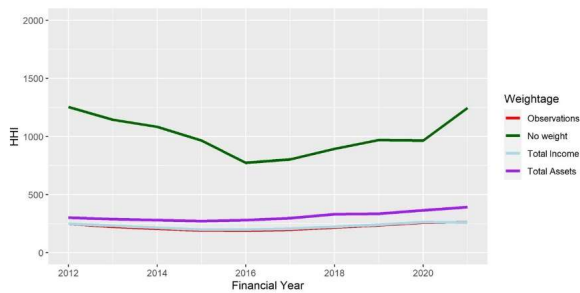*Figure 8: Aggregate HHI using total assets (under 4 weighting schemes)*

Table 5 and

Table 6 gives us aggregate HHIs where we use business income and income plus assets respectively. The trends are the same as that for assets. The numbers for income are smaller than for assets which could be because income is usually smaller than assets for most firms in brick and mortar businesses. However, firms in Information Technology (IT) and platforms tend to be light in assets. In this case assets will not be a good indicator of size. One should also not exaggerate trends. Looking at Figure 8, Figure 9 and

 we see that aggregate HHI is almost a straight line. Of course slopes will depend the units chosen but even then there doesn't seem to be a significant increase or decrease in concentration as measured by HHI.

Table 5: Aggregate HHI using total income (under 4 weighting schemes)

YEAR	SIMPLE AVERAGE	WA_J	WA_{TI}	WA_{TA}
2012	1365.46	324.56	290.66	292.14
2013	1103.98	263.89	264.86	269.36
2014	976.39	233.05	245.66	260.88
2015	782.21	208.33	217.40	247.92
2016	736.84	192.67	197.35	241.81
2017	763.32	189.75	197.09	251.62
2018	832.71	205.43	206.73	281.23
2019	923.26	227.39	229.00	290.06
2020	918.48	234.09	235.43	308.07
2021	1025.86	268.62	253.48	327.23

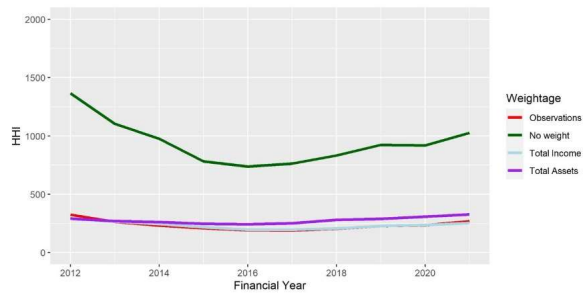*Figure 9: Aggregate HHI using total income (under 4 weighting schemes)*

Table 6: Aggregate HHI using total assets & income (under 4 weighing schemes)

YEAR	SIMPLE AVERAGE	WA_J	WA_{TI}	WA_{TA}
2012	1253.06	248.65	248.38	301.95
2013	1145.12	224.73	233.26	288.14
2014	1083.33	207.76	215.83	279.73
2015	965.08	191.93	198.48	272.08
2016	774.02	190.08	198.30	280.71
2017	802.84	198.40	206.25	297.57
2018	894.66	218.40	223.22	331.86
2019	969.45	237.40	240.86	336.47
2020	965.72	259.85	264.29	365.35
2021	1245.86	262.67	259.82	391.59

Figure 10: Aggregate HHI using total assets & income (under 4 weighing schemes)
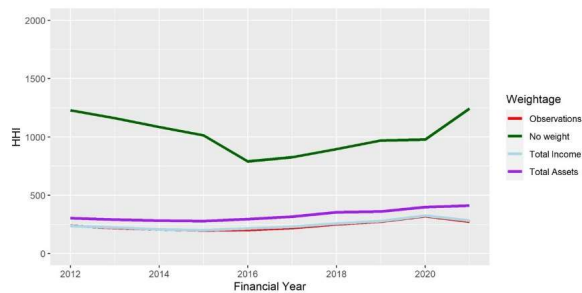
6   Conclusion

To date, we have collected and organized a decade's worth of data from Prowess. We have then used these data to estimate HHIs using three variables (total income, total assets, and total income + total assets) at three levels of aggregation (NIC Sections, NIC Divisions, and NIC Classes).

We have found results across the three variables to be similar in the sense that it is the same group of industry sectors that have high HHIs across the three variables. We have also found some evidence of falling HHIs over time, but this trend reverses in the COVID-19 year of FY2020-21.

There is further work to be done in expanding and tightening this analysis. First, we are yet to estimate market concentration using other measures like K-concentration ratios and the Lerner Index. Second, we still need to grapple with the nature of Prowess data. As we have observed in the case of the NIC section "Real Estate Services", a high HHI appears to be an artefact of missing data. Many real estate service providers like property brokers are small and unorganized and will not appear in a database like Prowess. In other cases, like in manufacturing of tobacco, high HHIs may indeed indicate market concentration. How are we to distinguish between the two cases? It is likely that our intuition and sectoral knowledge will need to guide us.

Third, there is much supplementary analysis that can done, given the scale and density of our results. These include robustness checks and merger simulation exercises.

References

1. Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and Innovation: An Inverted-U relationship. Quarterly Journal of Economics 120, 701–728.
2. Bai, Chong-En, Mao, Jie and Zhang, Qiong, (2014), Measuring Market Concentration in China: The Problem with Using Censored Data and its Rectification, China Economic Review, 30, 432-447.
3. Bain, J.S. (1959) Industrial Organization: A Treatise. 2nd Edition, John Wiley, London.
4. Benkard, C. Lanier, Yurukoglu, Ali and Zhang, Anthony Lee (2021), Concentration in Product Markets.
5. Borenstein, Severin and Bushnell, James (1999), An Empirical Analysis of the Potential for Market Power in California's Electricity Industry, The Journal of Industrial Economics, Vol. 47, No. 3, pp 285-323.
6. Breshnahan, Timothy F, (1989), Empirical Studies of Industries with Market Power, Handbook of Industrial Organization
7. Carbo, Santiago, Humphrey, David, Maudos, Joquin and Molyneux, Philip (2009), Cross-country Comparisons of Competition and Pricing Power in European Banking, Journal of International Money and Finance
8. Covarrubias, Matias, Gutierrez, German and Philippon (2019), From Good to Bad Concentration? US industries over the past 30 years, NBER Macroeconomics Annual 2019
9. De Loecker, Jan, Eeckhout, Jan and Unger, Gabriel (2020), The Rise of Market Power and the Macroeconomic Implications, The Quarterly Journal of Economics, Vol. 135, No.2, 561-643.
10. Hausman, Jerry A. and Sidak, Gregory, J. (2007), Evaluating Market Power Using Competitive Benchmark Prices instead of the Herfindahl-Hirschman Index, 74, No. 2, Antitrust Law Journal, 387-408.
11. International Monetary Fund (2019), The Rise of Corporate Market Power and its Macroeconomic Effects in World Economic Outlook, April 2029.
12. Jacob, Tinu Iype, Paul, Sunil and Sen, Neelanjan, (2021), Aggregate Markups, Market Power and Redistribution of Income
13. Kim, Han E. and Singal, Vijay (1993), Mergers and Market Power: Evidence From the Airlines Industry, Volume 83, No. 3, pp 549-569.
14. Mitton, Todd, (2008), Institutions and Concentration, Journal of Development Economics, 86, 367-394.
15. Nevo, Aviv, (2001), Measuring Market-power in the Ready-to-eat Cereal Industry, Econometrica, Vol 69, No. 2 pp 307-342.
16. Saraswathy, Beena, (2019), Economic Reforms and Market Competition in India: An Assessment, Working Paper 216, ISID.
17. Saraswathy, Beena, (2022), Product Market Concentration with Ownership Characteristics: En Exploration into Consolidation Activities of Business Groups, Working paper 245, ISID.
18. Syversion, Chad (2019), Macroeconomics and Market Power, Journal of Economic Perspectives, Vol. 33, No. 3.
19. White, Lawrence J, (2013), Market Power: How does it arise? How is it measured? In The Oxford Handbook of Managerial Economics, Thomas, Christopher R. and Shugart, William F., Oxford.
20. Wright, Joshua D., (2018), Market Concentration-Note, OECD.
21. Central Statistical Organization. (2008). *National Industrial Classification (All Economic Activities).* New Delhi: Ministry of Statistics and Programme Implementation.
22. Centre for Monitoring Indian Economy. (2022). *Prowessdx India: Database Dictionary.* Mumbai, India: Centre for Monitoring Indian Economy.
23. US DoJ; US FTC. (2010). *Horizontal Merger Guidelines.* Washington DC: US Department of Justice; Federal Trade Commission.