Recent trends in firm-level markups in the United Kingdom 2008-2019

Authors:

Diane Coyle a*    John McHale b*    Ioannis Bournakis c*    Jen-Chung Mei a,d*

a Bennett Institute for Public Policy, University of Cambridge, UK
b School of Business and Economics, University of Galway, Ireland
c SKEMA Business School and Université Côte d’Azur, France
d Westminster Business School, University of Westminster, UK

Date:

January 2024

* We thank Joshua Abramsky, Matthew Agarwala, Bart Van Ark, Jonathan Haskel, Richard Heys, Giordano Mion, Raquel Ortega-Argiles, Skevi Pericleous, Tony Venables, ONS colleagues, and participants in a Productivity Institute Brown Bag Seminar and Productivity Institute 2022 Postdoc Workshop for their helpful comments. Coyle is grateful for funding from the Productivity Institute, ESRC grant number ES/V002740/1, and Mei for funding from the Gatsby Charitable Foundation and the ONS small grant research support ONS2014988.
Recent trends in firm-level markups in the United Kingdom 2008-2019

Abstract

UK manufacturing firms have experienced sharp declines in productivity since 2008, whether measured by real revenue per worker or estimated total factor productivity. Less is known about trends in firms’ markups, which is important for understanding productivity dynamics. The estimation of markups is challenging without direct access to price and cost data, but they can be inferred using microdata on firms’ revenues and input use. In this paper we use two approaches to infer the evolution of aggregate markups for UK manufacturing firms. Both use estimates of the elasticity of substitution within industry subsectors. Our principal approach involves assumptions about the structure of competition between firms at various levels of industry aggregation, while the second tests the robustness using the influential Hall/DeLoecker approach to infer the markup based on variable input cost shares. Both approaches show large declines in estimated UK manufacturing markups since the financial crisis, estimating a decrease in industry-level gross markups of approximately eight percentage points between 2008 and 2019. There are significant contributions from both within-firm declines and declines due to reallocation. As markup declines are associated with an adverse shift in the distribution of firm-level manufacturing productivity, our results indicate that structural dynamics in manufacturing industry likely play a large part in the UK’s productivity puzzle.

Keywords: Markups, monopolistic competition, manufacturing
JEL: D22, D24, D42, D43, E24, O47

This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.
1. Introduction

In recent decades, economists have documented significant markup heterogeneity even within narrowly defined industries. As markups represent the difference between price and marginal cost, their evolution provides insights into an array of topics with important welfare implications such as market structure, consumer welfare, and productivity. Of particular interest is the unresolved inquiry into the connection between markups and productivity.

Recent evidence shows a rise in markups in the US, mainly driven by firms in the upper quantile of the markup distribution implying productivity differentials among firms (De Loecker and Eeckhout, 2021; De Loecker, Eeckhout, and Mongey, 2021). According to Autor, Dorn, Katz, Patterson, and Van Reenen (2020), the increase in markups in the US can be attributed to the expansion of highly productive 'superstar' firms after considering firm-level characteristics. A similar finding about the role of firms in the top 5% of the distribution is also documented in Calligaris et al. (2018) for EU firms during the period 2001-14.

Yet the evidence remains far from unified as there are substantial differences across countries in the empirical results regarding the trend in markups; for instance, van Heuvelen, Bettendorf, and Meijerink (2021) for the Netherlands and García-Perea, Lacuesta, and Roldan-Blanco (2021) for Spain, find stable rather than rising markups from 2006 – 2016, while Weche (2018) attributes the evidence of decreasing markups in manufacturing firms of six European countries to FDI competition. In the analysis of Weche and Wambach (2018), EU markups fell significantly during the financial crisis of 2007-08 and since then they have not returned to the pre-crisis average. Eslava and Haltiwanger (2023) find no relationship between markups and productivity dispersion at the firm level in Colombian manufacturing industry, using detailed data on input and output quantities, as well as prices. Apart from the diversified evidence regarding the trend of markups in different geographical regions, there is also disagreement about the macroeconomic and welfare interpretations of the observed trends in markups (Berry, Gaynor, and Scott Morton, 2019; Syverson, 2019; Autor, Dorn, Katz, Patterson, and Van Reenen, 2020; De Loecker et al., 2020).

In this paper, we contribute to the ongoing debate of understanding markup heterogeneity by exploring the evolution of markups in UK manufacturing since the financial crisis. Manufacturing has been identified as a major contributing sector of the
national productivity slowdown (Coyle and Mei, 2023; Coyle, McHale, Bournakis, and Mei, 2023). However, less is known about the contemporaneous changes in mark-ups. To identify the evolution of markups, one must first address systematically the challenge associated with estimating firm-level production functions in the absence of data on unit prices and quantities (Griliches and Mairesse, 1995; Klette and Griliches, 1996). The literature on markup measurement has made significant strides in advancing structural estimation techniques, employing various assumptions that enable the appropriate identification of markups (De Loecker and Warzynski, 2012; De Loecker, Goldberg, Khandelwal and Pavcnik, 2016; De Loecker, Eeckhout, and Unger, 2020; Bond, Hashemi, Kaplan, and Zoch, 2021; and Sampi, Jooste and Vostroknutova, 2021).¹

Here we build a heterogeneous-firm model in which markups can be directly recovered from an estimated firm revenue production function. The tractability of our framework lies in deriving input estimates and markups by utilising available data on deflated industry revenues and input expenditures. This approach mitigates limitations related to the frequent absence of unit prices and output volumes in firm-level data. In the context of monopolistic competition, our model yields an estimate of markup that is uniform across firms in the industry and constant over time. To derive this, we estimate the elasticity of substitution at the industry level (see, for instance, Klette and Griliches, 1996; De Loecker, 2011) incorporating CES industry aggregates of quality-adjusted firm outputs nested within a Cobb-Douglas utility function for the representative consumer, alongside a Cobb-Douglas firm-level production functions featuring heterogeneous Hicks-neutral levels of technical efficiency.

We extend our analytical framework by modifying the monopolistic competition assumption instead of either Cournot or Bertrand oligopolistic competition, which

¹ De Loecker and Warzynski (2012) develop a framework that estimates firm- or establishment-level markups by using the first-order conditions on variable factors and its output elasticity (respect to labour and materials) to revenue to derive the ratio estimator makeup (to its cost share of revenue). While it has been the most prominent framework in the literature, it also makes strong assumptions about functional forms (i.e., cost minimization and assumes that input adjustments are costless and that firms in the same industry face different input prices; also see section 2.3 for more details). Bond et al. (2021) and Sampi, Jooste, and Vostroknutova (2021) point out that the methodology is sensitive to: (1) the functional form of the production function; (2) the omission of demand shifters; (3) the absence of price information; (4) the violation of the Markov process for productivity; and (5) misspecification when marginal costs are excluded in the estimation process. Also see De Loecker, Goldberg, Khandelwal and Pavcnik (2016) for an alternative framework proposed that implements quantity and prices data at the firm-product level. Furthermore, see Raval (2023) for one of the most recent markup studies based on the Hall (1986) and De Loecker and Warzynski (2012) approach with data of Chile, Colombia, India, Indonesia, the U.S., and Southern Europe.
implies markups that are positively related to firms' shares in industry revenues, thus potentially varying across firms and over time. We also explore the robustness of our results using an alternative framework that allows markups to be derived using revenue share of variable inputs and the output elasticity of variable inputs, but without making a strong market structure assumptions (Hall, 1986; De Loecker and Warzynski, 2012). Therefore, we use both methods to infer the evolution of aggregate markups for UK manufacturing firms, finding broadly similar results.

Our findings show that under the assumption of a monopolistic market structure, industry-level gross markups have decreased by approximately eight percentage points in our demand-side approach and four percentage points in Hall/DeLoecker (i.e., the supply-side) approach from 2008 to 2019. In the markups derivation under assumptions of Cournot and Bertrand competition, results reveal that the average revenue-weighted markup for manufacturing declines by a similar amount as in the monopolistic market structure, although the results under SIC4 specification reveal non-negligible variation in markup estimates compared to those results under the more aggregated SIC2 and SIC3 specifications.

Our findings contrast with the rising markup trend in the US estimated by Autor et al. (2020) and De Loecker et al. (2020). For the US, for example, De Loecker et al. (2020) report sizeable increases in sales weighted markup estimates over the period 1950 to 2016. Identifying markups as the ratio of the variable factor’s output elasticity to its cost share in revenue in a long panel of manufacturing establishments, Foster, Haltiwanger, and Tuttle (2022) also find increased average sales-weighted markup in the US. However, they note that the size of the increase in markups declines systematically moving from 2-digit to 4-digit industry level.

Our results contribute to the literature indicating markup trends vary across countries. For example, Gibbon and Schain (2022) show relatively flat weighted average markups for European manufacturing firms from 2006 to 2016. Recent working papers by Ciapanna, Formai, Linarello, and Rovigatti (2022) and Jacob and Mion (2022) analysing markup evolution with European countries and the UK show similar dynamics. van Heuvelen, Bettendorf, and Meijerink (2021) also find relatively stable markups in the

---

2 Also see, Raval (2023) for one of the most recent markup studies based on the Hall (1986) and De Loecker and Warzynski (2012) approach with data of Chile, Colombia, India, Indonesia, the U.S., and Southern Europe.
Netherlands for the period 2006–2016. In contrast, Díez, Fan, and Villegas-Sánchez (2021) find average markups in services in advanced economies increased from 1.22 to 1.29 (in logarithms).

Methodologically, our paper is, perhaps, most closely related to Forlani, Martin, Mion, and Muûls (2022), Jacob and Mion (2022), De Ridder, Grassi, and Morzenti (2022), Foster, Haltiwanger, and Tuttle (2022). Using separate firm-level price and quantity data, Forlani, Martin, Mion, and Muûls (2022) develop a novel framework to recover heterogeneity in firm-level productivity, demand, and markups across Belgian firms. They find that firm-level productivity, demand and markups are negatively correlated. Jacob and Mion (2022) find weakening demand and decreasing quality-based productivity pushing down markups. De Ridder, Grassi, and Morzenti (2022) apply firm-level administrative production and pricing data to derive expressions for the biases that arise in the standard (quantity-based) production function and markup estimation with the assumption that firms are price makers. They find that markup estimates from revenue data are biased but remain correlated with ‘true’ markups using information on quantity and price data.

The rest of this paper proceeds as follows. Section 2 sets out the model framework and estimation methodology. Data are discussed in section 3 and estimates are presented in section 4. Section 5 discusses how our findings contribute to this literature and concludes.

2. Estimation framework

2.1 Firm-level revenue function

In this section we follow the method of our earlier paper, where more detail can be found (Coyle et al., 2023). The representative consumer has a Cobb-Douglas utility function over an index of manufactured goods, \( Z_t \), and an index of services, \( X_t \):

\[
U_t = Z_t^\alpha X_t^{1-\alpha}.
\]

(1)

Since utility is assumed to be homothetic, we can sum (1) over consumers to get the aggregate output index, \( Y_t \), and can define the aggregate price index, \( P_t \), such that \( P_t Z_t + P_t X_t = P_t Y_t \). The prices of a unit of the \( Z_t \) index and a unit of the \( X_t \) index are \( P_{zt} \) and \( P_{xt} \) respectively. Maximising their utility, the representative consumer allocates their nominal income over the two aggregates, to yield expenditure shares:

\[
P_{zt} Z_t = \alpha P_t Y_t
\]

(2)
\[ P_{Xt}X_t = (1 - \alpha) P_t Y_t. \]  
where \( P_t Y_t \) is nominal income.

We next derive the demand curves facing individual firms. Aggregate output is a CES function of the quality-adjusted goods produced by the \( N \) firms in the industry:

\[ Z_t = \left[ \sum_{i=1}^{N} (A_{it} Q_{it})^{\eta} \right]^{\frac{1}{\eta-1}}, \]

where \( A_{it} \) is a measure of the quality of the good produced by firm \( i \) at time \( t \) (where quality improvements result from changes to product characteristics that are valued by consumers), \( Q_{it} \) is the volume output produced by firm \( i \) at \( t \) and \( \eta \) is the elasticity of substitution between the \( N \) goods in the output index. We thus incorporate both a representative consumer with a preference for variety and vertical differentiation based on quality between products that enter into the industry output index. We denote quality-adjusted output as \( Q_{it}^* = \Lambda_{it} Q_{it} \).\(^3\) We assume that \( \eta > 1 \) and that each firm produces a single product variety. Given the allocation of income to manufacturing goods, we can derive the demand function facing a given firm producing a good with quality \( A_{it} \) as:

\[ Q_{it} = \Lambda_{it}^{-\eta} \left( \frac{P_{it}}{P_{Zt}} \right)^{-\frac{\eta}{\eta-1}} Z_t \]

\[ = \Lambda_{it}^{-\eta} \left( \frac{P_{it}}{P_{Zt}} \right)^{-\frac{\eta}{\eta-1}} \left( \frac{P_t Y_t}{P_{Zt}} \right)^{\frac{1}{\eta-1}}, \]

where the price index for the industry, \( P_{Zt} \), is given by:

\[ P_{Zt} = \left[ \sum_{i=1}^{N} \left( \frac{P_{it}}{A_{it}} \right)^{-\frac{\eta}{\eta-1}} \right]^{\frac{1}{\eta-1}}. \]

From (6), we can see that quality improvements are reflected in a lower industry price index. Moreover, the effect of a change in quality on the cost of achieving a particular level of \( Z_t \) is equivalent to a price change of equal proportion but opposite in sign.\(^4\)

---

\(^3\) Quality change thus enters the utility function in a “better is more” form (for a related analysis in the context of combining different vintages of capital in a capital aggregate, see Fisher (1965) and Hulten, (1992). Fisher and Shell (1972, Essay I) is a classic reference on the restrictions on utility functions required for quality change to be represented as the equivalent of “repackaging” and thus a “better is more” formulation. Muellbauer (1975) identifies the importance of homothetic preferences for this formulation. Deaton and Muellbauer (1980) discuss the relationship between hedonic methods of measuring quality change and utility-based measures of quality change (see also, Fisher and Griliches, 1995). We make the strong assumption that quality changes are accurately captured in the calculation of industry price indexes and that relative as well as absolute quality changes are captured. In addition to being reflected in the price index, the key requirement is that the quality change acts as a shift factor for the representative consumer’s inverse demand function for the good.

\(^4\) Fisher and Shell (1972) consider the case where a quality improvement for a given good affects the utility of other goods – for example, improvements in the quality of refrigerators also affects the utility from consuming ice cream. They show that where the ‘qualities’ of other goods are affected, the correct accounting for the effect of the initial quality change on the cost of living (here the cost of achieving a given
Turning to production side of the economy, we assume each firm \( i \) has the Cobb-Douglas production function technology:

\[
Q_{it} = \Omega_{it} L_{it}^{\beta_l} K_{it}^{\beta_k} M_{it}^{\beta_m},
\]

where \( \Omega_{it} \) is a (firm-specific) measure of Hicks-neutral technical change, \( L_{it} \) is labour, \( K_{it} \) is fixed capital and \( M_{it} \) is materials.

In order to derive the revenue function, we write the demand function (5) in inverse form as:

\[
\frac{P_{it}}{P_{Zt}} = A_{it}^{\frac{\eta-1}{\eta}} Q_{it}^{\frac{1}{\eta}} \bigg( \frac{P_{Zt}}{P_{Zt}} \bigg)^{\frac{1}{\eta}} L_{it}^{\frac{\eta-1}{\eta} \beta_l} K_{it}^{\frac{\eta-1}{\eta} \beta_k} M_{it}^{\frac{\eta-1}{\eta} \beta_m},
\]

where the quality indicator, \( A_{it} \), is a shift factor for the inverse demand function. As noted above, such shifts in quality can reflect relative as well as absolute changes in quality that correspond to changes in the representative consumer's marginal willingness to pay.

Using (2), (7) and (8), total deflated firm revenue is:

\[
R_{it} = \frac{R_{it}}{P_{Zt}} = \frac{P_{P_{Zt}}}{P_{Zt}} = (A_{it} \Omega_{it})^{\frac{\eta-1}{\eta}} L_{it}^{\frac{\eta-1}{\eta} \beta_l} K_{it}^{\frac{\eta-1}{\eta} \beta_k} M_{it}^{\frac{\eta-1}{\eta} \beta_m},
\]

where industry revenue is \( R_{Zt} = P_{Zt} Z_t = \alpha P_t Y_t \).

From (9), total revenue varies with the increased use of factors of production for two reasons. First, an increase in the use of a factor of production (say labour) leads to an increase in physical output; and second, the firm must lower its price to sell this increased level of output given that it faces a downward sloping demand curve. The coefficient on each input is the revenue elasticity of the input, \( (\eta - 1/\eta) \beta_f \) for \( f \in \{ l, k, m \} \), where the revenue elasticity will be lower than the output elasticity given our assumption that \( \eta > 1 \).

Taking natural logs of (9) and rearranging we obtain:

\[
r_{it} - p_{Zt} = \frac{1}{\eta} (r_{Zt} - p_{Zt}) + \frac{(\eta-1)\beta_l}{\eta} L_{it} + \frac{(\eta-1)\beta_k}{\eta} K_{it} + \frac{(\eta-1)\beta_m}{\eta} M_{it} + \frac{\eta-1}{\eta} (\lambda_{it} + \omega_{it})
\]

where lower case letters represent the natural log of a variable. A critical feature of (10) is that identification of \( \eta \) is possible from the estimated coefficient on the deflated-industry-revenue variable in the estimated revenue equation (Griliches and Klette, 1996).

\( Z_t \) will require adjustments in the equivalent prices of the other goods affected (here adjustments in the relevant "quality" levels of the other goods affected). This will also apply where the quality change for one good causes a reduction in the utility from other goods. For example, when the improvement in the quality of one brand of ice cream reduces the utility from the unimproved brands that are also being consumed. A change in a particular good's \( A_{it} \) captures relative as well as absolute changes in quality.
2.2 Estimating markups

This section derives the variable firm-level markups under the assumptions of both Cournot and Bertrand competition (see Atkenson and Burstein, 2008; Burstein et al., 2020; and the useful exposition of the solution technique in Forlani et al., 2022). We have shown in Coyle, McHale, Bournakis, and Mei (2023) how it is possible, under our functional form assumptions, to derive estimates of firm-level quality-adjusted total factor productivity (TFP') from the estimated revenue function outlined in Eq. (10), where the latter depends on being able to obtain an estimate of the elasticity of substitution, $\eta$, from the deflated industry revenue. Up to this point, however, we have not made any assumption about the market structure under which firms are operating and therefore have said nothing about how firms set mark up prices over their marginal costs. We thus next extend the framework in Coyle et al. (2023) to estimate firm-level markups, considering three possibilities: monopolistic competition, and Cournot and Bertrand competition.

2.2.1 Monopolistic competition

The markup behaviour is simple under monopolistic competition (M). Given that firms set prices and quantities taking industry prices and quantities as given, the (absolute value of the) price elasticity of demand is equal to $\eta$, i.e., the negative of the elasticity of substitution. In our framework, the markup in a monopolistically competitive market (M) is common across firms and constant over time, and is given by the standard formula:

$$\mu^M = \frac{\eta}{\eta - 1}$$  \hspace{1cm} (11)

2.2.2 Cournot competition

Under Cournot competition (C), firms will consider the effect of their output choices on the industry’s output. We adapt method for solving for the Cournot equilibrium in Atkeson and Burstein (2008) for our particular nested form of the demand function (i.e., CES industry aggregates nested within a Cobb-Douglas utility function), to obtain firm- and time-varying markups.
We assume that an industry is in a Cournot equilibrium with a given set of firm outputs (and associated firm prices and revenue shares). Under Cournot, each firm makes its output choice taking the output choices of the other firms in the industry as given; in contrast to monopolistic competition, each firm is sufficiently large taking into account the effect of its output choice on aggregate industry output. This effect on industry output must be taken into account in identifying the firm’s perceived price elasticity of demand and consequent optimal markup.

Although each firm is sufficiently large in relation to the respective industry, we assume that each firm is sufficiently small in comparison to the total economy, hence no effect on aggregate total output ($Y_t$) and the aggregate total economy price index ($P_t$). Using that $P_{Zt}Z_t = \alpha P_t Y_t$, we substitute the industry price index out of the inverse demand curve (Eq. 8):

$$\frac{P_{it}}{P_t} = \Lambda_{it}^{\frac{n-1}{n}} Q_{it}^{\frac{1}{n}} Z_t^{\frac{1}{n}} Y_t.$$

Using Eq. (4) we can write the elasticity of the industry output index with respect to firm output choice as:

$$\frac{dZ_t}{dQ_{it}} Z_t = \left[ \sum_{i=1}^{N} (\Lambda_{it} Q_{it})^{\frac{n-1}{n}} \right]^{-1} (\Lambda_{it} Q_{it})^{\frac{n-1}{n}}.$$

Multiplying both sides of Eq. (8) by $Q_{it}/Z_{it}$, we obtain the revenue share of firm $i$, $s_{it}$, in industry revenue as:

$$s_{it} = \frac{P_{it} Q_{it}}{P_{Zt} Z_t} = \left[ \sum_{i=1}^{N} (\Lambda_{it} Q_{it})^{\frac{n-1}{n}} \right]^{-1} (\Lambda_{it} Q_{it})^{\frac{n-1}{n}}.$$

Comparing (13) and (14) we therefore have:

$$s_{it} = \frac{dZ_t}{dQ_{it}} Z_t.$$

Taking natural logs of (12) and rearranging we have:

$$p_{it} = p_t + \left( \frac{n-1}{n} \right) \lambda_{it} - \frac{1}{n} q_{it} + \frac{1}{n} z_t + \ln \alpha + y_t - z_t,$$

where $\lambda_{it} = \ln \Lambda_{it}$. Now taking the partial derivative of $p_{it}$ with respect to $q_{it}$ (and taking $p_t$, $y_t$, $\lambda_t$ and $\alpha$ as given), we obtain:

$$\frac{dp_{it}}{dq_{it}} = -\frac{1}{n} \left( 1 - \frac{dZ_t}{dQ_{it}} \right) - \frac{dZ_t}{dq_{it}}$$

$$= -\frac{1}{n} (1 - s_{it}) - s_{it}.$$

---

5 See Vives (1999) for sufficiency conditions for the existence and uniqueness of equilibria under Cournot and Bertrand competition in markets with differentiated products.
Taking the inverse of Eq. (17) and multiplying through by $-1$ gives an expression for the absolute price elasticity of demand:

$$\epsilon_{it} = \frac{dq_{it}}{dp_{it}} = \frac{1}{\eta(1-s_{it}) + s_{it}}. \quad (18)$$

Finally, using the standard formula for the (gross) markup, we obtain a formula for the variable markup under Cournot competition:

$$\mu_{it}^C = \frac{\epsilon_{it}}{\epsilon_{it} - 1} \left( \frac{1}{\eta} \right) \left( \frac{\eta}{\eta - 1} \right). \quad (19)$$

where $\epsilon_{it}$ is the price elasticity of demand under Cournot and $s_{it}$ is firm $i$’s share of industry revenue. A firm’s markup under Cournot is therefore a multiple of the markup under monopolistic competition, with the size of the multiplier intuitively depending on the firm’s revenue share in total industry revenue. Higher firm revenue shares are thus associated with higher firm markups both in the cross section of industries and over time. Improvements in product quality (i.e., a higher $\Lambda_{it}$) or improvements in technical efficiency (i.e., a higher $\Omega_{it}$) will increase a firm’s revenue share, and thus be associated with increases in the markup as well as increases in $\text{TFPQ}^*$. 

Crucially, given that the markup depends only on the elasticity of substitution between products in an industry and the revenue shares, we can calculate the distribution of markups using our estimate of the relevant industry elasticity of substitution and the revenue shares that are available in our data.

2.2.3 Bertrand competition

Finally, turn to a Bertrand equilibrium with a given set of firm prices (and associated firm outputs and revenue shares). Under Bertrand (B), each firm makes its pricing choice taking the price choices of the other firms in the industry as given; however, each firm is assumed to be sufficiently large that it takes into account the effect of its pricing decision on the aggregate industry price index. To model a Bertrand firm’s markup behaviour, this effect on the industry price index must be taken into account to identify their perceived price elasticity of demand, and consequent choice of optimal markup.

We again assume that each firm is sufficiently small in relation to the total economy that it ignores any effect on aggregate total output ($Y_t$) and the aggregate total economy price index ($P_t$). Using Eq. (8), we can write the elasticity of the industry price index with respect to firm price as:
\[
\frac{dP_{zt} P_{it}}{dP_{it} P_{zt}} = \left[ \sum_{i=1}^{N} \left( \frac{P_{it}}{V_{it}} \right)^{1-\eta} \right]^{-1} \left( \frac{P_{it}}{V_{it}} \right)^{1-\eta}.
\]  

(20)

Dividing both sides of Eq. (5) by \( Z_t \) and multiplying the result by \( P_{it}/P_{zt} \), we obtain the revenue share of firm \( i \), \( s_{it} \), in industry revenue as:

\[
s_{it} = \frac{P_{it} Q_{it}}{P_{zt} Z_t} = \left[ \sum_{i=1}^{N} \left( \frac{P_{it}}{V_{it}} \right)^{1-\eta} \right]^{-1} \left( \frac{P_{it}}{V_{it}} \right)^{1-\eta}.
\]  

(21)

Comparing (20) and (21) we therefore have:

\[
s_{it} = \frac{\partial Z_t P_{it}}{\partial P_{it} P_{zt}}.
\]  

(22)

Taking natural logs of Eq. (5):

\[
q_{it} = (\eta - 1) \lambda_{it} - \eta p_t + (\eta - 1) p_{zt} + ln\alpha + p_t + y_t
\]  

(23)

Now taking the partial derivative of \( q_{it} \) with respect to \( P_{it} \) (again assuming given values of \( p_t, y_t, \lambda_t \) and \( \alpha \)) and multiplying through by \(-1\), we obtain an expression for the absolute price elasticity of demand under Bertrand competition:

\[
\xi_{it} = -\frac{\partial q_{it}}{\partial P_{it}} = \eta (1 - s_{it}) + s_{it},
\]  

(24)

where use is made of Eq. (11) to substitute for the elasticity of the industry price index with respect to the firm’s choice of price. Finally, again using the standard formula for the (gross) markup, we obtain a formula for the variable markup:

\[
\mu_{it}^B = \frac{\xi_{it}}{\xi_{it-1}} = \left( \frac{1}{1 - s_{it}} \right) \left( 1 - s_{it} \frac{(\eta - 1)}{\eta} \right) \left( \frac{\eta}{\eta - 1} \right),
\]  

(25)

where \( \zeta_{it} \) is the absolute value of the price elasticity of demand under Bertrand for firm \( i \) in time \( t \).

### 2.2.4 Comparing monopolistic, cournot, and Bertrand competition

Given a positive revenue share \( s_{it} > 0 \) and an elasticity of substitution greater than one \( (\eta > 1) \), it follows that:

\[
1 < \frac{1 - s_{it} \left( \frac{\eta - 1}{\eta} \right)}{1 - s_{it}} < \frac{1}{1 - s_{it}}.
\]  

(26)

which we can again calculate based solely on an estimate of the relevant elasticity of substitution and the revenue shares. For any positive revenue share, we can see that the markup under Bertrand will be lower than under Cournot but higher than under monopolistic competition. For any given revenue share, \( s_{it} \), we therefore have the following ordering of the implied markups:

\[
\mu^M < \mu_{it}^B < \mu_{it}^C.
\]  

(27)
Given $\eta > 1$, it is also easy to confirm from Eqs. (8) and (14) that the markup is more responsive to changes in the revenue share under Cournot than under Bertrand (where, of course, the markup is unresponsive to the revenue share under monopolistic competition):

$$
\frac{\partial \mu^C_{it}}{\partial s_{it}} = \frac{1}{(1-s_{it})^2} \left( \frac{\eta}{\eta-1} \right) > \frac{\partial \mu^B_{it}}{\partial s_{it}} = \frac{1}{\eta(1-s_{it})^2} \left( \frac{\eta}{\eta-1} \right).
$$

However, given that there are many firms in manufacturing industry and so the market share of each is expected to be small, markups will be similar in monopolistic and Bertrand competition. We therefore present the results based on Monopolistic and Cournot competition; the Bertrand results can be found in Appendix I.

### 2.3 Alternative approach to estimating markups

When it comes to estimating markups, there are two main methods: demand-side and supply-side approaches. The former is the most commonly used and involves, as we have done above, defining consumer utility functions and making assumptions about the market structure of firms to estimate markups (also see, Berry, Levinsohn, and Pakes 1995; Goldberg 1995; Feenstra and Weinstein 2010). The alternative is the supply-side (or cost production) approach, pioneered by Hall (1986) and more recently employed by De Loecker and Warzynski (2012), De Loecker et al. (2020), De Loecker and Eeckhout (2020), and Díez, Fan, and Villegas-Sánchez, (2021). While the latter does not require assumptions about preferences or market structure, it relies on cost minimization and assumes that input adjustments are costless and that firms in the same industry face different input prices.

Based on De Loecker and Warzynski (2012), the markup estimate can then be derived from the first order conditions for a single input in a cost function with generally positive fixed costs. The firm-level variable markup is then the ratio estimator of the output elasticity for a flexible input to that input’s cost share in total revenue across firms in the industry:

$$
\mu^R_{it} = \beta_{it}^p \left( \alpha_{it}^p \right)^{-1}.
$$

where $\beta_{it}^p$ is the output elasticity of the variable input, and $\alpha_{it}^p$ is the revenue share of variable input. Hence, with our revenue function, we can infer estimates of the implied output elasticities of inputs (materials) given our estimate of the elasticity of substitution.
to derive the ratio estimates at the firm level. This alternative way to generate markups serves as a robustness check on our results from demand-side estimation.

2.4 Estimation strategy

To reduce concerns related to unobserved heterogeneity and simultaneity that are common in the revenue production function literature (see, e.g., Griliches and Marisse, 1996), we allow for the possibility of adjustment costs in the setting of all inputs, productivity shocks that are serially correlated (which we model as AR(1)), and unobserved heterogeneity in productivity across firms. Letting $\theta_{it} = \frac{(\eta-1)}{\eta}(\lambda_{it} + \omega_{it})$, we assume $\theta_{it} = \theta_i + \nu_{it}$, where $\nu_{it} = \rho \nu_{it-1} + \xi_{it}$ and $\xi_{it}$ is a zero mean random shock that is potentially correlated with input choices, assuming $0 < |\rho| < 1$. Lagging (10) by one period, multiplying the resulting equation through by $\rho$, and subtracting the result from (10) gives the quasi-differenced equation:

$$r_{it} - p_{zt} = \rho(r_{it-1} - p_{zt-1}) + \frac{1}{\eta}(r_{zt} - p_{zt}) - \rho(r_{zt-1} - p_{zt-1}) + \frac{(\eta-1)\beta_l}{\eta}(l_{it} - \rho l_{it-1}) + \frac{(\eta-1)\beta_k}{\eta}(k_{it} - \rho k_{it-1}) + \frac{(\eta-1)\beta_m}{\eta}(m_{it} - \rho m_{it-1}) + (1 - \rho)\theta_i + \xi_{it}.$$  

(30)

As is well known, the presence of firm fixed effects leads to a correlation between the lagged dependent variable and the error term $\xi_{it}$ (Nickell, 1981). Input variables in the revenue equation will also be correlated with the error term where there are contemporaneous input responses to productivity shocks. One option for consistently estimating (30) is to take first differences and to instrument for the potentially endogenous right-hand-side variables. Blundell and Bond (1998, 2000) identify relatively mild initial conditions that allow lagged levels of the endogenous variables to be valid instruments for the endogenous first differences. However, Blundell and Bond (2000) also find that the lagged levels are weak instruments in a production-function-estimation setting. Alternatively, they suggest estimating a System GMM that includes the estimating equation in first differences and that equation in

---

6 Note that the ratio approach (or production approach) is referred in Bond et al. (2021) and Foster, Haltiwanger, and Tuttle (2022).

7 See Bond and Söderbom (2005).
levels. Given its documented good performance in production function estimation, we therefore adopt the Blundell-Bond System GMM estimator for estimation of the revenue function in order to obtain consistent estimates of $\eta$.\(^8\)

2.5 Markup growth decomposition

At the industry level, we first show the evolution of an index of revenue-share-weighted markup that is set equal to 1.0 in the first year of our sample. Moreover, following the approach of De Loecker et al. (2020), we decompose the evolution of that index into the product of an index of within-firm markup, an index of reallocation effects, and an index of entry and exit effects.

Letting this index be $x_t$ (measured in logs), we express it as a revenue-share-weighted average of the corresponding firm-level measures:

$$x_t = \sum_i x_{it} s_{it},$$

where $s_{it}$ is the share of firm $i$ in the total revenue of the industry in time $t$. Using the DeLoecker et al. (2020) decomposition, we can write the growth rate of the aggregate (approximated as the log difference) as sum of a number of components:

$$\Delta x_t = \sum_i \Delta x_{it} s_{it-1} + \sum_i \hat{x}_{it-1} s_{it} + \sum_i \Delta x_{it} \Delta s_{it} + \sum_{i \in \text{Entry}} \hat{x}_{it} s_{it} + \sum_{i \in \text{Exit}} \hat{x}_{it-1} s_{it-1},$$

where $\hat{x}_{it} = x_{it} - x_{t-1}$ and $\hat{x}_{it-1} = x_{it-1} - x_{t-1}$.\(^9\)

The first term on the right is the effect of within-firm productivity growth on the aggregate growth rate. The next two terms capture the reallocation effects between firms in the industry, and the final two terms capture the effects of firm entry and exit. We term the sum of the second two terms the reallocation effect and of the final two terms the entry/exit effect. The final four terms can be collectively thought of as a broad reallocation effect. Finally, setting the relevant level of the index equal to 1 in the first year of the sample, we use the relevant calculated weighted growth rates to infer the evolution of the level of the index over the sample. We present these index evolutions at the sector level.

---

\(^8\) As demonstrated in Diane et al. (2023), the results are not dependent on the method we employed and are robust to the ACF control function method (Ackerberg, Caves, and Frazer, 2015).

\(^9\) Following Haltiwanger (1997) and De Loecker et al. (2020), we demean by the appropriate aggregate (revenue weighted) level in order to correctly identify the role of the reallocation term.
3. Data
We construct a firm-level dataset that includes non-financial business firms in the UK in the Annual Business Survey, (ABS) covering the period 2008–2019 (to exclude the pandemic period). The ABS covers approximately two-thirds of the UK non-financial businesses, including firms’ revenue, employment costs, capital expenditure and intermediates purchases.

We use the smallest local unit in the data – the firm-level.\textsuperscript{10} We checked for duplication and removed 94 units from the sample. Building on the findings of Coyle and Mei (2023), we focus on firms in manufacturing as one of the industries that account the most for the post-2008 productivity slowdown. This gives us an unbalanced panel of data with 148,962 observations throughout the period 2008-2019.

For each firm, there is data on total revenue, total employment, capital stock, and purchases of inputs. As all values are in nominal terms, we employ the 2-digit industry-level ONS producer output price deflator and input price indices (manufacturing PPI and non-manufacturing SPPI) and annual estimates of gross and net capital stocks and consumption of fixed capital in the UK to adjust (deflate) the nominal value at 2015 prices (in £ thousand). We also construct firm-level capital stocks using the Perpetual Inventory Method (PIM).\textsuperscript{11}

4. Results
4.1 Markups under different market structures
We estimate firm-level markups based on Eqs. (11) and (19). We estimate our model specification using both SIC2- and SIC3-digit industry data to obtain industry revenue elasticities and then back out the firm-level markups. As firm revenue shares ($s_{it}$) are small at both 2 and 3-digit levels (even at the SIC3-digit level the revenue share is only

\begin{footnotesize}
\textsuperscript{10} Following Harris and Robinson (2002), we proceed at the firm-level by using the Annual Business Survey (ABS) 2008–2019, also used by other studies (Oulton, 1997; Griffith, 1999; Harris, 2002; Harris, 2005a; Harris and Moffat, 2015; and Harris and Moffat, 2017). While the establishment unit is also available, an establishment is not an economic unit but an accounting unit that often gains and loses local units as enterprises respond to ONS requests for information (Harris and Robinson, 2002). As companies open and/or close, or buy and/or sell firms, the number of local units covered by an establishment, and firms’ sizes and vintages, would change over time, which makes it difficult to undertake certain types of analysis in an economically meaningful way. This issue has been highlighted in Harris and Drinkwater (2000), Harris and Robinson (2002), and Harris (2002) in which the authors provide evidence on how unstable establishments, in terms of compositional changes, are over time.

\textsuperscript{11} See Coyle, McHale, Bournakis, and Mei (2023) and Martin (2002), and Harris and Moffat (2017) for more details.
\end{footnotesize}
around 0.00098 on average, with standard deviation 0.0092), so the difference in markups for each firm (under any assumption about competition) will be small. Hence, to provide more sensible analysis, we also estimate our model specification using detailed SIC4-digit industry data. Note that when moving from the SIC2-digit to the SIC4-digit level, we extend from 38 broad SIC2 sectors to 385 subsectors.13

Summary statistics are provided in Table 1. A clear pattern emerges. First, the table shows that average markup estimates are around 1.080 to 1.083 under SIC2, 1.056 and 1.063 under SIC3, and 1.046 to 1.063 under SIC4. These numbers vary with different market structure assumptions; in all cases, the monopolistic competition markup (i.e., time-invariant within industry) is the lowest, whereas the Cournot assumption gives the highest.

As flagged by Autor et al. (2020) and De Loecker et al. (2020), however, changes in aggregate markups are due to changes in unweighted markups and reallocation of economic activities, and so the average numbers do not fully capture the underlying distributional change. As in Haltiwanger (1997) and De Loecker et al. (2020), the unweighted markup change captures the average change attributed to a change in revenue-weighted markup while keeping the market shares (\(s_{it-1}\)) unchanged from last period, but the reallocation then captures the change in market share (\(\Delta s_{it}\)) and markup across firms (\(\Delta x_{it}\)). As noted in De Loecker et al. (2020), we also include two additional

<table>
<thead>
<tr>
<th>Panel A SIC2</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Eta</td>
<td>0.074</td>
<td>0.028</td>
<td>148,962</td>
</tr>
<tr>
<td>(\mu^M)</td>
<td>1.080</td>
<td>0.032</td>
<td>148,962</td>
</tr>
<tr>
<td>(\mu^C)</td>
<td>1.083</td>
<td>0.034</td>
<td>148,962</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B SIC3</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Eta</td>
<td>0.051</td>
<td>0.037</td>
<td>148,962</td>
</tr>
<tr>
<td>(\mu^M)</td>
<td>1.056</td>
<td>0.040</td>
<td>148,962</td>
</tr>
<tr>
<td>(\mu^C)</td>
<td>1.063</td>
<td>0.057</td>
<td>148,962</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C SIC4</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Eta</td>
<td>0.036</td>
<td>0.031</td>
<td>148,962</td>
</tr>
<tr>
<td>(\mu^M)</td>
<td>1.046</td>
<td>0.063</td>
<td>148,962</td>
</tr>
<tr>
<td>(\mu^C)</td>
<td>1.063</td>
<td>0.099</td>
<td>148,962</td>
</tr>
</tbody>
</table>

Notes: This table reports summary statistics through levels for each industry.

---

12 We provide the distribution figures for the revenue weights between \(t\) and \(t-1\) in Appendix II Figures AII 1 and AII 2. We trim all right tail observation (i.e., revenue weight less than 0.000001 and 0.000002, respectively) in order to show the clear pattern (skewed to the left) through the weights.

13 Foster, Haltiwanger, and Tuttle (2022) refer “less detailed” and “more detailed” estimates to “2-digit” and “4-digit” level information.
terms to capture firms’ entry and exit (see Eq. 34). Decomposing the average into within and reallocation effects helps us to understand better if firms’ market power is changing over time. Figure 1 plots the cumulative growth rates of average markups, taking into account separately firms’ entry, exit, and reallocation (we set the initial level at 1 in 2008 throughout).

Panel A SIC2

Panel B SIC3

Panel C SIC4

Figure 1. Revenue weighted Cumulative Markups Growth Overtime - Manufacturing

Notes: Number of observations are reported in Table 1.
Based on the results under monopolistic competition, we find that average revenue-weighted markups in manufacturing declined from 2008, but with a rise in 2010 and 2011. This rise is due to the reallocation term. After 2011, we find that markups once more decline over time again driven mainly by the reallocation term.

The within term is flatter. For the case of monopolistic competition, it is flat is by assumption, as firms within the same industry set prices and quantities taking industry prices and quantities as given, which leads the markup in a monopolistically competitive market to be common across firms and constant over time. This finding is similar to Foster, Haltiwanger, and Tuttle (2022) who find that the reallocation and net entry terms dominate the change (increase) in markups in the US firms.\footnote{By contrast, under the Bertrand competition (see, Figure AI 3 in Appendix I), the flat within term is due to a small revenue share across each firm. See also Figures AI 1 and 2 in Appendix I where we provide the distribution figures for the revenue weights between t and t-1.}

Figure 1 the suggests that UK manufacturing firms experienced a significant downturn in markups after 2011, with a substantial reduction in productivity among productive firms. Additional insight is provided by the density distribution of the unweighted TFP measures for 2008 and 2019, shown in Figure 2 for manufacturing. We find that there is a leftwards shift in the TFP distributions in 2019 (grey) compared to 2008 (red) in manufacturing. This adverse shift has occurred despite the flat-to-positive contribution to markups of net entry shown in Figure 2.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{manuf_TFPQ_dist.png}
\caption{The Shift of Overall TFPQ* Distribution in 2008 and 2019}
\textit{Notes:} The x-axis represents here ln TFPQ*.
\end{figure}
### 4.2 Robustness analysis: Markups under no market structure assumptions

Figure 3 shows results based on the alternative supply-side approach. The trends are broadly consistent with our demand side approach. First, we find that the within-term changes from relatively constant at the SIC2 level to a more dynamic one at SIC4. Second, the reallocation contribution term again increases between 2010 and 2012. Third, the firm net entry term also exhibits a similar contribution to the overall markup growth.

For the SIC2 level, we find that the within-term changes from relatively constant to more dynamic between 2010 and 2012. The reallocation contribution term and the firm net entry term also exhibit a similar contribution to the overall markup growth.

Panel A SIC2

Panel B SIC3

Panel C SIC4

Figure 3. Revenue weighted Ratio Markups Growth Overtime

Notes: The number of observations is reported in Table 1.

Fourth, the overall markup decreases in a similar way but with a less steeply declining trend compared to our demand approach; throughout the period 2008-2019, the ratio markups fall from 1 in 2008 to below 0.98 in 2019, whereas our demand-side approach above finds the markups fall from 1 in 2008 to 0.90 in 2019. One potential explanation for the difference between the two sets of estimates is that the implied output elasticity of variable input ($\beta_{it}^v$) with revenue share of variable input ($\alpha_{it}^v$) is used in the ratio markup, whereas the elasticity of industry revenue ($\eta$) with firm-level market share ($s_{it}$) is used in our demand approach’s markups through market assumptions. As different
variables are employed to capture markup dynamics, the slight differences are in line with expectations. Nevertheless, the overall trends captured here are broadly similar to our previous results, and these consistent trends across these results suggest the robustness of the estimates for the manufacturing sector.

5. Discussion
By the end of 2019, the level of aggregate labour productivity in the UK was about a fifth lower than it would have been if the 1990-2007 trend had continued (ONS). There has been surprisingly little consideration of trends in markups in this context. We provide estimates of markups for UK manufacturing firms since the financial crisis using techniques that are consistent with recent approaches to the measurement of productivity trends over this period. Manufacturing was selected as previous work has indicated its declining productivity growth has been a significant contributor to the overall productivity slowdown.

We find that in all cases estimated average revenue-weighted markups trended down from 2008 in UK manufacturing, although rising temporarily in 2010 and 2011 due to reallocation effects (reallocation among forms in the sector and net entry). While the within component has trended down, the reallocation between firms has been the main driver of the dynamics.

Our results contrast with results for the US that generally find rising markups, in manufacturing as well as other industries. The decline in markups in UK manufacturing is aligned with broader evidence highlighting the challenges and weaknesses within the UK manufacturing sector.

The literature is divided both in terms of the implications of markup trends for productivity and in terms of results as between the US and Europe/UK. As discussed above, most of the US results find rising markups and link this to higher productivity (including quality improvements) driving markups and market power. There is a mixed picture for European economies, but broadly the results suggest either lesser increases in markups, or declines. In this paper we find a large decrease in industry-level gross markups of approximately eight percentage points between 2008 and 2019, associated with declining firm-level productivity, for UK manufacturing. The constellation of results from this literature, including this contribution, point to underlying structural changes such as intangibles investments and/or trade and investment dynamics across OECD
economies as drivers of both markups and productivity trends. Our robust finding of declining markups, using different assumptions and methods, and combined with evidence on an adverse shift in firm-level total factor productivity, puts dynamics within UK manufacturing at the heart of the country’s productivity puzzle.
References


22


Forlani, E., Martin, R., Mion, G. and Muûls, M., 2022. Unraveling firms: Demand, productivity and markups heterogeneity.


Goodridge, P. and Haskel, J., 2022. *Accounting for the slowdown in UK innovation and productivity* (No. 022).


Appendix I

Figure AI 1. Revenue Weighted Distribution

Notes: The distribution is based on the 148,962 observations. We trim all right tail observation (revenue
weight > 0.000001).
Sources: ONS ABS dataset and authors’ own calculations.

Figure AI 2. Revenue Weighted Distribution

Notes: The distribution is based on the 148,962 observations. We trim all right tail observation (revenue
weight > 0.000001).
Sources: ONS ABS dataset and authors’ own calculations.
Figure AI 3. SIC2-SIC4 Revenue weighted Cumulative Markups Growth Overtime – Manufacturing
Sources: ONS ABS dataset and authors’ own calculations.