ONLINE APPENDICES

Appendix A. Data

Detailed description of data preparation for Historical Orbis

Our main data set is the UK component of Bureau Van Dijk's Historical Orbis. HO is a compilation of the active and dead incorporated firms in BVD's various databases. The UK component of this is FAME and has been provided to BVD over several decades. The main balance sheet and profit and loss accounting variables are kept (in a standardised way across countries in US$). There are about 2.9 million firms (rows) in 2016 (see Table A1), which is the last full year of data in our version (we are currently using the December 2018 vintage of HO, but we will update as more becomes available). The column headed 'Firms' is the number of companies in the raw data. There is clearly an increase in the number over time, starting with only 362,473 in 1996. The numbers further decline as one goes back in time to 1982. The main reason for the growth in the number of firms is likely that BVD has not kept all the inactive firms in the early years of HO. The number of firms has grown in the UK over this period, but not by this much. This is because the UK part of HO is from FAME and the data provider did not keep all the exiting firms before a certain year. This is one of the main problems with HO and its predecessors (such as Amadeus).

The columns labelled 'Revenue', 'Wage bill', 'Employment', 'COGS' and 'EBITDA' give the number of firms with non-missing values on each variable. It is clear that most firms do not report these items, due to accounting requirements being tougher for larger firms than smaller ones. The non-reporting is more severe for employment (115,183 observations in 2016) than for revenues (241,375 firms in 2016). Note, however, that the number of firms does not grow so much when conditioning on these non-missing values, as larger firms (which have mandatory reporting) are less likely to drop out of the sample (e.g. the number of firms reporting employment in 1996 is 92,644). Taking advantage of this, our analysis samples A and B are defined in the main text and below. These implement the data cleaning, focus on the market sector and condition on different sets of non-missing variables. The number of observations is much more stable for these samples. From Table A2, we see that Sample A has just under 31,000 firms in the first year and just over 33,000 in the last year and that Sample B has around 25,000 in the first year and 31,000 in the last year. Clearly, there is not the massive increase in firm numbers from the raw data after cleaning.

Table A1. Raw number of firms reporting data items

Year	Firms	Revenue	Wage bill	Employment	COGS	EBITDA
1996	362,473	149,275	99,604	92,644	109,887	137,448
1997	381,597	152,012	97,821	90,935	111,596	140,079
1998	540,335	188,774	103,039	94,939	127,816	165,741
1999	930,412	282,374	121,761	108,306	167,155	233,551
2000	1,065,781	303,775	122,855	108,167	178,675	247,289
2001	1,143,639	304,835	123,012	107,186	181,139	247,127
2002	1,209,337	308,068	127,438	108,045	183,811	247,196
2003	1,323,169	317,277	131,607	109,690	191,612	253,959
2004	1,519,083	356,252	134,088	103,387	223,841	258,830
2005	1,643,464	354,841	135,719	101,504	220,836	255,544
2006	1,775,753	367,416	133,209	101,013	224,239	248,524
2007	1,865,511	347,070	131,899	102,835	197,983	240,281
2008	1,920,397	304,655	128,927	102,584	172,411	228,738
2009	1,939,020	285,098	127,103	101,494	161,180	206,084
2010	2,004,953	295,114	152,120	107,763	151,445	184,133
2011	2,100,978	262,612	113,881	108,097	132,134	160,244
2012	2,233,100	248,193	107,859	107,139	124,457	151,521
2013	2,382,811	238,784	103,418	106,947	118,111	146,164
2014	2,548,551	229,638	102,215	107,136	113,119	139,173
2015	2,735,659	234,449	104,855	108,299	109,836	135,500
2016	2,891,300	241,375	109,659	115,183	108,822	136,546

Note: This table shows the raw number of observations before further sample restrictions. 'Firms' gives raw number of firms per year in Historical Orbis. The next column is the subsample where a firm reports a non-missing revenue number. Similarly, the other columns report numbers of firms with non-missing values for wage bill, employment, COGS or EBITDA, respectively.

Table A2. Analysis samples: number of observations

Year	Sample A	Sample B	Listed	Currently listed
1996	29,483	23,681	955	721
1997	29,762	23,707	1,069	797
1998	30,072	23,669	1,137	856
1999	30,356	23,462	1,190	840
2000	30,542	23,289	1,245	945
2001	31,140	23,346	1,326	1,001
2002	32,199	23,259	1,356	1,031
2003	32,891	22,998	1,347	992
2004	28,079	20,804	1,392	1,051
2005	26,566	19,925	1,438	1,131
2006	26,606	19,727	1,460	1,163
2007	27,436	19,924	1,452	1,157
2008	27,417	20,007	1,369	1,075
2009	26,687	23,023	1,306	969
2010	28,350	24,162	1,322	913
2011	28,803	24,550	1,326	887
2012	29,220	24,955	1,303	847
2013	29,536	25,202	1,309	855
2014	30,143	25,848	1,286	879
2015	30,876	26,983	1,254	867
2016	30,510	28,280	1,221	847

Note: The columns show the number of observations in the samples we use for the analysis. Samples A and B are defined in the text. They restrict the sample to firms in the 'market economy', remove duplicates, use observations at the highest level of aggregation, and drop firms that have missing values on employment and/or have fewer than 10 employees. Sample A also conditions on non-missing wage bill and EBITDA (but allows missing observations on sales and COGS). Sample B conditions on non-missing sales and COGS (but allows missing observations on wage bill and EBITDA). 'Currently listed' are firms that are currently on the UK stock market. 'Listed' includes all those that we know have been publicly listed at some point in time.

Duplicates. Historical Orbis contains multiple observations per year for some firms. The first reason for this is multiple 'filing types'. Some firms have both an observation stemming from the 'Annual Report' and one from 'Local Registry Filing'. In these cases, the Annual Report typically contains more information on the key variables for our analysis – revenue, COGS, the wage bill and employment. Second, HO contains some seemingly identical observations, which only differ in a few variables (e.g. revenue might be missing in one observation but not the other). This is particularly the case in the early years of the sample (pre-2002). Third, the filing period may change, which can lead to multiple observations per firm in a given year. Finally, some firms submit both a consolidated account and an unconsolidated account, as can be identified via the consolidation codes ('U2' and 'C2').

We remove duplicates by applying the following steps. First, conditional on firm ID, year and consolidation code:

1. we remove Local Registry Filing whenever there is also an Annual Report;
2. we take the observations with fewer missing values;
3. we remove remaining duplicates by taking the first observation that appears in the data set.

The only duplicates that are left after this procedure are those that differ in their consolidation code. Whenever a firm submits both an unconsolidated and a consolidated account, we select the consolidated ones and drop the unconsolidated components from subsidiaries.

Excluding subsidiaries. Historical Orbis contains the consolidated accounts for business groups in addition to the accounts of their subsidiaries. This requires some additional attention to rule out double counting the subsidiaries, since their information is also included in the accounts of their parent companies.

We use the HO Ownership files to construct the ownership hierarchy for each firm. We consider all majority links where a firm owns more than 50% of another firm and then repeatedly merge the direct owners to construct the full hierarchy. BVD records the ownership links at certain dates throughout the year and there may be some links missing in some of the yearly files. We impute the information in these cases by assuming that an ownership link remains valid until there is a new one. In addition, since the historical ownership information only goes back to 2002, we assume that the ownership structure remains constant in previous years (i.e. the structure is the same in 2000, 2001 and 2002 in our main analysis sample).

Based on the hierarchy, we focus on the highest level, which (1) is an industrial firm (rather than a bank or investment trust) and (2) submits a consolidated account, and exclude subsidiaries that are owned by these firms. We drop all holding companies as defined by those firms whose primary SIC code is 617 or have 'Holding' in their company names. This might be too conservative, as some of these do appear to be genuine industrial companies.

Sample restrictions based on sectors. For our main analysis, we focus on a sample of firms from the 'market sector'. This means we exclude industries with close links to the public sector, financial sector and oil-related industries. More precisely, we use the three-digit US 1987 SIC codes ('USSICCOR') and exclude the following industries:

- codes between 10 and 97 (agriculture, forestry and fishing);
- codes between 100 and 149 (mining);
- 1311 (manufacture of petroleum products);

- 430 and 431 (postal service);
- codes between 490 and 495 (including utilities such as electricity and water);
- codes between 600 and 679 (finance);
- codes between 800 and 809 (health);
- codes between 820 and 835 (education);
- 841, 863, 864, 865 and 866 (museums/art galleries, labour union, civic/social/fraternal associations, political and religious organisations respectively);
- 880 and 881 (private households);
- codes greater than or equal to 900 (public sector).

Foreign firms. Although we include subsidiaries of foreign firms, we exclude the consolidated accounts of non-UK firms. When a non-UK firm opens a branch in the UK, it may be required to register the accounts of the full parent company. As a result, the Orbis UK file includes some consolidated accounts for non-UK firms that can be identified through additional letters in the firm ID, which starts with either 'GBFC', 'GBSF' or 'GBNF' in these cases.

Sample selection. The coverage of HO for the UK should be all incorporated UK firms, both living and inactive. However, there appears to be some tail-off in earlier years. Part of this may be that the underlying data from FAME do not keep all firms. However, it may also be that BVD has not actually kept all inactive firms. The data set contains very few observations for the late 1970s and early 1980s. Consequently, we do not use the early years of data. Sample coverage appears good since the mid 1990s and we chose our main analysis sample to begin in 1996 when coverage seems to stabilise. Note that while the number of firms still increases substantially after 1996, many of these firms report only their assets. As a result, the number of firms that actually report revenue, wage bill, employment, COGS or EBITDA is much more stable.

Missing values on accounting variables. A major issue is that firms do not need to report all accounting items in all years. Broadly, in the UK, large firms have to report full accounts on the balance sheet and the profit and loss (P&L). Medium-sized firms also have to report full accounts on the balance sheet, but can report abbreviated P&L accounts. Finally, micro firms only need to report abbreviated balance sheets. Thus, (almost) all firms in Historical Orbis report basic balance sheet – current assets and liabilities. However, most firms have some accounting variables reported but not others. The definitions of large, medium and micro enterprises are mainly based on thresholds depending on the balance sheet assets, sales revenue and employment in the previous two years. The exact thresholds change over time. In addition to these mandatory rules, firms can voluntarily choose to report items and many firms do this.

Main analysis samples. We apply the cleaning rules above to the HO data. We then create three broad analysis samples. This is in order to avoid working with too many different samples.

First, we use publicly listed firms. We define whether a firm is listed based on the information on firms' legal status provided by Orbis, which contains whether a firm is unlisted, listed or delisted in the latest year of the data, as well as the dates of listing and delisting. Our main sample focuses on firms that are currently publicly listed on the UK stock market, but we also consider a version of these data on firms that we know have ever been listed on the UK stock market. We compare this sample with Worldscope (see De Loecker and Eeckhout (2018) for more details), an alternative database of listed firms. Worldscope has better coverage of firms in earlier years.

Second, 'Sample A' contains all firms that report the wage bill, EBITDA (earnings before interest tax, dividends and amortisation) and employment. We construct our main value added measure as the sum of wage bill and EBITDA and keep observations with positive value added. Furthermore, we restrict the sample to firms with 10 or more employees. This last restriction is because smaller firms change whether or not they report employment substantially over time, so if we do not make this restriction the sample numbers change a lot because of this reporting change. This is the sample we use to look at productivity and labour shares.

Our third sample ('Sample B') requires that firms report revenue, COGS and employment and have at least 10 employees. This is the sample we use to look at COGS shares and calculate markups. Note that we drop outliers in terms of markups (i.e. markups smaller than 0.5 or larger than 10), which comprise 2.7% of the sample. Alternative cut-offs made little difference. We also considered a number of other alternative samples (such as the intersection of Samples A and B) and obtained the same qualitative results. Table A2 shows the number of observations in each of the samples. The summary statistics are shown in Tables A3 and A4.

Table A3. Sample A: summary statistics

	Mean	Median	SD	N
Wage bill	9,852,885.73	1,742,703.81	93,940,404.68	616,674
EBITDA	5,691,654.52	422,575.25	108,403,376.28	616,674
Employment	394.49	67.00	5,061.42	616,674
Productivity	46,105.40	32,392.68	342,190.32	616,674
Labour share	0.79	0.79	0.26	599,919
Foreign	0.46	0.00	0.50	616,674
Listed	0.03	0.00	0.18	616,674

Note: This table shows the summary statistics for Sample A. Note that all years (starting with 1996) are pooled. The number of observations for the labour share is slightly lower because we exclude implausible labour shares (≥ 2).

Table A4. Sample B: summary statistics

	Mean	Median	SD	N
COGS	42,099,786.86	6,356,109.64	385,945,270.13	490,801
Revenue	60,122,183.13	9,401,648.58	553,919,732.45	490,801
Employment	441.20	69.00	5,591.89	490,801
Foreign	0.49	0.00	0.50	490,801
Listed	0.04	0.00	0.20	490,801

Note: This table shows the summary statistics for Sample B. Note that all years (starting with 1996) are pooled.

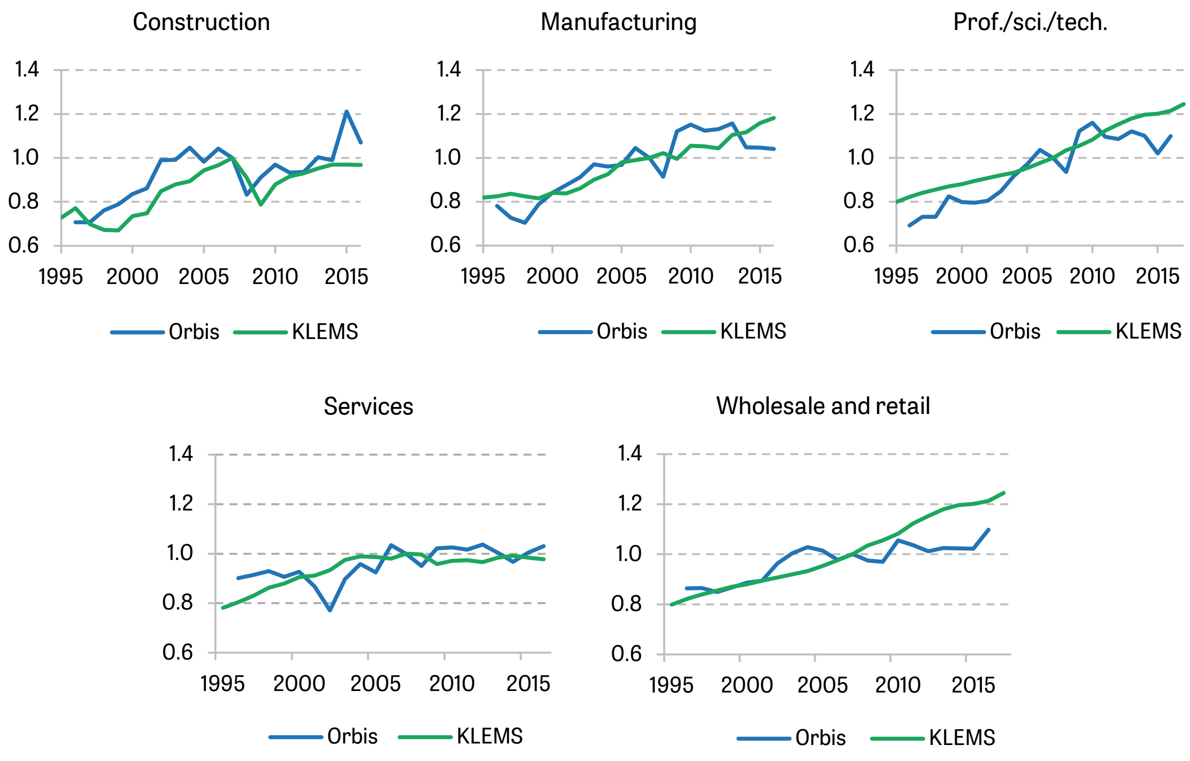
Appendix B. Further analysis

In this appendix, we present some further analysis of the data that we left out of the main paper due to space constraints. We look at productivity, markups, labour shares, concentration and business dynamism in turn in each subsection.

Productivity

A further way to compare the aggregated productivity trends across HO and administrative data sources is to look by industry. Figure A1 breaks down value added per worker in HO by the five broad sectors in the market economy and compares it with KLEMS data (KLEMS uses publicly available ONS administrative sources for the UK). The trends look broadly similar across the two data sets, with a bit more choppiness in the HO data (note that these are indexed to be 1 in 2007 and the levels of the productivity can be different). Productivity has grown fastest in manufacturing, wholesale and retail, and the professional, scientific and technological sector. Construction has also had fast, but volatile, productivity growth. Services have had slower growth. The only large difference between data sets is in the (small) professional, scientific and technological sector, where the productivity level jumped up faster in HO than in the KLEMS data after the financial crisis.

Figure A1. Productivity: comparing administrative data and Orbis by industry



Note: This figure compares productivity trends for different sectors between the ORBIS and KLEMS data. Productivity is normalised to 1 in 2007.

Source: Market sector from Historical Orbis; KLEMS data derived from ONS.

On the micro level, an alternative source for examining trends in productivity to our main HO database is to use ABI/ABS data from the ONS. We draw on the tabulated versions of these micro data produced by the Office for National Statistics (2020), which are also discussed in Oliveira-Cunha et al. (2021). These have the distribution of employment-weighted firm value added per worker, so the same concept that we have in panel B of Figure 4 in the main text.

As noted in Section 3 of the main text, there are many reasons why there could be differences between our main HO database and the ONS ABI/ABS and both sources have advantages and disadvantages. Before getting into these, we first discuss what the trends look like in the two data sets. Figure A2 displays the change in the overall distributions. We base this in 1998 – even though our HO data can go back to at least 1996 as in the main text – because this is the first year when the (re-designed) ABI/ABS is reliable. We end in 2016 as this is the last year of reliable HO data (the published ABI/ABS tables run through 2018[1]). Note that Figure A2 is comparing two cross-sectional distributions and displaying how they have changed – it is *not* comparing just the same firms over time. We choose to show the 10th to 95th percentiles as the ONS does not publish all the tails and these are more sensitive to outliers.[2]

Panel A of Figure A2 has the absolute change in (inflation-adjusted) real value added per worker. For example, the 87th percentile of the HO productivity distribution has seen a £30,760 growth in productivity (up from £91,391 in 1998 to £122,151 in 2016). The qualitative picture from this figure is very clear in both data sets: the vast majority of the distribution has seen a near-zero increase in productivity, whereas there has been substantial growth of productivity at the upper tail. In the HO data, productivity starts to increase at the 60th percentile and rises almost monotonically as we rise up the distribution. In the ONS data, productivity is stagnant until we reach the 75th percentile whereupon it also rises monotonically. To put it another way, three-quarters of workers have been in firms that have only managed a £800 increase in productivity over 18 years.

Quantitatively, the rise in upper-tail inequality is more dramatic in the HO data than in the ONS data. The growth at the 95th percentile is over twice as large in HO as in the ONS data. The most likely explanation for this is that HO includes the global consolidated accounts of firms. Some of these UK incorporated firms have substantial overseas sales and employment from their foreign affiliates in other countries, and these will be particularly strong at the top of the distribution. By contrast, the ONS data are only those of establishments located in the UK, so will not take into account the faster growth of subsidiary activity.[3]

Nonetheless, the broad message is clear: there has been a substantial increase in upper-tail inequality regardless of data source. This is broadly consistent with existing work on the ABI/ABS data in Bahaj et al. (2017) and Office for National Statistics (2019).

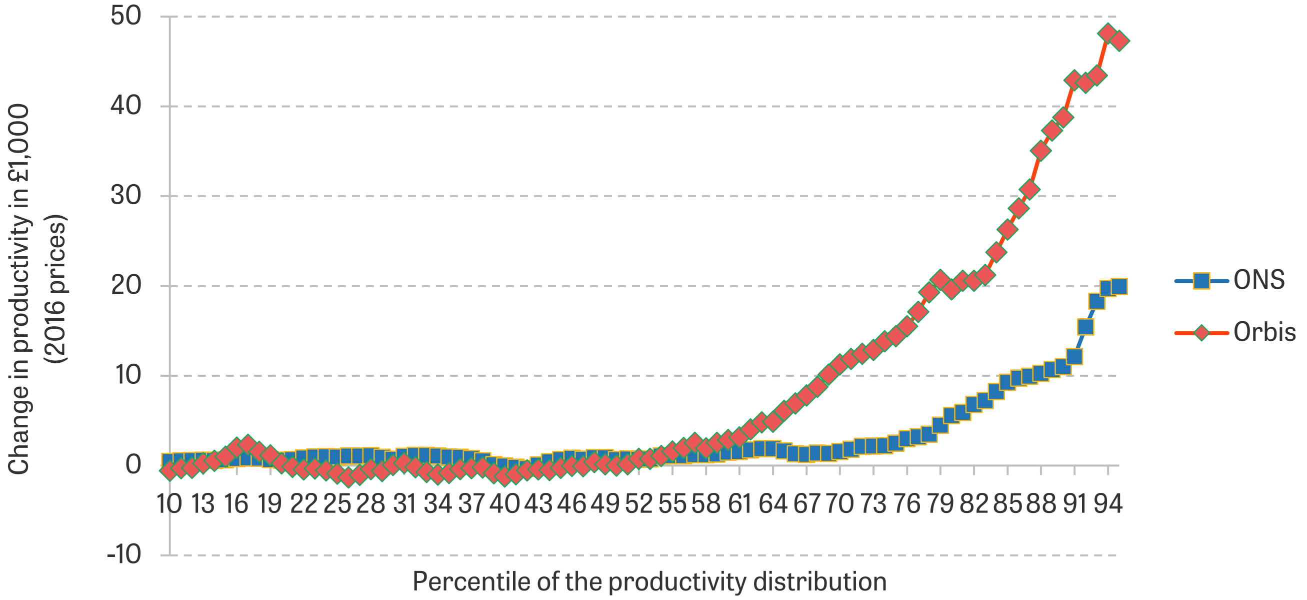
1 Including the last two years makes little difference – see below.

2 For example, there are significant numbers with negative productivity at the lower tail of the ABI/ABS distribution, which suggests measurement error: not just negative gross profits, but such large negatives that they exceed the wage bill.

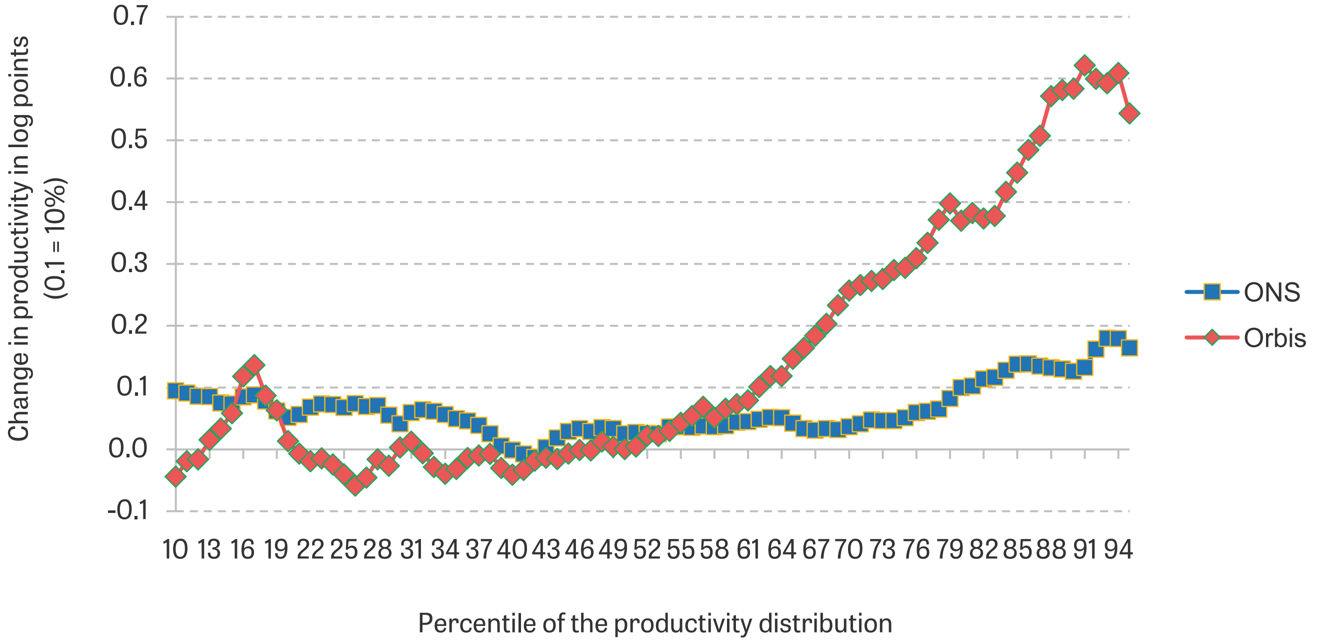
3 This will increase productivity in the upper quantiles because it effectively gives more weight to the larger firms (which have generally higher productivity). Moreover, if productivity of overseas establishments is rising faster than in the UK, this will magnify this effect.

Figure A2. Change in firm productivity at different points of the distribution 1998–2016 in Orbis and ONS data

Panel A. Absolute productivity change



Panel B. Percentage productivity change



Note: 'Orbis' is Sample A from our analysis in Figure 4 in the main text except (in order to be consistent with ONS, which only reports the full distribution including firms with negative productivity) we keep in the (few) firms with negative productivity. None of the firms has negative productivity in this figure, but the location of the percentiles is affected by this. The panels compare the distribution of value added per worker (labour productivity) over the 10th to 95th percentiles for both datasets in 1998 and 2016. The horizontal axis is the percentile and the vertical axis shows the change in 2016 (compared with 1998). Panel A has this in absolute terms (£), whereas panel B is in logs, so it approximates the percentage change (e.g. 0.1 is about a 10% increase). Note these are cross-sections of firms, so the figure is not looking at the change over time within a firm. The productivity distribution is weighted by firm employment.

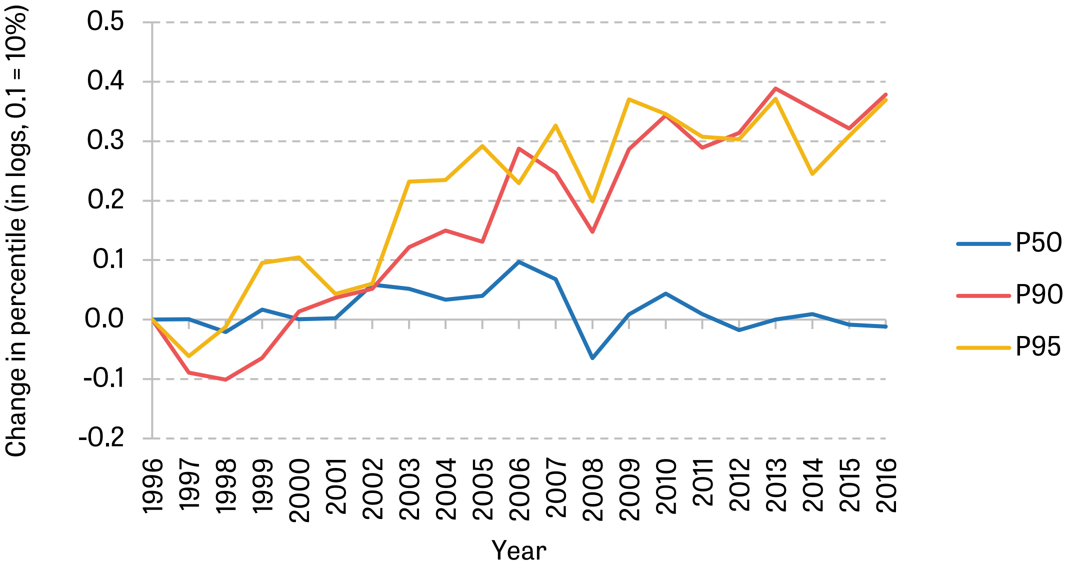
An alternative way to present the data is in terms of proportional growth, as in panel B of Figure A2 which uses log(productivity). This takes into account the fact that the lower percentiles have (by definition) lower productivity, so the same absolute change will mean a larger percentage change. This makes little difference to the qualitative picture for HO data: the growth at the bottom half of the distribution hovers around zero, whereas the top of the distribution has substantial increases of the order of 60% for those at the 90th and 95th percentiles. Although the ONS data also show the largest gains at the top albeit smaller in magnitude (e.g. about 20% at the 95th), the bottom sees faster growth than the middle. The 10th percentile has about 10% growth whereas the median has near-zero growth. This is mechanically due to the fact that absolute productivity growth is the same in the lower tail, so there is a bigger percentage increase in the bottom than the top. A focus on the 90:10 log productivity distribution would give the misleading impression that there had not been much change in the quantiles over time, whereas we can clearly see there are different things happening in the top of the distribution where the top is pulling away from the middle.

To look at the time-series patterns over the two-decade period, we can present the quantiles in a similar way to Figure 4 in the main text. Given the above discussion, we split the analysis into looking at upper-tail inequality separately from lower-tail inequality. Figure A3 presents the 95th, 90th and 50th percentiles for HO (panel A) and for ABI/ABS (panel B). As in the main text, we do this in logs, so these are percentage growths at different points of the distribution, and we do it for all years that we have reliable data (so ABI/ABS goes through 2018 and HO starts in 1996). Consistently with panel B of Figure A2, we see increased dispersion in both data sets, with the upper quantiles growing faster than the median. Note that this is mainly happening in the later years – after 2002 in HO and 2004 in ABI/ABS. In addition, as before, the magnitude of the trends is stronger in HO than in the ONS data. Nonetheless, the clear pattern is of increased upper-tail dispersion.

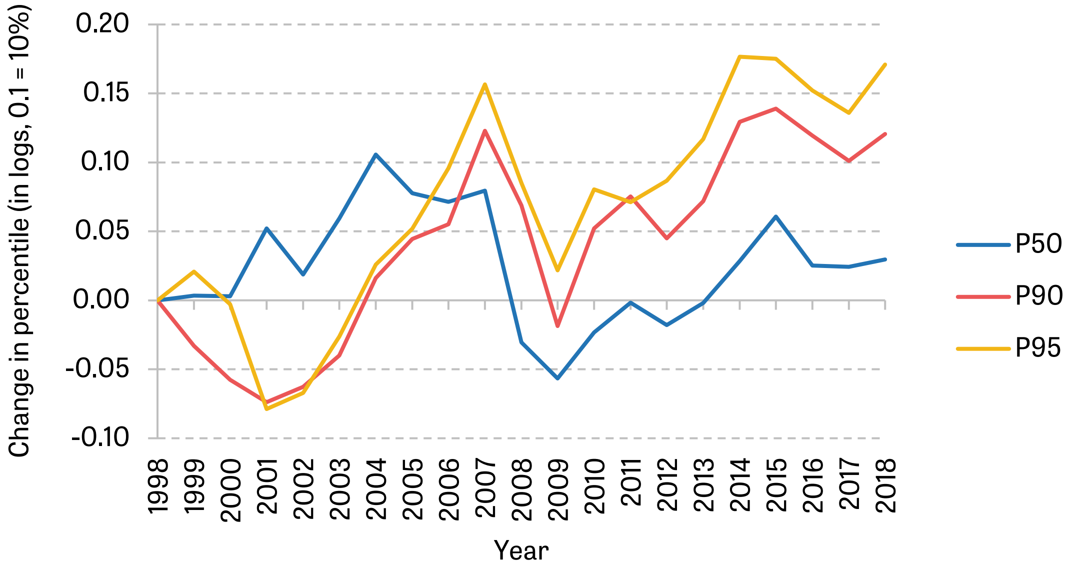
We turn to lower-tail inequality in Figure A4. Here the patterns (at least in the percentage terms shown here) do exhibit different trends. Over the period as a whole, both data sets agree that median productivity is broadly flat. However, in panel A, Orbis shows a fall of around 20% for the 10th percentile firms, whereas the ONS data show a 13% increase. Looking at the year-by-year changes, the data sets broadly agree that there was widening lower-tail inequality from the start of the period until the Great Recession. The 10th percentile recovers to some extent in the HO data, but not by enough to offset the 1996–2008 fall. By contrast, the ONS data in panel B indicates an enormous increase in productivity in one year for the 10th percentile: about 30 log points in 2009–10 (about a 35% increase). This is followed by another 20-log-point increase in 2012–13 that is enough to more than reverse the losses in the decade leading up to 2009.

Figure A3. Changing upper-tail inequality in firm productivity

Panel A. Historical Orbis



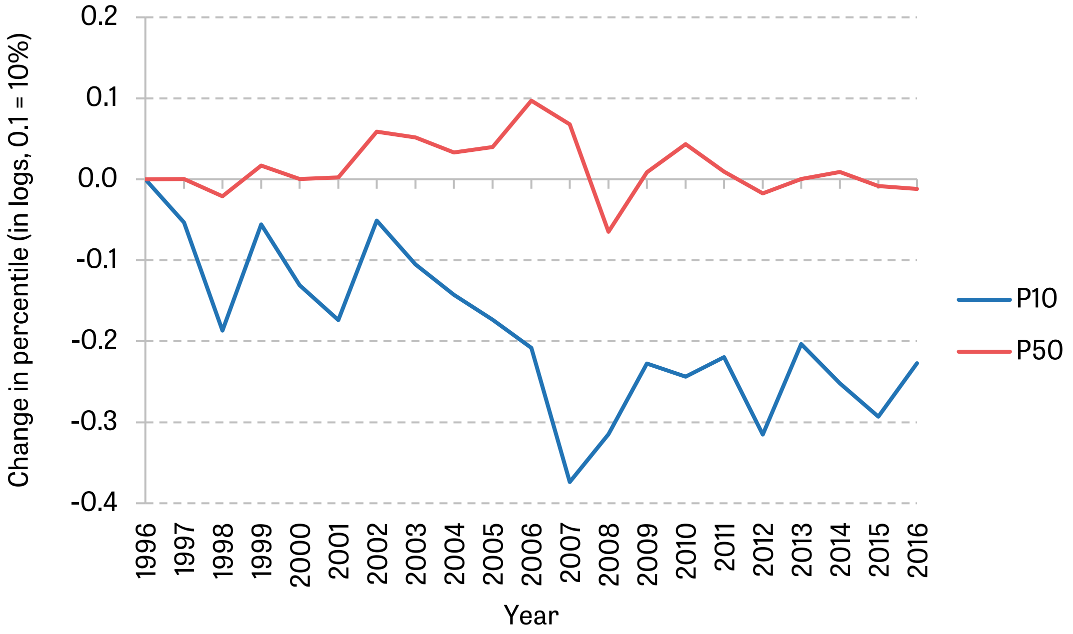
Panel B. ONS ABI/ABS



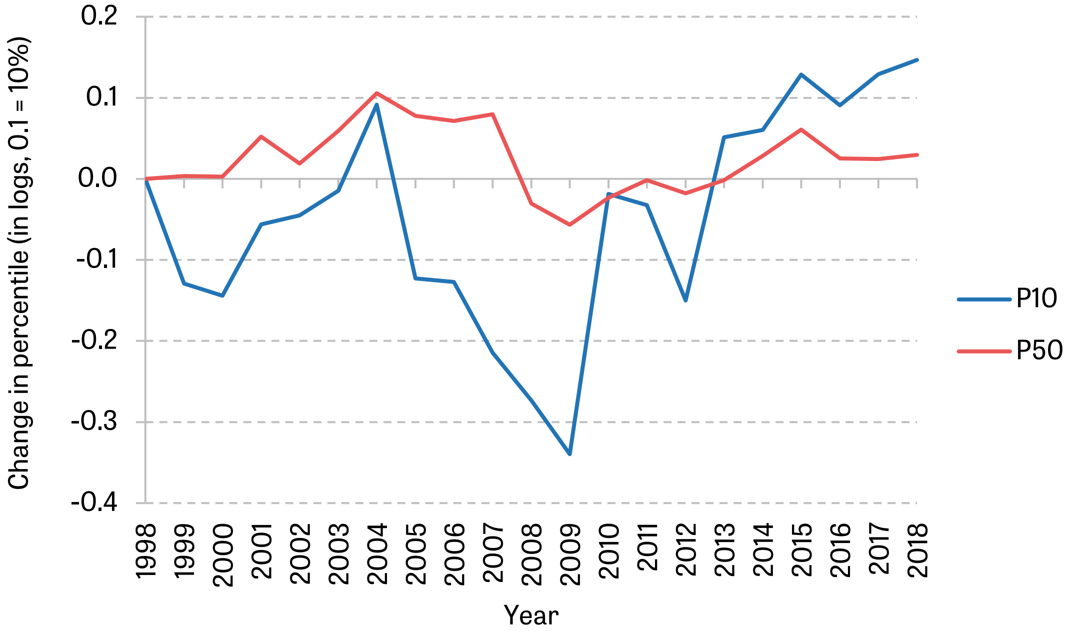
Note: Same samples as in Figure A2. The panels compare quantiles of employment-weighted log(real labour productivity) for the 95th, 90th and 50th percentiles (P95, P90 and P50, respectively) (1996–2016 for HO and 1998–2018 for ABI/ABS). The vertical axis shows the percentage change (in logs) for every year compared with the base year (normalised at zero). Panel A has the results for HO and Panel B for the ONS ABI/ABS.

Figure A4. Changing lower-tail inequality in firm productivity

Panel A. Historical ORBIS



Panel B. ONS ABI/ABS



Note: Same samples as in Figure A2. The panels compare quantiles of employment-weighted log(real labour productivity) for the 10th and 50th percentiles (P10 and P50, respectively) (1996–2016 for HO and 1998–2018 for ABI/ABS). The vertical axis shows the percentage change (in logs) for every year compared with the base year (normalised at zero). Panel A has the results for HO and Panel B for the ONS ABI/ABS.

Summary

In summary, we find that the ONS data are qualitatively in broad agreement with the HO data over the longer run in that there has been a substantial increase in upper-tail inequality from the mid 1990s until the late 2010s. The 95th and 90th percentiles of the employment-weighted productivity distribution have pulled away from the median. Since this covers about half the workforce, it is a major phenomenon.

What is less clear is what is happening in less productive firms where the other half of employees work. In our Orbis data the 10th percentile diverges from the median and inequality increases, whereas in the ONS data the 50:10 shrinks with the 10th percentile having a faster proportionate growth rate. In addition, there are quantitative differences in the magnitude of the changes, with HO showing larger shifts than the ONS ABI/ABS. We detail possible reasons for the differences below, but our broad sense is that understanding what is happening in the lower tail is much harder because the data coverage is much poorer. For large firms, both data sets have near-universal coverage. However, for the lower half, there is highly incomplete coverage and the methods we have taken to correct for this (sampling weights in ABI/ABS, restricted samples in HO) are imperfect. This is clearly an area where more work is needed.

Some possible reasons for divergence between HO and ONS ABI/ABS data

- Measurement of productivity is different. In HO we use accounting pre-tax and pre-depreciation profits (EBITDA) plus the wage bill to measure value added, whereas the ONS uses sales less purchases from the survey responses in ABI/ABS. Employment is conceptually similar, but one is accounting and the other survey responses.

- The units are different. HO are incorporated firms, whereas ABI/ABS are establishments ('reporting units', RUs). Many firms contain multiple reporting units, so the ONS data are effectively at a lower level of aggregation.

- In the ONS data, RUs can be collapsed to the enterprise unit (or enterprise group) level (although this is not how it is reported in the public use tables). Even then, the numbers in the ONS data refer to activity in the UK, whereas companies in HO use consolidated accounts that report worldwide activity. Consequently, the productivity in HO is the UK firm's global productivity, whereas in ABI/ABS it will be the domestic productivity. This affects multinationals. Note that goods and services produced in the UK and exported are captured in both data sets; the difference is the activities of foreign affiliates, which are in HO but not in ABI/ABS. Both are interesting of course.

- Coverage is different. The ABI/ABS is a stratified random sample. Firms with 250 or more employees (about half of all employees) are included every year. There is a probability sample for smaller firms. Sampling weights are known and results are reweighted to be representative. However, our analysis of the ONS micro data shows some problems with the weights, especially in the early years of the ABI/ABS.[4] HO is theoretically a 100% sample, but there are also concerns in the earlier years with the capture of all firms (e.g. those that exit). In addition, there are the issues we discussed earlier, that accounting regulations allow SMEs to not report all the items needed for construction of productivity. Sample A drops firms with under 10 employees where this is particularly a problem, but this means that the ONS data will

[4] It is likely that the ONS has some internal weights series for the early years, but these are not, to our knowledge, released to researchers. The ONS uses imputations in order to account for non-response of large reference units but not (as far as we know) for differential response rates with the known cell stratification for small units.

represent a lot more of these micro enterprises than HO. The weighting by employment in Figures A2–A4 helps mitigate this.

- The industries covered are different. Our analysis of HO has dropped some hard-to-measure and volatile sectors such as the extraction industries (oil and gas) and agriculture as well as public-sector-related industries such as education and health, whereas some of these are included in the ONS data.

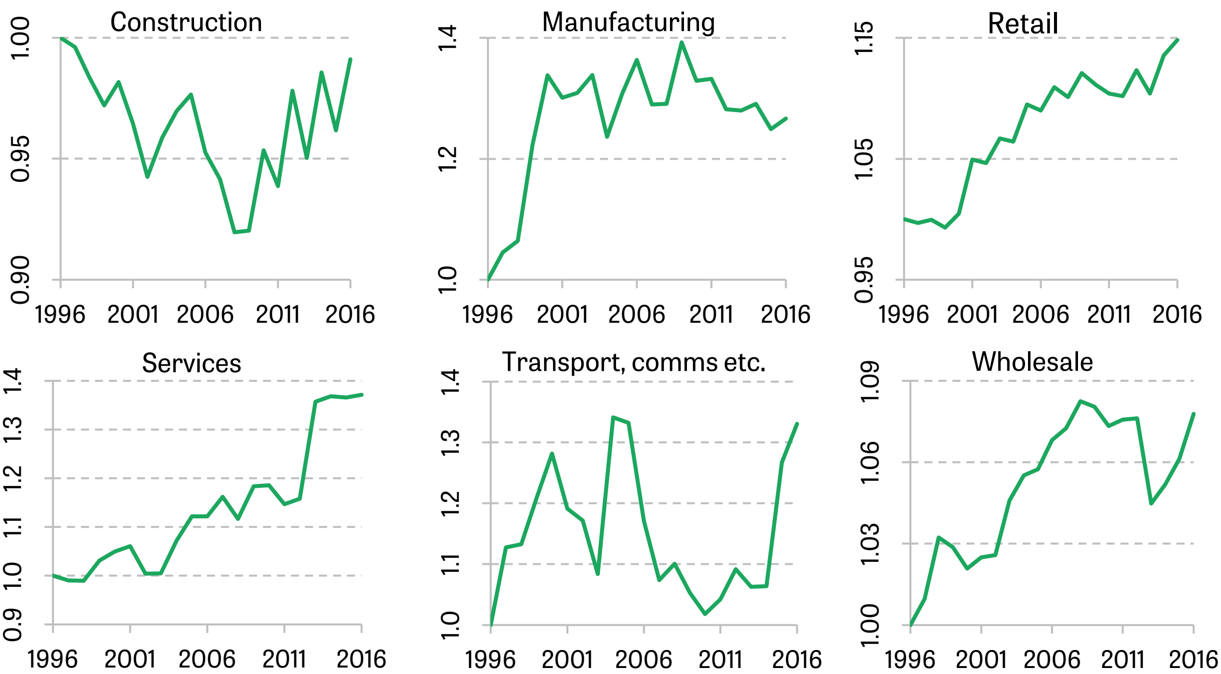
The bottom line is that there are a large number of reasons why different trends could be observed in the two data sets. The qualitative similarity for upper-tail inequality is therefore quite reassuring.

Markups

Figure A5 shows the markup changes by broad industry. The markup has risen in all sectors, although it is least clear in the volatile construction sector.

Figure A6 compares the weighted markup for publicly listed firms in HO with that in Worldscope. Worldscope has good coverage from the late 1980s onwards and Orbis since the mid 1990s (see Appendix A). We see that the markup has risen in both series.[5]

Figure A5. Mean markup by broad sector (1996 = 1)

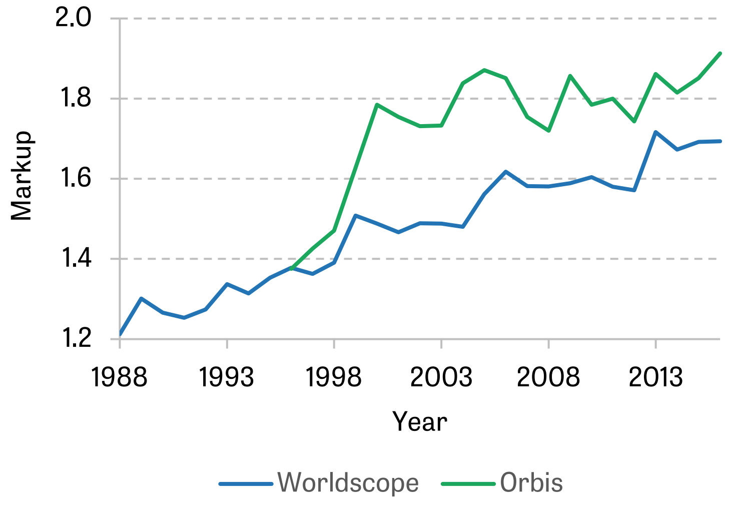


Note: Markups are computed with a constant output elasticity of 0.85 and the means are weighted by turnover.

Source: Market sector from Historical Orbis using Sample B.

[5] The levels are different: although the COGS share starts at broadly the same level in both series in 1996, the fall in Orbis is greater than in Worldscope. However, the trends are consistent. We are still investigating the exact reasons for differences between the data sources. The main reason seems to be subtle differences in the definition of net versus gross turnover measures, although there are also differences in what counts as a 'publicly listed' firm between data sets.

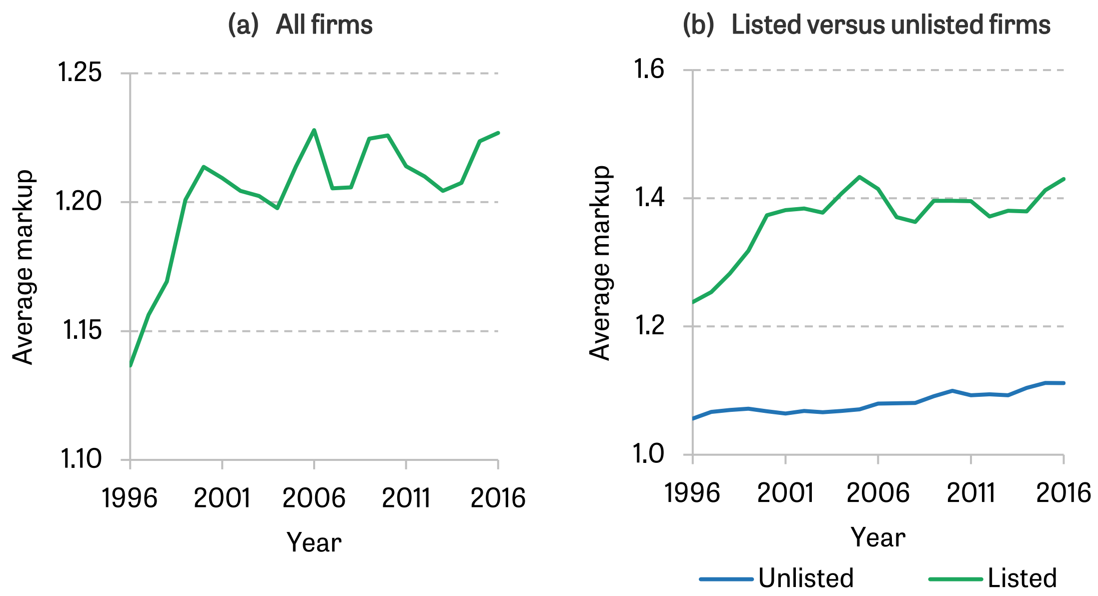
Figure A6. Mean markup for publicly listed firms, Orbis versus Worldscope



Note: Markups are computed with a constant output elasticity of 0.85 and the means are weighted by turnover.

Source: Market sector from Historical Orbis using Sample B; Worldscope.

Figure A7. Mean markups (weighted by COGS)



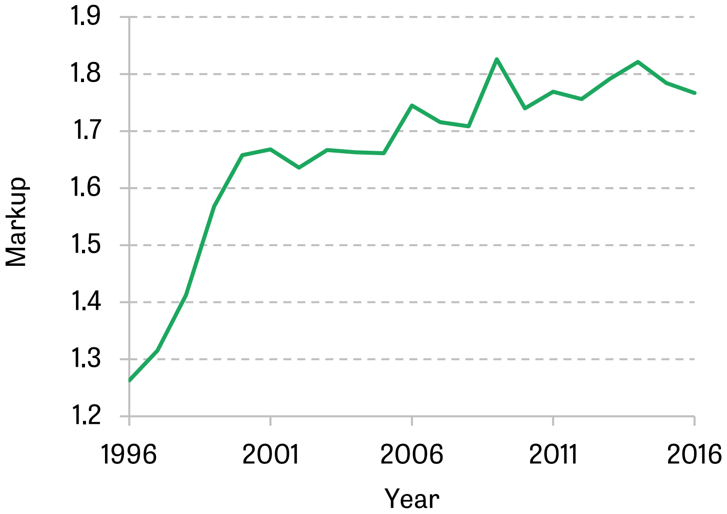
Note: Markups are computed with a constant output elasticity of 0.85. Note that the aggregation uses COGS as weights, whereas in the main text we use turnover weights.

Source: Market sector from Historical Orbis using Sample B.

Figure A7 uses input weights (COGS) to aggregate up to economy-wide markups instead of the output weights that we use in the main text. The qualitative trends are similar with input weights to the main analysis, with a substantial increase that is stronger for the listed firms. However, the magnitudes of the level and increase in markups are smaller.[6]

Recall that the main results (and robustness checks such as Figure A7) drop extreme values of firm-level markups (below 0.5 and above 10). As we noted in Appendix A, this only represents 2.7% of the sample, but we checked many alternative ways of dealing with outliers. For example, Figure A8 winsorises the top and bottom 1% of the firm markup distribution. It is clear that the main results are robust, with (if anything) a larger increase in the aggregate markup over time.

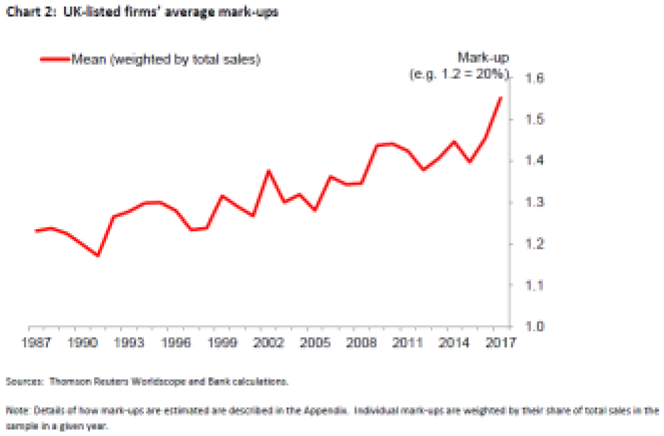
Figure A8. Aggregate markup, all Orbis firms (Winsorised sample)



Note: The markup calculation assumes an output elasticity of 0.85 and markups are weighted by turnover. See Appendix C for markup calculations. Outliers are winsorised using the 1% and 99% quantiles.

Source: Market sector from Historical Orbis using Sample B.

Figure A9. Our findings are consistent with other analyses (for listed firms)



Source: Aquilante et al., 2019.

6 Note that we use the same sample for firm markups dropping the outliers (see Appendix A). This makes little difference to Figure A7, however.

Figure A10. Our findings are consistent with other analyses (for listed firms)

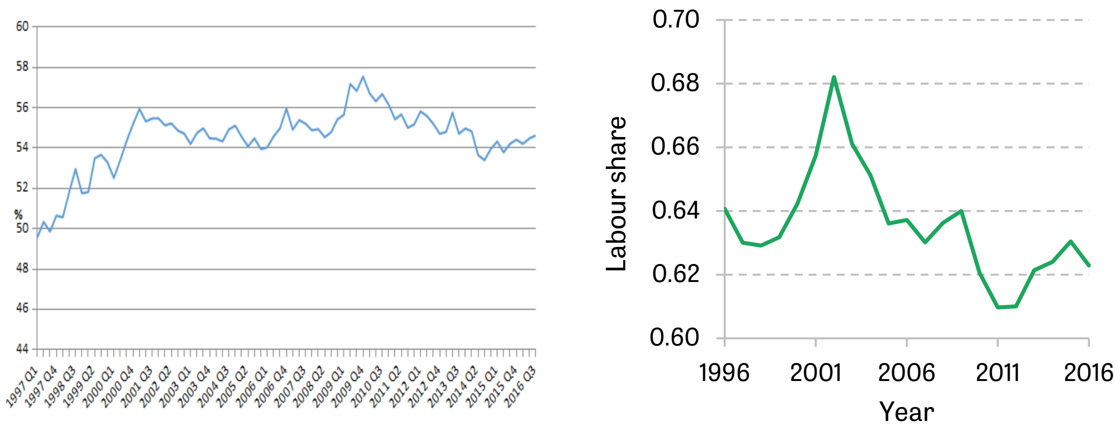
Source: Díez, Leigh and Tambunlertchai, 2018.

Finally, Figures A9 and A10 show that the trends in markup dispersion we have documented are also consistent with other analyses of UK data, although note that these papers only use listed firms.

Labour shares

Starting with the aggregates, the left-hand panel of Figure A11 shows the ONS series over roughly the same period as our HO data, from 1997 onwards. The right-hand panel of the figure gives the labour share trends in HO. Recall that there are large sampling and measurement differences between the administrative data and company accounts. Qualitatively, the trends are not so different from the macro data with a fast rise from the late 1990s through 2002. There is a fall after 2002, however. Nevertheless, the change between 1996 and 2016 is not very large: a fall from 64% to 62%.

Figure A11. Macro UK labour share of GDP: comparison between ONS data and Historical Orbis

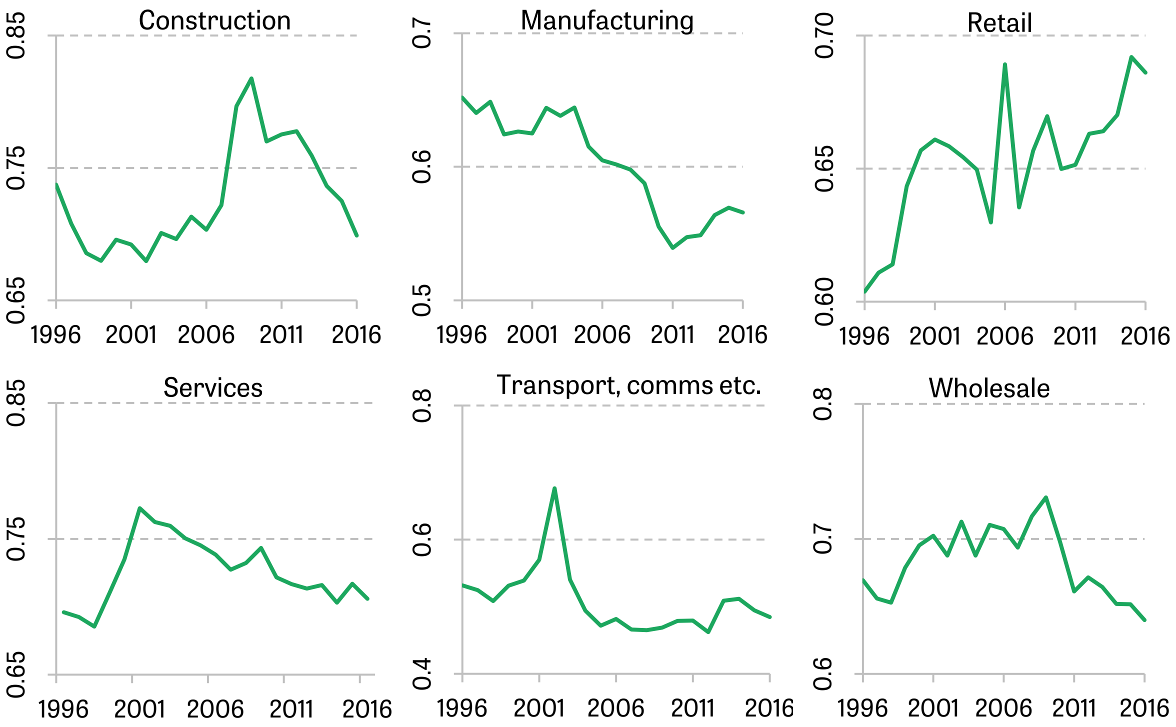


Note: Total wage bill divided by value added.

Source: ONS (macro labour share); market sector from Historical Orbis using Sample A.

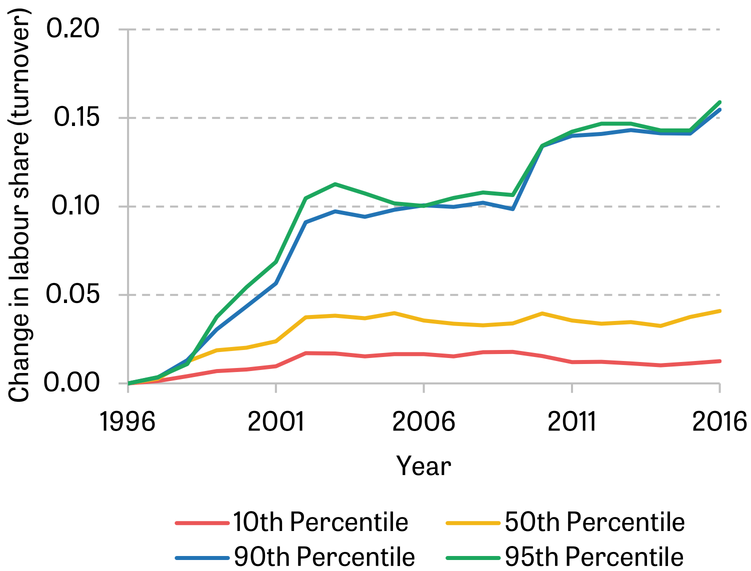
Next, we split the changing labour share by broad sector from HO in Figure A12. The labour share fell most in manufacturing, consistent with administrative data. Services and wholesale had larger falls in the 2000s than other sectors.

Figure A12. Historical Orbis labour shares by broad sector



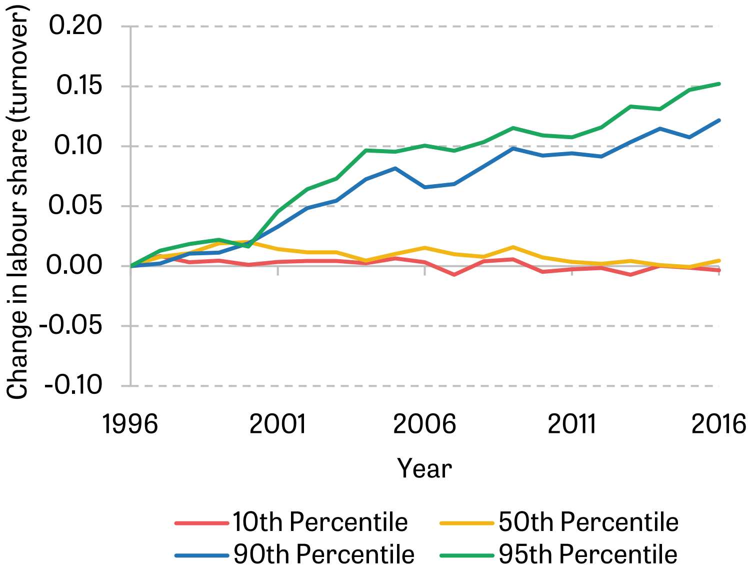
Source: Market sector from Historical Orbis using Sample A.

Figure A13. Changes in unweighted quantiles of the labour share of revenue



Source: Market sector from Historical Orbis using Sample A.

Figure A14. Changes in weighted quantiles of the labour share of revenue



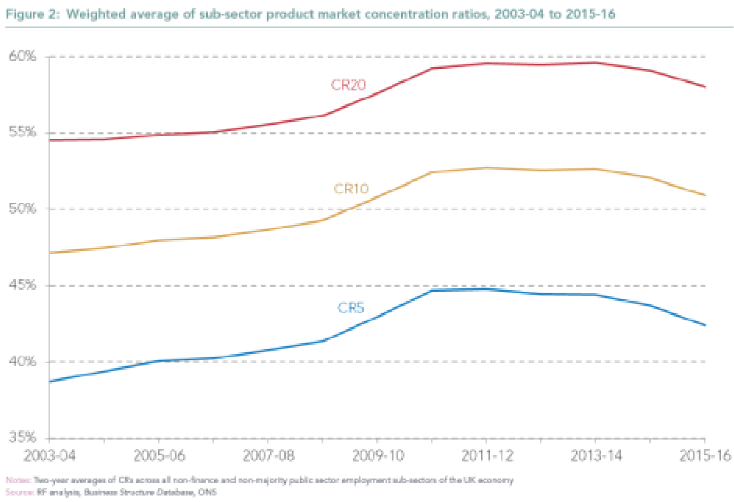
Source: Market sector from Historical Orbis using Sample A.

We look at dispersion of the labour share of revenue in Figures A13 and A14. As with productivity and markups, there appears to be some increase in dispersion in the unweighted and weighted quantiles.

Firm size and concentration

Figures A15 to A18 are analyses of the administrative data from other sources and also show increases in concentration. Figure A19 compares trends in concentration across countries.

Figure A15. Share of industry sales (traditional concentration)

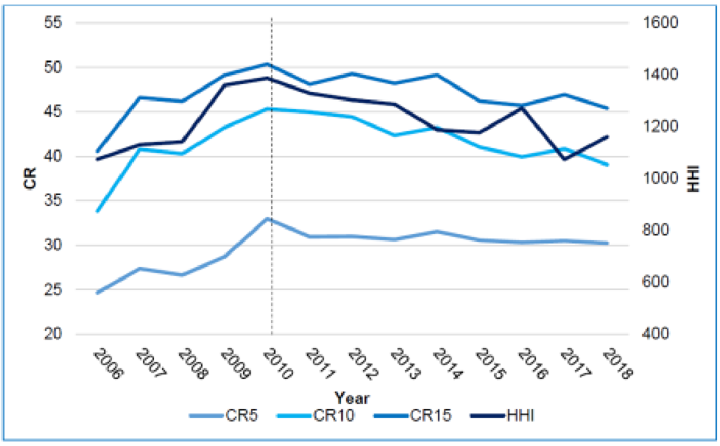


Note: Finance, public sector and wholesale of fuels are all dropped. The figure includes 608 five-digit SIC subsectors.

Source: Bell and Tomlinson (2018) using BSD.

Figure A16. Concentration (BEIS)

Figure 2: Weighted[10] average sector CRs and HHI, UK (2006-2018)

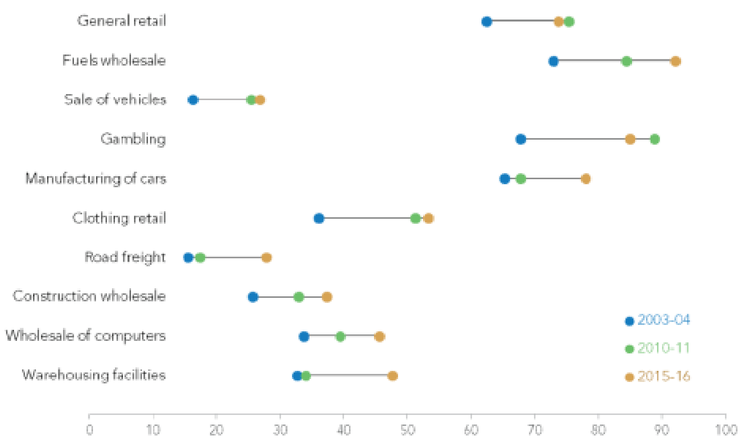


Source: BEIS analysis of the IDBR

Source: BEIS, 2020.

Figure A17. Concentration by sector

Figure 9: Changes in product market concentration, 10 largest sub-sectors with increases bigger than 10ppts



Source: RF analysis, Business Structure Database, ONS

Source: Bell and Tomlinson (2018) using BSD.

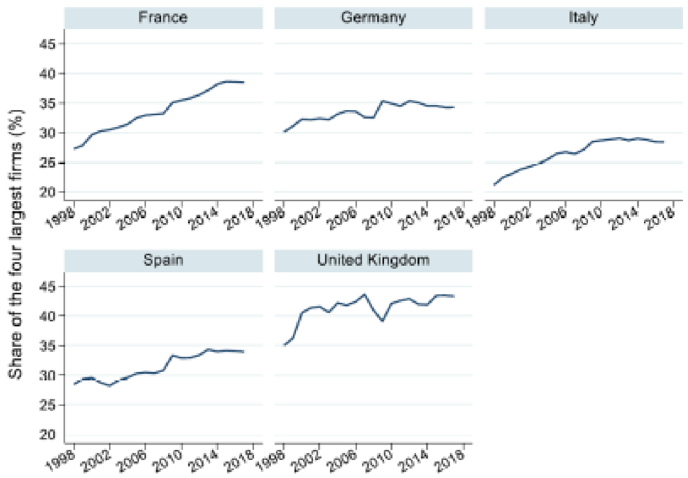
Figure A18. Concentration by industry in the UK

Note: The figure computes for each industry in the UK (at a highly disaggregated level) the share of turnover accruing to the top 5 and top 20 companies and then presents, as a time series, the average across industries in each year.

Source: Bahaj, Key and Piton, 2019.

Figure A19. Trends in industry concentration across countries

Figure 9 *Evolution of average industry concentration in France, Germany, Italy, Spain and the UK, by country, 1998-2017*



Source: Authors' calculations based on Euromonitor International's Passport Industrial database.

Source: Koltay, Lorincz and Valletti, 2020.

Business dynamism

Figures A20 to A22 have administrative data on rates and totals of entry and exit. These show no downward trends, but they are counts, so do not reflect the (un)importance of these firms. A better measure is the share of activity in firms of different ages. Figure A23 shows the share of employment in firms of different age and Figure A24 does the same for turnover. These show the declines in dynamism discussed in the main text.

Figure A20. Birth and death rates

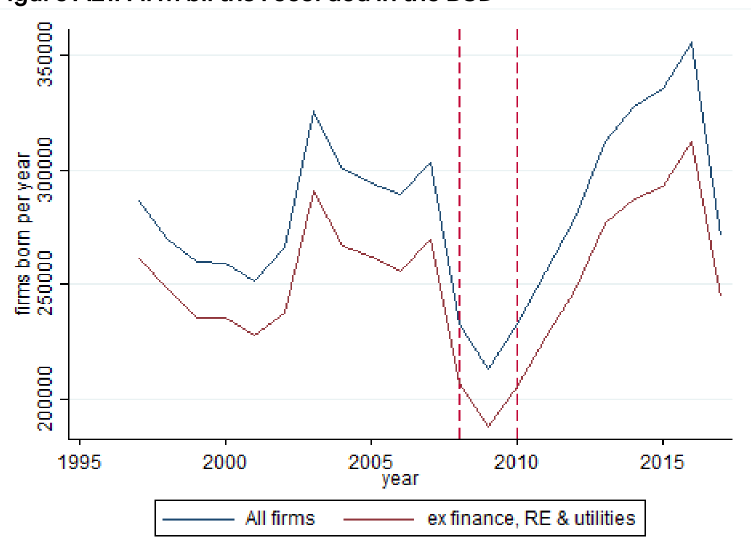
Business birth and death rates, UK

	Births (000s)	Birth Rate	Deaths (000s)	Death Rate
2001	243	12%	212	10%
2002	243	11%	213	10%
2003	267	12%	232	11%
2004	280	13%	244	11%
2005	275	13%	228	10%
2006	256	12%	207	9%
2007	281	12%	224	10%
2008	267	11%	223	10%
2009	236	10%	277	12%
2010	235	10%	249	11%
2011	261	11%	230	10%
2012	270	11%	252	11%
2013	346	14%	238	10%
2014	351	14%	246	10%
2015	383	14%	282	11%
2016	414	15%	281	10%
2017	382	13%	362	12%
2018	381	13%	336	11%

Source: ONS, *Business Demography*, 2019; House of Commons Library calculations

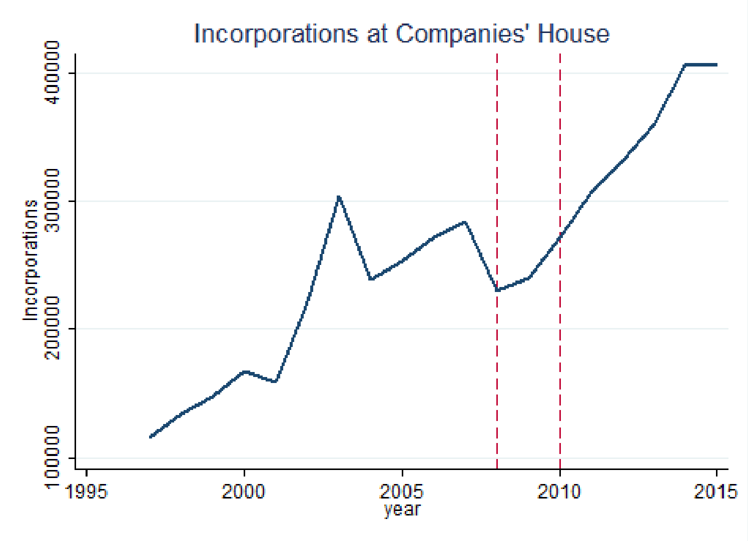
Notes: excludes the non-VAT registered businesses with no employees; Birth rate = New businesses as a % of active businesses; Death rate = Businesses that ceased trading as a % of active businesses

Figure A21. Firm births recorded in the BSD



Source: Davies, 2021.

Figure A22. New incorporations at Companies House



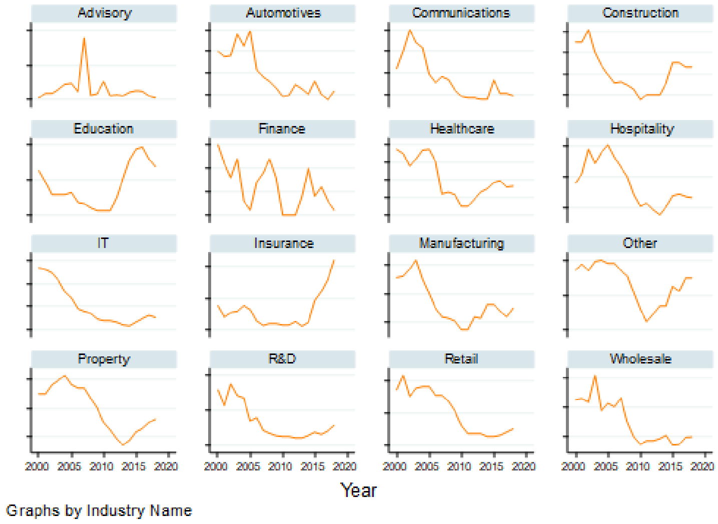
Source: Bahaj, Key and Piton, 2019.

Figure A23. Share of employment by firm age

Note: Initial decline is probably spurious as driven by need to calculate age.

Source: Bahaj, Key and Piton, 2019.

Figure A24. Share of turnover in young firms



Source: Davies, 2021.

Appendix C. Markups and decomposing the labour share

Estimating markups. Take the perspective of an individual cost-minimising firm i. For any variable factor input, we can write the demand for the factor V in the form of a share of revenue:

$$S_i^V = \frac{\theta_i^V \lambda_i^V}{\mu_i} \quad \text{(A1)}$$

where θ_i^V is the elasticity of output with respect to factor V, μ_i is the markup defined as the ratio of product market price to marginal cost and λ_i^V measures input market power which is the ratio of the marginal revenue of the factor to its factor price. This illustrates the idea that we can decompose the factor share into three terms: technology (θ_i^V), product market power (μ_i) and input market power (λ_i^V). Note that perfect competition is the boundary case where the market power terms will be equal to unity, so the output elasticity is equal to the factor share.

To estimate markups in the simplest way, we assume that variable factors are supplied perfectly elastically to the firm so that there is no input market power. We can rewrite equation A1 as

$$\mu_i = \frac{\theta_i^V}{S_i^V} \quad \text{(A2)}$$

To calculate economy-wide markups, we must aggregate the firm-specific markups weighting by the firm's relative size, ω_i:

$$\mu = \sum_i \omega_i \frac{\theta_i^V}{S_i^V} \quad \text{(A3)}$$

Our calculations calibrate an output elasticity with respect to variable inputs of 0.85 (i.e. θ^V = 0.85) and the main text uses firm-year turnover weights (both of these assumptions follow De Loecker, Eeckhout and Unger (2020)). Appendix B also considers input weights instead of output weights (so weighting by the cost of goods sold, COGS, instead of turnover), in which case the aggregate markup expression simplifies further to

$$\mu = \frac{\theta^V}{S^V} \quad \text{(A4)}$$

where the denominator is the aggregate sum of variable costs (which we proxy by COGS) divided by aggregate revenues.

Implied labour shares of value added. To compare our results with aggregate national statistics, we sometimes need to work in terms of value added (GDP). To consider this, we start from the first-order condition for labour inputs for the firm in equation A1, and (for brevity) drop the firm index i:

$$\frac{wL}{PQ} = \frac{\theta^L}{\mu} \quad \text{(A5)}$$

Note that the output elasticity of labour (θ^L) is not equal to the output elasticity on COGS (θ^V). If the production function takes the Cobb–Douglas form, $\Omega(K^{\theta^K}(L+M)^{\theta^V})$, where M = intermediate inputs and $\Omega > 0$ is a Hicks neutral efficiency term, the elasticity is

$$\theta^L = \frac{L}{L+M}\theta^V \tag{A6}$$

To back out the implied aggregate labour share ($\widetilde{LS}$), we again fix the technology parameters across firms, and aggregate using revenue shares to obtain (using firm i and time t)

$$\widetilde{LS}_t = \theta^L \sum_i \frac{R_{it}}{R_t}\frac{E(V)_{it}}{\theta^V R_{it}} \tag{A7}$$

$$= \frac{\theta^L}{\theta^V}\frac{COGS_t}{R_t} \tag{A8}$$

where V is total variable costs (COGS).

Now to convert the labour share in terms of gross output to one in value added, we define $a = \frac{VA}{R}$ in the macro data in a base year zero. This means that the implied (value-added-based) labour share can be computed as

$$S_{VA}^L = \frac{\theta^L}{a\theta^V}\frac{COGS_t}{R_t} \tag{A9}$$

This suggests a normalisation in year $t = 0$: choose $\frac{\theta^L}{a\theta^V}$ such that $\widetilde{LS}_{t=0} = LS_{t=0}$. The difference between the actual labour share and the implied labour share is a manifestation of three (aggregated) terms:

1. technological change in both the mix of labour and materials costs in COGS (L and M in V) and the parameters;
2. change in the value-added–gross-output conversion rate a;
3. any model mis-specifications, or 'wedges' in the first-order conditions such as
   - substitution between labour and materials;
   - adjustment costs in labour;
   - imperfect competition in the labour market (e.g. monopsony).

Labour shares over longer periods. We want to extrapolate over longer periods than just that which is available for our markup data from company accounts. In particular, we will look over the 1981–2019 period where we can get consistent aggregate data on key items (see Teichgraeber and Van Reenen (2021)).

We start from the compound growth (g) formula for number of years n between starting value, Y_0, and ending value, Y_1:

$$g = \left(\frac{Y_1}{Y_0}\right)^{\frac{1}{n}} - 1$$

In Orbis, we use $Y = COGS/R$ between 1996 and 2016 to calculate g. So

$$g = \left(\frac{Y_{2016}}{Y_{1996}}\right)^{\frac{1}{20}} - 1$$

$$g = \left(\frac{0.6859}{0.7482}\right)^{\frac{1}{20}} - 1 = -0.004337$$

We find a fall in the COGS to revenue ratio of about 0.44 percentage points a year. We want to estimate the change in $COGS/R$ between 1981 and 2019 to match the aggregate data. Rewriting the compound growth formula,

$$Y_1 = Y_0(1+g)^n$$

$$Y_{2019} = Y_{1996}(1+g)^{23}$$

$$Y_{1981} = Y_{1996}(1+g)^{-15}$$

So, since we find the COGS to revenue ratio in HO is 74.82% in 1996 ($Y_{1996} = \frac{COGS}{R} = 0.7482$),

$$\begin{aligned} Y_{2019} - Y_{1981} &= Y_{1996}[(1+g)^{23} - (1+g)^{-15}] \\ &= 0.7482[(1-0.004337)^{23} - (1-0.004337)^{-15}] \\ &= 0.7482[(0.9957)^{23} - (0.9957)^{-15}] \\ &= -0.1216 \end{aligned}$$

Hence the COGS to revenue ratio is predicted to fall by 12.2 percentage points between 1981 and 2019. The implied fall in the labour share of GDP will depend on the calibration for the technology parameters:

$$S_{VA}^L = \frac{\theta^L}{\alpha\theta^V}\left(\frac{COGS}{R}\right) = b\left(\frac{COGS}{R}\right)$$

In 1996, $S_{VA}^L = 0.4338$ (see Figure 14) and hence b = 0.4338/0.7482 = 0.5798. The predicted fall in the labour share of GDP is

$$\begin{aligned} \Delta\tilde{S}_{VA}^L &= 0.5798\,\Delta\left(\frac{COGS}{R}\right) \\ &= 0.5798(-0.1216) \\ &= -0.071 \end{aligned}$$

So we predict a fall of 7.1 percentage points in the labour share of GDP, compared with an actual fall of 3.5 percentage points. We discuss reasons why markup trends over-predict the fall of the labour share in the main text. Broadly, this could be due to falls in monopsony power, offsetting technical change and/or aggregation issues.

References

Aquilante, T., Chowla, S., Dacic, N., Haldane, A., Masolo, R., Schneider, P., Seneca, M., and Tatomir, S. (2019), 'Market Power and Monetary Policy', Bank of England Staff Working Paper 798.

Bahaj, S., Foulis, A., Gal, P., and Pinter, G. (2017), 'Productive and Unproductive Leverage', Bank of England.

Bahaj, S., Key, T., and Piton, S. (2019), 'Evidence on UK Business Dynamism', Bank of England.

BEIS (2020), 'Business Population Estimates for the UK and Regions 2020', Department for Business, Energy and Industrial Strategy.

Bell, T., and Tomlinson, D. (2018), 'Is Everybody Concentrating? Recent Trends in Product and Labour Market Concentration in the UK', Resolution Foundation, Briefing.

Davies, R. (2021), 'Trends in UK Firms' Dynamics', mimeo, Centre for Economic Performance, LSE.

De Loecker, J. and Eeckhout, J. (2018), 'Global Market Power', National Bureau of Economic Research, Working Paper 24768.

De Loecker, J., Eeckhout, J., and Unger, G. (2020), 'The Rise of Market Power and the Macroeconomic Implications', *Quarterly Journal of Economics*, 135(2), 561–644.

Díez, F. J., Leigh, D., and Tambunlertchai, S. (2018), 'Global Market Power and Its Macroeconomic Implications', International Monetary Fund, Working Paper 18/137.

Koltay, G., Lorincz, S., and Valletti, T. (2020), 'Concentration and Competition: Evidence from Europe and Implications for Policy', Chief Economist Team, DG-COMP, European Commission.

Office for National Statistics (2019), 'Firm-Level Labour Productivity Measures from the Annual Business Survey, Great Britain: 2017'.

Office for National Statistics (2020), 'Firm-Level Labour Productivity Measures from the Annual Business Survey, Great Britain: 1998 to 2018', https://www.ons.gov.uk/economy/economicoutputandproductivity/productivitymeasures/datasets/firmlevellabourproductivityestimatesfromtheannualbusinesssurveyabssummarystatistics.

Oliveira-Cunha, J., Kozler, J., Shah, P., Thwaites, G., and Valero, A. (2021), 'Business Time: How Ready Are UK Firms for the Decisive Decade?', The Economy 2030 Inquiry, Resolution Foundation.

Teichgraeber, A., and Van Reenen, J. (2021), 'Have Productivity and Pay Decoupled in the UK?', *International Productivity Monitor*, 41, Fall, 31–60.