Market Power in the UK

Kyung In Hwang   Anthony Savagar   Joel Kariel *

September 13, 2022

Abstract

We study product market power in the UK using administrative data from the UK business survey. Our data covers 1998-2014 with 46,000 firms per year accounting for 80% of UK output. We estimate firm-level markups, and present results on the aggregate and sectoral trends. We show evidence of rising markups, and increasing markup dispersion. We also show that markups and productivity are negatively related.

JEL: D4, D2, L1, E6.

Keywords: Markups; Market Power; Production Function Estimation; Productivity, UK Economy.

Word Count: 6,825

*a.savagar@kent.ac.uk. This research is funded under ESRC project reference ES/S000089/1. Thanks to the following people for comments: Chiara Criscuolo, Sara Calligaris, Tommasso Aquilante, Patrick Schneider, Jonathan Haskel, Jan de Loecker, Steve Davies, Miguel Leon Ledesma. Thanks to Yannis Galanakis for research assistance.

1 Introduction

There is evidence of rising market power in many advanced economies (Calligaris, Criscuolo, and Marcolin 2018; Diez, Leigh, and Tambunlertchai 2018; De Loecker and Eeckhout 2021). Documenting market power is important because market power affects efficiency which determines welfare (Baqaee and Farhi 2020), and it can explain trends in other aggregate variables such as productivity, investment and labour income shares in output (Autor, Dorn, Katz, Patterson, and Van Reenen 2017; Eggertsson, Robbins, and Wold 2021). In turn, up-to-date measures of market power directly affect policymaker decisions and public understanding.

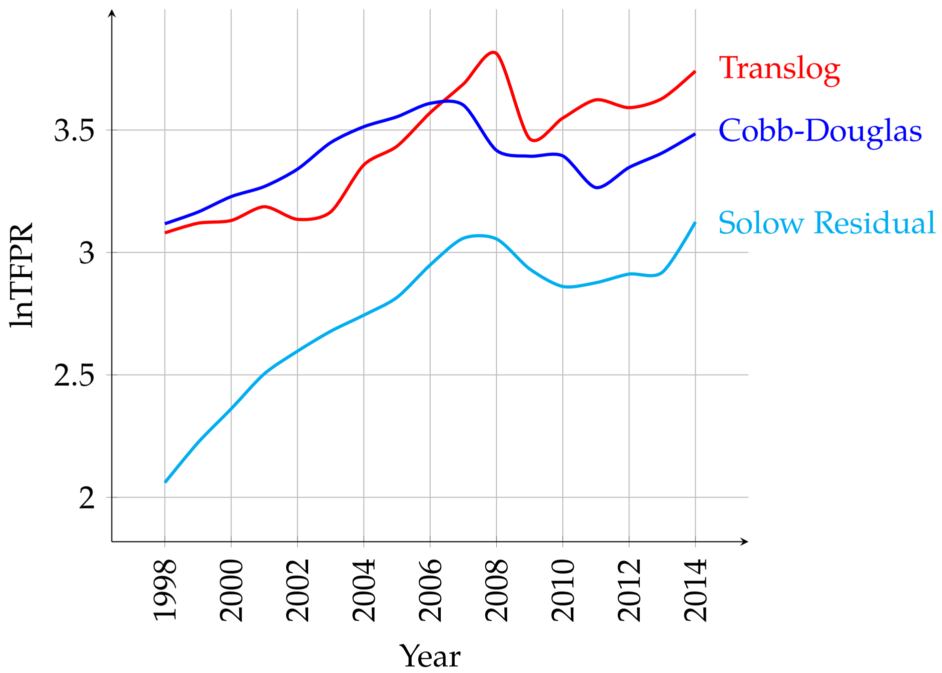
In this paper we document the state of market power in the UK economy by measuring firm-level price markups. Additionally, we study the relationship between aggregate market power and total factor productivity. Beyond the direct UK policy benefits, it is helpful to provide robust measurement on markups to add to the current global debate. The emerging consensus across many studies is that markups and other measures of market power are rising in advanced economies, but many ambiguities exist. Not all advanced economies observe rising market power; for example, the Netherlands has experienced stable markups from 2006 - 2016 (van Heuvelen, Bettendorf, and Meijerink 2021). Different indicators of market power such as concentration, profits, and markups, can contradict each other both theoretically (Syverson 2019) and empirically (Gutiérrez and Philippon 2020). Furthermore, there are many methodological pitfalls in the process of acquiring markups (Raval 2020; Bond, Hashemi, Kaplan, and Zoch 2021; De Ridder, Grassi, and Morzenti 2021; De Loecker 2021). On top of this, the dynamics behind rising markups differ. Spain has rising markups, but driven by small and unproductive firms (García-Perea, Lacuesta, and Roldan-Blanco 2021), whereas the US has rising markups driven by high-markup firms – so-called superstar firms (Autor, Dorn, Katz, Patterson, and Van Reenen 2020).

Given this complex background to the current market power literature, our main contribution is to robustly document markups in the UK economy. That means using secure firm-level data from the ONS (rather than proprietary datasets of larger firms), analysing the effects of different assumptions under-pinning markup estimates, and breaking-down aggregate trends into sub-sector trends and distributional trends.

Our results show that aggregate price markups in the UK are rising, and this is robust to many methodological choices. Markup levels on the other hand are more sensitive to underlying assumptions. The rising trend in aggregate markups comes from firms at the 90th percentile pulling-away, which conforms to the superstar firms hypothesis and similar empirical evidence for the US. We also explain that the UK economy is uniquely services-dominated and this propagates to aggregate results. Much of what we observe in the aggregate trends and levels of markups is driven by the behaviour of firms in the services sector. Finally, we link the rise in UK markups to low levels of productivity in the UK since the financial crisis.

The UK suffered a sharp fall in productivity following the 2007 financial crisis, and a slow recovery (Figure 1). This 'productivity puzzle' is usually expressed as a labour productivity puzzle, but our results in Figure 1 confirm it is also a Revenue Total Factor Productivity (TFPR) puzzle. The UK productivity puzzle has many potential explanations including globalisation, technology, reallocation, credit constraints on small firms, the omission and slowdown in intangible investment, among others (Wales, Black, Dolby, and Awano 2018; Harris and Moffat 2017; Criscuolo, Haskel, and R. Martin 2004; Miller and Barnett 2015; Crawford, Jin, and Simpson 2013; Goodridge, Haskel, and Wallis 2013). In this paper we show that the UK's productivity puzzle coincides with growth in market power.

Figure 1: Aggregate TFPR



Note: Weighted averages of firm-level TFPR from production function estimation with translog and Cobb-Douglas. Solow Residual is weighted average of log-linearised Cobb-Douglas production function with constant returns to scale.

To document trends markups we combine UK business survey data over the period 1998 - 2014 with production function estimation. This long time horizon allows us to estimate production functions at the firm level, and covers an era of productivity growth, slowdown, and mild recovery. We use control function approaches to estimate production functions (Olley and Pakes 1996; Levinsohn and Petrin 2003; Ackerberg, Caves, and Frazer 2015; Gandhi, Navarro, and Rivers 2018; Demirer 2020). Given elasticities from production function estimation, we obtain markups and TFPR. We use the Annual Respondents Database X (ARDx), which is a representative UK business survey data that covers over 46,000 firms each year. It samples smaller firms and covers the universe of larger firms.

Related Literature

De Loecker, Eeckhout, and Unger (2020) document rising market power in the US economy which has led to widespread interest in the topic across many countries. The literature has been widely surveyed, and we refer the reader to Syverson (2019) for comprehensive coverage. We will discuss work that is relevant for the UK context.

To the best of our knowledge, there is no recent research on the UK economy comprehensively documenting markups using recent methodological techniques in markup estimation. Aquilante, Chowla, Dacic, Haldane, Masolo, Schneider, Seneca, and Tatomir (2019) present initial analysis on UK market power using these techniques, but with an interest in monetary policy effects and with a small sample of publicly-listed firms in Thompson Reuters Worldscope data. They find modest increases in concentration and larger increases in markups. Several papers use an earlier version of the ARDx to measure proxies of market power, typically as control variables in broader applied studies of R&D, competition and investment. Aghion, Blundell, Griffith, Howitt, and Prantl (2004) and Griffith, Harrison, and Simpson (2010) use average profitability to proxy market power using the ARD dataset, whilst Griffith, Harrison, and Macartney (2007) use a similar measure nut with OECD data. Haskel and C. Martin (1992) and Griffith (2001) construct price-cost margins, with the latter constructing a Lerner Index under the assumption of constant *variable* returns to scale (*i.e.* that average variable cost proxies marginal cost). Their markup averages around 18%. This is close to estimates by Martins, Scarpetta, and Pilat (1996) of 16%, which follows the methodology of Roeger (1995) to directly estimate the Lerner Index with a first-difference equation. Haskel, C. Martin, and Small (1995) implement the Hall (1988) estimation approach which tests the equality of price and marginal cost under the assumption of constant returns to scale. They find an average markup *in manufacturing* from 1968 - 1989 of 2.0 and a positive correlation with industry concentration.

The most up-to-date estimates on UK markups from the Competition & Markets Authority (CMA) find an increase over the last two decades from around 1.22 to 1.34, which is not too far from the translog estimates we present in this paper Competition and Markets Authority (2022). The CMA report uses data on around 4,000 large companies (with more than 250 employees) from the FAME database. In contrast, we provide estimates from a dataset with over ten times as many firms, including a representative sample of medium- and small-sized businesses. Furthermore, the CMA analysis uses the 'cost share approach' to estimating markups, which assumes firms have constant returns to scale and the cost minimisation first order conditions hold for all inputs in a given year (even those with adjustment costs, e.g. capital). In contrast, we use the production function approach, despite recent criticisms (Bond, Hashemi, Kaplan, and Zoch 2021).

A number of recent papers analyse TFP in the UK (Goodridge, Haskel, and Wallis 2016; Harris and Moffat 2017; Schneider 2018). These papers decompose the UK productivity slowdown using firm-level data, and characterise the reallocation across firms and between labour and capital, the role of firm size, and the distribution across industries. Recent work by Jacob and Mion (2020) explains the UK's productivity puzzle (measured in TFPR terms) can be attributed to falling demand and falling Quantity TFP (TFPQ). Their study also uses the ARDx, and their main focus is manufacturing firms which can be merged with Prodcom data on prices to get an accurate measure of TFPQ.

There has been recent debate about the estimation of production functions using the current methods in the literature. One particular problem that biases estimates of output elasticities (and thus markups) is the lack of availability of price data, requiring the use of revenue as the dependent variable (Bond, Hashemi, Kaplan, and Zoch 2021; De Loecker 2021). However, emerging work suggests markup trends using the production function approach are more reliable than levels (De Ridder, Grassi, and Morzenti 2021). We predominantly focus on markup trends, but also report levels in the Appendix.

2 Data

To estimate UK markups and TFPR at the firm level, we use the ARDx dataset accessible from the UK Data Service and complied by the ONS (the Office of National Statistics). The ARDx is a time series dataset of UK firms, complied from three sources. From 1998

- 2008 it contains information from the Annual Business Inquiry (ABI). Subsequently it combines Annual Business Survey (ABS) and the Business Register Employment Survey (BRES). All three surveys are administrative surveys of UK businesses. The ARDx is the largest ONS dataset with 62,000 businesses and over 600 variables. In addition, firms' representatives are legally required to complete it (Gobey and Matikonis 2021).

The ABI/ABS contains financial information on around 62,000 UK businesses each year (J. Martin and Baybutt 2022), and the BRES provides employment-related variables. These data cover the non-financial business economy of the UK, with excellent coverage. The ABS typically has a response rate of around 80%, which therefore includes in excess of 80% of turnover of sampled units, since the ONS puts more effort into ensuring larger businesses respond to the survey (J. Martin and Baybutt 2022).

The ARDx dataset is a time series version of the annual ARD database which contains information on UK firms. As a firm-level panel dataset, the ARDx is designed for external researchers. In the ARDx data, the smallest unit is a local unit (LU) also known as an establishment, for example a plant, shop or warehouse. Typically researchers use the reporting unit (RU) which is the smallest unit that can respond to the questions on the survey, since an individual LU, such as a warehouse, is unlikely to collect the data necessary to respond to the questions. Notably 95% of reporting units consist of one local unit. That is, most reporting units are single establishment firms.

The ARDx runs from 1998 to 2014 and covers on average 46,000 firms per year once it is compiled, as some observations are dropped in the process. It provides production data on revenue, value added, capital expenditure, material purchases and employment. The sampling framework of the ARDx is the Inter-Departmental Business Register (IDBR) which is a population of UK firms. More specifically, the ARDx data includes all UK large firms and contains a stratified (by industry, region and employment size) sample of small businesses with less than 250 employees. As a result, the ARDx consists of a full balanced panel of information on large firms and a unbalanced panel of small firms. In order to provide some continuity, the majority of medium-sized firms are sampled on a two-year rotation, before being dropped from the sample (J. Martin and Baybutt 2022).

The ARDx is a good data source to estimate markups and TFPR in the UK for two reasons. First, the ARDx provides a micro panel dataset of production information with the widest coverage of UK firms. Even though the ARDx dataset contains a representative sample of UK businesses, particularly for those which have employees with less than 250, the ARDx is originally designed to construct UK national accounts. According to Haldane (2017), it accounts for 80% of UK GDP. Second, it covers a comprehensive of UK sectors, both manufacturing sectors and service sectors. Further details on the construction of capital stock, variable deflators, and data cleaning are contained in the Appendix.

3 Methodology

3.1 Methodological Issues on Production Estimation

It is important to understand how the methodology for estimating markups relates to the methodology for estimating production functions. We estimate markups by the *production approach*. This recent literature on the production approach to markup estimation relies on a long-established literature to estimate production functions known as the *control function approach*.

1. Production approach to markup estimation: The methodology for deriving firm-level markups defined as the ratio of output elasticity with respect to a flexible input (e.g., labour, material) to its revenue share. To do this we first need to estimate the output elasticity from production function estimation. The production approach to markup estimation has been widely implemented since De Loecker and Warzynski (2012). It adapts Hall (1988) to the firm level.

2. Control function approach to production function estimation: The methodology for estimating production functions and therefore output elasticities of factors.[1]

[1]The control function approach is one of three broad methodologies used to estimate production functions. The others are fixed effects (FE) and instrumental variables (IV).

> The control function approach is developed in a series of papers by Olley and Pakes (1996), Levinsohn and Petrin (2003), Wooldridge (2009), Ackerberg, Caves, and Frazer (2015), and Gandhi, Navarro, and Rivers (2018).

Therefore, measures of markups depend on production function estimation. We produce two different estimates of production functions: Cobb-Douglas and translog. The key difference is that the former returns *constant* coefficients (i.e. output elasticities), while the latter estimates *firm- and time-varying* elasticities. To obtain markups, we combine these coefficients with firm-level information on flexible input shares.

In our analysis, we use material input to control for unobserved productivity following Levinsohn and Petrin (2003) and Ackerberg, Caves, and Frazer (2015) production function estimation. We also use material input as a flexible variable to identify markups by the production approach. For comparison, De Loecker and Warzynski (2012) use material input to identify unobserved productivity using Ackerberg, Caves, and Frazer (2015) production function estimation to estimate output elasticities, but they use labour input as the flexible input to identify markups. Raval (2020) replicates their production function estimation approach, but tries both labour and materials as the flexible input to acquire the markup.

3.1.1 Material for Flexible Input

A flexible input refers to an input in the production function that is chosen by the firm without any intertemporal consideration, such as materials or labour, but not capital. In the production approach to markup estimation, the markup is calculated as the ratio of any flexible input's output elasticity over its expenditure share in total revenue. Typically papers use the labour input, materials input, or a composite of the two, as the variable input. Raval (2020) compares the alternatives. We use material input as the flexible input to identify markups. We provide estimates using labour as the variable input in the Appendix. The labour input in our dataset is recorded as number of employees, not hours worked. It is significantly less variable than materials input.[2]

There is one potential limitation with using the materials input to estimate markups when applying this method to the services sector. The materials share is significantly lower in services than other sectors, and some services firms use very little or no materials. For example, the average materials expenditure in revenue is almost 57% in manufacturing, but just 41% in services. Furthermore there is a big skew in services, with the tenth percentile of firms spending below 9% of revenue on materials. The corresponding value is around 30% in manufacturing. Further descriptive statistics are contained in Table 6 in the Appendix. The concern is that services use other flexible inputs more intensely, and the variation of those inputs would be more informative for estimating the markup. For this reason, we also provide estimates where the labour input is considered flexible. Further details on this will be discussed in Section 3.2.

3.1.2 Production Functions and Dependent Variable

To estimate production functions, we need to choose the functional form (Cobb-Douglas or translog) and the measure of the dependent variable (gross output or gross value-added). Our baseline estimates use both Cobb-Douglas and translog production functions with gross output. We choose gross output as we use materials as the flexible input in estimating markups.[3] We include estimates with gross value-added in the Appendix D.1, where labour is instead the flexible input.

When comparing markups over time, these two production functions have one significant difference: the constancy of output elasticities. Consider the equation for markups, which we derive from cost minimisation in the next section:

$$\mu_{it} = \theta_M \left(\alpha_{it}^{M}\right)^{-1},$$

[2]Some literature criticises the sensitivity of markup estimates to choice of free variable. Raval (2020) shows that markups estimated using labour input as the free variable can be negatively correlated with markups estimated using the material input. In De Loecker and Eeckhout (2017), they use an accounting variable known as cost of goods sold (COGS) as the free variable, but Traina (2018) shows the results are sensitive to a broader definition of variable costs which leads to a different markup trend. Both authors use financial accounts data available from Compustat.

[3]If we chose gross value-added, which is gross output less material costs, we would not obtain an estimate of the materials output elasticity, and hence could not use materials to compute markups.

where θ_M is the output elasticity with respect to the material input and α_{it}^M is the expenditure share of materials in total revenue. The elasticity is different under the two production functions:

1. **Cobb-Douglas**: θ_M does not vary across firms or over the estimation period. This embeds the assumption that all firms have production functions with the same shape, and differ only in idiosyncratic productivity shocks. The response of output to input adjustments is constant over time. Thus, it is simply changes in the materials input share that drive changes in the estimated markup over time and across firms.

2. **Translog**: θ_M can vary across firms and over time. As shown in Demirer (2020), this can have important implications for the evolution of markups over time, if the flexible elasticity is not constant.

3.1.3 Control Function Approach

The key identification issue for estimating the production function is that unobserved firm-specific productivity enters into the first-order condition that determines firm's optimal choice of input. This gives rise to an endogeneity issue that makes it hard to estimate unbiased estimates for the output elasticity. To address the endogeneity issue, Levinsohn and Petrin (2003) use the material input as a control variable that proxies for the unobserved productivity. In estimating a gross output production, even Levinsohn and Petrin (2003) has an identification issue. The control function approach exploits lagged input variables as a source of identification variation, meaning that output elasticities with respect to the inputs can be identified by exploiting the adjustment frictions. Thus, output elasticities with respect to free variables cannot be estimated. Therefore our baseline estimates use the Ackerberg, Caves, and Frazer (2015) approach, which sidesteps this identification problem by adjusting the timing assumptions, so labour is chosen before material inputs, but after investment.

3.2 Estimation Approach

Consider a firm with a production function $Y_{it} = F_{it}(L_{it}, K_{it}, M_{it}, A_{it})$, where Y_{it} is gross output for firm i and time t, L_{it} is labour, K_{it} is capital, M_{it} is material, and A_{it} is firm-specific productivity. The firm's cost minimization problem with respect to material shows that a firm employs materials M_{it} until the marginal revenue product of material input is driven down to equal the material input price:

$$P_{it}^{M} = \lambda_{it} \frac{\partial F_{it}}{\partial M_{it}}$$

where P_{it}^{M} is the material price, $\partial F_{it} / \partial M_{it}$ is the marginal product of materials, and λ_{it} is the Lagrange multiplier which is the marginal cost (which equals marginal revenue by profit maximisation). We define the markup as the price divided by marginal cost: $\mu_{it} \equiv \frac{P_{it}}{\lambda_{it}}$. The above condition yields

$$\mu_{it} = \theta_{M} (\alpha_{it}^{M})^{-1}$$

where θ_M ($= \frac{\ln \partial F_{it}}{\ln \partial M_{it}}$) is the output elasticity with respect to the material input and α_{it}^{M} ($= \frac{P_{it}^{M} \cdot M_{it}}{P_{it} \cdot Y_{it}}$) is the expenditure share of material inputs in total revenue. When considering a structural value-added production function, as we do for comparison in Section 4.1.1 and in Appendix D.1, we follow De Loecker and Scott (2016) and estimate the markup with

$$\mu_{it} = \frac{1}{\left(\mu_{it}^{X}\right)^{-1} + \alpha_{it}^{M}}$$

where μ_{it}^{X} is the markup if we treat input X (typically labour) as the flexible input in the original markup formula. The full derivation for this is in Appendix D. Estimates of the markup following this approach are contained in Appendix D.1.

To obtain the output elasticity of the material input, we estimate a production function. TFPR is the residual of the production function. We use the control function approach to estimating production functions, introduced by Levinsohn and Petrin (2003), and then extended by Ackerberg, Caves, and Frazer (2015). We estimate the following production function

$$y_{it} = \theta_m m_{it} + \theta_l l_{it} + \theta_k k_{it} + \omega_{it} + \epsilon_{it}$$

where the lower case letters denote logs. ω_{it} is unobserved productivity of firm i in period t which is anticipated by the firm. ϵ_{it} is an ex-post shock. The main econometric issue is to control for unobserved ω_{it}, which is correlated with input choices, and therefore causes the so-called simultaneity problem. Ackerberg, Caves, and Frazer (2015) show that firm cost minimization implies materials choice as a function of productivity and capital (a state variable): $m_{it} = m_{it}(\omega_{it}, k_{it}, l_{it})$. Then, assuming this function is monotonically increasing in ω_{it} and is invertible, productivity can be written as a function of materials and capital: $\omega_{it} = h(m_{it}, k_{it}, l_{it})$. Therefore we can proxy unobserved productivity with a function of observable variables. If we substitute this proxy into the original gross output production function we have

$$y_{it} = \theta_m m_{it} + \theta_l l_{it} + \theta_k k_{it} + h(m_{it}, k_{it}, l_{it}) + \epsilon_{it} = \Phi(m_{it}, k_{it}, l_{it}) + \epsilon_{it} \tag{1}$$

where $\Phi(m_{it}, k_{it}, l_{it}) \equiv \theta_m m_{it} + \theta_l l_{it} + \theta_k k_{it} + h(m_{it}, k_{it}, l_{it})$. Since $h(\cdot)$ is unknown, we run a partially non-parametric regression to obtain $\hat{\Phi}_{it}$.

The second major assumption of Ackerberg, Caves, and Frazer (2015) is that productivity, ω_{it}, evolves following a first-order Markov process. That is, all information about current productivity can be inferred from productivity last period. Formally,

$$\omega_{it} = E(\omega_{it} \mid \omega_{i,t-1}) + \xi_{it} \equiv g(\omega_{i,t-1}) + \xi_{it} \tag{2}$$

where ξ_{it} is the innovation in productivity which is uncorrelated with ω_{it}. Substituting this into the production function yields

$$y_{it} = \theta_m m_{it} + \theta_l l_{it} + \theta_k k_{it} + g(\hat{\Phi}_{i,t-1} - \theta_m m_{i,t-1} - \theta_k l_{i,t-1} - \theta_k k_{i,t-1}) + e_{it} \tag{3}$$

where $e_{it} = \xi_{it} + \epsilon_{it}$. Using a GMM estimator, the moment conditions $E[e_{it} k_{it}] = 0$,

$E[e_{it}l_{i,t-1}] = 0$ and $E[e_{it}m_{i,t-1}] = 0$, yield consistent estimates of $\hat{\theta}_m$ and $\hat{\theta}_l$. We use $\hat{\theta}_m$, the output elasticity with respect to the material, to estimate markups. Finally, log productivity (TFPR) is calculated by

$$\hat{\omega}_{it} = \hat{y}_{it} - (\hat{\theta}_m m_{it} + \hat{\theta}_l l_{it} + \hat{\theta}_k k_{it})$$

In order to obtain firm- and time-varying coefficients, we also estimate a translog production function, as in De Loecker and Warzynski (2012)

$$y_{it} = \beta_0+\beta_k k_{it}+\beta_l l_{it}+\beta_m m_{it}+\beta_{ll} l_{it}^2+\beta_{kk} k_{it}^2+\beta_{mm} m_{it}^2+\beta_{lk} l_{it} k_{it}+\beta_{lm} l_{it} m_{it}+\beta_{mk} m_{it} k_{it}+\epsilon_{it}.$$

We then follow the same procedure as before to estimate the coefficients and productivity process. The time- and firm-specific output elasticities for the materials input are easily computed:

$$\widehat{\theta}_{it}^m = \widehat{\beta}_m + 2\widehat{\beta}_{mm} m_{it} + \widehat{\beta}_{lm} l_{it} + \widehat{\beta}_{mk} k_{it}.$$

4 Results

Our baseline markup estimates follow Ackerberg, Caves, and Frazer (2015) with gross-output, with either a Cobb-Douglas or translog production function. Aggregation from firm-level markups uses value-added weights, unless otherwise reported for comparison purposes. Alternative estimates using other control function approaches or value-added as the dependent variable are included in Appendix D.1.

4.1 Markup Results

Figure 2 presents the aggregate markup trend over time. The markup has risen from 1998 to 2014. Although the two production functions tell slightly different stories, the main result is the same. There is a steady rise in the markup until the financial crisis, followed by a plateau or fall. The markup began to increase again around 2011. The overall rise is around 18% for translog, and 100% for Cobb-Douglas. The levels are included in Figure 9 in the Appendix. The 1998 estimates are 1.32 and 1.28 for each production function, and rise to 1.56 and 2.60 by 2014, respectively.

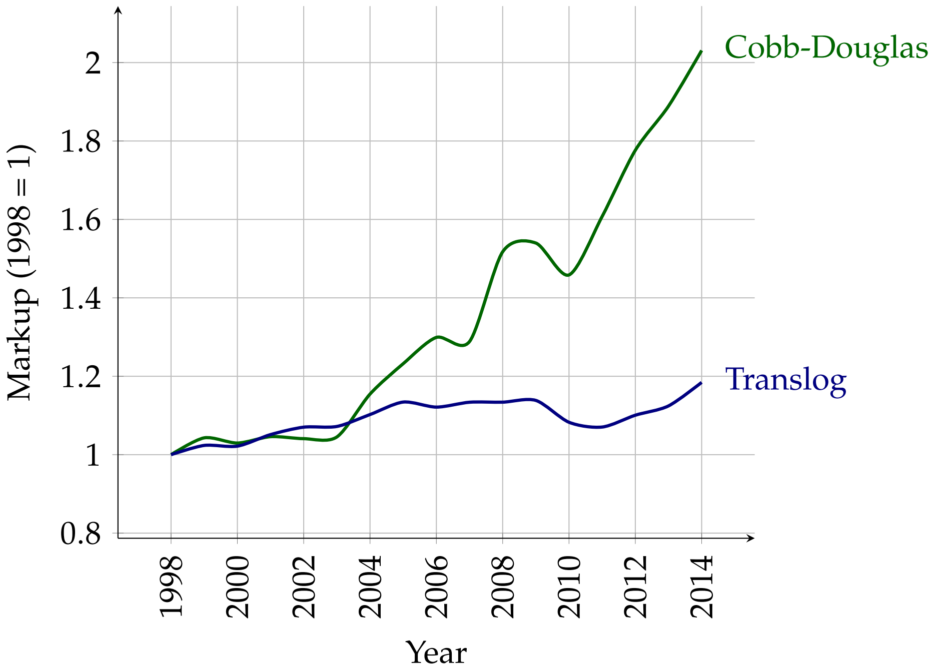


Figure 2: Aggregate Markup, 1998=1

With our firm-level markup estimates, we can look deeper into the underlying trends. Figure 3 presents trends in various percentiles along the markup distribution over time. These data show the markup of the firm ranked at tenth, twenty-fifth, the median, seventy-fifth, and ninetieth out of all firms in each year. A clear pattern emerges. The markup of firms in the top half of the distribution grew substantially until 2009, before a dip, and then a sharp recovery. The increasing gap between the 90th and 75th percentiles shows that the markup distribution is becoming more unequal at the top. For the bottom half of firms by markup, we see a slight increase before plateauing, and then a steady drop from 2007 to 2011. Firms in the bottom quarter of the distribution have markups no higher than the level of equivalent firms in 1998. Divergence between the 25th and 10th percentile only began around 2009. We find that *markup dispersion* between firms has been rising at the top of the distribution for much longer, and the variation is greater, than at the bottom. The levels of markup percentiles are plotted in Figure 10 in the Appendix.

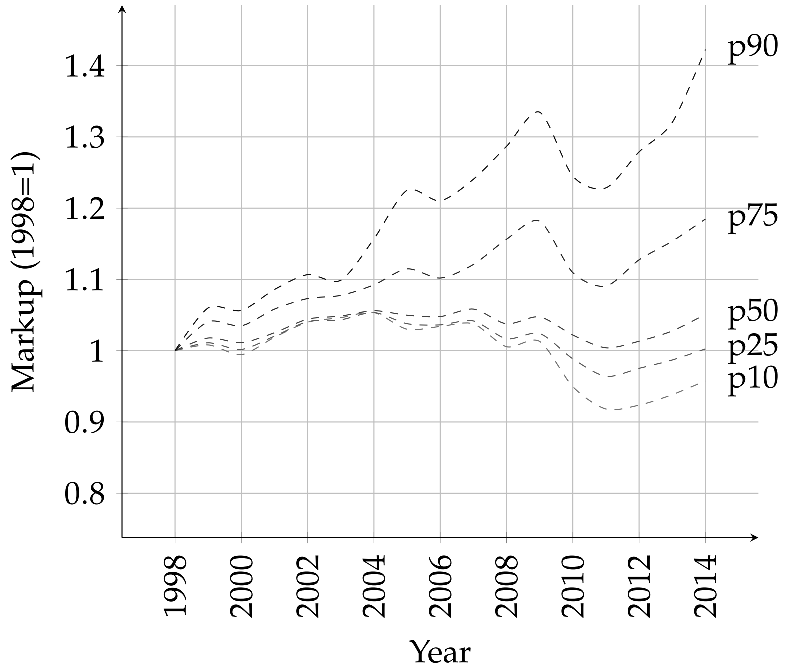


Figure 3: Translog Markup Percentile Trends, 1998=1

Figure 4 shows sectoral markups indexed in 1998. We find significant growth in Cobb-Douglas markups for service and construction sectors in the latter half of the sample, whereas the translog markup growth in these two sectors occurs earlier and falls dramatically during the financial crisis. In either case, it is services and construction that seem to contribute most to aggregate markup trends. Cobb-Douglas markups rise steadily over the sample, albeit it at different rates across sectors. On the other hand, translog markups experience greater fluctuations. In general translog markups do rise, but the fall during the crisis is more pronounced. The levels of sectoral markup are plotted in Figure 11 in the Appendix.

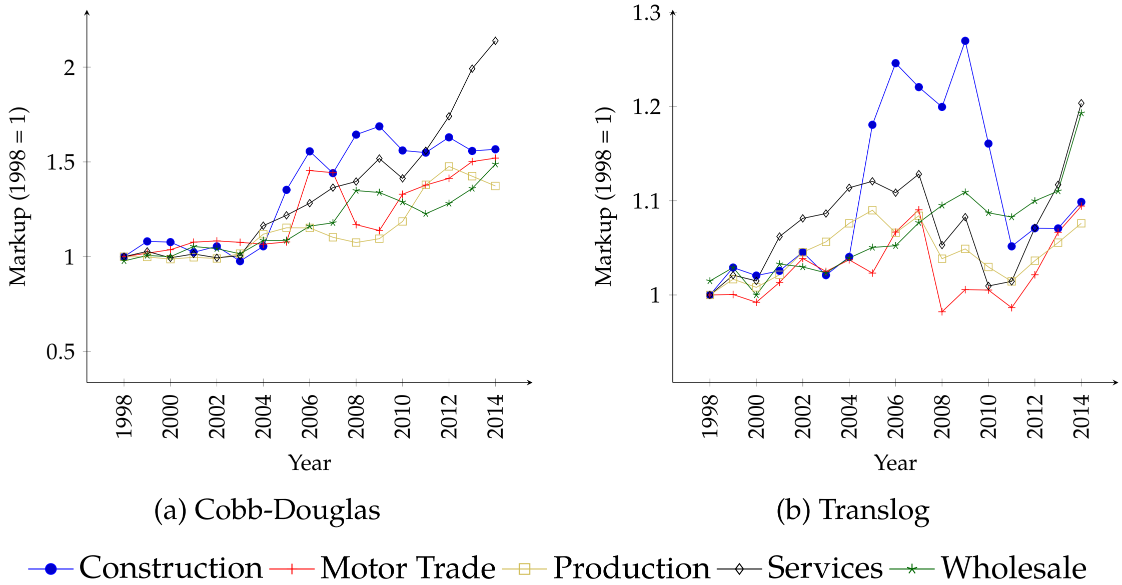


Figure 4: Markup Index, by Sector, 1998=1

The previously described pattern of *markup dispersion* does not hold for all sectors. The markup percentiles by sector are presented in Figure 5. The greatest rise in markups at the top of the distribution occurs for firms in construction and motor trade. The most significant decline at the bottom of the markup distribution occurs in wholesale, services and construction. The timing of the increased variation in markups varies across industries. Significant markup dispersion begins to show in construction, production, and wholesale from around 2004, and steadily grows in these sectors over the next decade. For services and motor trade, markup variation is noticeable even earlier in the sample, but actually falls around the financial crisis, before increasing sharply from around 2010. The levels of sectoral markup percentiles can be found in Figure 12 in the Appendix. These plots highlight the *levels* of markup variation by sector. The greatest difference between the top and bottom tenth percentiles of markups are found in services (over 3 compared to below 1), whereas the smallest difference is found in motor trade and wholesale (around 1.4 compared to 0.8).

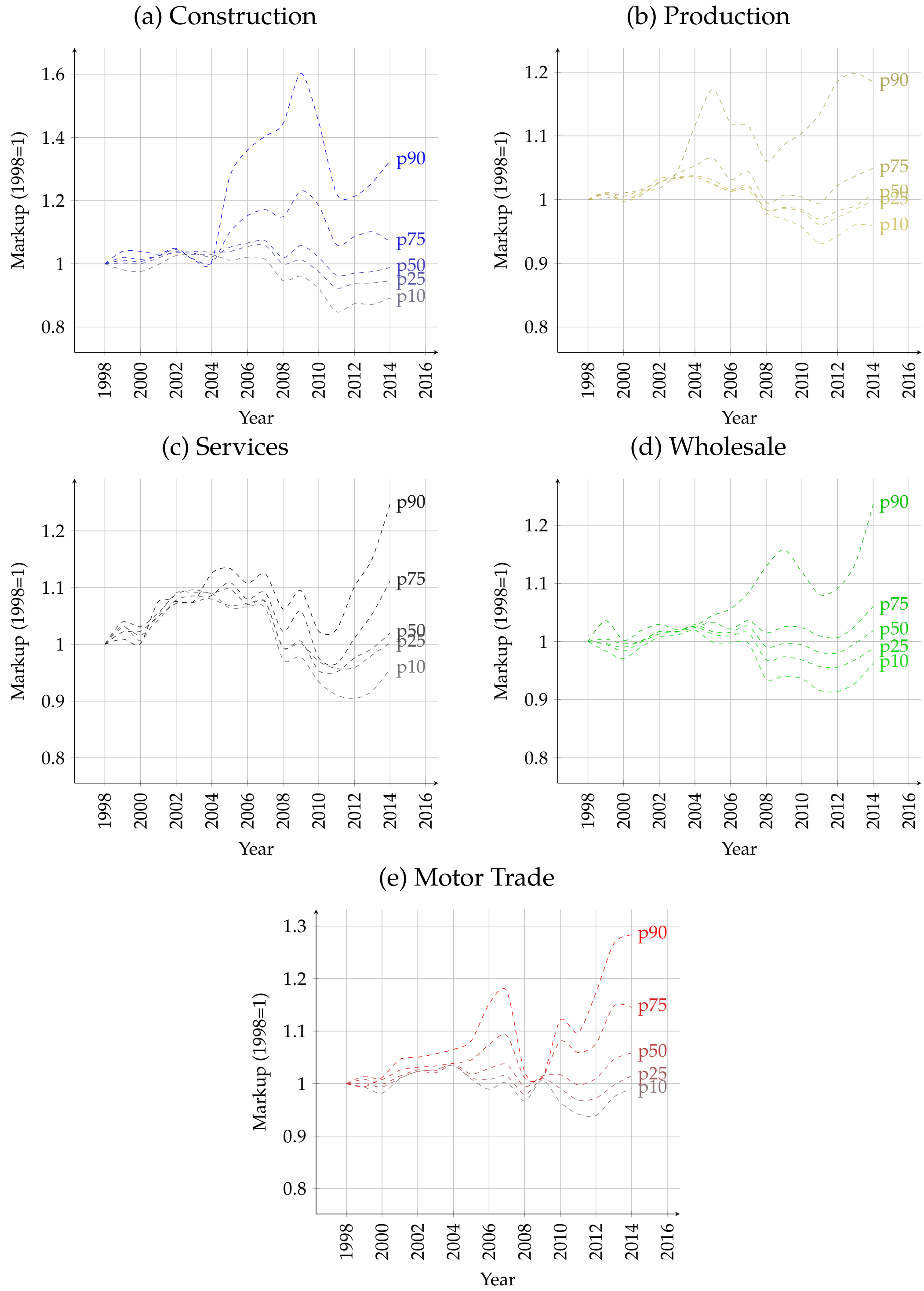
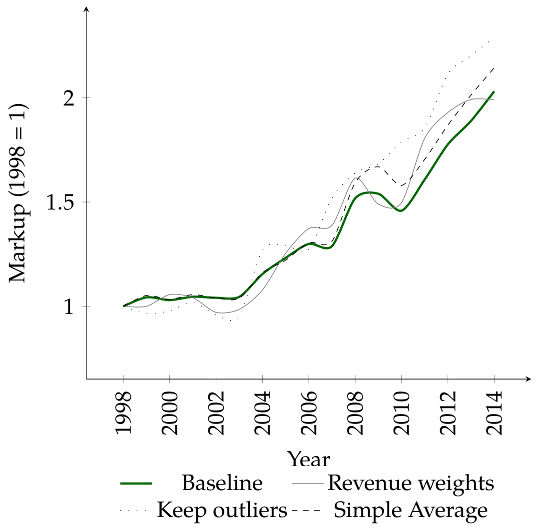


Figure 5: Markup Percentile Trends, by Sector, 1998=1

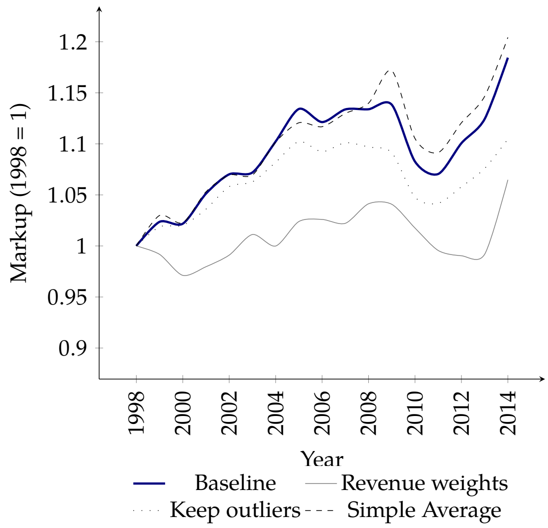
4.1.1 Robustness: Weighting and Flexible Input

Two important decisions in the process of calculating aggregate markups are which variable to use as a flexible input in the production function estimation approach and how to aggregate individual firm-level markups into a final aggregate measure.

In Figure 6, we present results for the aggregate markup with different weighting schemes, indexed to 1 in 1998. Our baseline uses value-added weight, the alternatives are: (1) simple average markup, (2) weighted average markup without dropping outliers, (3) revenue-weighted average markup. Aggregate markup trends are similar irrespective of different aggregation schemes. The markup computed from a Cobb-Douglas production function in panel 6a demonstrates an increasing trend, accelerating post-crisis. The translog markup plotted in panel 6b is rising until the financial crisis, dropping sharply before recovering from around 2011. For both aggregate markups, our baseline is below the simple average, and above revenue-based weights. Keeping outlier markup firms (top or bottom 1% by year) doesn't have a huge effect on our aggregate markup trends.



(a) Markup (Cobb-Douglas)

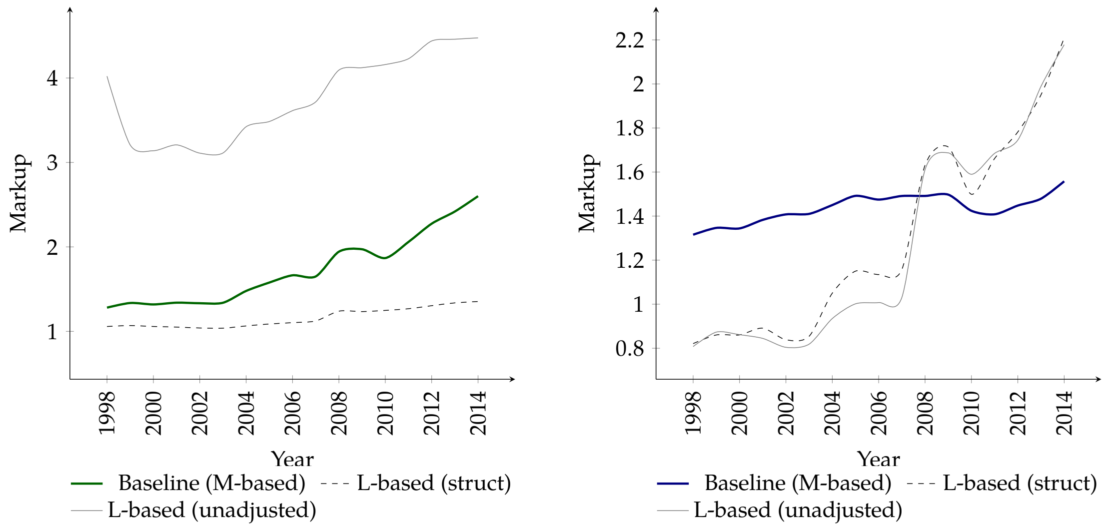


(b) Markup (Translog)

Figure 6: Aggregate Markup with Different Weights, 1998=1

In Figure 7, we show the comparison between two different approaches to computing markups. We show the levels to highlight the significant divergence. Our baseline approach uses *materials input* as the flexible factor, where we assume a gross output production function. An alternative is to use a structural value-added production function and use labour as the variable input as in De Loecker and Scott (2016). The solid lines illustrate our baseline aggregate markup estimates which rely on the materials input. The grey lines display the markup when we simply divide the estimated labour elasticity (from value-added production function estimation) by the labour share in revenue. The dashed lines plot the markup when the labour input is the flexible factor from a structural value-added production function, with the adjustment presented in De Loecker and Scott (2016).

For the Cobb-Douglas production function, the labour-based markup is very high, and follows a U-shape trend over the estimation period. Once the De Loecker and Scott (2016) adjustment is applied, the labour-based markup is closer to the materials-based markup, rising steadily from around 1.06 to 1.35. For the translog production function, the labour-based markup doesn't vary much whether it is adjusted or not. It starts below one but quickly rises, before experiencing a jump around 2008. It continues to rise post-crisis to over two. In terms of levels, the labour-based markup yields high markups relative to existing research on aggregate markups in the US and the UK. We present the indexed version of this plot in the Appendix in Figure 14.



(a) Markup (Cobb-Douglas) (b) Markup (Translog)

Figure 7: Aggregate Markup Levels with Different Flexible Input

4.2 Productivity and Markup Relationship

In Figure 8, we illustrate annual trends in both the aggregate TFPR and the aggregate markup. In particular, we draw their indices (1998=1) rather than absolute levels to focus on trends. During the pre-crisis period, both TFPR and markup have an increasing trend. For Cobb-Douglas estimates, both series rise at similar rates until around 2005, before TFPR slows down while markup growth accelerates. For translog estimates, markup growth always outstrips productivity growth, and both drop around the crisis before bouncing back. Markup growth post-crisis is fast, while TFPR grows more slowly.

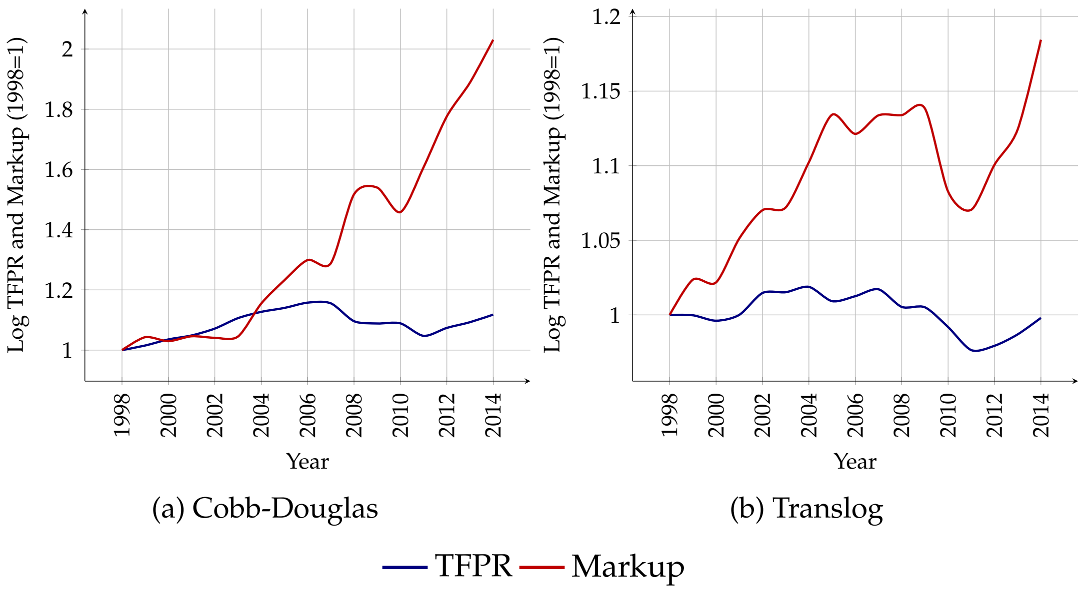


Figure 8: Aggregate Markups and TFPR, 1998 =1

To investigate the relationship between TFPR and the markup, we run a simple regression as follows:

$$\ln(\text{TFPR})_{it} = \alpha + \beta \ln(\text{Markup})_{it} + \eta_i + \zeta_t + \epsilon_{it}$$

where i implies firms (reporting units) and t indicates year. We add firm fixed effects (η_i) and year fixed effects (ζ_t) to control for firm-specific time-invariant characteristics and annual macroeconomic conditions respectively. Therefore, the coefficient β is the partial correlation between markups and TFPR, given time-invariant firm factors and annual time events are held constant.

In columns (1)-(3) of Table 1, all estimated coefficients for β are significant and negative. The relationship weakens when we add firm fixed effects (column (2)-(3) vs. column (1)). Note that both TFPR and markup trends in the aggregate level move upward before 2008 whereas they diverge somewhat after 2008. We run further regressions with the subsample for the pre-crisis period (1998-2008) as well as the subsample for the post-crisis period (2009-2014) to examine whether the relationship differs between two periods. Columns (4)-(9) imply that TFPR and markups are negatively related, but the relationship has become weaker since the financial crisis.

Table 1: Relationship between TFPR and Markups (Cobb-Douglas)

Dep Var:	Full period (1998-2014)			Pre-crisis (1998-2008)			Post-crisis (2009-2014)		
Log TFP	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Markup	-0.206***	-0.084***	-0.084***	-0.234***	-0.094***	-0.095***	-0.151***	-0.049***	-0.050***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)
Firm FE		✓	✓		✓	✓		✓	✓
Year FE			✓			✓			✓
Obs.	503,567	298,185	298,185	365,075	214,749	214,749	138,492	61,756	61,756
R^2	0.001	0.911	0.912	0.003	0.939	0.94	0.0001	0.952	0.953

Note: (1) ***p-value < 0.01; (2) Standard errors in parenthesis are clustered at the firm level.

We repeat this exercise with our translog estimates. The results are in Table 2. The main qualitative and quantitative findings are the same. We similarly find a strong negative and statistically significant relationship between TFPR and markups, which is strengthened within firms. Furthermore, the magnitude of the relationship shrinks towards zero in the post-crisis period, suggesting a weakening of the association between productivity and markups. There are two noticeable differences. The first is that there is no statistically significant relationship when looking at a simple correlation of translog estimates (see columns (1), (4), and (7)). The second difference is that the estimated coefficients are substantially lower for translog compared to Cobb-Douglas.

Table 2: Relationship between TFPR and Markups (Translog)

Dep Var:	Full period (1998-2014)			Pre-crisis (1998-2008)			Post-crisis (2009-2014)		
Log TFPR	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Markup	-0.300***	-0.255***	-0.266***	-0.364***	-0.253***	-0.266***	-0.181***	-0.274***	-0.281***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)
Firm FE		✓	✓		✓	✓		✓	✓
Year FE			✓			✓			✓
Obs.	500,361	296,967	296,967	362,763	214,064	214,064	137,598	61,409	61,409
R^2	0.000	0.964	0.965	0.000	0.962	0.964	0.0001	0.982	0.983

Note: (1) ***p-value < 0.01; (2) Standard errors in parenthesis are clustered at the firm level.

4.2.1 Discussion: Why are markups and TFPR negatively related?

The main purpose of our paper is to robustly document market power trends in the UK given recent interest in the topic from both policy and academic angles. A helpful by-product of the production estimation process that we use is to produce firm-level productivity measures. We show that our productivity measures re-enforce existing evidence for a UK producivity puzzle. This provides further evidence of the UK productivity puzzle from a different dataset and different methodological standpoint to existing literature. We document a negative relationship between our estimates of TFPR and markups which warrants further discussion.

The UK has a well-documented problem of productivity laggards: the probability distribution of firm's productivities has a dense left-hand tail of small, unproductive, firms that decreases average productivity (Haldane 2017; Awano, Ardanaz-Badia, and Wales 2017). An open question is how are these laggard firms able to survive? One explanation is that the low-interest rate environment of the 2010s has allowed 'zombie firms' to roll-over debt. Our evidence is indicative of low productivity firms surviving by charging high markups to offset their low output, thus generating sufficient revenues to break-even. Therefore, high-markup, low-productivity, laggards are core to the negative relationship between productivity and markups. A follow-up question is: *how are the productivity laggards able to charge high markups*? Under many demand systems – such as the demand curve arising from quadratic preferences or Feenstra (translog) preferences – a small (low demand), unproductive, firm has more elastic price elasticity of demand. This means it cannot charge a high price without a strong decline in demand. Similarly, in models of oligopolistic competition larger, more productive, firms have greater price setting ability due to higher market share. One answer to this question is a demand system that allows low output firms to charge high markups. A demand system like this arises if consumers have superconvex preferences (Mrázová and Neary 2017). Another possibility is that these small firms face weak competition due to barriers to entry and hence have inelastic demand and higher markups. For small, unproductive, firms these barriers to entry might take the form of geographic location or customer loyalty. The degree of market power depends on the definition of the anti-trust market. For example, a suburban convenience store, or a rural car mechanic, can have a large market share if the market definition depends on geography.

5 Conclusion

We construct firm-level measures of markups and TFPR for the UK. We apply recent techniques in production function estimation to a large-scale, confidential dataset of UK firms covering all sectors of the economy. This is a significant extension to existing work that uses proprietary data on publicly-listed companies. We show that markups are rising in the UK and that markup dispersion has also grown. While services have the highest markups, there is evidence of growth in markups across sectors. Finally, we present evidence on firm-level productivity, which is a by-product of the production function estimation process, and find that it is consistent with existing evidence of a so-called productivity puzzle in the UK. We note that there is a negative relationship between markups and productivity.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer (2015). "Identification properties of recent production function estimators". In: *Econometrica* 83.6, pp. 2411–2451 (cit. on pp. 4, 9, 11–14, 36).

Aghion, Philippe, Richard Blundell, Rachel Griffith, Peter Howitt, and Susanne Prantl (2004). "Entry and productivity growth: Evidence from microlevel panel data". In: *Journal of the European Economic Association* 2.2-3, pp. 265–276 (cit. on p. 5).

Aquilante, Tommaso, Shiv Chowla, Nikola Dacic, Andrew Haldane, Riccardo Masolo, Patrick Schneider, Martin Seneca, and Srdan Tatomir (2019). *Market power and monetary policy*. Bank of England working papers 798. Bank of England (cit. on p. 5).

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 (cit. on p. 2).

— (2020). "The Fall of the Labor Share and the Rise of Superstar Firms*". In: *The Quarterly Journal of Economics* 135.2, pp. 645–709. eprint: https://academic.oup.com/qje/article-pdf/135/2/645/32994954/qjaa004.pdf (cit. on p. 2).

Awano, Gaganan, Anna Ardanaz-Badia, and Philip Wales (2017). "Understanding firms in the bottom 10% of the Labour Productivity Distribution in Great Britain: "the laggards", 2003 to 2015". In: *Office for National Statistics Articles* (cit. on p. 23).

Baqaee, David and Emmanuel Farhi (2020). "Productivity and Misallocation in General Equilibrium*". In: *The Quarterly Journal of Economics* 135.1, pp. 105–163. eprint: https://academic.oup.com/qje/article-pdf/135/1/105/32893883/qjz030.pdf (cit. on p. 2).

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch (2021). "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data". In: *Journal of Monetary Economics* 121, pp. 1–14 (cit. on pp. 2, 6).

Calligaris, Sara, Chiara Criscuolo, and Luca Marcolin (2018). "Mark-ups in the digital era". In: (cit. on p. 2).

Competition and Markets Authority (2022). *The State of UK Competition*. Tech. rep. (cit. on p. 5).

Crawford, Claire, Wenchao Jin, and Helen Simpson (2013). "Productivity, Investment and Profits during the Great Recession: Evidence from UK Firms and Workers". In: *Fiscal Studies* 34.2, pp. 153–177 (cit. on p. 3).

Criscuolo, Chiara, Jonathan Haskel, and Ralf Martin (2004). "Import competition, productivity, and restructuring in UK manufacturing". In: *Oxford Review of Economic Policy* 20.3, pp. 393–408 (cit. on p. 3).

De Loecker, Jan (2021). "Comment on (Un)pleasant ... by Bond et al (2020)". In: *The Journal of Monetary Economics* 121, pp. 15–18 (cit. on pp. 2, 6).

De Loecker, Jan and Jan Eeckhout (2017). *The rise of market power and the macroeconomic implications*. Tech. rep. National Bureau of Economic Research (cit. on p. 10).

— (2021). *Global market power*. Tech. rep. National Bureau of Economic Research (cit. on p. 2).

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. eprint: https://academic.oup.com/qje/article-pdf/135/2/561/32995291/qjz041.pdf (cit. on p. 4).

De Loecker, Jan and Paul T Scott (2016). *Estimating market power Evidence from the US Brewing Industry*. Working Paper 22957. National Bureau of Economic Research (cit. on pp. 12, 19, 20, 40).

De Loecker, Jan and Frederic Warzynski (2012). "Markups and firm-level export status". In: *American Economic Review* 102.6, pp. 2437–71 (cit. on pp. 8, 9, 14).

De Ridder, Maarten, Basile Grassi, and Giovanni Morzenti (2021). *The Hitchhiker's Guide to Markup Estimation*. Working Papers 677. IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University (cit. on pp. 2, 6).

Demirer, Mert (2020). *Production Function Estimation with Factor-Augmenting Technology: An Application to Markups*. Tech. rep. (cit. on pp. 4, 11).

Diez, Mr Federico, Mr Daniel Leigh, and Suchanan Tambunlertchai (2018). *Global market power and its macroeconomic implications*. International Monetary Fund (cit. on p. 2).

Eggertsson, Gauti B., Jacob A. Robbins, and Ella Getz Wold (2021). "Kaldor and Piketty's facts: The rise of monopoly power in the United States". In: *Journal of Monetary Economics* 124. The Real Interest Rate and the Marginal Product of Capital in the XXIst Century October 15-16, 2020, S19–S38 (cit. on p. 2).

Gandhi, Amit, Salvador Navarro, and David A Rivers (2018). *On the identification of gross output production functions*. Tech. rep. CHCP Working Paper (cit. on pp. 4, 9).

García-Perea, Pilar, Aitor Lacuesta, and Pau Roldan-Blanco (2021). "Markups and cost structure: Small Spanish firms during the Great Recession". In: *Journal of Economic Behavior & Organization* 192, pp. 137–158 (cit. on p. 2).

Gobey, Matthew and Karolis Matikonis (2021). "Small business property tax reductions and job growth". In: *Small Business Economics* 56, pp. 277–3874 (cit. on p. 7).

Goodridge, Peter, Jonathan Haskel, and Gavin Wallis (2013). "Can Intangible Investment Explain the UK Productivity Puzzle?" In: *National Institute Economic Review* 224, R48–R58 (cit. on p. 3).

— (2016). "Accounting for the UK Productivity Puzzle: A Decomposition and Predictions". In: *Economica* (cit. on p. 6).

Griffith, Rachel (2001). *Product market competition, efficiency and agency costs: an empirical analysis*. Tech. rep. IFS Working Papers, Institute for Fiscal Studies (IFS) (cit. on p. 5).

Griffith, Rachel, Rupert Harrison, and Gareth Macartney (2007). "Product market reforms, labour market institutions and unemployment". In: *The Economic Journal* 117.519, pp. C142–C166 (cit. on p. 5).

Griffith, Rachel, Rupert Harrison, and Helen Simpson (2010). "Product Market Reform and Innovation in the EU". In: *Scandinavian Journal of Economics* 112.2, pp. 389–415 (cit. on p. 5).

Gutiérrez, Germán and Thomas Philippon (2020). "How EU markets became more competitive than US markets: A study of institutional drift". In: *Working Paper* (cit. on p. 2).

Haldane, Andrew (2017). *Productivity Puzzles*. Bank of England – Speech at London School of Economics (cit. on pp. 8, 23).

Hall, Robert E (1988). "The relation between price and marginal cost in US industry". In: *Journal of political Economy* 96.5, pp. 921–947 (cit. on pp. 5, 8).

Harris, Richard and John Moffat (2017). "The UK productivity puzzle, 2008–2012: evidence using plant-level estimates of total factor productivity". In: *Oxford Economic Papers* 69.3, pp. 529–549 (cit. on pp. 3, 6).

Haskel, Jonathan and Christopher Martin (1992). "Margins, concentration, unions and the business cycle : Theory and evidence for Britain". In: *International Journal of Industrial Organization* 10.4, pp. 611–632 (cit. on p. 5).

Haskel, Jonathan, Christopher Martin, and Ian Small (1995). "Price, Marginal Cost and the Business Cycle". In: *Oxford Bulletin of Economics and Statistics* 57.1, pp. 25–41 (cit. on p. 5).

Jacob, Nick and Giordano Mion (2020). *The UK's great demand and supply recession*. CEP Discussion Papers dp1737. Centre for Economic Performance, LSE (cit. on p. 6).

Levinsohn, James and Amil Petrin (2003). "Estimating production functions using inputs to control for unobservables". In: *The Review of Economic Studies* 70.2, pp. 317–341 (cit. on pp. 4, 9, 11, 12).

Martin, Josh and Cain Baybutt (2022). *The 'F Words': Why Surveying Businesses About Intangibles is so Hard*. Discussion Paper. ESCoE (cit. on pp. 7, 8).

Martin, Ralf (2002). "Building the capital stock". In: *Notes* 514.q516, p. 1000 (cit. on p. 42).

Martins, Joaquim Oliveira, Stefano Scarpetta, and Dirk Pilat (1996). *Mark-Up Ratios in Manufacturing Industries: Estimates for 14 OECD Countries*. OECD Economics Department Working Papers 162. OECD Publishing (cit. on p. 5).

Miller, Helen and Alina Barnett (2015). "Capital Allocation and Productivity". In: *Bank of England Conference. Understanding the Great Recession: from micro to macro* (cit. on p. 3).

Mrázová, Monika and J. Peter Neary (2017). "Not So Demanding: Demand Structure and Firm Behavior". In: *The American Economic Review* 107.12, pp. 3835–3874 (cit. on p. 24).

Olley, G. Steven and Ariel Pakes (1996). "The dynamics of productivity in the telecommunications equipment industry". In: *Econometrica* 64.6, pp. 1263–1297 (cit. on pp. 4, 9).

Raval, Devesh (2020). "Testing the Production Approach to Markup Estimation". In: *SSRN 3324849* (cit. on pp. 2, 9, 10).

Riley, Rebecca, Chiara Rosazza-Bondibene, and Garry Young (2015). *The UK productivity puzzle 2008-13: evidence from British businesses*. Bank of England working papers 531. Bank of England (cit. on p. 32).

Roeger, Werner (1995). "Can imperfect competition explain the difference between primal and dual productivity measures? Estimates for US manufacturing". In: *Journal of political Economy* 103.2, pp. 316–330 (cit. on p. 5).

Schneider, Patrick (2018). *Decomposing differences in productivity distributions*. Bank of England working papers 740. Bank of England (cit. on p. 6).

Syverson, Chad (2019). "Macroeconomics and market power: Context, implications, and open questions". In: *Journal of Economic Perspectives* 33.3, pp. 23–43 (cit. on pp. 2, 5).

Traina, James (2018). "Is aggregate market power increasing? production trends using financial statements". In: *Production Trends Using Financial Statements (February 8, 2018)* (cit. on p. 10).

UKDS (2017). *ARDx Capital Stock User Guide*. Tech. rep. The UK data service, ARDx documentation (cit. on p. 42).

van Heuvelen, Gerrit Hugo, Leon Bettendorf, and Gerdien Meijerink (2021). "Markups in a dual labour market: The case of the Netherlands". In: *International Journal of Industrial Organization* 77, p. 102762 (cit. on p. 2).

Wales, Philip, Russell Black, Ted Dolby, and Gaganan Awano (2018). *UK trade in goods and productivity: New findings*. Economic Statistics Centre of Excellence (ESCoE) Discussion Papers ESCoE DP-2018-09. Economic Statistics Centre of Excellence (ESCoE) (cit. on p. 3).

Wooldridge, Jeffrey M (2009). "On estimating firm-level production functions using proxy variables to control for unobservables". In: *Economics Letters* 104.3, pp. 112–114 (cit. on p. 9).

— (2013). *Introductory Econometrics: A Modern Approach (Fifth Edition)*. Cengage Learning (cit. on p. 33).

A Production Data Construction

A.1 Variable Construction

The ARDx dataset contains firm-level (i.e. reporting unit) production data that allows us to estimate production functions and in turn estimate firm-level markups. We use gross revenues to measure gross output of firms.[4] We use the number of employees to measure labour input. We use a variable called 'total purchase of energy, goods and service' to measure material input. The ARDx dataset does not provide a capital stock variable, but it reports capital expenditure so we construct capital stock using the Perpetual Inventory Method.[5]

We convert gross output (i.e. firm-level revenues) into real values using ONS experimental industry deflators. It contains 2-digit SIC industry level deflators (2010=100). We merge the ONS deflators to the ARDx dataset using firm-specific SIC 2-digit codes, and calculate real gross output with the deflators. The ONS industry level deflators provide 3-digit level information for some industries. We merge the deflator data to the ARDx with SIC 3-digit codes for those industries. We deflate the material input variable with ONS producer price inflation time series data, available at https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/producerpriceindex. We also deflate the capital stock variable with industry-invariant gross fixed capital formation deflator.

A.2 Sector Classifications

The ARDx dataset is categorized into eight sectors. Each sector receives a different questionnaire. Table 3 shows how these sectors aggregate from SIC 2-digit industries. While there are eight sectors (i.e., Production, Construction, Catering, Motor Trade, Retail, Wholesale, Property and Other Services) in the ARDx, we present results for

[4]In fact, there are two variables that indicate firms' gross revenues in the ARDx: IDBR-based revenues and ABS-based revenues. We use the ABS-based revenues as it is directly comparable to input variables, all of which are extracted from the ABS, not the IDBR.

[5]See below for details of capital stock construction.

five sectors. ARDx has no information about firm-level value added for the retail sector and the catering sector. We drop the property sector which only comprises a single 2-digit SIC sector that we remove in our econometric estimation (see cleaning section). We refer to 'Other Services' as 'Services'. The most important two sectors are Production and Services. These account for the majority of 2-digit subsectors and they also account for a majority of turnover. The remaining sectors all correspond to one or two 2-digit SIC codes.

Table 3: Sector Classification in ARDx

Sector	Abbr.	SIC 92/03 codes	SIC07 codes
Production	PD	1/2/10-41	1-39
Construction	CN	45-45	41-43
Catering	CA	55-55	55-56
Motor trade	MT	50-50	45-45
Retail	RE	52-52	47-47
Wholesale	WH	51-51	46-46
Property	PR	70-70	68-68
Other Services	ST	60-93	49-96

A.3 Cleaning Data

We estimate production functions by SIC 2-digit level industry. We drop industries that have less than 100 observations over the sample period (1998-2014) due to potentially providing imprecise estimates. In addition, we exclude industries in which it is difficult to define output and inputs in production and thus the elasticities become difficult to interpret. Consequently, we drop financial and insurance (SIC, 64-66), Agriculture, Forestry and Fishing (SIC, 1-3), Education, Human Health and Social Work Activity (SIC, 85-88), Mining and Quarrying (SIC, 5-9), Public Administration and Defence (SIC, 84), Utilities (SIC, 35-39) and Real Estate Activities (SIC, 68).[6]

Next, we exclude extreme values for variables in the production function. We drop observations in which the output, capital, material and labour values are equal to or less than one. For output, capital and materials this implies values below £1,000. For

[6]This is consistent with other ARDx users Riley, Rosazza-Bondibene, and Young (2015).

labour this implies firms with only one employee (including the director).[7] In other words, we remove very small firms. In addition, we remove input factor shares share (input expenditure share in revenue) outliers. We drop observations with input factor shares in the top and bottom 0.1% in each year. For our translog estimates, we remove all firm observations with negative flexible output elasticities, else those firm markups would be negative. Table 4 contains the number of firms at each stage of the data cleaning process.

Table 4: Data Cleaning: Firms Dropped

	# Firms
All ARD firms	854,732
Drop if no 2-digit sector	852,424
Drop if < 100 firms in sector	852,331
Drop non-market sectors	761,348
Take logs of regression variables	539,368
Drop outlier factor shares	503,567
Remove $\widehat{\theta}_{it}^{m} < 0$ (translog only)	500,361

[7]Removing values below 1 also avoids logarithms returning negative or undefined values. Typically this is addressed by transforming variables to $\ln(1+x)$ and leads to small, but acceptable, approximation error. However, given that all variables would need to be transformed and there are complications in interpreting the resulting elasticities $\ln(1+y)$ will not be normally distributed Wooldridge (2013, p. 193)), it is an advantage to drop.

B Summary Tables

Table 5: Descriptive Statistics of Regression Variables for Full Sample

	Mean	SD	p10	p50	p90	No. Obs
Revenue	39,736	675,183	92	1,458	42,797	503,567
Labour	224	2,213	2	20	349	503,567
Capital	7,696	150,007	22	351	7,915	503,567
Materials	29,651	636,176	32	703	26,255	503,567
Materials Share	0.55	-	0.17	0.58	0.87	503,567
Labour Share	0.26	-	0.04	0.23	0.52	503,567

Table 6: Descriptive Statistics of Regression Variables by Broad Sector

	Mean	SD	p10	p50	p90	No. Obs
Manufacturing						
Revenue	36,005	235,437	336	4,294	58,896	125,737
Labour	192	576	8	54	431	125,737
Capital	10,362	75,776	148	1,498	16,154	125,737
Materials	24,954	178,528	122	2,400	38,999	125,737
Materials Share	0.57	-	0.30	0.58	0.81	125,737
Labour Share	0.28	-	0.11	0.27	0.47	125,737
Construction						
Revenue	17,812	108,789	111	1,414	48,782	51,784
Labour	103	395	2	11	214	51,784
Capital	2,309	41,523	11	104	2,210	51,784
Materials	12,467	89,027	18	343	16,896	51,784
Materials Share	0.51	-	0.17	0.52	0.81	51,784
Labour Share	0.25	-	0.00	0.24	0.49	51,784
Trade, Wholesale, Transport						
Revenue	62,673	1,102,305	111	1,414	48,782	182,814
Labour	256	3,404	2	14	244	182,814
Capital	7,092	103,075	20	245	5,667	182,814
Materials	52,666	1,044,112	61	929	26,219	182,814
Materials Share	0.69	-	0.37	0.74	0.92	182,814
Labour Share	0.16	-	0.02	0.13	0.35	182,814
Services						
Revenue	25,276	284,335	65	728	28,673	179,028
Labour	249	1,627	2	17	403	179,028
Capital	8,821	228,905	20	218	5,435	179,028
Materials	14,417	209,297	15	242	11,263	179,028
Materials Share	0.41	-	0.09	0.38	0.77	179,028
Labour Share	0.34	-	0.06	0.32	0.68	179,028

C Additional Figures

All figures in this section contain markups estimated with a production function following Ackerberg, Caves, and Frazer (2015) with gross output.

C.1 Markup Levels

Figure 9 contains the average economy-wide markup. The average markup has an increasing trend from 1998 to 2014. For Cobb-Douglas it increases from 1.28 in 1998 to 2.60 in 2014 and for translog is increases from 1.32 to 1.56.

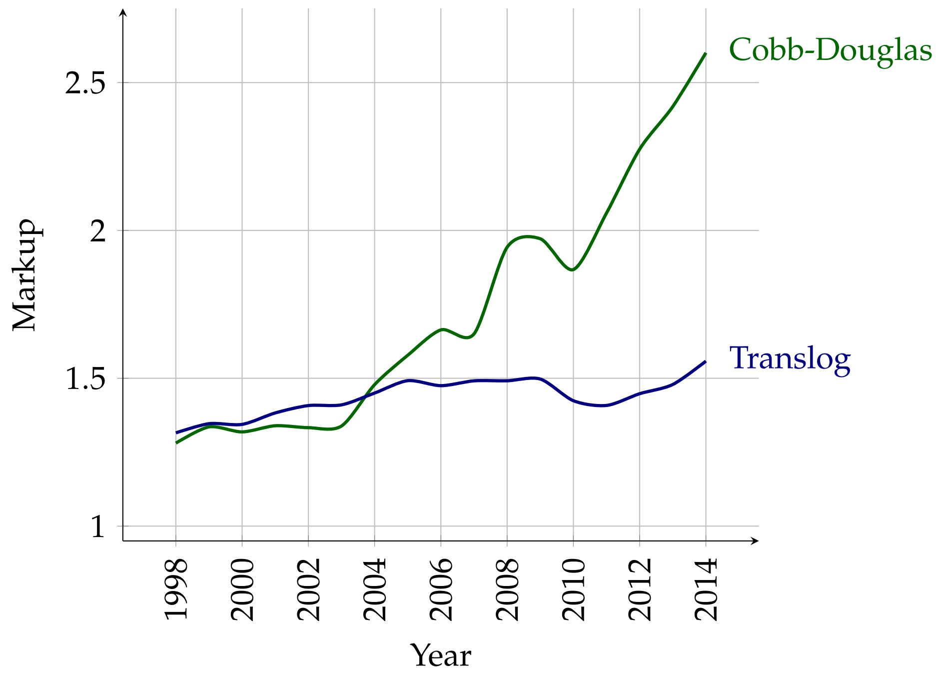


Figure 9: Aggregate Markup Levels

Figure 10 plots the percentiles across the markup distribution over time. The widening gap between the markups at the top and bottom of the distribution is apparent, and is driven mostly by a rise for the high markup firms.

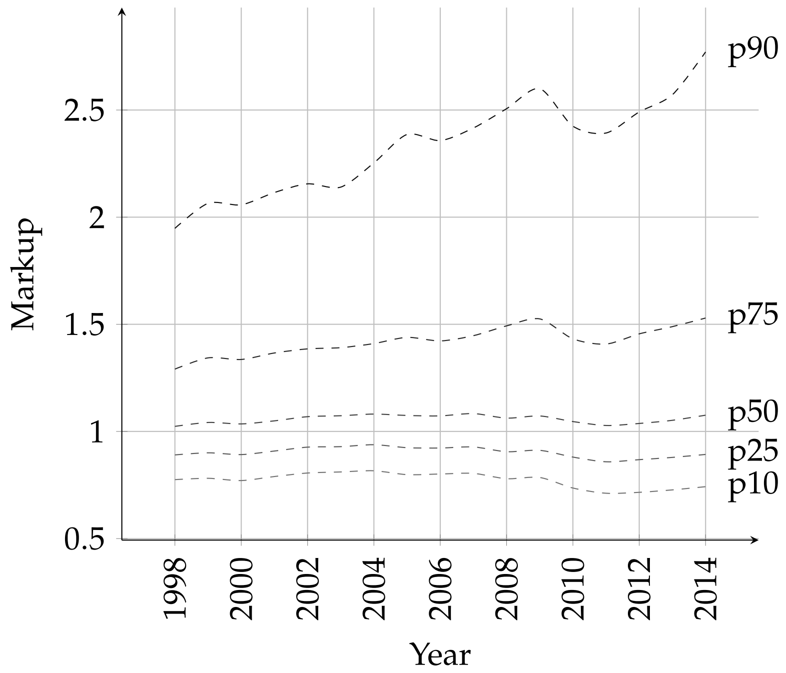
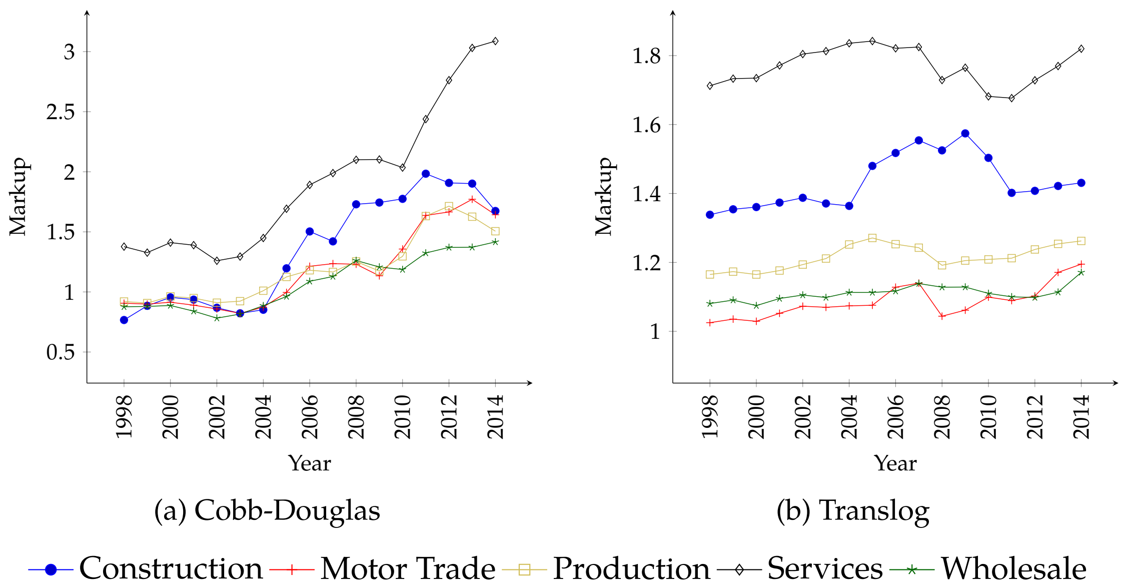


Figure 10: Markup Levels Percentile Distribution

Figure 11 presents sector-level markups in levels. The sector with the highest markup is the services and construction sectors, while motor trade and wholesale are the lowest. This result is consistent across the production functions.



(a) Cobb-Douglas (b) Translog

Figure 11: Markup Levels, by Sector

Figure 12 shows the percentiles across the markup distribution over time for each sector.

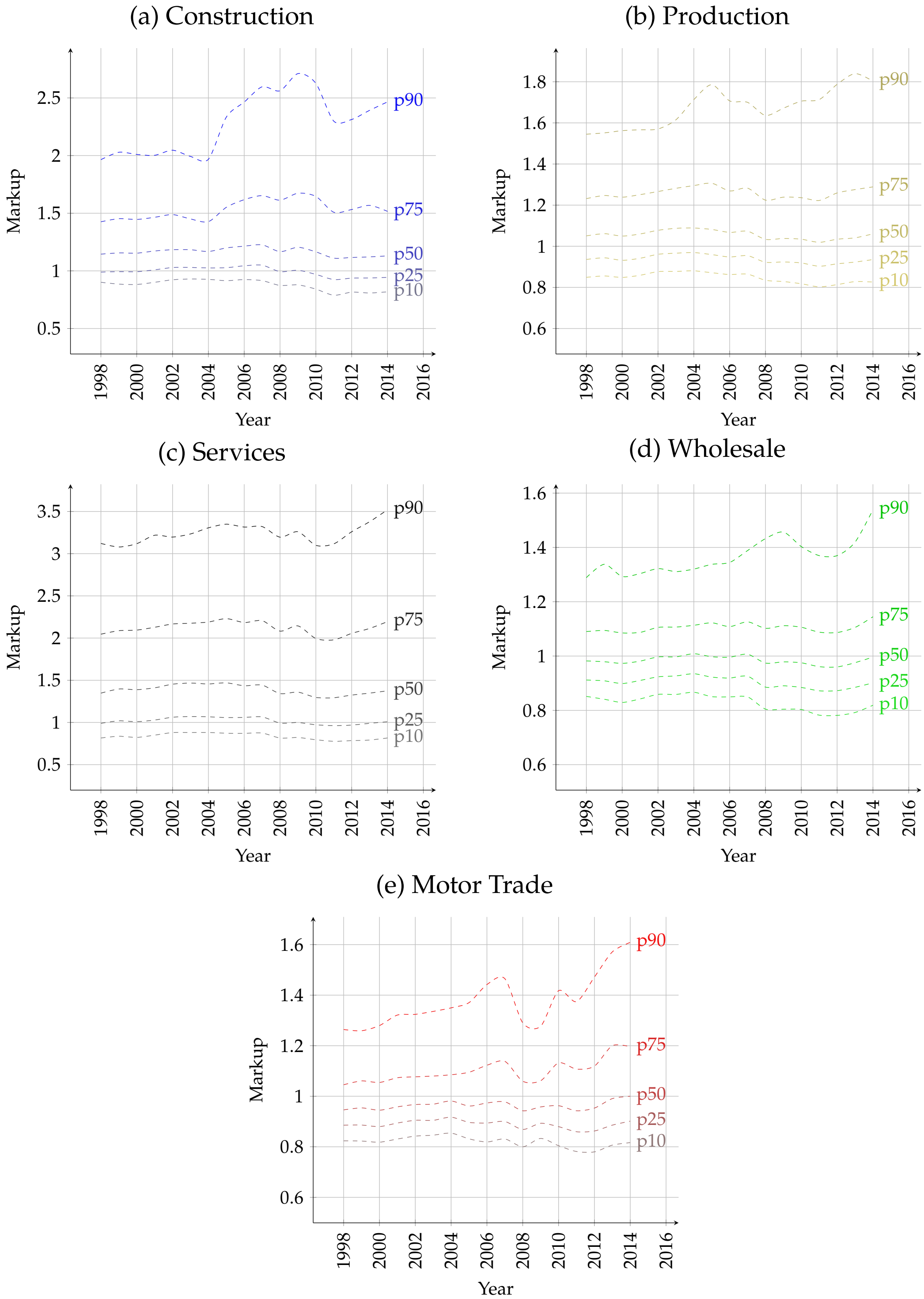
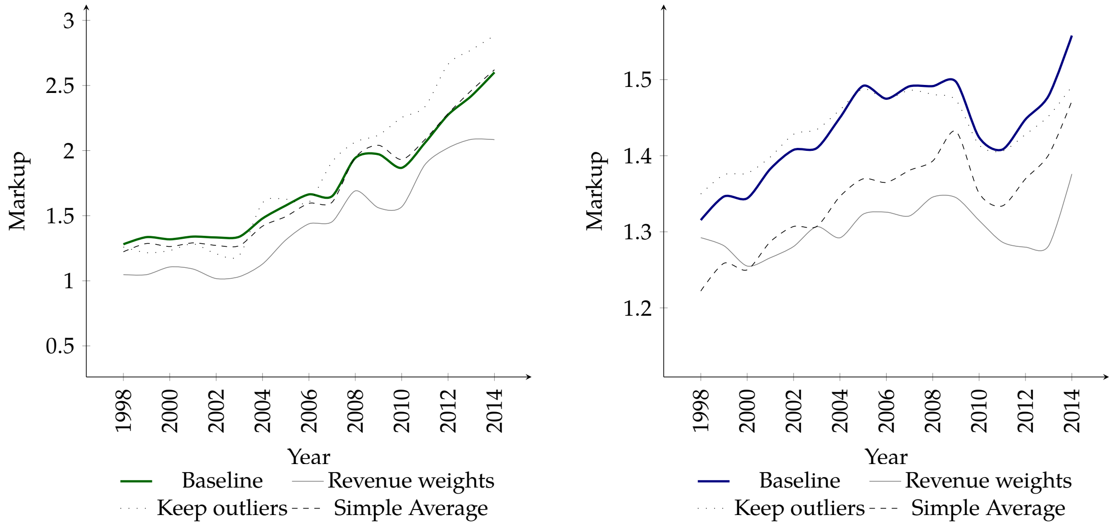


Figure 12: Markup Levels Percentile Distribution, by Sector

C.2 Robustness

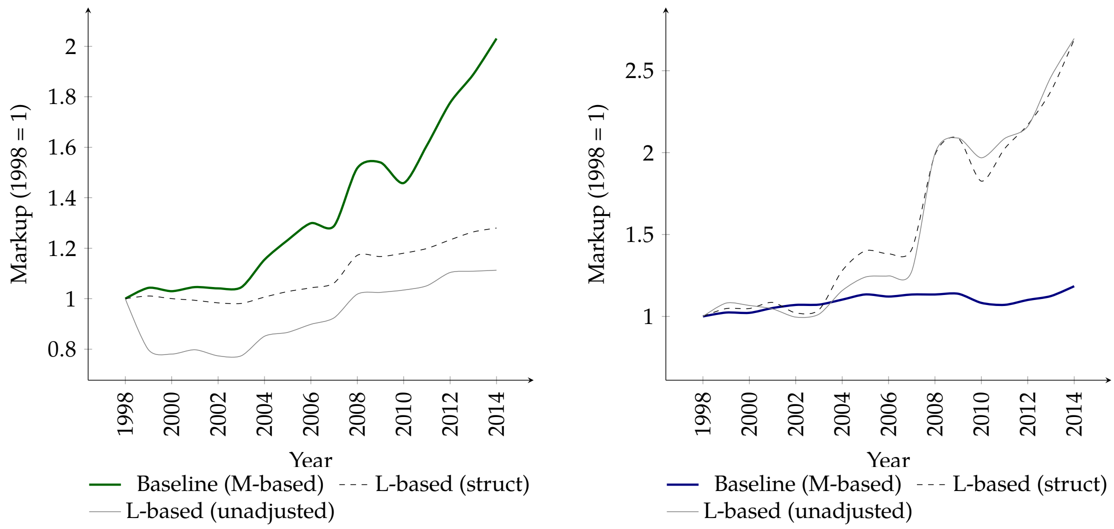
The markup *levels* of each weighting scheme is in Figure 13. The main observation of note is that the revenue-weighted aggregate markup with the Cobb-Douglas production function is below other weighting procedures.



(a) Markup (Cobb-Douglas) (b) Markup (Translog)

Figure 13: Aggregate Markup Levels with Different Weights

Figure 14 contains the indexed markup estimates when comparing the gross output production function estimation with the value-added approach.



(a) Markup (Cobb-Douglas) (b) Markup (Translog)

Figure 14: Aggregate Markup Levels with Different Flexible Input (1998=1)

D Structural Value-Added Production Function

Our baseline markup estimates take the ratio of the materials elasticity to the materials share in revenue. This is possible because we estimate production functions with materials on the right hand side. If instead we estimate value-added production functions, then the regressors are only labour and capital. De Loecker and Scott (2016) show how to estimate the markup in such a situation, with a structural value-added production function.

Consider output with a Leontief structure with a production function of capital and labour, and the materials input:

$$Y_{it} = \min\{F(K_{it}, L_{it}, A_{it}), \beta_m M_{it}\}$$

Cost minimisation ensures that:

$$Y_{it} = \beta_m M_{it} = F(K_{it}, L_{it}, A_{it})$$

Marginal revenue of the flexible input X_{it} is equal to the flexible input price, plus the costs of employing materials inputs in equal proportion:

$$\lambda_{it} \frac{\partial F_{it}}{\partial X_{it}} = P^X_{it} + \frac{P^M_{it}}{\beta_m}$$

Computing the markup as the output price divided by the marginal cost λ_{it}, and noting that $\frac{\partial F_{it}}{\partial X_{it}} = \frac{\partial \ln F_{it}}{\partial \ln X_{it}} \frac{Y_{it}}{X_{it}}$ and $\beta_m = \frac{Y_{it}}{M_{it}}$ (from cost minimisation above), we obtain:

$$\begin{aligned}
\mu_{it} &= \frac{P_{it}}{P^X_{it} \frac{\partial X_{it}}{\partial F_{it}} + \frac{P_{itM}}{\beta_m}} \\
&= \frac{1}{\frac{P^X_{it}}{P_{it}} \frac{\partial \ln X_{it}}{\partial \ln F_{it}} \frac{X_{it}}{Y_{it}} + \frac{P^M_{it}}{P_{it}} \frac{M_{it}}{Y_{it}}} \\
&= \frac{1}{\alpha^X_{it} \left(\theta^X_{it}\right)^{-1} + \alpha^M_{it}}
\end{aligned}$$

D.1 Markups from Value-Added Production Functions

Figure 15 plots the average economy-wide markup in levels, when estimated with a structural value-added production function. As with markup estimates from gross output production functions, we see a rise in the level over time. However, the rise is greater for translog than Cobb-Douglas, which is the opposite result from our baseline estimates in Section 4. The Cobb-Douglas markup rises from around 1.06 to 1.35, while the translog markup increases from 0.82 to 2.21.

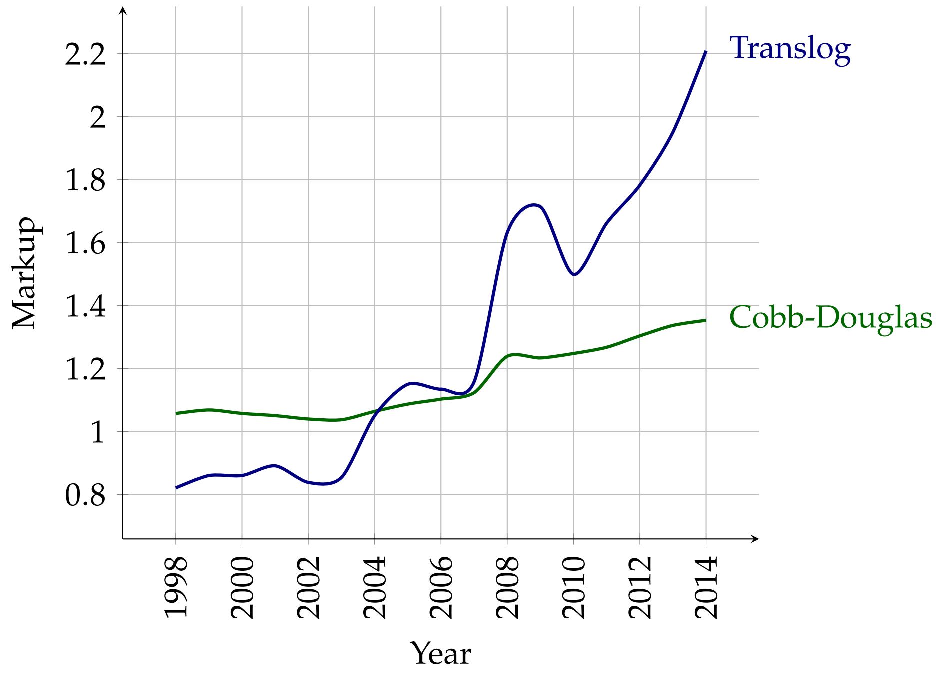


Figure 15: Structural Value-Added Aggregate Markup

Figure 16 presents sector-level markups in levels, when estimated with a structural value-added production function. As with gross output estimates, the markups in services and construction have the highest levels and drive the rise over time. Similarly, the markups levels in wholesale and motor trade are lowest. Once again, we see a greater rise in Cobb-Douglas markups due to the constancy of the estimated elasticity, while the translog markups rise less quickly.

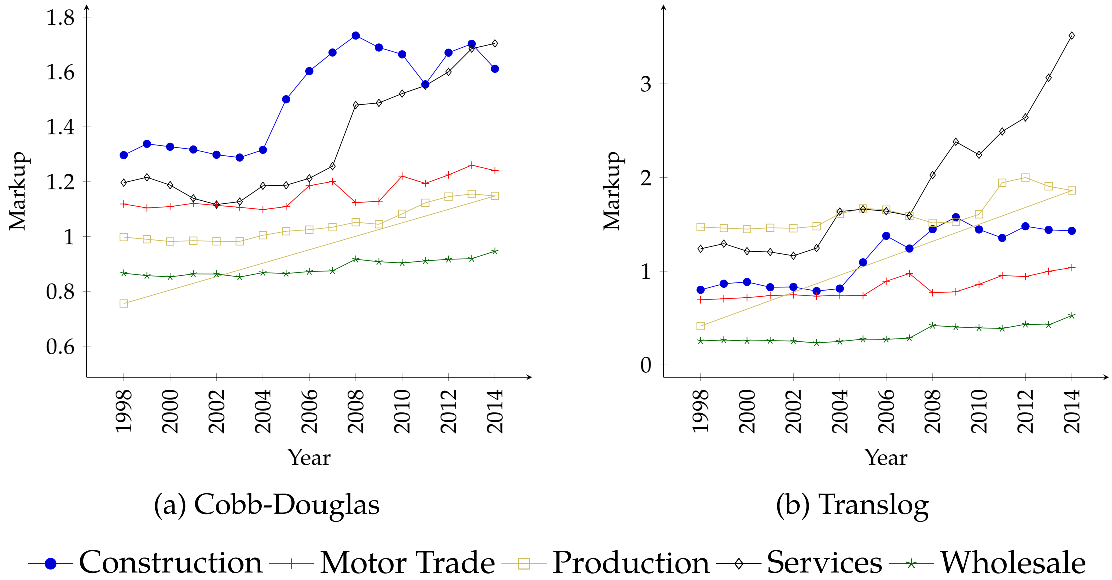


Figure 16: Structural Value-Added Markup Levels, by Sector

E Capital Stock Construction

We construct firm-level capital stocks by the Perpetual Inventory Method (PIM). Our capital construction rests on R. Martin (2002) and UKDS (2017). The PIM is formulated by

$$K_t = K_{t-1}(1-\delta) + i_t$$

where K_t implies the capital stock at year t and i_t is the investment at year t. The ARDx dataset holds the information of firm-level investments[8]. There are two issues when applying the PIM. First, the ARDx dataset does not offer the initial capital stock (K_0) which is essential for constructing the capital stock by the PIM. Second, the investment variable are missing in many firms recorded in the ARDx data.[9]

We establish firm-level capital stocks with a focus on dealing with the two issues. For the starting capital, we identify the initial year of each firm by the first year when each firm appears in the ARDx dataset. Then, we allocate industry-level aggregate capital stocks to each firm according to revenue-based weights. For the missing values

[8]Following UKDS (2017), we define the investment as the sum of expenditures on land and building, vehicle and other fixed capital, all of which are accessible from the ARDx dataset

[9]As of 2014, 66.5% of firms have either missing or zero values in the investment.

in the investment, we impute the missing investments by the linear interpolation. The details on the construction procedure of capital stock are as follows.

Step 1: Merge the ARDx data with the register panel

There is a file called 'register panel data' in the ARDx folder of UKDS secure lab. It includes a complete universe of firms (reporting units) no matter whether they are in the survey (ARDx) scope for each year. We merge the ARDx data with the register panel so that we can identify firm-year-specific observations that do not appear in the ARDx even if they are actually active in the market. Those observations are excluded in the ARDx data not because they exit the market but because they are not selected in the annual ARDx sample.

Step 2: Interpolate the missing investment

In the merged data, the investment variable becomes having even more missing values as many as the ARDx data is augmented. In other words, all of observations added from Step 1 must have missing values in the investment. We interpolate the missing investment cells on the basis of the employment variable. We use the merged data in interpolation, because by doing so we can take into account firms that are active in the market but they are excluded in the ARDx due to the ARDx' annual sampling. Unless they are not taken into account, the accumulation of the capital stock will not be done in those firms, leading to the capital stock being measured lower than the true value.

Step 3: Allocate aggregate capital stock to each firm, and generate the firm's initial capital stock

In order to apply the PIM to building the capital stock, it is necessary to know the initial capital value of each firm. We first define the initial year of each firm as the first year when a firm appears in the augmented ARDx data. Then we generate the initial capital of the firm defined as

$$\text{Initial K}_{ist} = \text{Aggregate K}_{st} \times \frac{\text{Total Observed Investment}_{st}}{\text{Aggregate Investment}_{st}} \times \frac{\text{Y}_{ist}}{\text{Aggregate Y}_{st}}$$

Suppose that the initial year of firm i is year t. Then, we allocate the aggregated capital of sector s to which firm i belongs is the firm according to revenue share of the firm in total sectoral revenue.[10] Basically, we use it as the proxy for the initial capital stock of the firm. However, it is unlikely that the aggregate capital is exhausted by the firms that are included in the ARDx dataset. For the reason, we normalize the aggregate capital by the share of total observed investment from the ARDx in aggregate capital expenditure.

[10]The ONS provides a (Stata) code by which one can establish sector-level aggregate capital stock. We use the code to get sector-level aggregate capital stock.