Market Concentration and Productivity: Evidence from the UK *

Anthony Savagar†
University of Kent

Oluwaseun Aguda
OPS, Ontario

Yannis Galanakis
King's College London

Jingwei Wu
University of Kent

March 2024

Abstract

We measure product market concentration and business dynamism in the UK from 1997 to 2020 and study the relationship with productivity. Our results show that concentration in the UK is increasing among narrow industries on average, but for a broad market definition, concentration and business dynamism are stable. We find a negative relationship between concentration and productivity for the average firm, but a positive relationship for the average worker. This occurs because higher industry concentration is associated with better allocative efficiency, measured by the proportion of workers in higher productivity firms.

Key words: Product market concentration, productivity, business dynamism, UK economy, Business Structure Database (BSD).

JEL Classification: D2, D4, E2, L1, L4, O4

---

*Acknowledgements: The authors would like to thank the following people for helpful discussions: Steve Davies, Franco Mariuzzo, Chiara Criscuolo, Matej Bajgar, Sara Calligaris, Rebecca Riley, Patrick Schneider, Sophie Piton, Jakob Schneebacher, Kyung In Hwang, and participants of the ESCoE Conference 2020. Views are those of the authors and errors are the responsibility of the authors.
We provide a web app to analyse our data further https://asavagar.shinyapps.io/UK-market-structure/.

†**Corresponding author:** Anthony Savagar, E: a.savagar@kent.ac.uk

1 Introduction

In many advanced economies, product market concentration is rising. Rising concentration can indicate weaker competition, which can harm productivity, or rising concentration may reflect more efficient firms acquiring market share, which enhances productivity. The aim of our paper is to document product market concentration in the UK, and to examine the relationship between concentration and labour productivity.

The paper makes three contributions. First, we document product market concentration in the UK. We find that concentration for a broad market definition is stable over the sample period 1997-2020, but there was an increase in concentration up to 2016 for a sub-sample that excludes financial services. For a narrower market definition (SIC 5-digit), we find that concentration is increasing on average. Secondly, we document business dynamism facts on firm entry and exit levels and measures of allocative efficiency.[1] We find that levels of entry and exit are stable over the period 1997-2020. Allocative efficiency, which captures the extent to which workers are employed at more productive firms, improved until the mid 2010s but has declined since, particularly among high allocative efficiency industries. Finally, we analyse the relationship between concentration and labour productivity. We find a negative relationship between product market concentration and labour productivity for the average firm, but a positive relationship for the average worker. The difference occurs because high concentration is positively associated with allocative efficiency, indicating that in concentrated industries, workers are more densely distributed at high-productivity firms.

Beyond the UK policy setting, our paper advances understanding of aggregate concentration behaviour in several ways. First, our data is population-wide across all legal forms. This coverage is more representative of the whole economy than popular proprietary datasets that are biased towards limited liability businesses which submit full financial accounts.[2] Our data include any business with a tax record, either through value-added tax (VAT) or payroll tax (PAYE). This covers unlimited companies, self-employed individuals, NGOs, and public sector bodies, such as medical surgeries and academy trusts, provided that they have a staff member on payroll or sufficient turnover to register for VAT. This broader inclusion leads to a more accurate measure of concentration as the true market is captured.

A second area that we advance is understanding the effects of market definition on

[1]Generally, 'business dynamism' refers to a suite of economic indicators. Akcigit and Ates (2021) present ten measures for the US. We focus on those related to concentration, entry & exit and productivity. There is scope for future research on the UK to analyse employment reallocation, as initiated in Lui, Black, Lavandero-Mason, and Shafat (2020), and firm-level growth dispersion.

[2]Bajgar, Berlingieri, Calligaris, Criscuolo, and Timmis (2023) and Bajgar, Berlingieri, Calligaris, Criscuolo, and Timmis (2020) assess the advantages and disadvantages of Orbis data.

concentration measures. To our knowledge, we are the first to document the sensitivity of concentration measures to market definitions based on SIC industries. We show that in the UK narrow industry definitions of markets have become more concentrated, while broad industry definitions are stable. This is related to Rossi-Hansberg, Sarte, and Trachter (2021) who show that the *geographic* definition of a market is important for understanding concentration. They show that in the US, local markets have become less concentrated, but aggregate markets have become more concentrated.

Lastly, our article provides new insights into the complex relationship between product market concentration and productivity. At the firm level, we show a negative relationship between concentration and average-firm productivity. At the industry level, we decompose employment-weighted productivity, which represents average-worker productivity, into unweighted productivity, which represents average-firm productivity, and allocative efficiency, which captures the weight of employment in high-productivity firms. We find that while increased concentration decreases average-firm productivity, it increases allocative efficiency. Overall, the positive effect outweighs the negative effect, leading to a positive relationship between concentration and average-worker productivity.

Broadly, this contributes to current debates in the market power literature about whether concentration is 'good' or 'bad' (Covarrubias, Gutiérrez, and Philippon 2020), due to superstar firms or competition abuses. The negative channels are present in our work, but so are the positive channels. Ultimately, when we quantify these effects, we find the positive channel to be stronger. Overall, this supports the superstar firms 'winner-takes-all' hypothesis (Autor, Dorn, Katz, Patterson, and Van Reenen 2017), but does not ignore that there are negative effects associated with concentration, potentially due to antitrust abuses (Philippon 2018).

Theory shows that market concentration and productivity can be positively or negatively related, and this is also reflected in empirical work (Aghion, Bloom, Blundell, Griffith, and Howitt 2005). High concentration can decrease productivity if it reduces competition, raises barriers to entry, or encourages rent-seeking behaviour, such as political lobbying. On the other hand, high concentration can increase productivity through scale economies, network effects, or R&D investment, as in Schumpeterian growth literature. Furthermore, in various models concentration is ambiguously related to market power and competition. A Cournot model delivers a positive relationship between concentration and market power, as fewer firms raise concentration and increase price setting ability. However, in other frameworks, a decrease in competition increases substitutability due to less differentiation, reduces price-cost margins, and increases concentration (Asplund and Nocke 2006; Melitz and Ottaviano 2008). In Schumpeterian growth literature, the profits from market power are required to innovate and improve productivity (Aghion and Howitt 1992). Review articles by Holmes and Schmitz Jr (2010) and Syverson (2019) provide comprehensive analysis. Our paper emphasises the complexity in the relationship between concentration and productivity, providing evidence on some of the channels through which concentration can be positively or negatively related to productivity.

Related Literature: Recent research finds that product market concentration is rising in the US (Grullon, Larkin, and Michaely 2019; Autor, Dorn, Katz, Patterson, and Van Reenen 2017).[3] The evidence for Europe is mixed. Bajgar, Berlingieri, Calligaris, Criscuolo, and Timmis (2023) show rising concentration in Europe using Multiprod and Orbis data, while Gutierrez and Philippon (2022) suggest more stable concentration in Europe. The differences occur due to coverage differences in proprietary datasets.

The evidence for the UK is less established, but there is a growing number of policy reports and working papers on the topic. Bell and Tomlinson (2018) analyse UK concentration using BSD data. They find that the market share of the top 100 companies rose from 18% to 23% between 2004 and 2016. Additionally, the weighted-average CR5 across 5-digit sectors increased from 39% to 42%, and the weighted-average HHI increased from 880 to 940 units.[4] Aquilante, Chowla, Dacic, Haldane, Masolo, Schneider, Seneca, and Tatomir (2019) cover market concentration in a broader study of UK market power (markups) and monetary policy. They conclude that *'there is no clear trend [in aggregate concentration]'* based on the largest 100 firms from 1998-2016 using Worldscope data on large firms. This is consistent with our finding for broad market definitions, but we show that the conclusion is sensitive to the granularity of the market definition, which helps to reconcile with the Bell and Tomlinson (2018) result. Recent work by Davies (2021) reports a rise in UK concentration and high concentration levels for a subset of 4-digit industries in the UK 1997-2018. Cellan-Jones, Farook, Ferrari, Harris, Rutt, and Walker (2022) summarise the findings of the Competition and Markets Authority's (CMA) *State of Competition Report 2022* which reports a growth then decline in CR5 when averaged with revenue-weights across 4-digit industries (CMA 2022). Corfe and Gicheva (2017) focus on consumer industries and find rising concentration. In addition to studying alternative market definitions, and extesnive data and methodological background, we also extend the literature by analysing concentration distributions and the relationship to productivity.

To our knowledge, no recent research has examined the relationship between concentration and productivity in the UK, and international studies typically exclude

[3]Grullon, Larkin, and Michaely (2019) use CRSP-Compustat merged database on publicly-listed firms. They also incorporate information on private firms from the U.S. Census Bureau and the U.S. Bureau of Labor Statistics. Autor, Dorn, Katz, Patterson, and Van Reenen (2017) uses data from the U.S. Economic Census.

[4]They analyse 608 5-digit SIC sectors. They omit sectors in 'financial services', 'wholesale of fuels' and sectors with high public sector employment.

smaller firms. Bighelli, Di Mauro, Melitz, and Mertens (2023) document recent concentration trends in Europe excluding the UK. They find that concentration has risen since 2008 and that it is *positively* related to productivity. This evidence supports the competitive market ('winner takes all') hypothesis, where most efficient and innovative producers gain a higher market share (Van Reenen 2018). The positive correlation is at the sector level, and it is primarily driven by reallocation from less productive to more productive firms. This is similar to our work which shows the positive effect on allocative efficiency offsets the negative effect on unweighted productivity (average-firm productivity). Furthermore, unlike other work, we also provide firm-level regressions which suggest the negative effect on the average firm holds when we control for firm effects, ignoring distributional changes. In other work, Autor, Dorn, Katz, Patterson, and Van Reenen (2017) find there is a positive relationship regardless of the productivity measure (output per worker, value-added per worker, TFP, or patents per worker). An important distinction of our work is that we capture smaller firms in our dataset, which can help to explain the negative effect of concentration on the average firm, but the positive effect once allocative efficiency is considered.

In Section 2 we discuss our data. In Section 3 we present descriptive statistics on concentration and business dynamism. In Section 4 we analyse the relationship between labour productivity and concentration.

2 Data

Our data source is the Business Structure Database (BSD). The BSD is a firm-level dataset provided by the UK Office of National Statistics (ONS) to accredited researchers. It includes basic information on the near population of UK firms, approximately 2m per year, and is annual 1997-2020. The data is collected for tax purposes. A firm is on the BSD if it qualifies for value-added tax (turnover exceeds £85,000 in 2022) or has at least one payroll employee.

The advantage of the BSD is that it has near-universal coverage of UK firms across all legal forms, providing they have a record with the tax office. This makes it an ideal dataset for studying concentration, and entry and exit which require data on the entire market. This differs from studies of concentration that use proprietary datasets which cover larger, limited liability, firms (e.g. Orbis, Compustat and Worldscope) but are unrepresentative of sole proprietors who make up roughly half of the UK business population. The main variables of interest for us are employment, turnover, and industry, and we calculate entry and exit based on activity.

Annual observations in the BSD can correspond to a firm's economic activity for up to the previous two calendar years. This is because the snapshot is taken early in the calendar year and it summarises the most recent accounts the firm has submitted.

This explains the presence of the Great Recession with a lag in our below descriptive statistics.

2.1 Full Sample and Sub-sample

We present descriptive statistics for a full-sample and a sub-sample of the dataset. The sub-sample excludes sectors that are known to be poorly measured or in which using turnover to represent output is misleading.

1. *Full-sample*: Includes all one-digit sectors.

2. *Sub-sample*: Excludes nine one-digit sectors. The following are excluded: Financial sector; Agriculture; Mining; Electricity; Water; Real Estate; Public Administration and Defense; Education; Human Health and Social Work Activities.

The sub-sample contributes half of aggregate sales and a third of employment. The financial sector accounts for the largest turnover in aggregate UK turnover in the BSD followed by the Wholesale sector. In the case of employees, Education, Public Administration, Human Health represent a significant portion of employment and accounts for the difference between the full sample and sub-sample aggregate employees. For our regression analyses and descriptive statistics at a granular industry level, we do not omit sectors because they are controlled for either with fixed effects in regressions or because we are analysing at the industry level.

Section (SIC07 One-digit)	No. of Divisions (SIC07 Two-digit)	Full sample	Sub-sample
Agriculture, Forestry and Fishing	3	Yes	
Mining and Quarrying	5	Yes	
Manufacturing	24	Yes	Yes
Electricity, Gas, Steam and A/C	1	Yes	
Water Supply and Waste Management	4	Yes	
Construction	3	Yes	Yes
Wholesale, Retail and Motor Trade	3	Yes	Yes
Transport and Storage	5	Yes	Yes
Accommodation and Catering	2	Yes	Yes
Information and Communication	6	Yes	Yes
Financial and Insurance Services	3	Yes	
Property (Real Estate Activities)	1	Yes	
Professional, Science and Tech.	7	Yes	Yes
Administrative and Support Services	6	Yes	Yes
Public Administration	1	Yes	
Education	1	Yes	
Human Health and Social Work	3	Yes	
Arts, Entertainment and Recreation	4	Yes	Yes
Other Services	3	Yes	Yes
Household Production	2		
Extraterritorial Activities	1		
Total Sections / Divisions: 21 / 88		**19 / 85**	**10 / 63**

We always omit Household Production and Extraterritorial Activities.

Table 1: Full Sample and Sub-sample

3 Descriptive Statistics

In this section we aggregate the firm-level data and study how the broad trends compare to well-known aggregate trends.

Figure 1 shows an increasing trend in the number of firms (enterprises) and there are declines during the Great Recession. This growth in the business population is consistent with other proxies of business activity, such as the number of limited-liability companies registered with Companies House (Companies House 2023), and Business Population Estimates (BPE) from the Department for Business.[5] Neither external data

[5]The Companies House reference https://www.gov.uk/government/

source extends pre-2000, and during this period the growth in firm population is exceptionally strong, potentially reflecting coverage improvements after the dataset was initiated. In our analysis we exclude pre-2000 data, but report it for our remaining descriptive statistics.

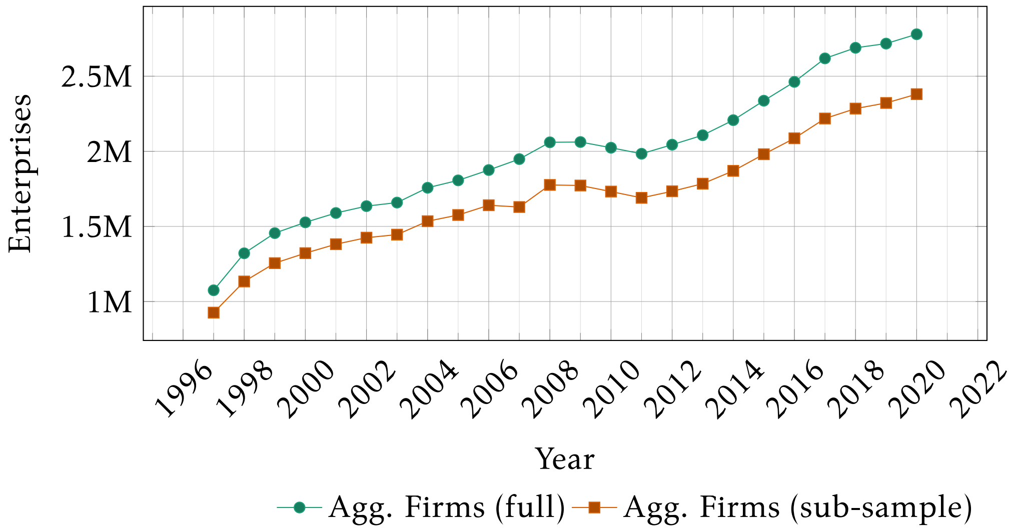


Figure 1: Aggregate Firms (BSD, 1997-2020)
Source: Authors' calculation based on BSD 1997-2020

Minor measurement changes increased the coverage of firms in 2008, 2012, 2014 and 2015. In 2008, the ONS added PAYE only firms. These are firms below the VAT turnover threshold but that have a registered employee. In 2012, 2014 and 2015 tax changes altered the coverage thresholds. These coverage events do not cause obvious breaks in the data.

3.1 Sales, Employees and Labour Productivity

First, we show the underlying components of labour productivity which is the ratio of sales to employees, and then labour productivity itself. These three plots show that the firm-level data captures the main trends in the aggregate data.

Figure 2 presents aggregate real sales (2016 prices) for the sub-sample and full sample.[6] Aggregate sales is the sum across all firms in a given year. A puzzling trend in the full sample is high and declining sales between 1997-2003, which disappears in the sub-sample when financial services and other sectors are dropped. This trend causes differences between the full and sub-sample for variables based on sales such

statistics/incorporated-companies-in-the-uk-july-to-september-2023/incorporated-companies-in-the-uk-july-to-september-2023 and the Department for Business data https://www.gov.uk/government/statistics/business-population-estimates-2022/business-population-estimates-for-the-uk-and-regions-2022-statistical-release-html.

[6]Price adjustment for the real series are applied at the two-digit level before aggregation. The price deflators for each two-digit industry are given by the ONS.

as concentration and labour productivity. Since sales is our proxy for output which is usually measured by GDP or GVA, it should broadly correspond to GDP trends.[7] This seems to be the case. Both samples show an upward trend in aggregate sales over the 2000s with a dip in 2010-2011 that, given the timing considerations of the BSD, corresponds to the recession period of 2008-2009.[8]

As a robustness check we can compare our total turnover figures to total turnover data from the Annual Business Survey (ABS) which is publicly available. The ABS is used in the construction of national accounts. It surveys a representative, random sample of firms, stratified by sector, region, and employment size, and then weights these strata to give aggregate figures. The data excludes financial services. We note that ABS analysis (Figure 1) shows an increase from 3T to 3.5T in nominal terms between 2008 and 2016.[9] This closely replicates the nominal sales (excl. finance) in our BSD data over the same period. We report nominal sales in Appendix Figure 17. An important reason to document this consistency with national accounts data is that alternative approaches to measuring concentration, which use proprietary data such as Orbis, can mismeasure concentration, as explained by Bajgar, Berlingieri, Calligaris, Criscuolo, and Timmis (2023), and may underestimate productivity dispersion between the median and tenth percentile firm due to under-representation of small firms (Bajgar, Berlingieri, Calligaris, Criscuolo, and Timmis 2020).

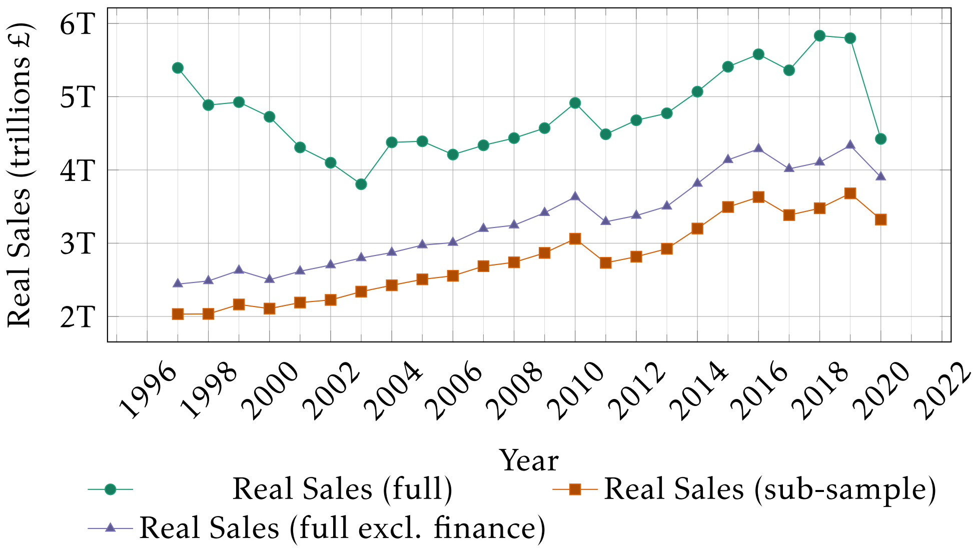


Figure 2: Aggregate Real Sales (BSD, 1997-2020)
Source: Authors' calculation based on BSD 1997-2020

[7]Sales are only a proxy for GDP because they include the value of intermediate goods. This creates double-counting that is exacerbated when there are firms with long value-chains. The BSD only includes sales data. There is no information on value-added.

[8]Due to BSD timings, we might expect a lag of up to 2 years which means economic activities in 2009 goes mostly to 2011 BSD records.

[9]The source of ABS analysis is https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/bulletins/uknonfinancialbusinesseconomyannualbusinesssurvey/uknonfinancialbusinesseconomy2016regionalresults.

Figure 3 shows aggregate employment data in the BSD. It also reflects documented aggregate trends. Comparing BSD employment data to aggregate UK employment data shows that firms in the BSD covers about 98% of total UK aggregate employment. As at the first quarter of 2018, UK official statistics report employment of 32.36 million, while BSD employment data captured in March 2018 was 31.64 million.[10] The trend in BSD employment is similar to aggregate employment with a lag of one year. Before 2010, aggregate UK employment was at its peak in 2008 while aggregate BSD employment was at its peak in 2009. This shows the firms in the BSD cover a significant portion of UK business activity.

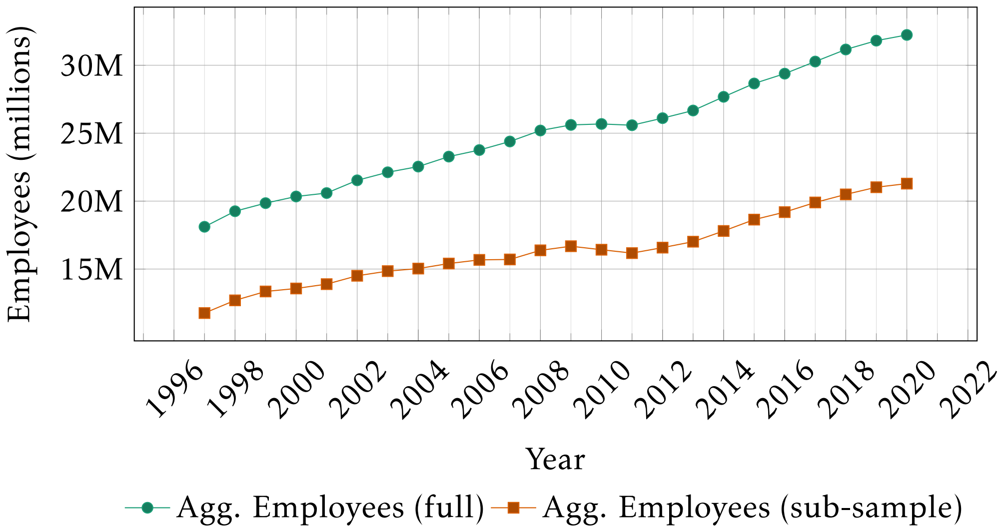


Figure 3: Aggregate Employees (BSD, 1997-2020)
Source: Authors' calculation based on BSD 1997-2020

Figure 4 plots aggregate labour productivity according to the BSD data. Labour productivity is calculated as aggregate real turnover divided by aggregate employees in a given year.[11] This is equivalent to measuring firm-level labour productivity (revenue per worker), and calculating the weighted-average using the firm's share in aggregate employment as the weight. Throughout the paper we use real revenue per employee as our measure of productivity. This is consistent with other work that takes a broad macroeconomic approach, such as Bartelsman, Haltiwanger, and Scarpetta (2013), Decker, Haltiwanger, Ron S Jarmin, and Miranda (2016), and Decker, Haltiwanger, Ron S. Jarmin, and Miranda (2020), due to the comprehensive coverage of revenue and employment data, and strong correlation with value-added per worker measures across industries.

[10]The source of the official data is the series 'Number of People in Employment (aged 16 and over, seasonally adjusted):000s' (series ID: MGRZ) https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/timeseries/mgrz/lms which is from the Labour Market Statistics (LMS) time series.

[11]In the appendix we plot average labour productivity across firms. That is, we calculate firm-level productivity and then take the simple mean across all firms. The trend is similar.

Both samples capture a peak in labour productivity in 2010 followed by a stark decline and subsequently low growth. The sub-sample captures the pre-crisis period better than the full sample which shows sharply declining productivity from 1997-2003.[12] In the sub-sample, there is steady growth in labour productivity over the 2000s, which declines in tandem with the Great Recession, and follows a slower growth path after 2011.[13] In both samples, the significant increase in labour productivity in 2010 is due to the increase in sales which was accompanied by a slight fall in the number of employees.

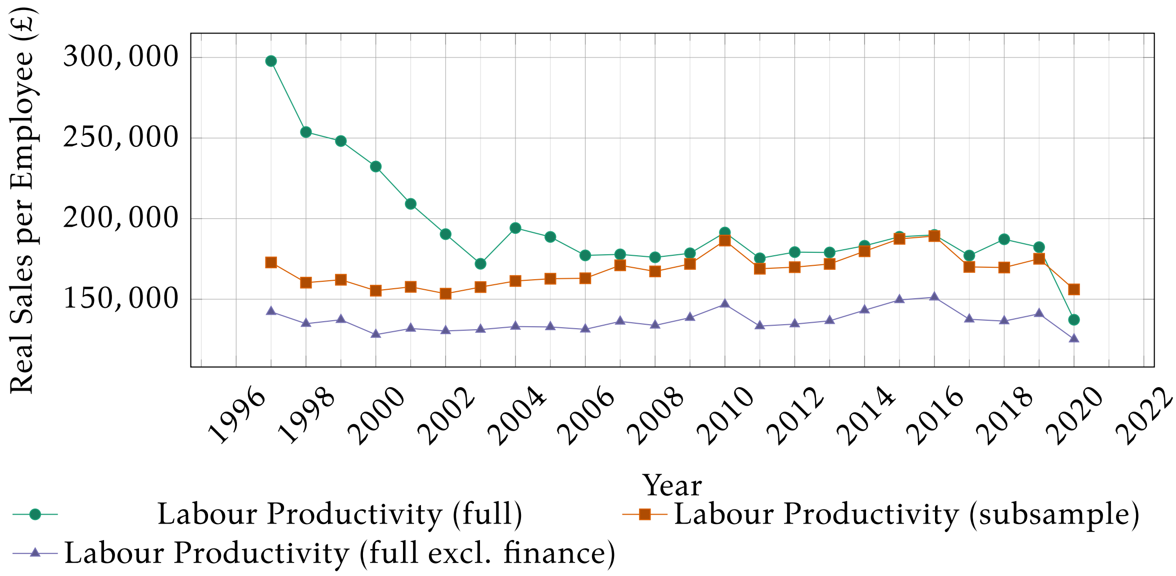


Figure 4: Aggregate Labour Productivity (BSD, 1997-2020)
Source: Authors' calculation based on BSD 1997-2020

3.2 Product Market Concentration

Concentration ratios (CRN) represent the sales share of the biggest N firms in a market. The market can be the whole economy or granular sectors. In this section we report concentration ratios for a broad market definition and more granular 5-digiti sectors. We also report different definitions in terms of the number of firms included in the numerator of the concentration ratio. Including fewer firms is more likely to capture a dominant group of firms that could engage in anti-competitive behaviour. For example, Cavalleri, Eliet, McAdam, Petroulakis, Soares, and Vansteenkiste (2019) and Bajgar, Berlingieri, Calligaris, Criscuolo, and Timmis (2023) use a variety of CR4, CR8 and CR20 measures, and studies such as Furman and Orszag (2018) use CR50,

[12]This early decline in labour productivity is because of the decline in aggregate sales over the same period (Figure 2), whereas employment over the same period (Figure 3) has little effect as it shows a consistent increasing trend. Further, the fall in sales, and thus labour productivity, comes from the finance sector from 1997-2003.

[13]Remember that 2011 in the BSD is capturing economic activity for 2009 and 2010.

whilst Aquilante, Chowla, Dacic, Haldane, Masolo, Schneider, Seneca, and Tatomir (2019) report CR100.

Figures 5 and 6 report concentration ratios for the aggregate economy, treating the whole economy as the market, similar to evidence in Aquilante, Chowla, Dacic, Haldane, Masolo, Schneider, Seneca, and Tatomir (2019). In this context CR5 represents the sales share of the largest five enterprise units in the dataset. Figure 5 shows that aggregate measures of concentration are stable to decreasing in the UK over the period 1997-2020. There is an increase in concentration from 2009-2010 which typically occurs when firms exit during recession. The CR5 measure fluctuates around the 5% level from 2008 onwards. The implication is that one twentieth of all sales in the UK go through the largest five firms.[14] Figure 6 shows that in the sub-sample concentration ratios increase up to 2016 but declined rapidly afterwards. CR5 more than doubled (4% to 10%) from 1998-2016.[15] The rapid decline in 2020 suggest that the size of the top 5 firms in the economy fell more than proportionally than total market size, likely due to the influence of COVID-19. As a comparison, Cavalleri, Eliet, McAdam, Petroulakis, Soares, and Vansteenkiste (2019, Figures 1 & 2) report similar levels of concentration for several European countries. They report CR4 and use various aggregation techniques.

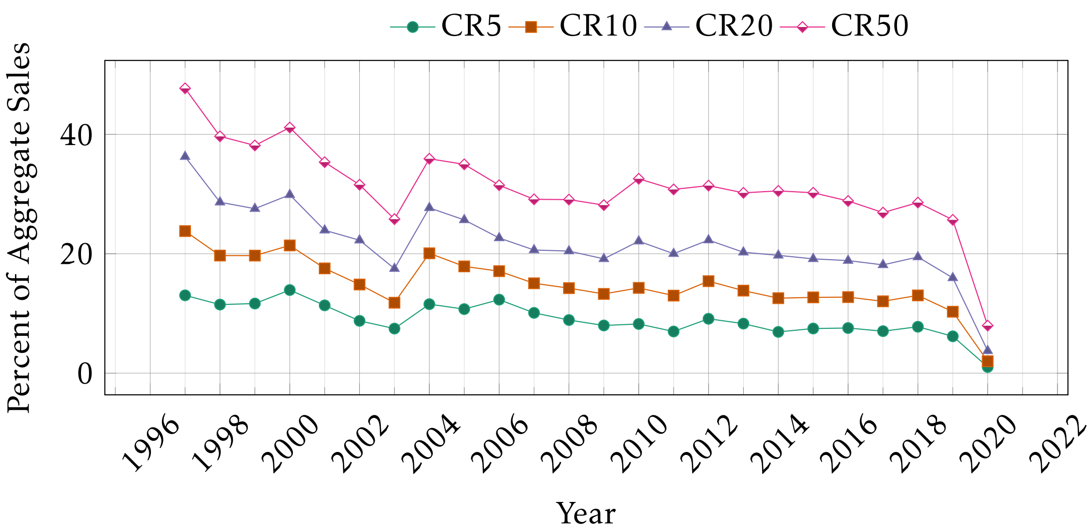


Figure 5: Aggregate Economy Concentration Ratios, Full Sample

Source: Authors' calculation based on BSD 1997-2020

[14]In the appendix we plot average CR5 across the main sectors. When we weight each sector by its revenue share, we find similar results.

[15]The spike in concentration in 1997 might be due to under-reporting of smaller firms. If the aggregate economy is missing smaller firms this reduces total sales and increases the relative size of large firms.

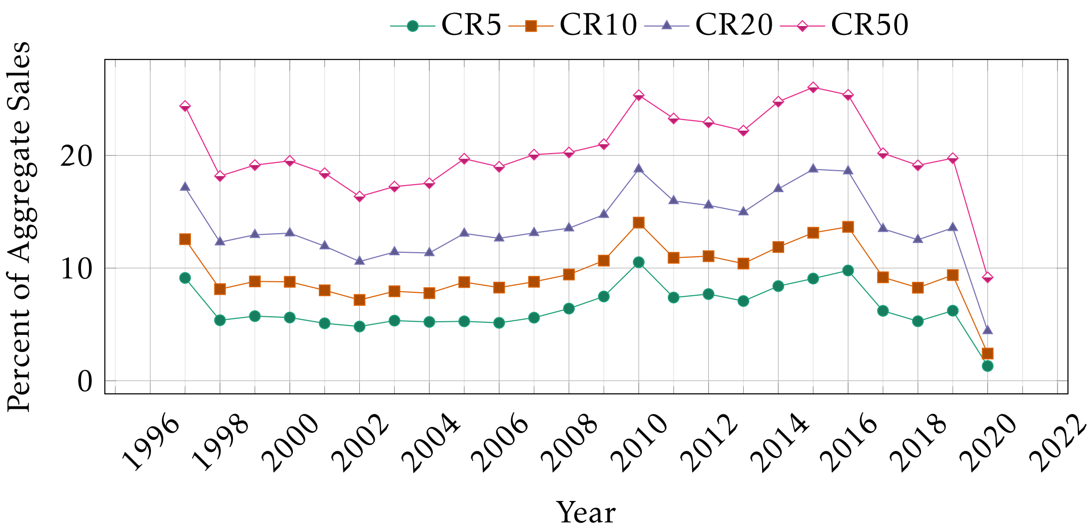


Figure 6: Aggregate Economy Concentration Ratios, Sub-sample
Source: Authors' calculation based on BSD 1997-2020

3.2.1 Concentration Aggregated from 5-digit SIC Sectors

Figure 7 plots concentration ratios aggregated from the most granular industry definition (5-digit SIC), similar to the preferred definition in Autor, Dorn, Katz, Patterson, and Van Reenen (2020) who use 4-digit industries. There are approximately 600 5-digit SIC industries each year, depending on whether an industry observes activity or passes disclosure rules within the secure lab. We present the median and mean CRN for the concentration level of these industries each year.

We observe an increasing trend across all measures over the period, and this is particularly strong for the median measure. In terms of levels, the mean always exceeds the median, which suggests that there is a tail of high-concentration industries. For CR5, the median 5-digit industry has a market share of 15 to 20% among the top 5 firms, and in the average industry the top 5 firms have 20 to 25% of market share. Both measures imply nearly one-fifth of the market held by a small group of firms. The remaining measures CR10, CR20 and CR50 each increase the amount of market share held as a wider number of firms is taken into consideration, and the increasing trends become starker. For CR10, the amount of market share for the median 5-digit industry increases from 20% to 25% over the period considered, and CR20 shows similarly sharp increases.

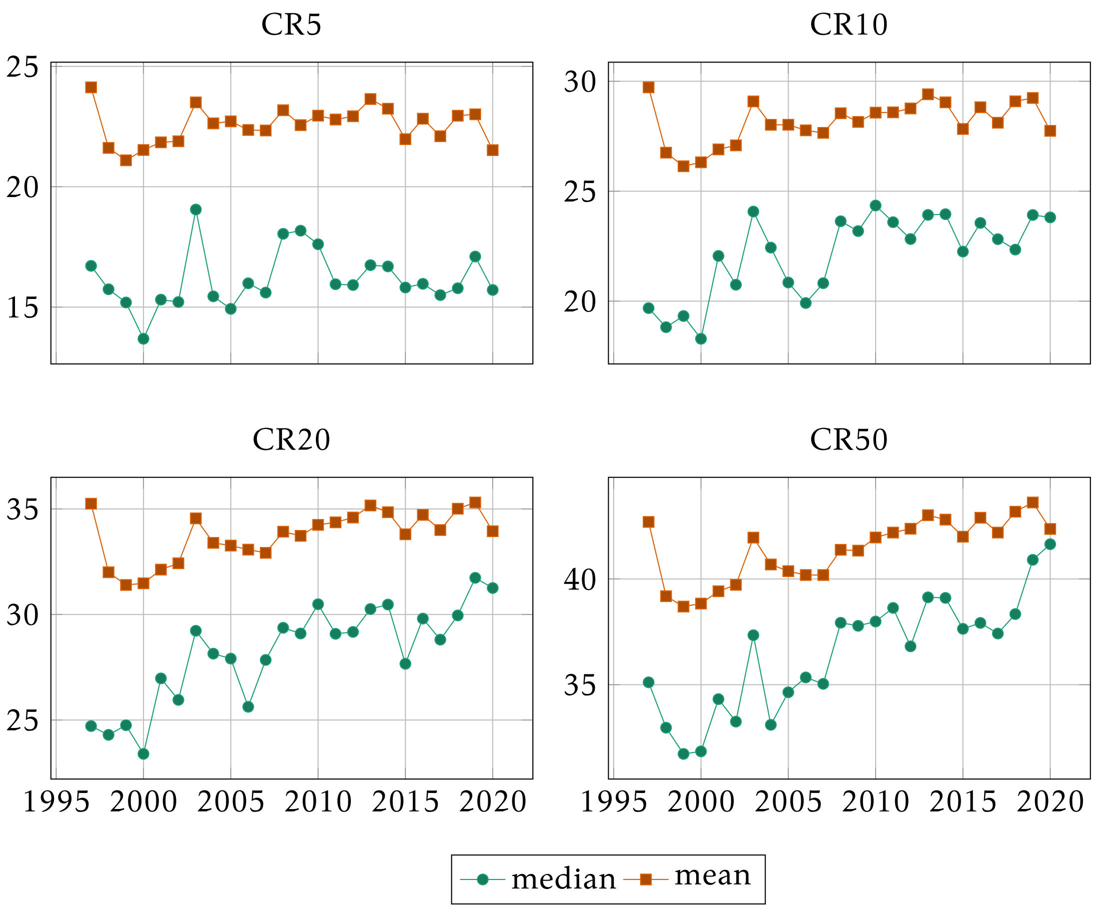


Figure 7: Aggregated Concentration Ratios for Full Sample from 5-digit SIC
Note: We calculate concentration at the the 5-digit SIC which is approximately 600 industries per year. Then, we take the mean and median across these industries each year. Each panel presents the top 5, 10, 20, or 50 firms.
Source: Authors' calculation based on BSD 1997-2020

3.2.2 Concentration Distribution across 5-digit SIC Sectors

Figure 8 shows the changing distribution of CR5 at the 5-digit level. We observe the density of the distribution shifting right over time. This reflects more 5-digit industries with higher levels of concentration. This is consistent with the growth in the mean and median concentration that we document in Figure 7.

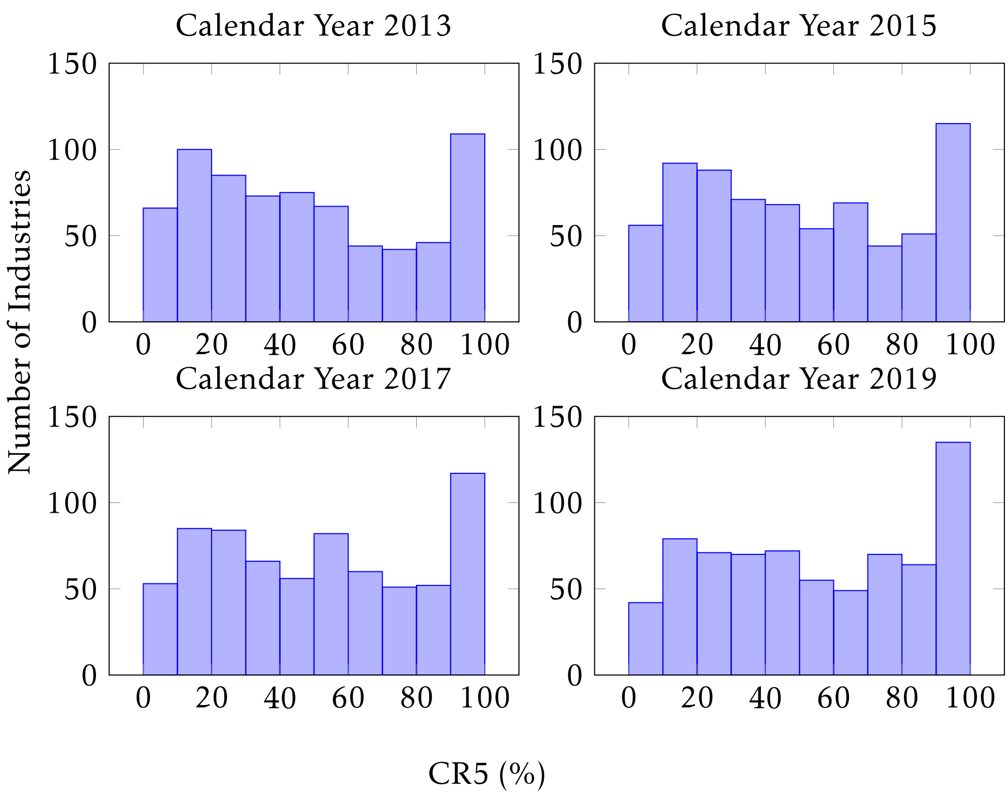


Figure 8: CR5 Distribution at 5 digit, by year

Each histogram shows the distribution of 5-digit industries across ten bins of CR5 concentration (0-10%, 10-20% ... 90-100%). In all years there are over 100 5-digit industries with a CR5 of 90-100%.

In figures 9 and 10 we classify sections of the distribution and plot their evolution over time. Figure 9 shows changes in the percentage of industries that have high and low CR5 from 1997-2020. We classify two parts of the concentration distribution CR5: 0-20% represents low concentration industries and CR5: 80-100% represents high concentration industries. Over the period, low concentration industries fall from over 20-25% of 5-digit industries to under 20% of 5-digit industries.

In 2007 there is a discontinuity leading to a spike in high concentration industries. In 2007 SIC industrial classifications were updated from SIC 2003 to SIC 2007. We matched 5-digit SIC 2003 classifications with their nearest SIC 2007 classification using ONS methodology. This re-classification is less important for higher degrees of aggregation such as at the two-digit where most industries will remain in the same two-digit classification after re-classification, whereas at the 5-digit level more industries may cease to exist and be reallocated to other industries.

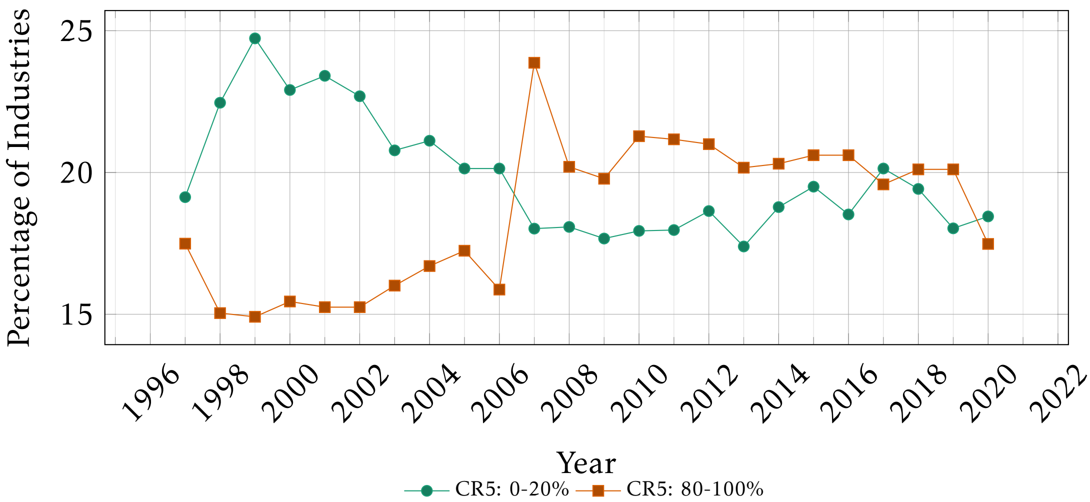


Figure 9: Percentage of high and low CR5 industries, 1997-2020, 5-digit

In antitrust cases, industries are often categorized based on the Hirfindal-Hirschman Index (HHI), which is calculated as the total of squared market shares. The HHI is then used to classify industries as follows: HHI values between 0 and 1000 indicate low concentration industries, HHI values between 1000 and 1800 indicate medium concentration industries, and HHI values between 1800 and 10000 indicate high concentration industries (Whish and Bailey 2021, p.43).[16] According to this classification, Figure 10 shows that between 25-30% of industries are high concentration and this has increased over the sample period. The level of moderately concentrated industries is stable at 15%. The remaining 50-60% of industries are low concentration.

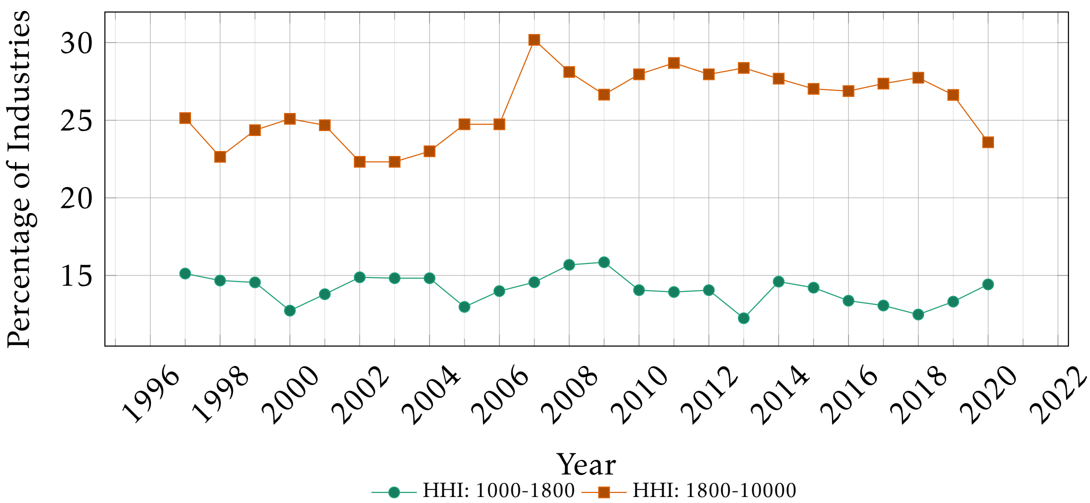


Figure 10: Percentage of high and medium HHI industries, 1997-2020, 5-digit

In the appendix we provide further analysis of concentration at the BSD sector level.

[16] Paragraph 16 of the European Commission's *Guidelines on the assessment of horizontal mergers*.

3.3 Business Dynamism: Entry, Exit and Net Entry

Entry statistics are one indicator of *business dynamism*. Net entry is an alternative indicator of competition, also called *business churn*. Unlike our other variables, we do not have data for 1997 for entry and exit. We determine entry as the first year that a firm is recorded as being active and records employees and turnover as non-zero or not missing. Exit is the first year the firm is recorded as being inactive having been active the previous year or the first year a firm records turnover and employees as zero.

Figures 11 and 12 show that aggregate entry statistics have a flat trend between 1998-2020, suggesting stable business dynamism.[17] The fluctuations we observe are consistent with well-known characteristics of the business cycle (Tian 2018). Entry and exit typically co-move, except in recessionary periods when entry declines and exit increases. Additionally, entry is more volatile than exit. Between 2008-2011, there was a fall in the number of firms entering and an increase in the number of firms exiting, so net entry became negative. This supports a common mechanism for countercyclical markups in business cycle theory. For example, in business cycle models with Cournot oligopolistic competition such as Savagar (2021), as the stock of firms declines during recession, an individual firm's ability to affect industry output rises, price elasticity of demand becomes more inelastic, which raises price setting ability, so price markups rise. This apparent decline in competition during recessions coincides with a higher concentration, as shown in our earlier figures.

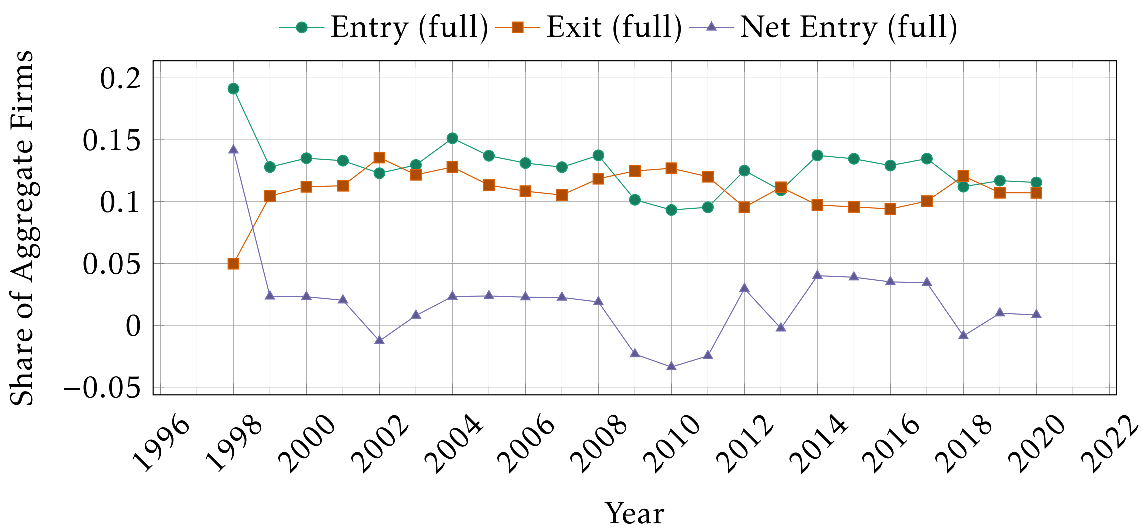


Figure 11: Aggregate Entry and Exit Rate for Full Sample (BSD, 1998-2020)
Source: Authors' calculation based on BSD 1998-2020

Figure 12 shows that the entry and exit rates for the sub-sample are similar.

[17]These measures are for entry and exit in the whole economy, not averaging entry and exit statistics across different industries.

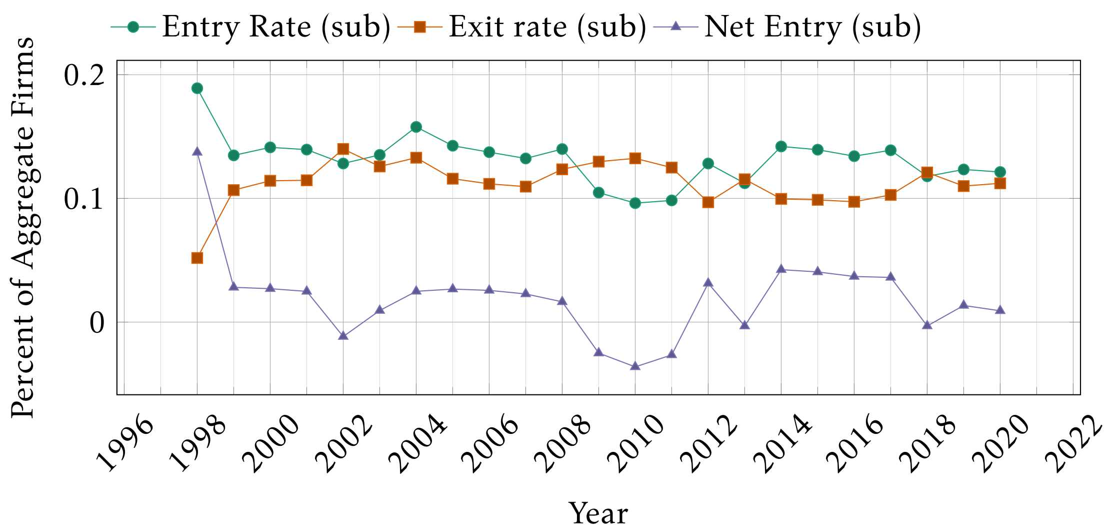


Figure 12: Aggregate Entry and Exit Rate for Sub-sample (BSD, 1998-2020)
Source: Authors' calculation based on BSD 1998-2020

In the appendix we provide further analysis of business dynamism at the BSD sector level.

4 Concentration and Productivity

We use reduced-form regression analyses to study the relationship between concentration and productivity. In Section 4.1, we provide firm-level regressions on this relationship, and in Section 4.2, we provide industry-level regressions on industry productivity, and its decompositions. Our aim is to provide controlled correlations on the relationship between productivity and concentration. This approach follows other literature studying concentration in the aggregate (Covarrubias, Gutiérrez, and Philippon 2020; Bighelli, Di Mauro, Melitz, and Mertens 2023). We lag concentration and net entry measures in order to mitigate reverse-causality, which is the most obvious form of endogeneity. However, this does not overcome the potential for omitted variables, which are correlated with contemporaneous concentration and future productivity, to create endogeneity weakening evidence of causality.

4.1 Empirical Methodology

The dependent variable is labour productivity and the main independent variable is concentration. We also include net entry as an alternative indicator of competition. Our regression specification is:

$$\begin{aligned}\text{Productivity}_{ijt} = \alpha_j + \alpha_t + \beta_1 \text{Concentration}_{jt-1} + \beta_2 \text{Net Entry}_{jt-1} \\ + \beta_3 (\text{Concentration}_{jt-1} \times \text{Net Entry}_{jt-1}) + \gamma^\top \mathbf{x}_{ijt} + \varepsilon_{ijt} \qquad (1)\end{aligned}$$

The subscript $\jmath$ indicates a sector and $\imath$ indicates a firm. We consider two-digit SIC sectors. There are 85 two-digit sectors. Concentration and net entry are sector-level variables, whereas $\mathbf{x}$ is a vector of firm-level control variables to account for other factors that may influence productivity. The firm-level controls are market share, firm size and firm age. The dependent variable, productivity, is also at the firm level. We use industry ($\alpha_\jmath$) and year (α_t) fixed effects.

We use a 2-digit definition in our regression analyses for several reasons. First, it appears to capture distinct industry functions well (e.g. veterinary or postal activities).[18] At a higher level there is little distinction between industry activities, whereas if we go more granular it becomes more likely that some firms will operate across multiple industries, but their activity will only be recorded in their primary industry. Second, at the 2-digit level few firms change SIC definitions from pre-2007 SIC classification to post-2007 SIC classifications. Lastly, we are consistent with comparable studies such as Furman and Orszag (2018) for the US and Cavalleri, Eliet, McAdam, Petroulakis, Soares, and Vansteenkiste (2019), Bighelli, Di Mauro, Melitz, and Mertens (2023), and Bajgar, Berlingieri, Calligaris, Criscuolo, and Timmis (2023) for Europe.

Table 2 reports the measure that we use for each variable in equation (1). Our main measure of concentration is CR5. We use CR5, rather than CR10, CR20, CR50 or HHI, because a lower number of firms better captures weak competition. That is, it is easier to sustain anti-competitive practice with high market share among fewer firms.[19]

Variable	Measure	Unit
Productivity	$\log\left(\frac{\text{Real Sales}_{\imath\jmath t}}{\text{Employees}_{\imath\jmath t}}\right)$	Log
Concentration	$\text{CR5}_{\jmath t-1} = \frac{\sum_{\imath=1}^{5} \text{Real Sales}_{\imath\jmath t-1}}{\text{Real Sales}_{\jmath t-1}} \times 100$	% of market
Net Entry	$\frac{\text{Births}_{\jmath t-1} - \text{Deaths}_{\jmath t-1}}{\text{Firms}_{\jmath t-1}} \times 100$	% of active firms
Market Share	$\frac{\text{Real Sales}_{\imath\jmath t}}{\text{Real Sales}_{\jmath t}} \times 100$	% of market
Age	$\text{Birth Date}_{\imath\jmath} - \text{Death Date}_{\imath\jmath}$	Years
Size	$\log\left(\text{Employees}_{\imath\jmath t}\right)$	Log

Table 2: Variables and Corresponding Measure in the Data

[18] We provide a list of 2-digit industries in the appendix.

[19] In the appendix, we perform a sensitivity analysis of our results to alternative measures of concentration.

Estimation Strategy

We use first-difference estimation to estimate regression (1).[20] First-difference estimation accounts for unobservable firm-specific effects. We use first-differences instead of fixed effects due to computational restrictions in the secure lab. We account for within-cluster error correlation by clustering standard errors at the sector level. A disadvantage of using both first-difference estimation and lagged independent variables is that we lose at least two observations per firm in the regression analysis. This accounts for the difference in observations recorded in the summary statistics table and regression tables.

Productivity and concenttration are subject to reverse causality. Increased concentration causes productivity to decrease if it causes anti-competitive behaviour. However, increased productivity causes increased concentration if productive firms increase their market share. The former relationship implies a negative causal relationship from increased concentration to decreased productivity, whereas the latter relationship implies a positive causal relationship from increased productivity to increased market power. To mitigate the effect of this form of endogeneity we use a lagged measure of concentration.

We include 'Net Entry' as an alternative indicator of competition in an industry. Net entry should keep incumbent firms operating efficiently. Therefore, we would expect it to have a positive coefficient. Similarly to concentration, net entry also suffers from the reverse-causality problem. That is, net entry might increase productivity because it keeps incumbent firms operating efficiently. However, high-productivity industries may encourage more entry.[21] Both directions of causality imply a positive relationship. Similarly to concentration, we mitigate the reverse-causality between labour productivity and net entry by using lagged measures of net entry.

We include a 'Concentration × Net Entry' interaction term. We expect that when net entry is high then an industry does not face weak competition. Therefore the effect of concentration on productivity will be less negative when net entry is high and could be positive if the presence of concentration occurs when there is high net entry as concentration is increasing due to high-productivity firms out-competing entrants. Conversely, if net entry is low this implies little competition so the effect of concentration should be more negative.[22] In other words, we use net entry as an indicator of

[20]In the appendix we provide pooled OLS estimates.

[21]This could be because of a new technology. If technological innovations are industry-specific, this is controlled for in by the fixed effects. The problem would be more prominent with an aggregate industry-wide technology improvement.

[22]This interpretation focuses on the effect of concentration, given net entry behaviour. However, there are two interpretations of an interaction term: the relationship between productivity and net entry depends on concentration or the relationship between productivity and concentration depends on net entry. The alternative interpretation is that the effect of net entry will vary depending on the level of concentration. When concentration is high, net entry should have a stronger positive effect on

whether concentration is taking place for 'good' (productive frontier) or 'bad' (anti-competitive behaviour) reasons.

Regression Variable Summary Statistics

Variables	**Obs.**	**Mean**	**Median**
Productivity	40,610,710	75,360	77,650
CR5 (%)	41,313,405	16.95	14.05
Net Entry (%)	41,313,405	1.78	1.14
Market Share (%)	41,313,405	0.00	0.00
Firm Age	41,313,405	10.39	7
Employees	41,313,405	12.72	2

Table 3: Summary Statistics

Table 3 presents summary statistics of variables used in the regression. CR5, net entry and market share are percent of two-digit industry. Productivity, age and employees are firm level. The average labour productivity is £75,000, which implies a worker generates £75,000 sales for a firm in a year. This figure appears high and might be inflated because the dataset excludes firms that do not pay VAT (revenue below £85,000 in 2018) and have no employees.[23] For context, average real income over the period 2000-2018 is roughly £28,000 (2020 prices). This implies a 37% average share of wages in sales. CR5 shows average concentration ratios across two-digit industries. On average, the top 5 firms account for 17% of turnover. Average net entry at a two-digit industry is 1.8% of total firms in the industry, which implies a net increase in firms each year. Market share shows that on average firms are very small. An average firm accounts for 0.004% of sales in its two-digit industry. Median firm age is 7 years while the median number of employees (firm size) is 2. The difference between mean and median for firm size implies there is a large number of small firms (positive skew).

4.1.1 Firm-level Regression Results

The results of our regression are in Table 4[24].

productivity.

[23]Firms with sales below the VAT threshold are included if they have a PAYE employee.

[24]In the appendix we provide results for pooled OLS regressions. These do not control for firm FE. The results have a similar economic significance but are less statistically significant. The interpretation of the pooled-OLS and FD regression coefficients is identical

Table 4: The Effect of CR5 on Log Labour Productivity

	All	All	All	All	All	Serv.	Prod.
Δ CR5$_{t-1}$	-0.082***	-0.069**	-0.068**	-0.069**	-0.067**	-0.073*	0.005
	(0.002)	(0.027)	(0.027)	(0.028)	(0.027)	(0.037)	(0.035)
Δ Net Entry$_{t-1}$			0.053	0.033	0.042	0.092	-0.079
			(0.058)	(0.073)	(0.068)	(0.087)	(0.071)
Δ(CR5$_{t-1}$ $\times$ Net Entry$_{t-1}$)				0.128	0.097	0.024	0.029
				(0.204)	(0.182)	(0.249)	(0.158)
Δ Market Share$_t$					0.192***	0.298***	0.110***
					(0.032)	(0.044)	(0.028)
ΔFirm Age$_t$					0.007**	-0.001	0.005
					(0.003)	(0.004)	(0.003)
ΔFirm Age$_t^2$					-0.018**	-0.001	-0.012
					(0.008)	(0.011)	(0.007)
Δ ln (Firm Size$_t$)					-0.564***	-0.510***	-0.621***
					(0.038)	(0.022)	(0.056)
Δ ln (Firm Size$_t$)^2					-0.021	-0.529	0.853
					(0.538)	(0.777)	(0.992)
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
N	31,914,138	31,914,138	31,914,138	31,914,138	31,914,138	16,190,002	3,392,781
R^2	0.000	0.003	0.003	0.003	0.106	0.084	0.146
Clusters		85	85	85	85	39	37

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Regressions are in first difference.

Source: BSD data, 2000-2020.
Dependent variable is log labour productivity.
All independent variables are in levels, except firm size which is in logs.
CR5, net entry and market share are measured at two-digit SIC industry level.
Coefficients and standard deviations for the quadratic terms are multiplied by 100 (*i.e.* variable unit is divided by 100.)
CR5 and Net Entry variable units are shares not percentages.
Standard errors are clustered at two-digit industry level. Sales are constant 2016 values.
Industry FE are at the two-digit sector level.

The results in Table 4 show a statistically significant negative relationship between concentration and labour productivity in columns 1 to 3. Higher concentration levels are associated with lower labour productivity levels. The results in the full specification have the following economic interpretations: A 1 percentage point increase in CR5 decreases productivity by 0.07%. A 1 percentage point increase in net entry increases productivity by 0.04%. A 1 percentage point increase in market share increases productivity by 19.2%. An increase in firm age by 1 year increases productivity by 0.01%, but at a decreasing rate. An increase in firm size of 1% (employees) decreases productivity by 0.56%.

In terms of economic significance, our relationship suggests a 10 percentage point increase in CR5 corresponds to a 0.7% fall in productivity levels of the average firm. Since average labour productivity is approximately £80,000 sales per worker per year, then a five percentage point increase in CR5 corresponds to a £280 decrease in sales per worker per year.

Our firm-level results are supportive of hypotheses that concentration relates negatively to productivity, potentially due to lower competition or antitrust abuses, and are consistent with papers such as Covarrubias, Gutiérrez, and Philippon (2020) and De Loecker, Eeckhout, and Unger (2020). It is important to note that our firm-level results are for the average firm, conditional on the factors we control for, but do not consider the distributional implications of concentration reallocating resources across producers.

Our findings at the firm-level differ from other results in the literature, which find a positive association, more in favour of the 'winner takes all' hypothesis that productive firms have taken market share due to superior technologies. For example, Cavalleri, Eliet, McAdam, Petroulakis, Soares, and Vansteenkiste (2019) find a positive relationship between TFP growth and concentration in the high-tech sectors of Germany, France, Italy and Spain. And, Bighelli, Di Mauro, Melitz, and Mertens (2023) find a positive assocation between concentration and allocative efficiency in a study of 15 European countries.[25]

Specific Sector Regressions: Services & Production

The final two columns of Table 4 report the relationship between concentration and productivity by sector. We analyse the two largest BSD sectors *Production* and *Services*.[26] The results show opposite effects of concentration on labour productivity across the two sectors. In Services there is a negative and significant effect of concentration on productivity, whereas there is a positive but not significant effect in the production sector. Notably, the sample size falls sharply in these subsets, which will limit the precision of results.

In services the result appears the same as for our full specification. The relationship and magnitude are similar, indicating that the average firm has weaker productivity when there is greater concentration. Services reflecting the aggregate outcomes is common in UK analyses due to the size of the sector. We observe that this sub-sample represents over half of firms. The sector is also characterised by a large number of smaller (low employee) producers, which can help to rationalise why greater concentration may negatively relate to productivity of the average firm. For example, if higher concentration leads larger services firms to acquire the more productive workers, then the average firm will have lower productivity. This is why industry-wide measures of

[25]Belgium, Czech R., Finland, France, Germany, Italy, Lithuania, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Sweden, Switzerland.

[26]These are ONS-specific aggregations of two-digit sectors called 'BSD sectors'. They are used in the BSD and other business datasets because they correspond to sampling frames used in business surveys. *Services* is called 'Other Services' by the ONS definition, see Appendix. These two sectors are the sectors comprising most two-digit industries. In addition, they are the largest sectors, both in aggregate turnover and employment terms. Services accounts for 50% of total turnover and Production 15%.

productivity are important, which motivates our next section, because at the firm-level we overlook distributional effects that could counteract the average firms' experience.

Although the result for production is insigificant. There are plausible economic mechanisms which could explain the lack of a negative relationship and indicators of a positive relationship. Concentration may be positively related to productivity in the production sector due economies of scale. If production requires a large overhead cost, then distribution of economic activity across small firms will under-utilise the overhead and damage productivity. Hence greater concentration, increases utilisation, which increases productivity. In the production sector, net entry has a negative relationship with productivity. This also supports a theory that entry leads to replication of fixed overhead costs that are under-utilized and weaken productivity. In the production sector the interaction term indicates a negative relationship between concentration and productivity that is worse when net entry is high. Hence greater concentration increases the negative effect of entry.

4.2 Productivity Decomposition

Our firm-level regression results show that a higher concentration is associated with lower productivity for the mean firm with the characteristics for which we control. However, this firm-level approach does not capture the distribution of employment across firms. To understand the effect on average-worker productivity, we study the relationship between concentration and the weighted-average productivity in an industry.

4.2.1 Decomposition Methodology

In general, we can express a weighted average as an additive decomposition of the unweighted average and a dispersion term:[27]

$$x_j \equiv \sum_{i=1}^{N} \omega_i x_i = \bar{x}_N + \sum_{i=1}^{N} (x_i - \bar{x}_N)(\omega_i - \bar{\omega}_N). \tag{2}$$

This provides a cross-sectional decomposition at a point in time, and x_j represent industry j weight-average productivity. To improve clarity, we omit the time t and in-

[27]This general decomposition applied to productivity is referred to as a static OP decomposition (Olley and Pakes 1996). Bighelli, Di Mauro, Melitz, and Mertens (2023) use the decomposition to understand business dynamism and concentration. Their analysis focuses on the relationship between concentration and allocative efficiency, which is one component of the decomposition. Whereas we analyse the relationship with all components of the decomposition, and we show that the unweighted results are consistent with our firm-level results in the previous section. Furthermore, this highlights both the positive and negative potential of concentration for productivity. We are able to do this extension because we have firm-level data, rather than the 'micro-aggregated' data in their cross-country panel.

dustry j subscripts in equation (2) as the decomposition is applied to an industry in a given time period. We apply the decomposition to approximately 85 two-digit SIC industries each year 2005 to 2020. There are N firms in an industry and each firm is indexed by $i \in 1 \dots N$. The variable $x_i = sales_i/L_i$ is a firm's labour revenue productivity, and $\bar{x}_N = \frac{1}{N}\sum_{i=1}^{N} x_i$ is the unweighted-average of labour productivity. The weight $\omega_i \equiv L_i/\sum_{i=1}^{N} L_i$ is firm's labour share in total industry labour.[28] And, since shares sum to one, then the unweighted average share is $\bar{\omega}_N = \frac{1}{N}\sum_{i=1}^{N} \omega_i = \frac{1}{N}$.

The weighted-average on the left-hand side is the average-worker productivity, and the unweighted-average productivity is the average-firm productivity. The cross-product of deviations term measures the covariation of firm productivity and the employment share. A common economic interpretation is allocative efficiency (Olley and Pakes 1996). It captures the extent to which workers are allocated to high-productivity or low-productivity firms. It is positive when above-average productivity firms have above-average employment or vice-versa. It is negative when below-average productivity firms have above-average employment, or vice-versa.

In our case, we can write

$$\sum_{i=1}^{N}(x_i - \bar{x}_N)(\omega_i - \bar{\omega}_N) = \frac{sales_j}{L_j} - \frac{1}{N_j}\sum_{i=1}^{N_j}\left(\frac{sales_i}{L_i}\right)$$

In other words, allocative efficiency is the difference between average-worker productivity and average-firm productivity:

$$\text{Allocative Efficiency} = \text{Avg. Worker Productivity} - \text{Avg. Firm Productivity.}$$

Our application to labour (revenue) productivity, weighted by the labour share is the same as Bartelsman, Haltiwanger, and Scarpetta (2013) for the US, and Bighelli, Di Mauro, Melitz, and Mertens (2023) for Europe. An advantage of our work with UK data is that we have representation across firm size and legal forms. We are not restricted to larger firms or micro-aggregated data.

For the descriptive figures that follow, we drop industries 06 (oil), 07 (mining metal), 09 (quarrying), 12 (tobacco), 19 (petrol manufacturing), 35 (electricity), 39 (waste), 46 (wholesale), 64 (finance), 65 (insurance), 66 (other finance), and we restrict the time period to 2005-2020. This removes outliers which improves plot scaling. In our regression analyses, we do not omit industries, as we control for fixed differences

[28] Since our measure of productivity is labour revenue productivity ($x_i = sales_i/L_i$), then the weighted-average is equivalent to total sales divided by total employment:

$$\sum_{i=1}^{N}\omega_i x_i = \sum_{i=1}^{N}\left(\frac{sales_i}{L_i} \times \frac{L_i}{\sum_{i=1}^{N} L_i}\right) = \frac{\sum_{i=1}^{N} sales_i}{\sum_{i=1}^{N} L_i} = \frac{sales_j}{L_j}.$$

between industries. Whether we include or exclude this list of industries does not affect the signs or significance of the results. In fact, the significance increases if these industries are removed.

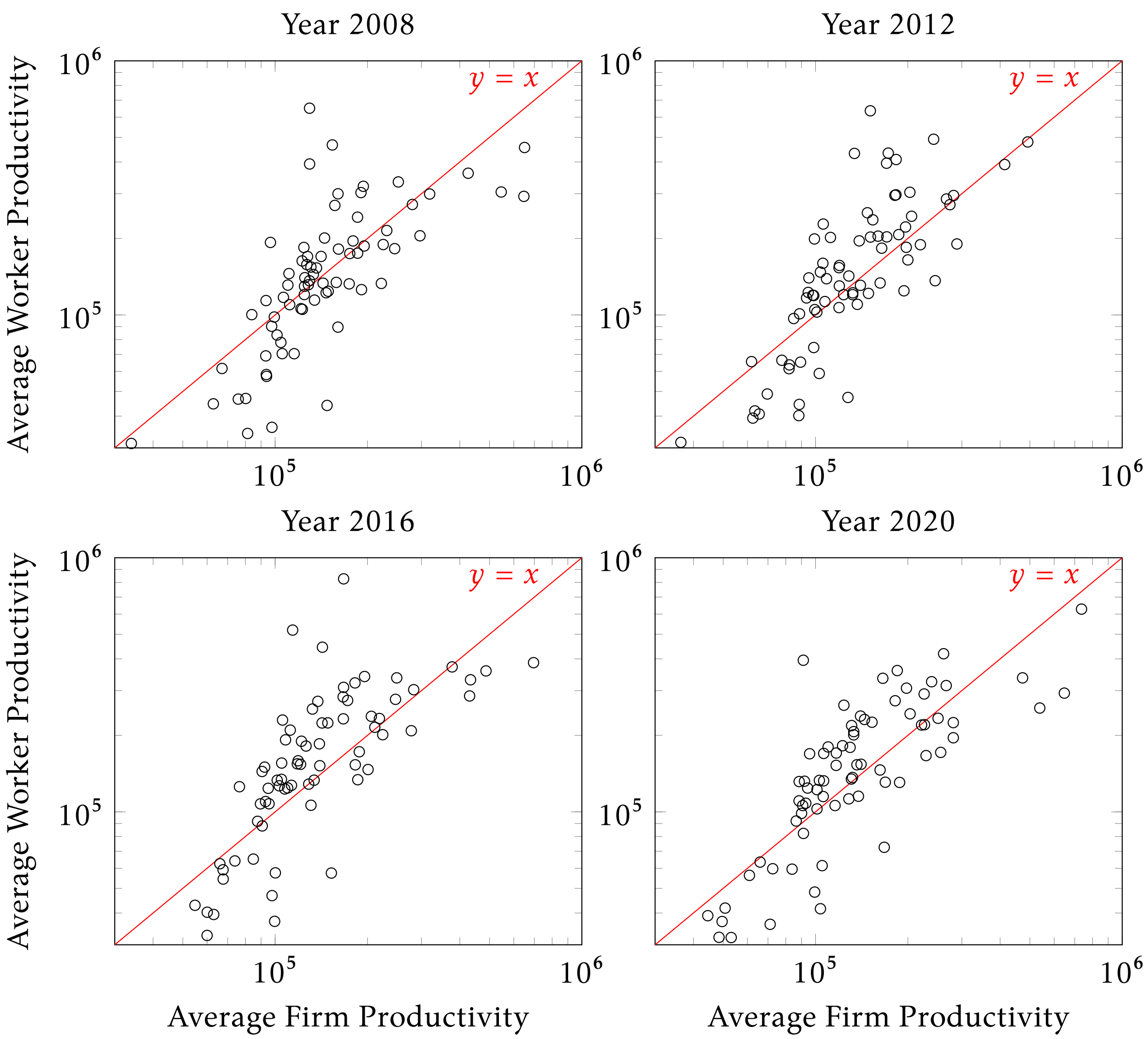
4.2.2 Decomposition Descriptive Statistics

The plots in Figure 13 show average-worker productivity against average-firm productivity for a selection of four years. In each plot a scatter point represents a two-digit industry. The $y = x$ line captures the case when the two measures coincide, which implies the allocation term is zero. Points above the line are industries where there is positive allocative efficiency, so above (below) average productivity firms have above (below) average employment, whereas points below the line represent industries where allocative efficiency is negative, so above (below) average size firms have below (above) average productivity. Figure 13 illustrates that most variation in our firm-year panel data occurs between industries, while variation over time is limited.

In Figure 14 we illustrate the cross-sectional variation in allocative efficiency across industries in 2015, which we take as a representative year. Industries 01-49 mostly have positive allocative efficiency, whilst industries 50-96 mostly have negative allocative efficiency. This division roughly represents production industries in the first half, such as agriculture, mining, manufacturing and utilities. The second half is services such as real estate, professional & scientific, education, health and arts. In the Appendix we present this break-down in terms of average-firm and average-worker productivity, rather than the allocative efficiency, which is the difference between the two. We also include a key to identify the industries from their two-digit code.

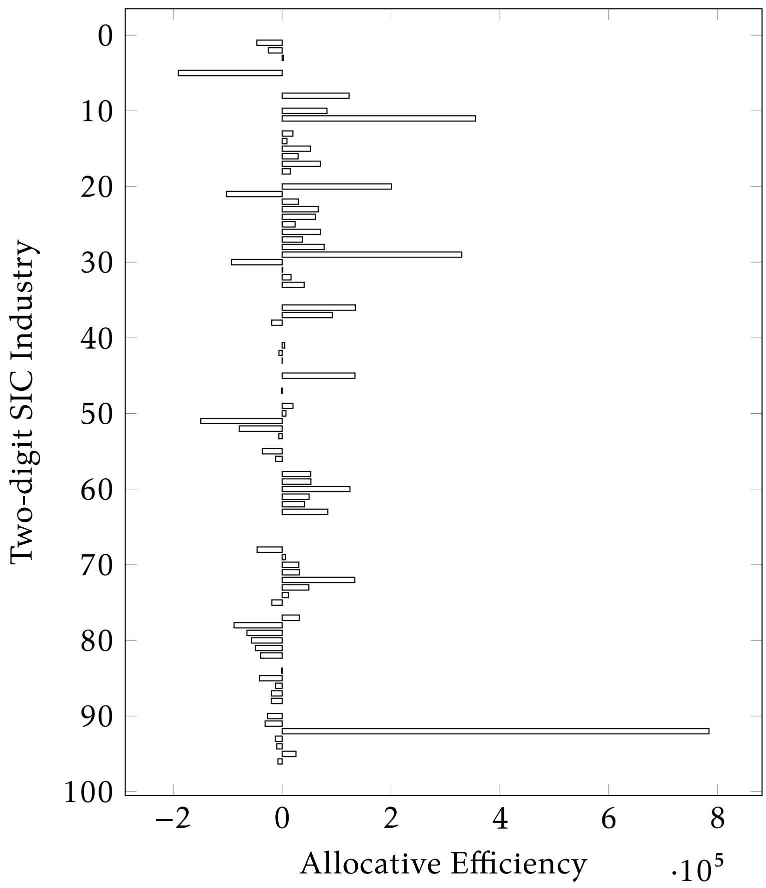
In Figure 15 we present the time-series variation in the data through the mean and quantiles across industries each year. Allocative efficiency broadly increased over the sample period, but has declined since 2016. On average allocative efficiency is positive, implying that average-labour productivity exceeds average-firm productivity.

Figure 13: Average Firm Productivity versus Average Employee Productivity



Note: The axes use logarithmic scaling. The domain is 30,000 to 1,000,000 for both axes in each plot. Each point is a two-digit sector. The $y = x$ line demarcates between positive and negative allocative efficiency. Above the line there is positive allocative efficiency, which implies that average worker productivity exceeds average firm productivity. Vice-versa below the line.

Figure 14: Allocative Efficiency by Industry in 2015



Note: In Appendix F.2, we provide a key for the SIC codes.

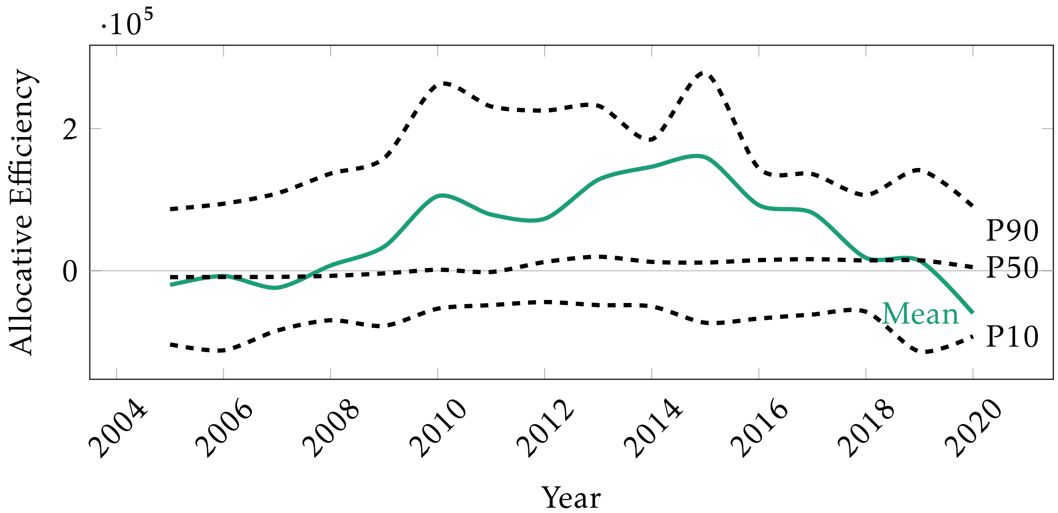


Figure 15: Allocative Efficiency over time, percentiles over industries

From our descriptive analysis of allocative efficiency we conclude that the bulk of industries have positive allocative efficiency. The more productive firms employ more labour. Allocative efficiency is more common in productivion industries, whereas in services below average productivity firms tend to have above average employment. Over time allocative efficiency has improved, but it has flattened or declined since 2016.

4.2.3 Decomposition Regression

We estimate the impact of lagged concentration (CR5) on each component of the decomposition, controlling for industry fixed effects (α_j) and time fixed effects (α_t).

$$y_{j,t} = \alpha_j + \alpha_t + \beta_1 \text{Concentration}_{j,t-1} + v_{j,t} \tag{3}$$

where y is the outcome of interest, namely the logarithm of weighted productivity (average-worker productivity), the logarithm of unweighted productivity (average-firm productivity), and the dispersion term (allocative efficiency), not logged due to negatives. The subscript j indicates a two-digit sector, while t is time in years.

Table 5: The Effect of CR5 on Decomposition Components, OLS

	ln(Worker Prod.)	ln(Firm Prod.)	Alloc. Eff.
CR5_{t-1}	0.54^{***}	-0.10	$507,735^{**}$
	(0.08)	(0.10)	(189,981)
Industry FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
N × T	1,250	1,250	1,250
N	85	85	85

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Source: BSD 2005-2020.
CR5 is measured at two-digit SIC industry level.
Industry FE are at the two-digit sector level.

The economic implications of the findings presented in Table 5 can be summarized as follows. An increase of 1 percentage points in CR5 realtes to an increase in average-worker productivity of 0.5%. However, it results in a decrease in average firm productivity of 0.1%, although this decrease has limited statistical significance; it is still one standard deviation from zero. Furthermore, this increase in CR5 relates to better allocative efficiency, specifically an increase of £5,077.[29] The magnitude is ap-

[29]We are interested in a 1/100 of a unit increase which cancels-out when interpreting the logarithmic changes in percentages $100 \times \beta_1 \times \Delta CR5$ where the change in CR5 is 1 percentage point $\Delta CR5 = 0.01$. For allocative efficiency, which is in levels, we have a 1/100-unit change $\Delta CR5 = 0.01$ in the absolute value of the coefficient $\beta_1 \times \Delta CR5$.

proximately half of the average allocative efficiency over the full, pooled, sample. The results in Table 5 for the average-firm effect (-0.1%) are consistent with our firm-level findings in Table 4 (-0.07%), which additionally controlled for firm fixed effects.

We conclude that higher concentration is associated with lower productivity for the average firm. But it is associated with a greater weight of employees in higher productivity firms, so average worker productivity is higher in more concentrated industries. Therefore, concentration is associated with greater allocative efficiency supporting the 'winner take all' hypothesis of efficient workers acquiring market share (Van Reenen 2018; Autor, Dorn, Katz, Patterson, and Van Reenen 2020).

Our empirical results reflect the nuanced theoretical relationship between productivity and concentration. Based on existing work, a plausible interpretation of our results is that in more concentrated industries, a dominant group of firms may benefit from scale effects, which benefits the disproportionate share of workers that they employ, raising average-worker productivity. But despite higher average-worker productivity, average-firm productivity suffers, potentially because the average-firm faces anti-competitive behavior, or it cannot operate at a 'minimum efficient scale' to exploit scale economies. This interpretation is consistent with existing evidence that highly efficient 'superstar' firms have driven market concentration (Autor, Dorn, Katz, Patterson, and Van Reenen 2020), and their dominance is closely related to intangible investment (Crouzet and Eberly 2019; Bessen 2020; Bajgar, Criscuolo, and Timmis 2021), but regulation has also weakened (Gutiérrez and Philippon 2017; Grullon, Larkin, and Michaely 2019), which could harm the average firm.[30] Furthermore, Lashkari, Bauer, and Boussard (2024) link intangible investment with scale economies, and Kariel and Savagar (2023) show that scale economies have increased in the UK.

5 Conclusion

We document product market concentration and entry dynamics in the UK using an administrative data set of all firms from 1997-2020. The dataset include comprehensive coverage of firm size, legal form and sectors. We show that the market share of the largest firms in the economy is stable over the period. However, the largest firms' market share increases for most of the period when finance is excluded, and for narrow industry definitions (5-digit SIC) we find that average concentration is rising. We show that trends in entry and exit are relatively stable over the period, differing from a well-documented decline in the USA.

Using measures of market concentration at the two-digit industry level, we find ev-

[30]Aghion, Bergeaud, Boppart, Klenow, and Li (2023) and De Ridder (2024) propose theories linking intangible investment to growing market power. Corrado, Haskel, Jona-Lasinio, and Iommi (2022) review literature on intangible investment.

idence that product market concentration and firm productivity are negatively related for the average firm. However, when we consider weighted-average productivity at the industry level, we find this result is overturned. This reflects that weighted-average measures include distributional effects. Our results at the industry-level show that although concentration remains negatively related to the average-firm productivity, it is positively related to the average-worker productivity. This arises because in industries with more concentration, workers locate at more productive firms. In other words, concentration is positively associated with allocative efficiency, and we confirm this relationship directly.

Future work should investigate more sophisticated measures of market power and productivity. For example, Hwang, Savagar, and Kariel (2023) investigate the relationship between price markups and TFP in the UK. Nonethless, it is a useful measurement exercise to document concentration trends in the UK, their sensitivity to market definition, and the nuanced relationship with productivity, particularly given the current prevalence of concentration studies in policy, the media, and for many other countries.

References

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li (2023). "A theory of falling growth and rising rents". In: *Review of Economic Studies* 90.6, pp. 2675–2702.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt (2005). "Competition and innovation: An inverted-U relationship". In: *The Quarterly Journal of Economics* 120.2, pp. 701–728.

Aghion, Philippe and Peter Howitt (Mar. 1992). "A Model of Growth through Creative Destruction". In: *Econometrica* 60.2, pp. 323–351.

Akcigit, Ufuk and Sina T Ates (2021). "Ten facts on declining business dynamism and lessons from endogenous growth theory". In: *American Economic Journal: Macroeconomics* 13.1, pp. 257–298.

Aquilante, Tommaso, Shiv Chowla, Nikola Dacic, Andrew Haldane, Riccardo Masolo, Patrick Schneider, Martin Seneca, and Srdan Tatomir (2019). "Staff Working Paper No. 798 Market power and monetary policy". In.

Asplund, Marcus and Volker Nocke (2006). "Firm turnover in imperfectly competitive markets". In: *The Review of Economic Studies* 73.2, pp. 295–327.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen (2017). "Concentrating on the Fall of the Labor Share". In: *American Economic Review* 107.5, pp. 180–85.

— (2020). "The fall of the labor share and the rise of superstar firms". In: *The Quarterly Journal of Economics* 135.2, pp. 645–709.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis (2020). "Coverage and representativeness of Orbis data". In: *OECD Science, Technology and Industry Working Papers* 2020/06.

— (Jan. 2023). "Industry concentration in Europe and North America". In: *Industrial and Corporate Change*. eprint: https://academic.oup.com/icc/advance-article-pdf/doi/10.1093/icc/dtac059/48751098/dtac059.pdf.

Bajgar, Matej, Chiara Criscuolo, and Jonathan Timmis (Sept. 2021). *Intangibles and industry concentration: Supersize me*. OECD Science, Technology and Industry Working Papers 2021/12. OECD Publishing.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta (Feb. 2013). "Cross-Country Differences in Productivity: The Role of Allocation and Selection". In: *American Economic Review* 103.1, pp. 305–34.

Bell, Torsten and Daniel Tomlinson (July 2018). "Is everybody concentrating? Recent trends in product and labour market concentration in the UK". In: *Resolution Foundation Publications*.

Bessen, James (2020). “Industry Concentration and Information Technology”. In: *The Journal of Law and Economics* 63.3, pp. 531–555.

Bighelli, Tommaso, Filippo Di Mauro, Marc J Melitz, and Matthias Mertens (2023). “European firm concentration and aggregate productivity”. In: *Journal of the European Economic Association* 21.2, pp. 455–483.

Cavalleri, Maria Chiara, Alice Eliet, Peter McAdam, Filippos Petroulakis, Ana Soares, and Isabel Vansteenkiste (Mar. 2019). *Concentration, market power and dynamism in the euro area*. Working Paper Series 2253. European Central Bank.

Cellan-Jones, Adam, Hussein Farook, Riccardo Ferrari, Maxwell Harris, Alex Rutt, and Mike Walker (2022). “Recent Developments at the CMA: 2021–22”. In: *Review of Industrial Organization* 61.4, pp. 381–403.

CMA (Apr. 2022). *The State of UK Competition*. Tech. rep.

Companies House (Oct. 26, 2023). “Incorporated companies in the UK: July to September 2023”. In: *UK Government Official Statistics*.

Corfe, Scott and Nicole Gicheva (2017). “Concentration not competition: the state of UK consumer markets”. In: *Social Market Foundation Reports*.

Corrado, Carol, Jonathan Haskel, Cecilia Jona-Lasinio, and Massimiliano Iommi (Aug. 2022). “Intangible Capital and Modern Economies”. In: *Journal of Economic Perspectives* 36.3, pp. 3–28.

Covarrubias, Matias, Germán Gutiérrez, and Thomas Philippon (2020). “From Good to Bad Concentration? US Industries over the past 30 years”. In: *NBER Macroeconomics Annual* 34.1, pp. 1–46.

Crouzet, Nicolas and Janice C Eberly (2019). *Understanding weak capital investment: The role of market concentration and intangibles*. Tech. rep. National Bureau of Economic Research.

Davies, Stephen (Dec. 2021). *Competition and Concentration: Charting the Faultlines*. Discussion Papers 21-11. Centre for Competition Policy (CCP), University of East Anglia.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger (2020). “The rise of market power and the macroeconomic implications”. In: *The Quarterly Journal of Economics* 135.2, pp. 561–644.

De Ridder, Maarten (Jan. 2024). “Market Power and Innovation in the Intangible Economy”. In: *American Economic Review* 114.1, pp. 199–251.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda (2016). “Declining business dynamism: What we know and the way forward”. In: *American Economic Review* 106.5, pp. 203–07.

— (Dec. 2020). “Changing Business Dynamism and Productivity: Shocks versus Responsiveness”. In: *American Economic Review* 110.12, pp. 3952–90.

Du, Jun and Karen Bonner (2016). “Decomposing UK aggregate labour productivity and growth: 1998-2013 using the ONS Business Structure Database data”. In: *ERC Research Paper* 48.

Furman, Jason and Peter Orszag (2018). “A firm-level perspective on the role of rents in the rise in inequality”. In: *Toward a just society: Joseph stiglitz and twenty-first century economics*. Columbia University Press, pp. 19–47.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely (2019). “Are US industries becoming more concentrated?” In: *Review of Finance* 23.4, pp. 697–743.

Gutierrez, German and Thomas Philippon (Dec. 2022). “How European Markets Became Free: A Study of Institutional Drift”. In: *Journal of the European Economic Association* 21.1, pp. 251–292. eprint: https://academic.oup.com/jeea/article-pdf/21/1/251/49159838/jvac071.pdf.

Gutiérrez, Germán and Thomas Philippon (2017). *Declining Competition and Investment in the US*. Tech. rep. National Bureau of Economic Research.

Holmes, Thomas J and James A Schmitz Jr (2010). “Competition and productivity: a review of evidence”. In: *Annual Review Economics* 2.1, pp. 619–642.

Hwang, Kyung-In, Anthony Savagar, and Joel Kariel (2023). “Market Power in the UK”. In: *Working Paper*.

Kariel, Joel and Anthony Savagar (2023). “Scale Economies and Aggregate Productivity”. In: *Working Paper*.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard (2024). “Information technology and returns to scale”. In: *American Economic Review*.

Lui, Silvia, Russell Black, Josefa Lavandero-Mason, and Mohammad Shafat (Oct. 2020). *Business Dynamism in the UK: New Findings Using a Novel Dataset*. Economic Statistics Centre of Excellence (ESCoE) Discussion Papers ESCoE DP-2020-14. Economic Statistics Centre of Excellence (ESCoE).

Melitz, Marc J and Gianmarco IP Ottaviano (2008). “Market size, trade, and productivity”. In: *The review of economic studies* 75.1, pp. 295–316.

Olley, G. Steven and Ariel Pakes (1996). “The dynamics of productivity in the telecommunications equipment industry”. In: *Econometrica* 64.6, pp. 1263–1297.

Philippon, Thomas (2018). “A Primer On Concentration, Investment and Growth”. In.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter (2021). “Diverging trends in national and local concentration”. In: *NBER Macroeconomics Annual* 35.1, pp. 115–150.

Savagar, Anthony (2021). “Measured productivity with endogenous markups and economic profits”. In: *Journal of Economic Dynamics and Control* 133, p. 104232.

Syverson, Chad (2019). “Macroeconomics and market power: Context, implications, and open questions”. In: *Journal of Economic Perspectives* 33.3, pp. 23–43.

Tian, Can (2018). “Firm-level entry and exit dynamics over the business cycles”. In: *European Economic Review* 102, pp. 298–326.

Van Reenen, John (Sept. 2018). *Increasing differences between firms: market power and the macro-economy*. CEP Discussion Papers dp1576. Centre for Economic Performance, LSE.

Whish, Richard and David Bailey (2021). *Competition Law*. Oxford University Press.

A Descriptive Statistics

A.1 Aggregate Time Series Statistics

Table 6 summarises the aggregate annual data. It also provides the share of the services and production sectors in total real sales. These are the largest two One-digit sectors in aggregate real sales.

Year	Enterprises	Entry	Exit	Employees	Real Sales (£m)	Services (%)	Production (%)
1997	1,075,090			18,110,502	5,393,074	67.98	14.00
1998	1,321,225	252,814	65,833	19,255,092	4,885,578	62.50	15.01
1999	1,455,407	186,261	152,167	19,851,018	4,925,996	60.58	15.45
2000	1,527,201	206,318	171,039	20,339,884	4,725,748	61.76	15.04
2001	1,589,395	211,538	179,358	20,593,166	4,307,477	56.56	18.05
2002	1,635,202	201,040	221,751	21,530,486	4,098,721	53.53	18.65
2003	1,658,749	215,015	201,896	22,122,568	3,804,810	49.03	18.99
2004	1,757,446	265,778	224,877	22,541,196	4,376,894	55.41	16.24
2005	1,806,696	247,662	204,808	23,275,768	4,391,159	53.91	16.32
2006	1,875,485	246,033	203,362	23,758,580	4,209,926	51.77	16.80
2007	1,948,286	249,099	205,239	24,387,332	4,336,612	50.00	17.78
2008	2,060,091	283,115	244,176	25,195,884	4,434,510	51.10	15.51
2009	2,061,921	209,237	257,277	25,607,450	4,571,046	49.35	16.55
2010	2,023,946	188,748	257,003	25,677,764	4,914,382	49.31	16.26
2011	1,983,789	189,355	238,329	25,584,662	4,487,267	51.14	16.21
2012	2,044,154	255,651	195,104	26,105,020	4,678,756	52.03	16.05
2013	2,107,104	229,967	235,073	26,667,278	4,773,772	51.57	16.70
2014	2,207,507	303,228	214,686	27,674,956	5,068,780	49.89	16.64
2015	2,336,886	314,610	223,760	28,656,576	5,408,977	48.31	17.67
2016	2,462,327	318,115	231,648	29,378,782	5,579,351	48.90	16.71
2017	2,618,596	352,843	262,859	30,270,348	5,360,783	52.35	15.39
2018	2,689,104	301,419	324,638	31,158,416	5,832,771	55.98	13.52
2019	2,717,045	317,733	291,165	31,808,436	5,798,924	52.49	14.59
2020	2,779,397	321,026	297,694	32,229,096	4,423,769	45.16	16.89

Table 6: Annual Aggregate BSD Data

A.2 Distributional Statistics for 2015

We present distributional statistics to show how sales and employment are distributed across firms in the BSD. The distributions are for the year 2015 which is representative of other years. We choose 2015 because it is recent but unlikely to be revised.

Table 7 shows distribution of firms across employees in 2015. The distribution is similar for other years. Single employee firms account for 45% of the total BSD firms. Over 99% of firms have under 100 employees.

Employees	Firms	Percent	Cum.	Real Sales (£m)
1	1,060,525	45.43	45.43	153,786
2	406,316	17.41	62.84	109,225
3	187,666	8.04	70.88	67,827
4	124,764	5.34	76.22	72,063
5	87,978	3.77	79.99	53,954
6 to 10	215,051	9.21	89.2	229,694
11 to 20	128,081	5.49	94.69	248,639
21 to 30	41,996	1.8	96.49	122,748
31 to 40	21,425	0.92	97.41	108,212
41 to 50	12,710	0.54	97.95	128,132
51 to 100	24,674	1.06	99.01	277,719
101+	23,134	0.99	100	3,501,423
Total	2,334,320	100		

Table 7: Distribution of Firm Size in Terms of Employees (2015)

Table 8 shows distribution of firms across sales in 2015. The distribution is similar for other years. Over 90% of firms in the BSD have sales under £1,000,000 in a year.

Sales (£)	Firms	Percent	Cum.	Real Sales (£m)	Employees
0-5,000	37,969	1.62	1.62	94	77,855
5,000-10,000	25,513	1.09	2.70	213	42,755
10,001- 20,000	45,527	1.94	4.64	728	75,843
20,001- 50,000	228,764	9.74	14.38	8,752	327,214
50,001-100,000	591,252	25.17	39.56	46,257	998,451
100,001-200,000	580,682	24.72	64.28	84,006	1,400,068
200,001-500,000	439,756	18.72	83.00	141,726	2,086,626
500,001-1,000,000	176,710	7.52	90.53	127,279	1,735,846
1,000,001-2,000,000	100,718	4.29	94.82	143,580	1,735,349
2,000,001-5,000,000	68,069	2.9	97.72	213,849	2,274,852
5,000,001-10,000,000	25,441	1.08	98.80	180,066	1,834,762
10,000,001-100,000,000	24,420	1.04	99.84	665,594	6,009,688
100,000,001 and above	3,801	0.16	100	3,507,418	10,455,858
Total	2,348,622	100			

Table 8: Distribution of Firm Size in Terms of Sales (2015)

B BSD Relative to ONS Aggregates

To examine if our analysis of BSD data reflects aggregate UK trends as reported by the ONS, we compare our UK entry and exit rates in the BSD with ONS 'Business Demography' data. Both datasets are derived from the IDBR, so we should not expect major differences.

Table 9 shows ONS and BSD entry and exit rates in percentages. In most cases, ONS and BSD rates are similar. There are occasional differences of 1-2%. According to ONS business demography, entry and exit rates are based on the Inter-Departmental Business Register (IDBR). Birth rate is calculated as the number of new registrations (VAT and PAYE) as a proportion of the active businesses. Active businesses are businesses that had either turnover or employment at any time during the reference period. Death rate is calculated using the number of deaths (de-registration of VAT and Pay As You Earn (PAYE)) as a proportion of the active businesses.

Year	ONS entry rate	BSD entry rate	ONS exit rate	BSD exit rate
2010	10.0	9.3	10.6	12.7
2011	11.2	9.5	9.8	12.0
2012	11.4	12.5	10.6	9.5
2013	14.1	10.9	9.7	11.1
2014	13.7	13.7	9.7	9.7
2015	14.3	13.5	9.4	9.5

Table 9: Entry and Exit rate (percentages) in ONS and BSD data

C Regression Sensitivity Analysis

The additional regression results are for the time period 1997-2018.

C.1 Pooled OLS Regression

Table 10 shows the results from OLS regressions of four specifications, not first-differenced. There are more observations than our main regression table as the regression is not first differenced. The first row shows that the lagged concentration ratio is negative, consistent with our main results, but not significant at the 90% level. Net entry has no effect on productivity except in column 4 where sector and year fixed effects are absent. The interaction term between concentration and net entry is also insignificant. Market share has strongly significant positive effect. Age increases productivity, but at a decreasing rate, and firm size (employment) decreases productivity at an increasing rate.

Table 10: The effect of CR5 on Log Labour Productivity; OLS

	1	**2**	**3**	**4**	**5**
$CR5_{t-1}$	-0.060	-0.0597	-0.0468	0.399	-0.0384
	(0.090)	(0.095)	(0.103)	(0.268)	(0.103)
Net $Entry_{t-1}$		-0.195	-0.179	0.186**	-0.197
		(0.207)	(0.209)	(0.080)	(0.207)
$CR5_{t-1} \times$ Net $Entry_{t-1}$			-0.0994	-0.408	-0.0858
			(0.160)	(0.294)	(0.164)
Market $Share_t$				0.216***	0.203***
				(0.0435)	(0.0403)
Firm Age_t				0.0156***	0.0096***
				(0.0048)	(0.0028)
Firm Age_t^2				-0.0003**	-0.0002***
				(0.0001)	(0.00006)
Firm $Size_t$				-0.137**	-0.147***
				(0.0625)	(0.0503)
Firm $Size_t^2$				0.0108	0.0169*
				(0.0106)	(0.0090)
Industry FE	Yes	Yes	Yes		Yes
Year FE	Yes	Yes	Yes		Yes
N	35,041,501	35,041,501	35,041,501	35,041,501	35,041,501
R^2	0.096	0.096	0.097	0.014	0.106
Clusters	85	85	85	85	85

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Note: Standard errors are clustered at industry level. Sales are constant 2016 values.
CR5, net entry and market share are measured at two-digit SIC industry level.
Coefficients and SE for CR5 are multiplied by 100 (*i.e.* the variable unit is divided by 100.)
Dependent variable: log labour productivity

C.2 Other Concentration Measures

Table 11 shows the effect of concentration on productivity using other measures of concentration. The concentration ratio (CR*N* measures) capture market share of the *N* top firms with the highest market shares while the HHI index represents concentration from all firms in the sector. The results show other measures of concentration have weak negative effects on productivity, usually greater than 1 standard deviation from zero, but not significant at the 90% level. In all specifications, we find that net entry and market share increase productivity, while firm size reduces productivity.

Table 11: The Effect of Different Measures of Concentration on Log Labor Productivity

	1	2	3	4
$\Delta CR10_{t-1}$	-0.0334			
	(0.0210)			
$\Delta(CR10_{t-1} \times \text{Net Entry}_{t-1})$	0.0051			
	(0.0376)			
$\Delta CR20_{t-1}$		-0.0321		
		(0.0216)		
$\Delta(CR20_{t-1} \times \text{Net Entry}_{t-1})$		0.0038		
		(0.0363)		
$\Delta CR50_{t-1}$			-0.0291	
			(0.0222)	
$\Delta(CR50_{t-1} \times \text{Net Entry}_{t-1})$			-0.0017	
			(0.0361)	
ΔHHI_{t-1}				-0.0073
				(0.0053)
$\Delta(HHI_{t-1} \times \text{Net Entry}_{t-1})$				-0.0037
				(0.00593)
Δ Net Entry_{t-1}	0.107**	0.107**	0.109**	0.112**
	(0.0494)	(0.0492)	(0.0492)	(0.0508)
Δ Market Share_t	0.188***	0.188***	0.188***	0.188***
	(0.0329)	(0.0329)	(0.0329)	(0.0329)
Δ Firm Age_t	0.00274	0.00275	0.00276	0.00274
	(0.00296)	(0.00296)	(0.00296)	(0.00296)
Δ Firm Age_t^2	-0.0001	-0.0001	-0.0001	-0.0001
	(0.00008)	(0.00008)	(0.00008)	(0.00008)
Δ Firm Size_t	-0.569***	-0.569***	-0.569***	-0.569***
	(0.0371)	(0.0371)	(0.0371)	(0.0371)
Δ Firm Size_t^2	-0.0007	-0.0007	-0.0007	-0.0007
	(0.0053)	(0.0053)	(0.0053)	(0.0053)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	28,845,009	28,845,009	28,845,009	28,845,009
R^2	0.11	0.11	0.11	0.11
Clusters	85	85	85	85

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. Regressions are in first difference.

Note: Standard errors are clustered at industry level. Sales are constant 2016 values. Concentration, net entry and market share are measured at two-digit SIC industry level. Coefficients and SE for CR are multiplied by 100 and HHI by 1000.
Dependent variable is log labour productivity.

D Aggregate Plots

D.1 Average Labour Productivity

Figure 16 shows that average labour productivity has a decreasing trend. Average labour productivity takes the unweighted arithmetic average of labour productivity at the firm-level. That is, $\frac{1}{N}\left(\frac{y_1}{l_1} + \frac{y_2}{l_2} + \ldots + \frac{y_N}{l_N}\right)$ where N is number of firms and y_i for $i \in 1, \ldots, N$ is revenue of a firm and l_i is employees at a firm. The initial fall in average labour productivity is likely due to greater coverage of small firms, but after 2005 the dynamic stabilizes and is close to the aggregate labour productivity in the main appendix.[31]

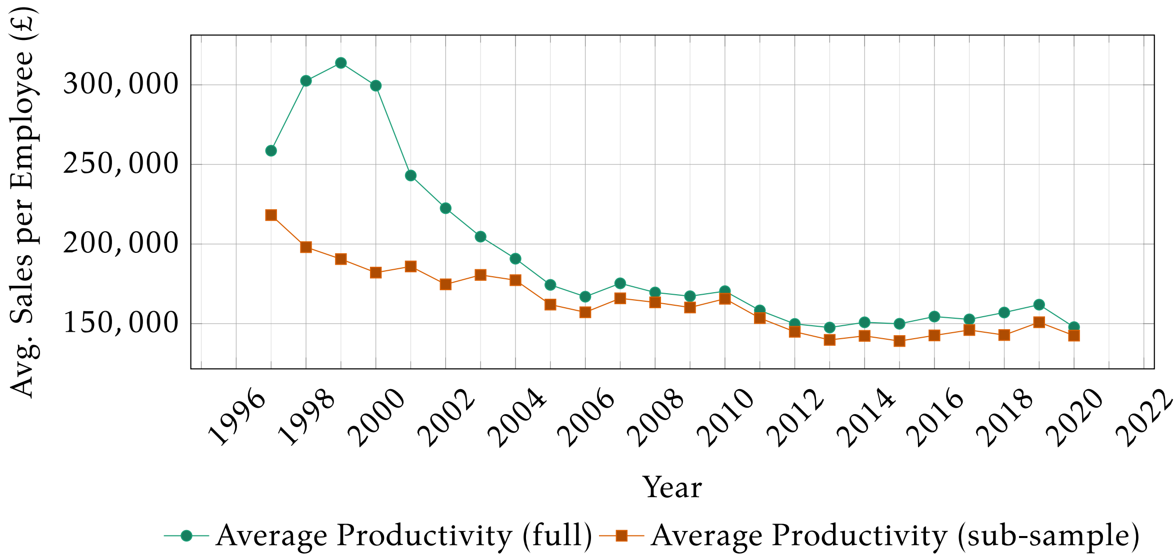


Figure 16: Average Firm Labour Productivity (BSD, 1997-2020)
Source: Authors' calculation based on BSD 1997-2020

D.2 Aggregate Nominal and Real Sales

Figure 17 shows aggregate nominal and real sales. This shows the effect of deflating nominal sales with the ONS 2016 deflator. Both nominal and real sales show an upward trend in aggregate sales over the 2000s with a dip in 2010-2011 which corresponds to the recession period of 2008-2009 given the timing considerations of the BSD. We use real sales throughout our analysis.

[31]Du and Bonner (2016) also suggest that the decreasing trend in average labour productivity is due to an increase in single-employee firms.

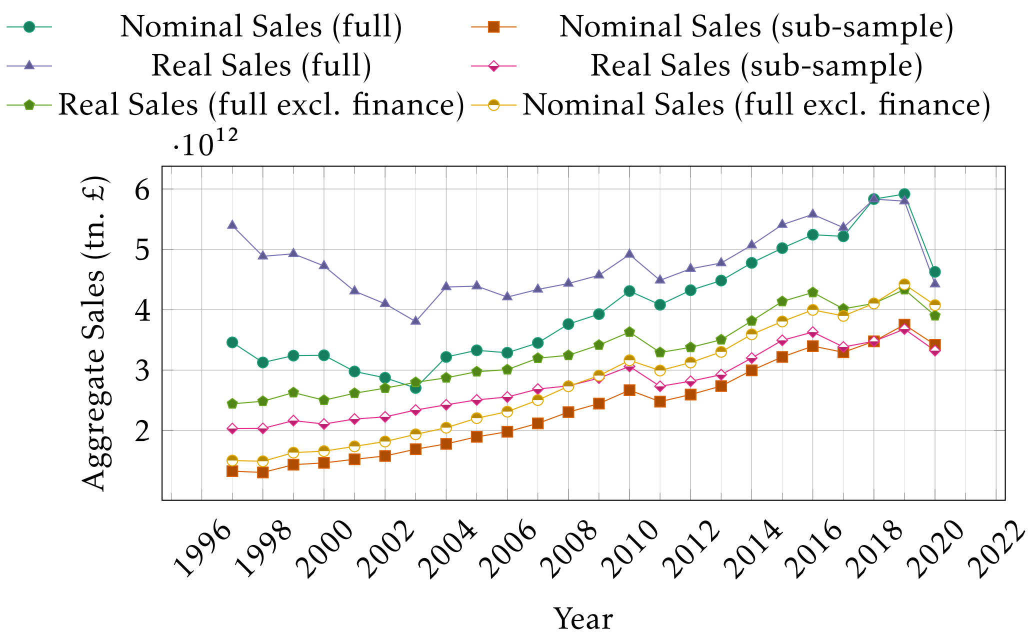


Figure 17: Aggregate Sales Real Vs Nominal (BSD, 1997-2020)
Source: Authors' calculation based on BSD 1997-2020

D.3 Aggregate Concentration Excluding Finance

Figure 18 shows an alternative, less restricted, sub-sample. It only excludes financial services from the full sample. The resulting figure shows a similar pattern to the sub-sample. Therefore the flat trend in Figure 5 is driven by the financial services sector. This emphasizes that the largest firms in the dataset are financial services firms. Therefore, we could re-interpret the trend in Figure 5 as a flat trend in the sales share of financial services firms in total sales.

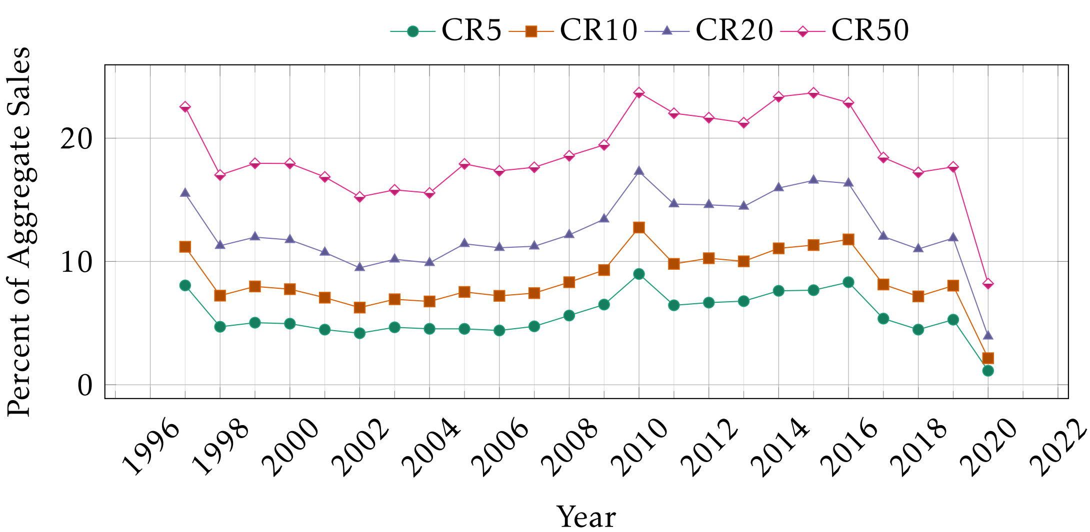


Figure 18: Aggregate CR5, CR10, CR20 and CR50, *excluding financial services*
Source: Authors' calculation based on BSD 1997-2020

E BSD Sector Plots

BSD Sectors are an ONS-specific aggregation of two-digit SIC industries. The BSD sector is used as a sampling frame for ONS surveys such as the Annual Business Survey (ABS). They do not correspond to SIC One-digit sectors.

BSD Sector	Two-digit	One-digit
Production	1-39	Agriculture, Forestry and Fishing (A), Mining and Quarrying (B), Manufacturing (C), Electricity, Gas, Steam and A/C (D), Water Supply and Waste Management (E).
Construction	41-43	Construction (F)
Motor Trade	45	Wholesale, Retail and Motor Trade (G)
Wholesale	46	Wholesale, Retail and Motor Trade (G)
Retail	47	Wholesale, Retail and Motor Trade (G)
Other Services	49-96*	Transport and Storage (H), Information and Communication (J), Financial and Insurance Services (K), Professional, Science and Tech (M), Administrative and Support Services (N), Public Administration (O) Education (P), Human Health and Social Work (Q), Arts, Entertainment and Recreation (R), Other Services (S).
Catering	55-56	Accommodation and Catering (I)
Property	68	Real estate Activities (L)

Table 12: Two-digit SIC, Sector and Sub-sector classification

Note: Other Services 49-96 excludes Catering (55-56) and Property (68).

E.1 Concentration BSD Sector Level

Figure 19 plots concentration ratios at the BSD sector level.

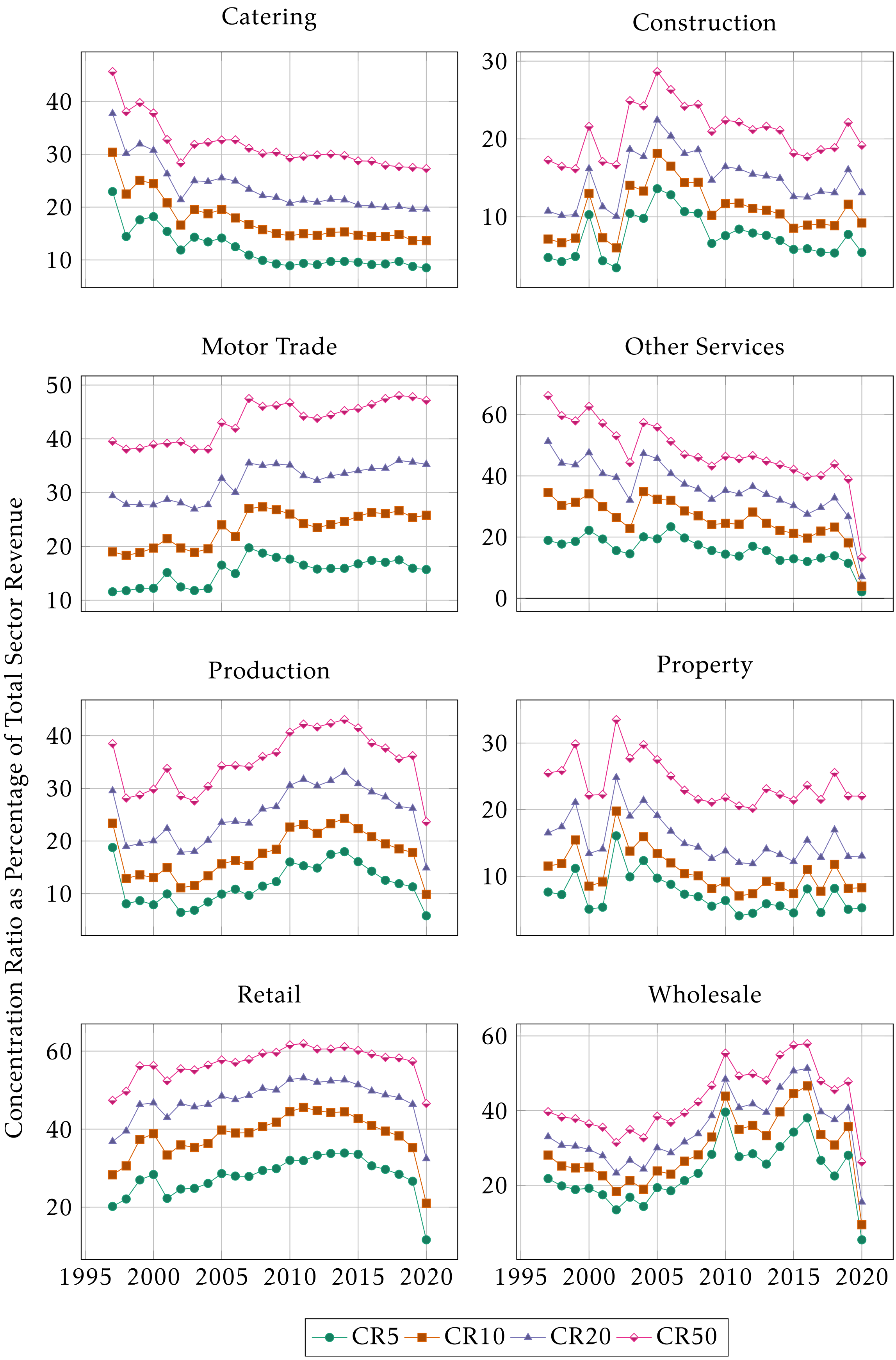


Figure 19: Concentration Ratios at BSD Sector Level
Source: Authors' calculation based on BSD 1997-2020

Figure 20 plots Hirfindahl-Hirschman Indices at the BSD sector level.

Figure 20: Herfindahl-Hirschman Indices at BSD Sector Level
Source: Authors' calculation based on BSD 1997-2020

E.2 Firm Entry & Exit BSD Sector level

Figure 21 plots entry and exit rates at the BSD sector level.[32] The plots are consistent with our aggregate observations. There are no clear long-run trends in entry and exit and the measures are responsive to the business cycle. Entry fluctuates more than exit. There is consistently high net-entry in services and low or negative net entry in production. All sectors observe more exit than entry during the Great Recession and high net entry throughout the 2010s.

[32]The 'BSD sector level' refers to an ONS-specific aggregation of two-digit industries.

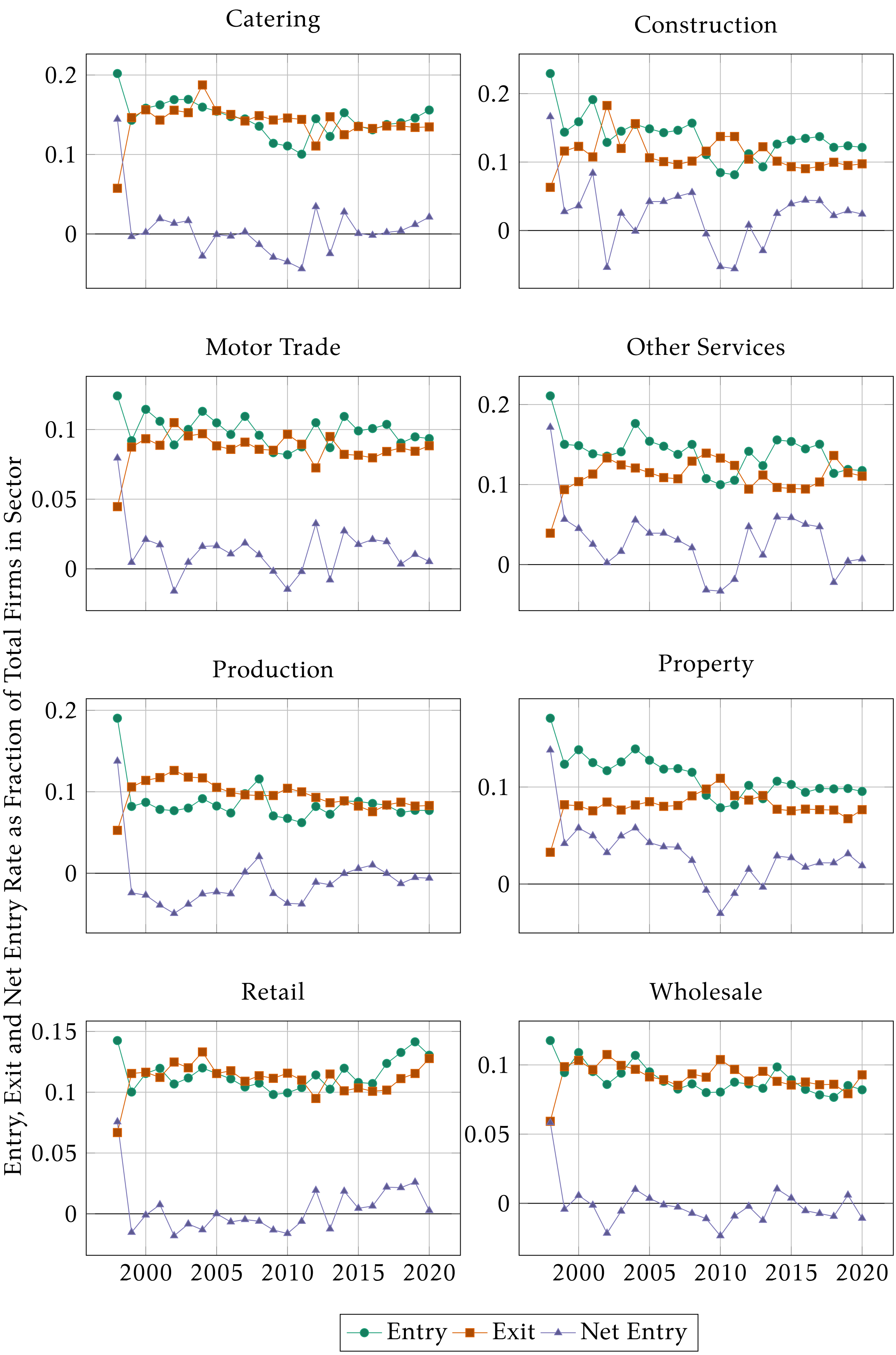


Figure 21: Entry, Exit and Net Entry Rates at BSD Sector Level
Source: Authors' calculation based on BSD 1997-2020

F Further Decomposition Theory and Empirics

F.1 Decomposition Theory

We present a standard statistical decomposition of a weighted-average into an unweighted-average and a covariance term. In our analysis, we apply this decomposition to productivity (employment weighted). In this context, the decomposition is commonly known as an Olley-Pakes decomposition following Olley and Pakes (1996).

We begin with a definition of the sample covariance between firm productivity x_i and the firm's labour share within an industry, given by $\omega_i = L_i / \sum_{i=1}^{N} L_i$. The unweighted average is defined as $\bar{x}_N \equiv \frac{1}{N}\sum_{i=1}^{N} x_i$. Since the shares sum to one, the unweighted average share is $\bar{\omega}_N = \frac{1}{N}\sum_{i=1}^{N}\omega_i = \frac{1}{N}$. There are N firms in the industry, subscripted with i, and, without loss of generality, we do not apply Bessel's correction to our covariance definition:

$$\begin{aligned}
\mathbb{C}ov(\boldsymbol{x}_N, \boldsymbol{\omega}_N) &= \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x}_N)(\omega_i - \bar{\omega}_N) \\
&= \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x}_N)\left(\omega_i - \frac{1}{N}\right) \\
&= \frac{\bar{x}_N}{N}\sum_{i=1}^{N}\left(\frac{x_i}{\bar{x}_N} - 1\right)\left(\omega_i - \frac{1}{N}\right) \\
&= \frac{\bar{x}_N}{N}\left(\sum_{i=1}^{N}\frac{x_i\omega_i}{\bar{x}_N} - \frac{1}{N}\sum_{i=1}^{N}\frac{x_i}{\bar{x}_N} - \sum_{i=1}^{N}\omega_i + \sum_{i=1}^{N}\frac{1}{N}\right) \\
&= \frac{\bar{x}_N}{N}\left(\sum_{i=1}^{N}\frac{x_i\omega_i}{\bar{x}_N} - 1 - 1 + 1\right) \\
&= \frac{\bar{x}_N}{N}\left(\sum_{i=1}^{N}\frac{x_i\omega_i}{\bar{x}_N} - 1\right) \\
&= \frac{1}{N}\left(\sum_{i=1}^{N}x_i\omega_i - \bar{x}_N\right).
\end{aligned}$$

Hence, we obtain the standard additive decomposition:

$$\sum_{i=1}^{N} x_i\omega_i = \bar{x}_N + N \cdot \mathbb{C}ov(\boldsymbol{x}_N, \boldsymbol{\omega}_N).$$

The covariance simplifies to give the decomposition we use in the paper:

$$\sum_{i=1}^{N} x_i \omega_i = \bar{x}_N + \sum_{i=1}^{N} (x_i - \bar{x}_N)(\omega_i - \bar{\omega}_N).$$

When the covariance is substituted out the second term is the sum of product deviations. This measures the covariability of the two variables. In this case, the extent to which firms with more employment share are more productive. The additive term in this decomposition adjusts the average-firm-productivity upward or downward to yield the average-worker-productivity. This is also a measure of worker concentration in high-productivity firms. The measure is relative to the average firm productivity $\bar{x}_N$. The additive decomposition highlights the importance of considering both the overall productivity level of firms and the distribution of workers across productivity levels within those firms when assessing overall workforce productivity.

F.2 Additional Decomposition Results

In Figures 22 and 23, we compare average-firm and average-worker productivity across industries in 2015. The figures give a sense of the level of average-worker productivity and average-firm productivity. For example, industry 45 (motor vehicles) has an average-worker productivity of £35,000 and an average-firm productivity of £20,000. The difference between the bars captures allocative efficiency, which we plot in Figure 14. If average-worker productivity exceeds average-firm productivity there is positive allocative efficiency, and vice-versa for negative allocative efficiency.

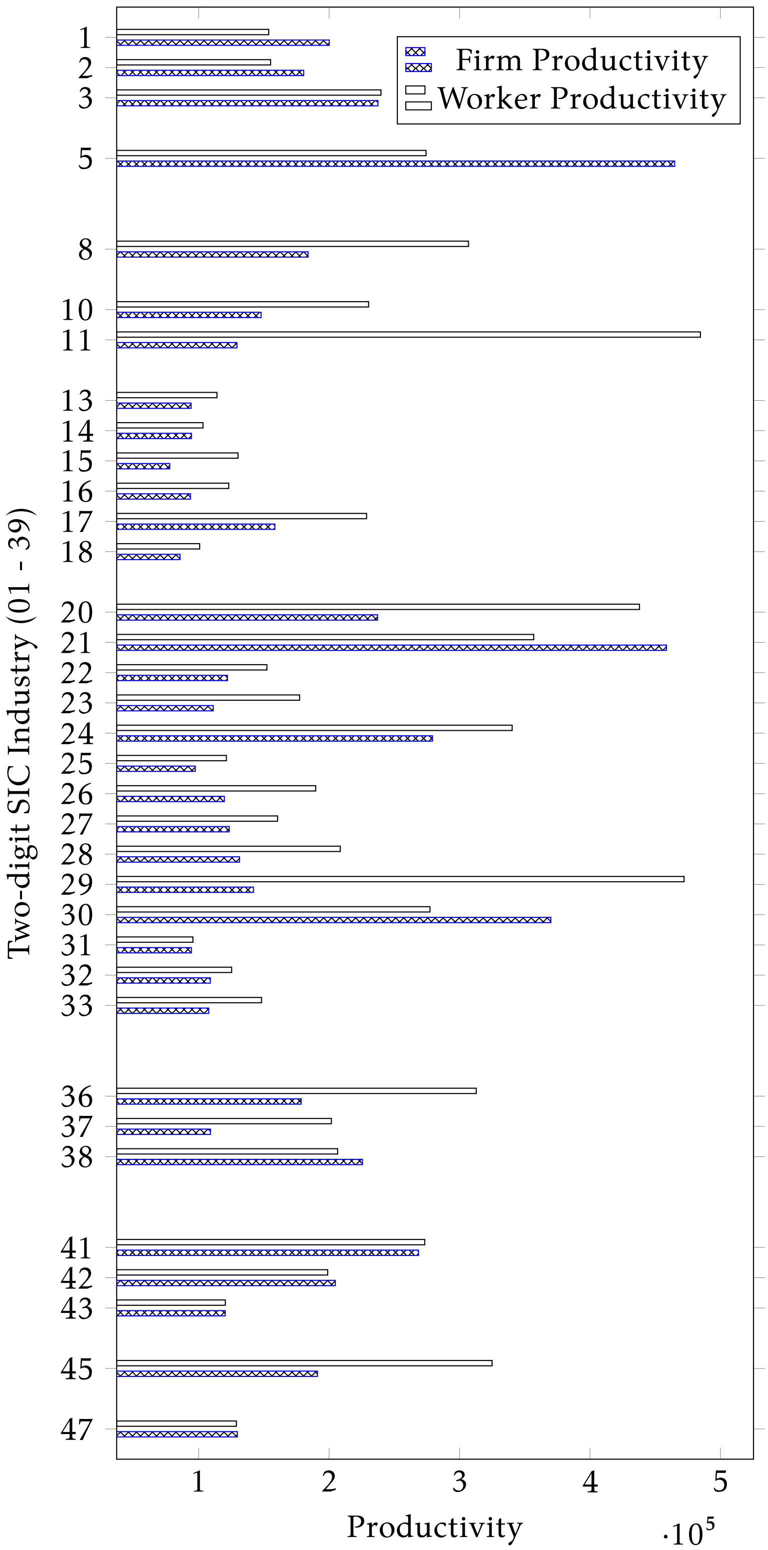


Figure 22: Firm Vs. Worker Productivity by Industry in 2015, Sectors 01-33

Note: We drop sectors 06 (oil), 07 (mining metal), 09 (quarrying), 12 (tobacco), 19 (petrol manufacturing), 35 (electricity), 39 (waste management), 46 (wholesale).

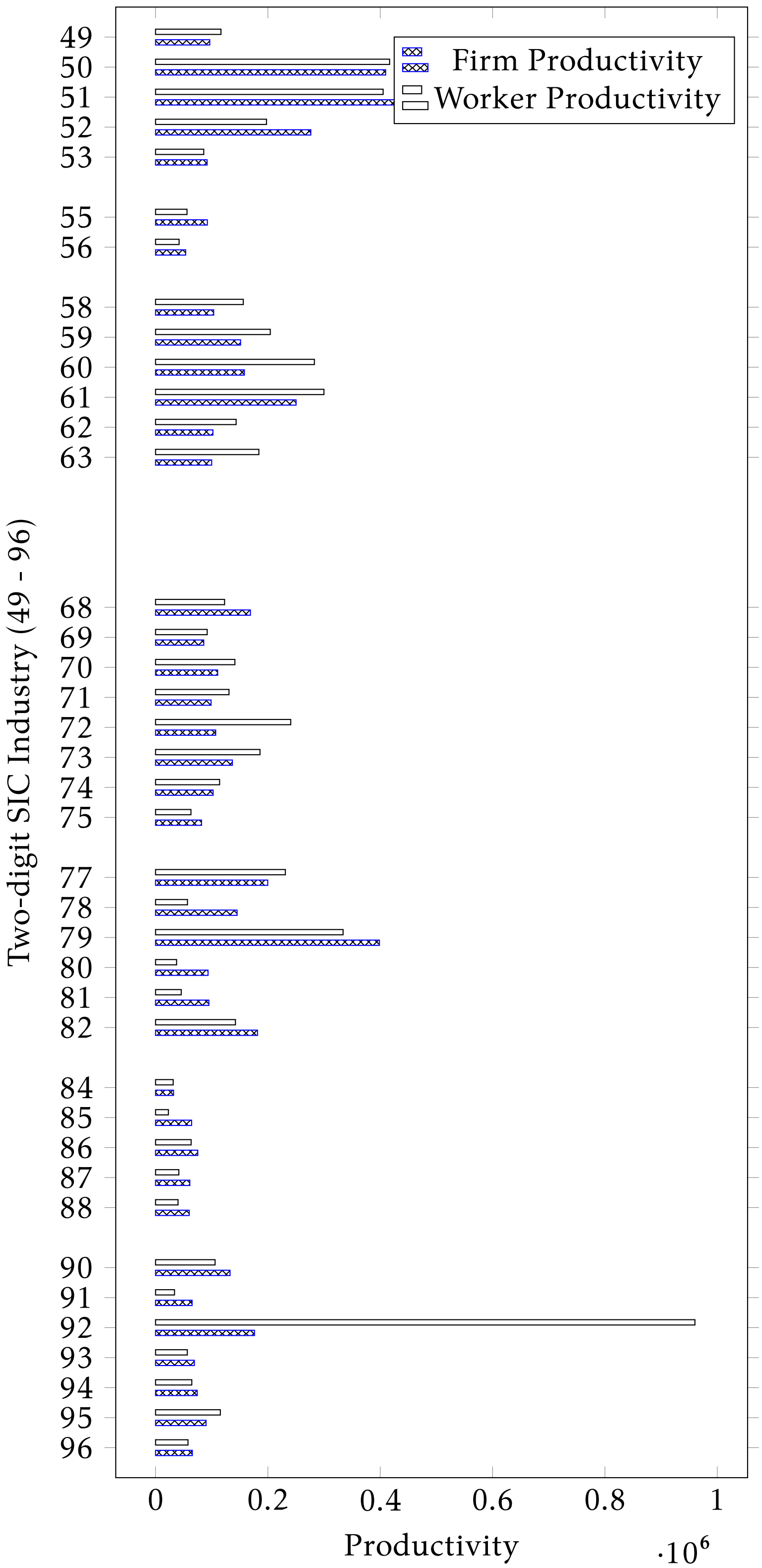


Figure 23: Firm Vs. Worker Productivity by Industry in 2015, Sectors 49-96

Note: We drop sectors 64 (finance), 65 (insurance), 66 (other finance), 92 (gambling).

01	Crop and animal production	A	Agriculture, forestry and fishing
02	Forestry and logging	A	Agriculture, forestry and fishing
03	Fishing and aquaculture	A	Agriculture, forestry and fishing
05	Mining of coal	B	Mining, quarrying and utilities
06	Extraction of petroleum and gas	B	Mining, quarrying and utilities
07	Mining of metal ores	B	Mining, quarrying and utilities
08	Other mining and quarrying	B	Mining, quarrying and utilities
09	Mining support service activities	B	Mining, quarrying and utilities
10	Manufacture of food products	C	Manufacturing
11	Manufacture of beverages	C	Manufacturing
12	Manufacture of tobacco products	C	Manufacturing
13	Manufacture of textiles	C	Manufacturing
14	Manufacture of wearing apparel	C	Manufacturing
15	Manufacture of leather	C	Manufacturing
16	Manufacture of wood	C	Manufacturing
17	Manufacture of paper	C	Manufacturing
18	Printing	C	Manufacturing
19	Manufacture of coke and petroleum	C	Manufacturing
20	Manufacture of chemicals	C	Manufacturing
21	Manufacture of pharmaceuticals	C	Manufacturing
22	Manufacture of rubber and plastic	C	Manufacturing
23	Manufacture of non-metallic minerals	C	Manufacturing
24	Manufacture of basic metals	C	Manufacturing
25	Manufacture of fabricated metal	C	Manufacturing
26	Manufacture of computers	C	Manufacturing
27	Manufacture of electrical equipment	C	Manufacturing
28	Manufacture of machinery	C	Manufacturing
29	Manufacture of motor vehicles	C	Manufacturing
30	Manufacture of other transport	C	Manufacturing
31	Manufacture of furniture	C	Manufacturing
32	Other manufacturing	C	Manufacturing
33	Repair and installation of machinery	C	Manufacturing
35	Electricity; gas; steam and AC	D	Electricity, gas, steam and AC
36	Water collection; treatment, supply	E	Water, sewerage, waste
37	Sewerage	E	Water, sewerage, waste
38	Waste collection	E	Water, sewerage, waste
39	Remediation; other waste management	E	Water, sewerage, waste
41	Construction of buildings	F	Construction
42	Civil engineering	F	Construction
43	Specialised construction	F	Construction
45	Motor vehicles	G	Wholesale and retail trade
46	Wholesale trade	G	Wholesale and retail trade
47	Retail trade	G	Wholesale and retail trade

Table 13: SIC Two-digit Sectors 01-47 (abbreviated definitions)

49	Land transport and pipelines	H	Transport and storage
50	Water transport	H	Transport and storage
51	Air transport	H	Transport and storage
52	Warehousing	H	Transport and storage
53	Postal and courier	H	Transport and storage
55	Accommodation	I	Accommodation and food services
56	Food and beverage	I	Accommodation and food services
58	Publishing	J	Information and communication
59	Video, TV, Music	J	Information and communication
60	Programming and broadcasting	J	Information and communication
61	Telecommunications	J	Information and communication
62	Computer programming	J	Information and communication
63	Information service activities	J	Information and communication
64	Financial service activities	K	Finance and insurance
65	Insurance	K	Finance and insurance
66	Other financial services	K	Finance and insurance
68	Real estate activities	L	Real estate activities
69	Legal and accounting activities	M	Professional, scientific and technical
70	Head offices; management consultancy	M	Professional, scientific and technical
71	Architectural and engineering	M	Professional, scientific and technical
72	Scientific research	M	Professional, scientific and technical
73	Advertising	M	Professional, scientific and technical
74	Other professional; scientific	M	Professional, scientific and technical
75	Veterinary activities	M	Professional, scientific and technical
77	Rental and leasing activities	N	Administrative and support services
78	Employment activities	N	Administrative and support services
79	Travel agency	N	Administrative and support services
80	Security	N	Administrative and support services
81	Services to buildings	N	Administrative and support services
82	Office administrative	N	Administrative and support services
84	Public administration and defence	O	Public administration and defence
85	Education	P	Education
86	Human health	Q	Human health and social work
87	Residential care	Q	Human health and social work
88	Social work	Q	Human health and social work
90	Creative; arts and entertainment	R	Arts, entertainment and recreation
91	Libraries; archives; museums	R	Arts, entertainment and recreation
92	Gambling and betting	R	Arts, entertainment and recreation
93	Sports; amusement; recreation	R	Arts, entertainment and recreation
94	Membership organisations	S	Other service activities
95	Repair of household goods	S	Other service activities
96	Other personal service	S	Other service activities
97	Domestic personnel	T	Activities of households as employers
98	Household subsistence	T	Activities of households as employers
99	Intergovernmental organizations	U	Intergovernmental organizations

Table 14: SIC Two-digit Sectors 49-99 (abbreviated definitions)